IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2

Dokumentation und Würdigung der Stellungnahmen zum Vorbericht

Auftrag A05-03
Version 1.0
Stand: 02.12.2008

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

22.02.2005

Interne Auftragsnummer:

A05-03

Anschrift des Herausgebers:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Dillenburger Str. 27
51105 Köln

Tel.: +49 221 35685-0
Fax: +49 221 35685-1
Berichte@iqwig.de
www.iqwig.de

Zu allen Dokumenten, auf die via Internet zugegriffen wurde und die entsprechend zitiert sind, ist das jeweilige Zugriffsdatum angegeben. Sofern diese Dokumente zukünftig nicht mehr über die genannte Zugriffsadresse verfügbar sein sollten, können sie im Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen eingesehen werden.

Inhaltsverzeichnis

Seite

1 Dokumentation des Stellungnahmeprozesses 1

2 Würdigung der Stellungnahmen..... 2

3 Darlegung potenzieller Interessenkonflikte der Stellungnehmenden 3

3.1 Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen 3

3.2 Interessenkonflikte von stellungnehmenden Privatpersonen 6

3.3 Interessenkonflikte von weiteren Teilnehmern an der wissenschaftlichen Erörterung (externen Sachverständigen, externen Reviewern) 6

4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll..... 9

4.1 Teilnehmerliste der wissenschaftlichen Erörterung 9

4.2 Tagesordnung der wissenschaftlichen Erörterung 10

4.3 Protokoll der wissenschaftlichen Erörterung..... 11

Anhang: Dokumentation der Stellungnahmen..... 82

1 Dokumentation des Stellungnahmeprozesses

Am 28.03.2008 wurde der Vorbericht in der Version 1.0 vom 18.03.2008 veröffentlicht und zur Anhörung gestellt. Bis zum 25.04.2008 konnten schriftliche Stellungnahmen eingereicht werden. Insgesamt wurden 13 Stellungnahmen form- und fristgerecht abgegeben. Diese Stellungnahmen sind im Anhang abgebildet.

Unklare Aspekte in den schriftlichen Stellungnahmen wurden in einer wissenschaftlichen Erörterung am 17.06.2008 im IQWiG diskutiert. Das Wortprotokoll der Erörterung befindet sich in Kapitel 4.

Eine Würdigung der in der Anhörung vorgebrachten Aspekte befindet sich im Kapitel „Diskussion“ des Abschlussberichts. Im Abschlussbericht sind darüber hinaus Änderungen, die sich durch die Anhörung ergeben haben, zusammenfassend dargestellt. Der Abschlussbericht ist unter www.iqwig.de veröffentlicht.

2 Würdigung der Stellungnahmen

Die einzelnen Stellungnahmen wurden hinsichtlich valider wissenschaftlicher Argumente für eine Änderung des Vorberichts überprüft. Die wesentlichen Argumente werden im Kapitel „Diskussion“ des Abschlussberichts gewürdigt. Neben projektspezifischen wissenschaftlichen Aspekten wurden auch übergeordnete Punkte, z. B. zur allgemeinen Institutsmethodik, zur Aufgabe des Instituts, zu rechtlichen Vorgaben für das Institut etc., angesprochen. Auf diese letztgenannten Punkte wurde im Rahmen der projektspezifischen Würdigung der Stellungnahmen nicht weiter eingegangen.

Der Abschlussbericht wurde unter www.iqwig.de veröffentlicht.

3 Darlegung potenzieller Interessenkonflikte der Stellungnehmenden

Im Folgenden sind die potenziellen Interessenkonflikte der Stellungnehmenden zusammenfassend dargestellt. Alle Informationen beruhen auf Selbstangabe der einzelnen Personen anhand des „Formblatts zur Darlegung potenzieller Interessenkonflikte“. Das Formblatt ist unter www.iqwig.de abrufbar. Die in diesem Formblatt aufgeführten Fragen finden sich im Anschluss an diese Zusammenfassung.

3.1 Interessenkonflikte von Stellungnehmenden aus Organisationen, Institutionen und Firmen

Organisation / Institution / Firma	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Bundesverband der Pharmazeutischen Industrie	Sickmüller, Barbara, Prof. Dr.	nein	nein	nein	nein	nein	nein	nein	nein
Bundesverband der Pharmazeutischen Industrie	Schnorpfeil, Willi, Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Bundesverband der Pharmazeutischen Industrie	Paar, Dieter, PD Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Bundesverband der niedergelassenen Diabetologen	Kusterer, Klaus, Prof. Dr. Dr.	nein	ja	ja	nein	nein	ja	nein	nein
Deutsche Diabetes-Gesellschaft (DDG)	Fritsche, Andreas, Prof. Dr.	nein	nein	ja	ja	nein	nein	nein	ja
Deutsche Diabetes-Gesellschaft (DDG)	Gallwitz, Baptist, Prof. Dr.	nein	ja	ja	ja	nein	nein	nein	nein
Deutsche Diabetes-Gesellschaft (DDG)	Klein, Harald, Prof. Dr.	nein	nein	ja	ja	ja	ja	nein	nein

Organisation / Institution / Firma	**Name**	**Frage 1**	**Frage 2**	**Frage 3**	**Frage 4**	**Frage 5**	**Frage 6**	**Frage 7**	**Frage 8**
Deutsche Diabetes-Gesellschaft (DDG)	Tschritter, Otto, Dr.	nein	nein	nein	nein	nein	nein	nein	ja
Deutscher Diabetiker Bund (DDB)	von Lilienfeld-Toal, Hermann, Prof. Dr.	nein	nein	ja	nein	nein	nein	nein	ja
Deutscher Diabetiker Bund (DDB)	Möhler, Dieter	nein	nein	nein	nein	nein	nein	nein	nein
Deutscher Diabetiker Bund (DDB)	Windisch, Heinz	nein	nein	nein	nein	nein	nein	nein	nein
Hoechst GmbH	Kurrle, Roland, Dr.	ja	nein	nein	nein	nein	ja	ja	nein
Hoechst GmbH	Landgraf, Wolfgang, Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Hoechst GmbH	Sandow, Jürgen, Prof. Dr.	ja	ja	ja	nein	nein	nein	nein	nein
Novo Nordisk Pharma GmbH	El-Haschimi, Karim, Dr.	ja	nicht bekannt	nein	nein	nein	nein	nein	nein
Novo Nordisk Pharma GmbH	Grundner, Martin, Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Novo Nordisk Pharma GmbH	Koenen, Christoph, Dr.	ja	nein	nein	nein	nein	ja	ja	nein
Sanofi-Aventis Deutschland GmbH	Frick, Markus, Dr.	ja	nein	nein	nein	nein	nein	nein	nein
Sanofi-Aventis Deutschland GmbH	Hessel, Franz, Dr.	ja	nein	ja	ja	nein	ja	nein	nein

Organisation / Institution / Firma	**Name**	**Frage 1**	**Frage 2**	**Frage 3**	**Frage 4**	**Frage 5**	**Frage 6**	**Frage 7**	**Frage 8**
Sanofi-Aventis Deutschland GmbH	Knollmeyer, Johannes, Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Verband Forschender Arzneimittelhersteller	Stammberger, Ingo, Dr.	ja	nein	nein	nein	nein	ja	nein	nein
Verband Forschender Arzneimittelhersteller	Thole, Henning	ja	ja	ja	nein	nein	nein	nein	ja
Verband Forschender Arzneimittelhersteller	Wahler, Steffen, Dr.	ja	ja	ja	ja	ja	ja	ja	nein

3.2 Interessenkonflikte von stellungnehmenden Privatpersonen

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Bradley, Clare, Prof.	ja	ja	ja	ja	nein	ja	ja	ja
Bramlage, Peter, PD Dr.	ja	ja	ja	nein	nein	nein	ja	nein
Bretzel, Reinhard, Prof. Dr.	nein	ja	ja	ja	nein	nein	nein	nein
Chantelau, Ernst, Prof. Dr.	nein	nein	nein	nein	nein	nein	nein	nein
Debusmann, Edeltraud und Horst	nein	nein	nein	nein	nein	nein	nein	nein

3.3 Interessenkonflikte von weiteren Teilnehmern an der wissenschaftlichen Erörterung (externen Sachverständigen, externen Reviewern)

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Richter, Bernd, Priv.-Doz. Dr.	nein	nein	nein	nein	nein	nein	nein	nein

Im „Formblatt zur Darlegung potenzieller Interessenkonflikte“ wurden folgende 8 Fragen gestellt:

Frage 1: Sind oder waren Sie innerhalb der letzten 3 Jahre bei einer Person, Institution oder Firma[1] abhängig (angestellt) beschäftigt, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut[2] finanziell profitieren könnte?

Frage 2: Beraten Sie oder haben Sie innerhalb der letzten 3 Jahre eine Person, Institution oder Firma direkt oder indirekt[3] beraten, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 3: Haben Sie abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre im Auftrag einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, Honorare für Vorträge, Stellungnahmen, Ausrichtung und / oder Teilnahme an Kongressen und Seminaren – auch im Rahmen von Fortbildungsveranstaltungen – oder für (populär-) wissenschaftliche oder sonstige Aussagen oder Artikel erhalten?[4]

Frage 4: Haben Sie und/oder die Einrichtung, für die Sie tätig sind, abseits einer Anstellung oder Beratungstätigkeit innerhalb der letzten 3 Jahre von einer Person, Institution oder Firma, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte, finanzielle Unterstützung für Forschungsaktivitäten, andere wissenschaftliche Leistungen oder Patentanmeldungen erhalten?

Frage 5: Haben Sie und / oder die Einrichtung, bei der Sie angestellt bzw. beschäftigt sind, innerhalb der letzten 3 Jahre sonstige finanzielle oder geldwerte Zuwendungen (z. B. Ausrüstung, Personal, Reisekostenunterstützung ohne wissenschaftliche Gegenleistung) von einer Person, Institution oder Firma erhalten, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

---

[1] Mit solchen „Personen, Institutionen oder Firmen“ sind im Folgenden alle Einrichtungen gemeint, die direkt oder indirekt einen finanziellen oder geldwerten Vorteil aus dem Ergebnis Ihrer wissenschaftlichen Arbeit für das Institut ziehen könnten. Hierzu gehören z. B. auch medizinische Einrichtungen, die eine zu bewertende medizinische Intervention durchführen und hierdurch Einkünfte erhalten.

[2] Mit „wissenschaftlicher Arbeit für das Institut“ sind im Folgenden alle von Ihnen für das Institut erbrachten oder zu erbringenden Leistungen und / oder an das Institut gerichteten mündlichen und schriftlichen Recherchen, Bewertungen, Berichte und Stellungnahmen gemeint.

[3] „Indirekt“ heißt in diesem Zusammenhang z. B. im Auftrag eines Institutes, das wiederum für eine entsprechende Person, Institution oder Firma tätig wird.

[4] Sofern Sie von einer Person, Institution oder Firma im Verlauf der letzten 3 Jahre mehrfach Honorare erhalten haben reicht es aus, diese für die jeweilige Art der Tätigkeit summiert anzugeben.

Frage 6: Besitzen Sie Aktien, Optionsscheine oder sonstige Geschäftsanteile (auch in Fonds) von einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 7: Haben Sie persönliche Beziehungen zu einer Person, Firma oder Institution bzw. Mitarbeitern einer Firma oder Institution, welche von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut finanziell profitieren könnte?

Frage 8: Gibt es andere bislang nicht dargestellte potenzielle Interessenkonflikte, die in Beziehung zu Ihrer wissenschaftlichen Arbeit für das Institut stehen könnten?

4 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

4.1 Teilnehmerliste der wissenschaftlichen Erörterung

Name	Organisation / Institution / Firma
Ebrahim, Susanne, Dipl.-Psych.	IQWiG
Filla, Christoph	Protokollant
Frick, Markus, Dr.	Sanofi-Aventis Deutschland GmbH
Fritsche, Andreas, Prof. Dr.	Deutsche Diabetes-Gesellschaft, Pharmakotherapieausschuss
Gallwitz, Baptist, Prof. Dr.	Deutsche Diabetes-Gesellschaft, Pharmakotherapieausschuss
Grundner, Martin, Dr.	Novo Nordisk Pharma GmbH
Jost, Marco, Dr.	IQWiG
Kaiser, Thomas, Dr.	IQWiG
Knollmeyer, Johannes, Dr.	Sanofi-Aventis Deutschland GmbH
Koenen, Christoph, Dr.	Novo Nordisk Pharma GmbH
Kurrle, Roland, Dr.	Hoechst GmbH
von Lilienfeld-Toal, Hermann, Prof. Dr.	Deutscher Diabetiker Bund e. V.
Paar, Dieter, Priv.-Doz. Dr.	Bundesverband der Pharmazeutischen Industrie e. V.
Richter, Bernd, Priv.-Doz. Dr.	Cochrane Metabolic and Endocrine Disorders Review Group, Medizinische Einrichtungen der Heinrich-Heine-Universität, Düsseldorf
Sandow, Jürgen, Prof. Dr.	Hoechst GmbH
Sawicki, Peter, Prof. Dr.	IQWiG (Moderation)
Schnorpfeil, Willi, Dr.	Bundesverband der Pharmazeutischen Industrie e. V.
Stammberger, Ingo, Dr.	Verband Forschender Arzneimittelhersteller e. V.
Sturtz, Sibylle, Dr.	IQWiG
Thole, Henning	Verband Forschender Arzneimittelhersteller e. V.
Vervölgyi, Volker, Dr.	IQWiG

4.2 Tagesordnung der wissenschaftlichen Erörterung

	Begrüßung und Einleitung
TOP 1	Zusätzliche relevante Studien/Zusätzliche Daten zu relevanten Studien
TOP 2	Unterschiede in der praktischen Anwendung zwischen langwirksamen Insulinanaloga und NPH-Insulin
TOP 3	Nicht schwere Hypoglykämien
TOP 4	Gemeinsame Betrachtung von Hypoglykämien und dem HbA1c-Wert
TOP 5	Sonstige Aspekte der Studienbewertung
TOP 5.1	Bewertung fehlender Werte bei Patient-reported Outcomes
TOP 5.2	Umgang mit fehlender Evidenz
TOP 5.3	Bewertung der morgendlichen Gabe von Insulin Glargin
TOP 6	Verschiedenes

4.3 Protokoll der wissenschaftlichen Erörterung

Moderator Sawicki: Guten Morgen! Ich darf Sie herzlich begrüßen. Die allermeisten Gesichter kommen mir bekannt vor. Daher kann ich die allgemeine Klärung kurzhalten. Wir werden heute die Stellungnahmen zum Vorbericht A05-03 „Langwirksame Insulinanaloga zur Behandlung von Diabetes mellitus Typ 2“ erörtern. Diese Erörterung ist mit vier Stunden einschließlich Pause, die wir gegen 12 Uhr machen werden, angesetzt. Wahrscheinlich werden wir kürzer tagen. Das heißt aber nicht, dass es nichts zu essen gibt. Sie können trotzdem gestärkt nach Hause fahren.

Zum Ablauf. Sie sehen hier diese Aufzeichnungsgeräte. Wir zeichnen auf, was hier gesagt wird. Deswegen ist es auch wichtig, dass Sie in der Hitze der Diskussion nicht vergessen, Ihren Namen zu nennen, bevor Sie etwas sagen, damit die Aussagen auch den richtigen Namen zugeordnet werden können. Wir machen daraus ein Wortprotokoll. Dieses wird zusammen mit den übrigen Dokumenten zum Abschlussbericht publiziert werden. Eigene Aufzeichnungen sind aufgrund der bekannten Probleme, die damit verbunden sind, nicht gestattet.

Haben Sie zur Gestaltung dieser Erörterung Fragen? – Das ist nicht der Fall. Dann können wir direkt in die Tagesordnung einsteigen.

Ich rufe auf:

Tagesordnungspunkt 1:
Zusätzliche relevante Studien /
Zusätzliche Daten zu relevanten Studien

Herr Vervölgyi, was möchten wir dazu erörtern?

Herr Vervölgyi: Zum ersten Aspekt der zusätzlichen relevanten Studien beziehe ich mich auf eine Stellungnahme vom VFA. Sie haben genannt, dass wir Studien, die in anderen systematischen Übersichten berücksichtigt wurden, nicht eingeschlossen haben. Sie haben jedoch keine systematischen Übersichten zitiert. Deswegen wollten wir gerne fragen, ob wir Ihrer Meinung nach andere systematische Übersichten, die zusätzliche relevante Studien enthalten, nicht identifiziert haben.

Herr Thole: Wir haben uns in der Stellungnahme auf die nicht in dem Bericht enthaltene Studie von Yokoyama bezogen.

Moderator Sawicki: Können Sie ein genaues Zitat angeben?

Herr Kaiser: Das ist bekannt. Das ist im Bericht enthalten.

Herr Thole: Sie ist nicht drin. Soweit ich es gesehen habe, verwenden Sie die Studie inhaltlich nicht.

Zu den Meta-Analysen, die hier im Raum stehen. Ich denke, diese müsste man diskutieren. Ich habe dazu nichts gefunden. Ich konnte mich auf die Frage nicht konkreter vorbereiten, weil ich vorher keine Tagesordnung und Frage hatte. Ich werde es aber gerne nachreichen, wenn es möglich ist.

Moderator Sawicki: Sicher.

Herr Thole: Ich habe dazu gleich eine präzisierende Frage: Wie ist es grundsätzlich mit der Nachreichmöglichkeit von Daten in diesem Verfahren, also weiteren Analysen, weiteren Studien, die wir eventuell haben? Wie ist das vorgesehen?

Herr Kaiser: Grundsätzlich bestand die Möglichkeit, im Rahmen des Stellungnahmeverfahrens alles einzureichen. Es gibt also keinen Grund, nachträglich irgendwelche Dinge einzureichen. In den Ausnahmefällen, in denen beispielsweise auf Studien hingewiesen wurde, die gerade eben abgeschlossen wurden oder zu denen gerade eben der Studienbericht fertiggestellt wurde, kann man eine Ausnahme machen. Das bleibt aber die einzige Ausnahme.

Herr Thole: Inwiefern tangiert das die Regelungen in den aktuellen Methoden 3.0, dass man Daten nachreichen kann, was in den Methoden 2.0 meiner Kenntnis nach nicht vorgesehen war?

Herr Kaiser: Ich weiß nicht genau, worauf Sie hinauswollen. Allerdings denke ich, dass man das allgemein diskutieren muss und dass es keinen Sinn macht, es hier für das Projekt zu diskutieren.

Herr Thole: Die Frage hat ja potenzielle Relevanz für das Projekt. Meine Frage wäre, ob man die Frage öffnet. Sie war an uns gerichtet. Ich denke, dass hier genügend Vertreter sitzen, die Bedarf haben, über potenziell nachzureichende Studien zu diskutieren. Ich komme auf die Frage, weil meiner Kenntnis nach in den Methoden 3 explizit gesagt wird, dass in den entsprechenden Anhörungsverfahren Daten nachgereicht werden können. Ich sehe hier das Problem – falls ich mich richtig erinnere –, dass die Frist für die Kommentierung im April ablief, also für diese Anhörung. Die neuen Methoden sind einen Monat danach entstanden. Wenn jetzt die neuen Methoden 3 hier gelten, dann hätten wir einen Verfahrensablauf gehabt, in dem dieses explizite Nachreichen von Studien innerhalb der Anhörung nicht vorgesehen war.

Daraus ergibt sich die Frage – Sie haben mir eben gesagt, dass Meta-Analysen zur Diskussion stehen, die möglicherweise zu berücksichtigen sind –, wie man mit diesen in diesem

Verfahren konkret umgeht. Wie wird also eine Nachfrist zur Nacheinreichung von Daten gesetzt?

Herr Kaiser: Die Darstellung in den Methoden ist lediglich eine Konkretisierung unseres bisherigen Verfahrens. Denn Bestand der Methoden waren immer auch die Leitfäden zu den Stellungnahmeverfahren, wie Sie wissen. Dort war in dem gültigen Formblatt der Leitfaden zum Stellungnahmeverfahren erwähnt, dass Sie also Studien einreichen können. Insofern haben die neuen Methoden damit überhaupt nichts zu tun. Sie haben das Formblatt auch unterzeichnet. Das heißt, Sie haben den Leitfaden zur Kenntnis genommen. Das heißt, Sie hatten die Möglichkeit, die Dinge einzureichen. Die Methoden ändern daran nichts.

Herr Thole: Eine Frage zur Präzisierung: Wie ich die Methoden 3.0 verstehe, werden ausdrücklich auch Nicht-RCT-Daten zur Nacheinreichung zugelassen.

Herr Kaiser: Das war auch vorher so. Es ist wie gesagt nur eine Konkretisierung, weil es offensichtlich ein Missverständnis war. Wenn Sie sich aber verschiedenste Abschlussberichte anschauen, werden Sie sehen, dass wir auch die Studien diskutiert haben, die nicht den Einschlusskriterien entsprochen haben, und dass wir geschaut haben, ob sich daraus nicht doch ein Nutzen ableiten lässt. Diese muss man aber diskutieren. Das ist bisher so gewesen. Also, die jetzige Konkretisierung in den Methoden ändert daran nichts. Sie hat deswegen auch keinen Einfluss auf dieses Projekt.

Herr Thole: Wie ist es dann mit der Yokoyama-Studie konkret in diesem Fall?

Herr Kaiser: Die Diskussion haben Sie in dem Vorbericht gelesen. Nur weil es die Studie gibt, heißt das nicht, dass sie einen Nutzen zeigt. Man muss diskutieren, ob sie in der Lage ist, die Fragestellung zu beantworten. Ich gehe davon aus: Wenn Sie konkret anderer Ansicht als die Yokoyama-Studie sind, dann haben Sie das in Ihrer schriftlichen Stellungnahme geäußert. Deswegen werden wir uns auf der Basis Ihrer Stellungnahme und natürlich auf der Basis der Stellungnahmen der anderen Vertreter anschauen, ob valide Argumente genannt werden, um mit der Yokoyama-Studie anders umzugehen. Das hat aber mit den Methoden gar nichts zu tun.

Moderator Sawicki: Herr Thole, hat das zur Klärung beigetragen?

Herr Thole: Nicht endgültig. Nein.

Moderator Sawicki: Was ist denn unklar geblieben?

Herr Thole: Der Punkt ist, ich sehe es nicht so, dass das Nacheinreichen von Studien da ist. Und die Tatsache, dass die Yokoyama-Studie als nicht aufgabenpflichtig bezeichnet wird, kann ich weder aufgrund der Ausführungen des Vorberichts noch jetzt nachvollziehen.

Moderator Sawicki: Das habe ich jetzt nicht verstanden.

Herr Thole: Ich sehe auch einen Widerspruch, weil Sie in andere Arbeiten …

Moderator Sawicki: Also, die Möglichkeit, Studien nachzureichen oder einzureichen aufgrund der Lektüre des Vorberichts, ist doch gegeben. Im Rahmen des Stellungnahmeverfahrens können Sie doch alle Studien einreichen.

Herr Thole: Wie Sie das gerade erklärt haben, ja.

Moderator Sawicki: Das war doch bisher auch schon möglich.

Herr Thole: Ich sehe einen Unterschied in der Fokussierung der Berichtspläne auf RCTs und auf der Nennung in den Methoden 3.0 darin, dass es auf alle Studiendaten eröffnet wird. Ich sehe einen qualitativen Unterschied, der in den Methoden 3.0 neu ist und der in der Anhörung für dieses Verfahren nicht berücksichtigt worden ist, weil zum Zeitpunkt der Anhörung die Methoden 2.0 gültig waren. Da wurden eine andere Formulierung und Darstellung gewählt. Das heißt, es ist in den Methoden der Umgang mit nachzureichenden Studien anders dargestellt, sodass sich für mich der Gehalt tatsächlich geändert hat.

Der Punkt ist der, dass die Yokoyama-Studie – das hatten wir in der Stellungnahme auch geschrieben – im Bericht nicht berücksichtigt wird, obwohl sie die Einschlusskriterien erfüllt. Diese Frage hat für mich nichts mit dem Nacheinreichen von Studien zu tun, sondern mit der Berücksichtigung einer vorliegenden Studie.

Herr Kaiser: Es ist sicherlich richtig, dass es zwei getrennte Felder sind. Sie haben nur die Methoden angesprochen. Deshalb habe ich auf Ihre Frage geantwortet.

Noch einmal: Das ist Ihre Auffassung. Wenn Sie sich die Abschlussberichte von anderen Berichten anschauen, dann sehen Sie, dass Ihre Auffassung so nicht richtig ist. Wir haben vielmehr auch vorher schon nicht randomisierte Studien diskutiert, wenn sie denn eingereicht worden sind.

Das Zweite zur Yokoyama-Studie. Sie erfüllt die Einschlusskriterien eben nicht. Sie erfüllt einige der Einschlusskriterien. Sie erfüllt aber gerade das Einschlusskriterium, dass es vergleichbare Blutzucker senkende Behandlungen sein müssen, nicht. Das ist im Vorbericht dargestellt.

Nichtsdestotrotz: Wir werden uns mit den Argumenten, die Sie beschrieben haben, noch einmal auseinandersetzen. Es ist natürlich Ihr Recht, in den Stellungnahmen darauf hinzuweisen, dass es aus Ihrer Sicht so nicht zu interpretieren ist, und mit den Stellungnahmen werden wir uns auseinandersetzen. Dafür ist das Stellungnahmeverfahren ja da.

Herr Knollmeyer: Es ergeben sich zwei ergänzende Punkte. – Wir hatten das Thema auch angesprochen. Zum einen verstößt die Yokoyama-Studie unserer Auffassung nach gegen

keines der Ausschlusskriterien. Sie erfüllt die Einschlusskriterien vollständig, sodass wir den Ausschluss dieser Studie für ungerechtfertigt halten.

Zum anderen hat sie Einfluss auf die Darstellung der Landkarte der Beleglage für langwirksame Insulinanaloga. Das ist eine Studie mit intensivierter Insulintherapie wie unsere Studie 3006 auch, die Sie auf dieser Landkarte als solche nicht ausgewiesen haben.

Moderator Sawicki: Danke schön.

Herr Knollmeyer: Gibt es eine Antwort auf die Frage?

Herr Kaiser: Ich kann gerne noch einmal darauf antworten. – Wir werden uns die Argumente anschauen und dann im Bericht entsprechend sehen, ob sich daran etwas ändert. Es kommt auf die Validität der Argumente an.

Herr Knollmeyer: Ich hätte gerne konkrete Auskunft. Sie haben sich unsere Stellungnahme angesehen. Sie haben unsere Argumente gelesen. Ich hätte gerne eine Einschätzung von Ihnen, wie Sie das zukünftig darstellen möchten.

Herr Kaiser: Gegenstand der Erörterung ist die Klärung unklarer Fragen. Für uns hat sich aus Ihrer Stellungnahme dazu nichts Unklares ergeben. Das Ergebnis wird dann im Abschlussbericht dargestellt werden.

Herr Knollmeyer: Für mich ist eine wissenschaftliche Erörterung immer bidirektional. Es kann sich nicht nur darum handeln, dass wir Ihre Fragen beantworten. Vielmehr muss Gegenstand einer Anhörung auch sein, dass die übrigen Teilnehmer Fragen an Sie richten, auf die wir dann eine Antwort erwarten.

Moderator Sawicki: Danke schön für Ihre Einschätzung. – Weitere Punkte zu diesem Aspekt?

Herr Paar: Erst einmal möchte ich ausdrücklich unterstreichen, was Herr Knollmeyer sagt. Ich finde es schade, wenn wir es nicht erörtern, sondern wenn Sie sagen, dass Sie die Sachen, die wir eingereicht haben, zur Kenntnis nehmen müssen. Ich fände es ausgesprochen günstiger, wenn man es wirklich erörtern könnte; so ist auch der Titel dieser Anhörung. Und ich würde auch darum bitten, dass Sie den Ausschluss der Meta-Analysen von Mullins und Rosenstock zurücknähmen. Wir haben auch dazu Argumente vorgetragen. Ich gehe davon aus – wenn ich die Spielregeln gerade richtig verstanden habe –, dass wir das jetzt auch nicht erörtern werden. Ich möchte trotzdem zu Protokoll geben, dass wir dafür plädieren, diese Studien nicht komplett auszuschließen. Sie haben sicherlich Gründe, sie unter Umständen eingeschränkt zu beurteilen. Das kann ich zwar nachvollziehen, aber ich finde es sehr bedauerlich, die Evidenz aus diesen Studien komplett auszublenden.

Herr Knollmeyer: Ich möchte noch einen Punkt zur Yokoyama-Studie und zum Nichteinschluss ansprechen. Mitarbeiter des Instituts sind ja an der Abfassung einer Cochrane-Analyse beteiligt gewesen. Im Rahmen dieser Cochrane-Analyse, die sehr ähnliche Ein- und Ausschlusskriterien wie die IQWiG-Nutzenbewertung aufweist, ist die Yokoyama-Studie berücksichtigt worden. Können Sie mir diese Diskrepanz aufklären?

Moderator Sawicki: Hier in diesem Rahmen?

Herr Knollmeyer: Gerne, ja.

Moderator Sawicki: Nein. – Herr Gallwitz.

Herr Gallwitz: Ich wollte auch auf den Punkt Cochrane-Report eingehen. Da ist es so, dass die Nutzenbewertung der langwirkenden Insulinanaloga zu einem anderen Schluss kommt als der IQWiG-Bericht. Die Cochrane-Analyse ist in dem Vorbericht überhaupt nicht berücksichtigt, obwohl im Cochrane-Bericht das IQWiG als externer Supporter genannt wird. Können Sie das kurz kommentieren?

Moderator Sawicki: Die Cochrane-Autoren, die Sie ansprechen, haben natürlich als externe Sachverständige hieran mitgearbeitet. Das ist die Überschneidung. Dass es aufgrund von unterschiedlichen wissenschaftlichen Prozessen und Stellungnahmen der Institute zu teilweise nicht hundertprozentig kongruenten Aussagen kommt, ist üblich. Das ist ja nichts Besonderes. Was wir jetzt hier machen, ist eine Evaluation nach den Methoden des Instituts.

Herr Paar: Wenn wir also die Gründe erörtert bekämen, warum zum Beispiel die primären und sekundären Zielkriterien Chochrane vs. IQWiG-Assessment unterschiedlich sind, müssten wir noch einmal schreiben. Wir werden es heute nicht besprechen; so habe ich es jetzt verstanden. Oder wie gehen wir jetzt weiter vor?

Moderator Sawicki: Wir werden es aufnehmen. Das, was Sie jetzt sagen, wird ja festgehalten. Erstens haben wir es gehört. Zweitens wird es im Protokoll festgehalten. Wir werden darauf selbstverständlich in der Diskussion eingehen.

Herr Paar: Sie werden es heute also nicht kommentieren, aber Sie werden es diskutieren in der …

Moderator Sawicki: Genau. Der Prozess des Institutes ist – vielleicht kann ich es noch einmal erklären –, dass Sie die Möglichkeit haben, Stellungnahmen einzureichen. Das haben Sie getan. Diese Stellungnahmen gucken wir uns an. Wir diskutieren jetzt unklare Aspekte dieser Stellungnahmen, und dann wird mit dem Abschlussbericht eine Würdigung dieser Stellungnahmen publiziert werden. Und wenn es wichtig ist, auf solche Aspekte, die Sie erwähnt haben, einzugehen, dann werden wir es im Rahmen der Diskussion tun.

Gibt es zu Tagesordnungspunkt 1 weitere Aspekte, die wir gerne geklärt haben möchten?

Herr Vervölgyi: Hinsichtlich des zweiten Aspekts „Zusätzliche Daten zu relevanten Studien“ beziehe ich mich auf die Stellungnahme der Firma Sanofi-Aventis. Sie haben uns zusätzliche Daten geliefert, um die Tabellen in dem Vorbericht aufzufüllen. Da habe ich einige Fragen, die ich gerne mit Ihnen durchgehen möchte.

Die erste bezieht sich auf die Hospitalisierung. In der Überschrift in der Stellungnahme steht, dass jeweils die Anzahl der Ereignisse angegeben ist. Ich denke, da beziehen Sie sich auf Tabelle 30 des Vorberichts, in der schwerwiegende unerwünschte Ereignisse zur Hospitalisierung geführt haben. Bei uns im Vorbericht sind es die Patienten mit Ereignissen. Deswegen wollte ich sicherheitshalber nachfragen, ob es sich tatsächlich um die Anzahl der Ereignisse handelt, die Sie angegeben haben.

Herr Knollmeyer: Also, bei der Hospitalisierung ist die Basis die Anzahl der Ereignisse, und das geht aus der Überschrift hervor. Üblich ist hier eine 1:1-Verteilung. Die Anzahl der Patienten ist auch die Anzahl der Ereignisse. Falls Sie dieses zusätzliche Detail interessieren sollte, könnten wir das separat recherchieren und die Angaben aus der Datenbank holen.

Herr Vervölgyi: Vielen Dank. – Die zweite Frage bezieht sich auf die Angaben zur Tabelle 40 zu schweren nächtlichen Hypoglykämien. Da sagen Sie zur Studie 3102, dass sich weder in der Glargin-Gruppe noch in der NPH-Gruppe schwere nächtliche Hypoglykämien ereignet hätten. Im Studienbericht steht aber, dass unter Glargin zwei Patienten eine schwere nächtliche Hypoglykämie gehabt hätten. Diese beiden Patienten erfüllen die Definition für eine nächtliche Hypoglykämie. Das ist eine Diskrepanz, die wir nicht ganz verstehen.

Herr Knollmeyer: Das muss ich nachrecherchieren. Das liefern wir Ihnen nach.

Herr Vervölgyi: Das Dritte: Zur Studie 4001 haben Sie sowohl für die schweren nächtlichen Hypoglykämien als auch für die nächtlichen Hypoglykämien gesamt jeweils nur einen Wert für die Glargin-Gruppe. Es waren ja zwei Glargin-Arme in der Studie 4001. Aber Sie haben jeweils nur einen Wert angegeben. Auch das möchten wir gerne erfragen. Gilt das für beide gleichermaßen, oder ist es eine Zusammenfassung der beiden Arme?

Herr Knollmeyer: Das muss ich auch prüfen. Eine Antwort werden Sie zeitnah kriegen.

Herr Vervölgyi: Als Letztes zur Tabelle „Hypoglykämien insgesamt“. Da geben Sie für die Studie 4012 null Ereignisse pro Patientenjahr jeweils für die Glargin- und NPH-Gruppe an. Aber es ergeben sich 682 zu 1019 Ereignisse insgesamt. Also, die Zahlen passen nicht ganz zusammen. Das wäre der letzte Punkt.

Herr Knollmeyer: Das sind die Daten, die ich von der Biometrie auf Nachfrage erhalten habe. Wir können das noch einmal klären. Ganz offenkundig – das wäre meine einzige Erklärungsmöglichkeit – ist es, dass hier zwar leichte, aber überhaupt keine schweren Hypoglykämien stattgefunden haben. Wir überprüfen das aber sicherheitshalber noch einmal.

Herr Vervölgyi: Vielen Dank.

Herr Kaiser: Allerdings bezieht sich die Zahl in Ihrer Tabelle auf Hypoglykämien insgesamt und nicht nur auf schwere. Das ist ja die Diskrepanz, die entstanden ist.

Herr Knollmeyer: Dies ist die Tabelle „Nächtliche schwere Hypoglykämien".

Herr Vervölgyi: Ihrer Stellungnahme ist eher zu entnehmen, dass Sie sich selber auf Tabelle 46 „Hypoglykämien insgesamt" beziehen. Deswegen gibt es die Diskrepanz.

Herr Knollmeyer: Tabelle 46 ist darunter. Da ist die Ergänzung zur Studie 4012.

Herr Vervölgyi: Genau. Und da stehen null gegen null Ereignisse pro Patientenjahr.

Herr Kaiser: Und das passt nicht zur Ereignisrate.

Herr Knollmeyer: Wie gesagt: Das waren die Angaben, die wir aus dem Datenbestand erhalten haben.

Moderator Sawicki: Zum Tagesordnungspunkt 1 gibt es keine weiteren Wortmeldungen. Damit kommen wir zu Tagesordnungspunkt 2:

Tagesordnungspunkt 2:
Unterschiede in der praktischen Anwendung
zwischen langwirksamen Insulinanaloga und NPH-Insulin

Was sind da die wesentlichen Aspekte?

Herr Vervölgyi: Da hätten wir eine Frage an die Deutsche Diabetes-Gesellschaft. Sie haben in Ihrer Stellungnahme auf die Unterschiede in der praktischen Anwendung zwischen den Analoga und NPH-Insulin verwiesen und sagen, dass die Vorteile der Analoga – sprich die geringe Injektionsfrequenz, dass keine Resuspendierung nötig ist und dass die freie Wahl des Injektionszeitpunktes besteht – in klinischen Studien schwer zu überprüfen seien. Unsere Frage zielt dahin, wo Sie die spezifischen Barrieren für diese Überprüfbarkeit sehen. Wo sehen Sie die Probleme, um das in einer klinischen Studie abzubilden?

Herr Fritsche: Ich sehe keine Notwendigkeit, in einer klinischen Studie zu prüfen, ob eine Injektion einmal pro 24 Stunden oder zweimal pro 24 Stunden durchgeführt wird. Das ist ein Faktum. Die einmalige Injektion ist ein Vorteil für den Patienten in der Anwendbarkeit. Das weiß man auch aus vielen anderen Darreichungsformen wie Tabletten. Eine einmalige Gabe beinhaltet die bessere Compliance. Auch hier zeigt die klinische Erfahrung, dass die Compliance mit weniger Injektionen besser ist und dass die Injektion weniger häufig vergessen wird.

Das Resuspendieren, also das Aufschütteln des Insulins, ist eine Sache, die in klinischen Studien schwierig zu prüfen ist. Da gibt es eben aus Ulm diese Untersuchung, die gezeigt hat, dass es bei NPH-Insulin ein 20-faches oder 25-faches Schütteln erfordert, bis die Suspension wirklich gleichmäßig durchmischt ist. Es ist äußerst fraglich, ob es die Patienten in der Praxis immer machen. Ich kann mir auch kein RCT vorstellen, bei dem über Monate gemessen wird, ob die Patienten tatsächlich dieses Schütteln in der praktischen Anwendung vornehmen.

Moderator Sawicki: Nein, aber es geht um die Auswirkungen des Schüttelns. Es muss doch eine Bedeutung haben, dass man schüttelt. Was ist denn die Bedeutung?

Herr Fritsche: Es geht um die Hypoglykämierate, die unterschiedliche Wirkung des Insulins, die unterschiedlichen Peak-Werte, die unterschiedlichen Wirkprofile, die unterschiedliche Pharmakokinetik. Darauf kommen wir im Bereich Hypoglykämie sicherlich zu sprechen. Das ist natürlich in den RCT Trials erfolgt.

Aber die Tatsache, ob ich ein Medikament in der Anwendbarkeit auspacken muss, ist ein Wert an sich, den ich nicht zu prüfen brauche. Das ist ein Vorteil, der auf der Hand liegt.

Herr Lilienfeld-Toal: Vonseiten des Betroffenenverbandes möchte ich das außerordentlich unterstreichen. Das ist eine intensiv diskutierte Tatsache, dass eine Injektion am Tag für den Betroffenen nicht so dramatisch ist wie zwei. Zwei sind dramatischer als eine.

Moderator Sawicki: Es war auch ein Argument, mit dem die langwirksamen Analoga vor dem Zweiten Weltkrieg eingeführt wurden. Man müsste dabei die Patientenzufriedenheit messen. Das könnte man in randomisierten kontrollierten Studien durchaus tun. Man könnte diese beiden Präparationen – Suspension so oder anders – bezüglich der Patientenzufriedenheit miteinander vergleichen. Man könnte auch zweifache Gabe vs. einfache Gabe miteinander vergleichen. Vielleicht ist die Anpassung für die Patienten besser, wenn sie zweimal spritzen. Vielleicht sind sie flexibler. Die einmalige Gabe pro Tag hat man vor dem Zweiten Weltkrieg gemacht. Da hat man die intensivierte Insulintherapie verlassen und auch die Arbeiten von Stolte quasi nicht berücksichtigt, dass die Patientenzufriedenheit mit einer einmaligen Gabe schlechter gewesen ist. Denn die Patienten, die flexibler waren, mussten die Diät darauf anpassen. Nur die Injektionen zu zählen, ist ein bisschen zu kurz gesprochen.

Herr Kaiser: Unsere Frage zielte ja darauf ab, welche Barrieren es für solche klinischen Studien gibt, um das genau zu untersuchen. Also, es erscheint zunächst einmal plausibel, dass es aufgrund des Unterschieds in den beiden verschiedenen Therapiealternativen zu einem Vorteil für den Patienten kommt. Aber es gibt viele Sachen, die zwar plausibel sind, sich aber nicht in einen echten Therapievorteil übersetzen.

Kurzer Punkt noch zu der einmaligen Gabe von Tabletten: Da gibt es aktuell etwas zu den Antihypertensiva, bei denen es sich doch nicht so deutlich zeigt. Also, es gibt verschiedene Evidenz dazu, ob es wirklich so ist mit einmaliger und zweimaliger Gabe.

Unsere Frage zielte darauf ab: Gibt es, wie Sie es in Ihrer Stellungnahme beschrieben haben, Barrieren, um das zu testen? – Diese Barrieren sehen wir so nicht. Unseres Erachtens kann man natürlich eine randomisierte Studie machen, in der man ein Präparat, das einmal am Tag gegeben wird, einmal gibt und ein Präparat, das eher häufiger am Tag – wie oft auch immer – gegeben wird, häufiger gibt. Ich denke, das ist keine Barriere, um einen solchen Vergleich in einer randomisierten Studie durchzuführen. Sollte sich durch die unterschiedliche Einfachheit der Gebräuchlichkeit – vielleicht kommt auch noch der Pen als ein möglicher Punkt mit dazu – dieses Medikaments ein Unterschied ergeben, dann würde man es in der randomisierten Studie doch auch sehen. Man muss es nicht anhand der Suspension prüfen. Führt es, wie Sie es beschrieben haben, dazu, dass tatsächlich weniger Hypoglykämien auftreten? – Denn es geht doch um das Ziel, das man beim Patienten erreichen will. Da sehe ich die Barriere nicht.

Herr Knollmeyer: Herr Prof. Sawicki, ich möchte noch mal das Beispiel aus der Vorkriegszeit aufgreifen, das Sie erwähnten. Die können wir mit der heutigen Zeit – hier reden wir ja über den Diabetes mellitus Typ 2 – überhaupt nicht vergleichen. Die langwirksamen Insulinanaloga in der einmal täglichen Gabe haben ihren Schwerpunkt in der gleichzeitigen Verordnung mit oralen Antidiabetika, die es nachweislich in der Ära des Zweiten Weltkriegs und davor gar nicht gab. Insofern reden wir über ein bestimmtes Patientenkollektiv, das mit einem bestimmten Behandlungskonzept behandelt wird. Dabei sehen wir, dass sich dieses Behandlungskonzept einer weiten Verbreitung erfreut, weil es für den Patienten einfach durchzuführen ist und weil es zu guten klinischen Ergebnissen führt. Insofern bitte ich darum, zu differenzieren, um welche Patientenkollektive es sich handelt und um welche Behandlungskonzepte es sich letztendlich in den klinischen Studien handelt.

Noch eine Bemerkung zur Frage der Häufigkeit der Injektionen. Es gibt durchaus Studien, in denen dem Arzt die Möglichkeit gegeben worden ist, von der mehrfach täglichen Injektion des NPH-Insulins Gebrauch zu machen; Sie kennen die Studien. Das sind Ergebnisse, die durchaus Vorteile des langwirksamen Insulinanalogons stützen.

Insofern sind diese klinischen Studien durchgeführt worden. Zu diesen Studien sind die Ergebnisse vorhanden. Teilweise sind sie veröffentlicht. Weitere werden in Kürze folgen.

Moderator Sawicki: Sie haben mich grundlegend missverstanden. Es geht nur darum: Ich wolle Ihnen klarmachen, dass das Zählen der Anzahl der Injektionen nicht unbedingt ein Vorteil für die Patienten sein muss. Dass man die Therapien von vor dem Zweiten Weltkrieg mit den jetzigen nicht vergleichen kann, ist selbstverständlich. Für diese Bemerkung hätten Sie unsere Zeit nicht in Anspruch nehmen müssen, Herr Knollmeyer. Es geht darum, nicht Patienteninjektionen zu zählen und zu sagen, dass das ein Wert ist. Es kann auch sein, dass

die Patienten mit mehreren Injektionen flexibler sind. Also, man muss die Patienten fragen, ob es ihnen besser geht.

Herr Thole: Ich habe eine Rückfrage zwecks Klärung. Sowohl Sie als auch Herr Kaiser haben gerade mehrfach gesagt, Sie hätten Annahmen, also etwas könnte so sein, etwas würde so sein, etwas könnte sich so darstellen. Herr Kaiser sagte auch – sofern ich mich richtig erinnere –, dass Sie aufgrund dieser Annahmen in die Fragestellung hineingehen. Ich finde das etwas verwirrend. Denn ich verstehe Ihren Auftrag so, dass Sie die vorhandene Evidenz bewerten sollen. Ich verstehe Ihre Ausführungen so, dass Sie eine Annahme machen, die Sie setzen, und dass Sie dann nur noch die dazu passenden Studien aufnehmen, dass Sie aber nicht die Studien betrachten, die zu dieser Fragestellung tatsächlich vorhanden sind. Es geht aus meiner Sicht – so habe ich es verstanden – darum, die vorhandene Evidenz zu bewerten und zu beschreiben und nicht durch eine Annahme zu sagen: Das halten wir für das Richtige, diesen Aspekt halten wir für den hier zu untersuchenden Aspekt, und genau dazu wählen wir die Studien aus.

Dann kommen Sie zu Ausschlusskriterien, die – sofern ich die Diskussion bislang richtig verfolgt habe – unter anderem dazu führen, dass die Yokoyama-Studie im Cochrane-Review behandelt wird, aber dass die Yokoyama-Studie, die einen Datenpool zur thematischen Fragestellung der Analoga liefert, hier nicht behandelt wird. Ich finde es zumindest interessant, dass Sie aufgrund von Annahmen zu Setzungen kommen, die dann auf die vorhandene Evidenz bei RCTs wirksam sind. Sie schließen eben Dinge ein und aus. Das finde ich, höflich gesagt, verwirrend.

Herr Kaiser: Herr Thole, ich sehe, dass Sie verwirrt sind, aber Ihre Annahme ist falsch. Denn meine Aussage war: Basierend auf der Stellungnahme der Deutschen Diabetes-Gesellschaft könnte man es so annehmen. – Das heißt, ich habe die Stellungnahme der Deutschen Diabetes-Gesellschaft aufgegriffen. Ich habe aber gleichzeitig gesagt: Diese Annahme muss man prüfen, und deswegen muss man solche Studien durchführen, die solche Annahmen prüfen.

Wir sind grundsätzlich offen. Wir haben Studien eingeschlossen, die sowohl einen Vergleich mit einer einmal als auch mit einer zweimal täglichen Gabe von NPH untersucht haben. Insofern sind wir diesbezüglich nicht eingeschränkt. Wir würden auch Studien einschließen, in denen es möglicherweise eine besondere oder keine besondere Instruktion zum Gebrauch von NPH gibt. Zunächst sind wir offen. Wir müssen es uns anschauen. Wir müssen uns anschauen, wie die Studien durchgeführt worden sind. Aber das ist ein grundsätzliches Missverständnis. Ich bin nur auf die Annahmen der Deutschen Diabetes-Gesellschaft eingegangen.

Herr Thole: Als Antwort: Erstens fürs Protokoll. Man sieht mich hier keineswegs verwirrt, sondern ich habe eine Frage gestellt.

(Herr Kaiser: Sie haben es selbst gesagt!)

Ich zittere hier nicht, und ich sitze hier nicht auf eine komische Art und Weise. Ich sage es, weil es hinterher veröffentlicht wird. Ich fühle mich hier eigentlich sehr wohl. Ich kann meine Fragen stellen, bekomme aber leider keine Antworten, weil es hier nicht diskutiert wird, wie wir vorhin gelernt haben.

(Herr Kaiser: Ich habe Ihnen doch geantwortet!)

Fakt ist, dass der Umgang mit Ein- und Ausschlusskriterien – das zeigen auch ganz viele Stellungnahmen – sehr wohl so ist, wie ich das beschrieben habe. Es kommt zu Annahmen und Setzungen. Denn sonst würden wir hier nicht über die Ein- und Ausschlusskriterien diskutieren. Es ist bislang in jedem Berichtsplan der Satz enthalten: Da die potenzielle Fragestellung in einem RCT erforscht werden kann oder könnte, fokussieren Sie sich nur auf den Einbezug dieser Studienform und lassen andere Evidenz aus. – Das ist das deutlichste von Ihnen schriftlich fixierte Zeichen dieser Setzung. Sie sagen, Sie haben eine gewisse Annahme, die dann zum Ausschluss vorhandener Evidenz führt. Das ist mein Statement, und das ist auch so.

Moderator Sawicki: Wenn Sie sich außerhalb der Rednerliste befinden, bitte ich, dass Sie sich kurz fassen. – Herr Fritsche.

Herr Fritsche: Ich möchte noch mal auf die Frage von Herrn Kaiser zur Stellungnahme der DGG mit den Zeitpunkten eingehen. Ein wichtiger Punkt ist unserer Meinung nach eben auch der frei wählbare Zeitpunkt, wann diese langwirksamen Insulinanaloga gegeben werden können. Das ist auch schlecht in RCT Trials prüfbar, weil Sie hier eine gewisse Reproduzierbarkeit haben müssen. In der praktischen klinischen Anwendung kommt es sehr deutlich zum Tragen: Man kann dem Patienten das Insulin dann geben, wenn er es zuverlässig injizieren kann.

Viele Patienten – gerade Typ-2-Diabetes-Patienten – vergessen die spätabendliche Injektion, und bei solchen Patienten ist es gut, das Insulin morgens oder mittags zu geben, also zu einem Zeitpunkt, an dem es sich der Patient gerne injizieren möchte.

Herr Frick: Ich möchte noch einmal auf die Frage zurückkommen, die Sie, Herr Sawicki, und Sie, Herr Kaiser, in die Runde gestellt hatten, wo es nämlich grundsätzliche Hindernisse gebe, in einer randomisierten Studie die Patientenzufriedenheit bei Einmal- vs. Zweimalgabe zu testen. Kurzer Einschub vorab: Natürlich ist nach den Standards der evidenzbasierten Medizin die vorhandene Evidenz nicht die theoretisch erreichbare Evidenz; das möchte ich nicht weiter ausführen.

Zur Frage, die Sie in den Raum gestellt haben, zurück. Es gibt natürlich ein methodisches Hindernis gerade bei Patient-reported Outcomes. Sie schreiben in Ihrem Papier von Double-

Dummy-Technik, die man nach Möglichkeit einsetzen sollte, um gerade bei Patient-reported Outcomes zu verhindern, dass Patienten, die sich selber gegenüber nicht verblindbar sind, nicht erfahren, welche Therapie sie bekommen. Hier haben Sie klassisch eine Konstellation, in der eine Nichtverblindbarkeit gegeben ist, wenn Sie zweimal gegen einmal geben. Und Sie haben klassisch bei einem Patient-reported Outcome eine Situation, bei der Sie auch keine verblindete Endpunkterhebung machen können. Sie können den Patienten nicht gegenüber seinem eigenen Wissen verblinden, und er ist derjenige, der das Maß festlegt. Das heißt, das würde bei Ihnen notwendigerweise in der Qualitätsbeurteilung der Studien einen leichten oder sogar schwerwiegenden wegen der fehlenden verblindeten Endpunkteerhebung Mangel bedeuten. Zu einem leichten Mangel würde es auf jeden Fall führen. Hinsichtlich der Validität und Aussagekraft würde es von Ihnen herabgestuft werden.

Es ist ein grundsätzliches methodisches Problem, dass hier etwas gefordert wird, was auch theoretisch nicht erreichbar ist, und dass das Nichterreichen eines theoretisch nicht erreichbaren Ziels zur Herabstufung der methodischen Qualität genommen wird. Das wird wiederum als Hinweis darauf genommen, dass diese Daten nicht als valide betrachtet werden und nicht gleichrangig in die Bewertung einfließen. Das ist ein grundsätzliches methodisches Problem.

Moderator Sawicki: Das tun wir nicht. Wir fordern nicht doppelblinde Studien mit Double-Dummy-Technik für nicht verblindbare Endpunkte. Das tun wir nicht. Da irren Sie sich.

Herr Frick: Herr Sawicki, ganz kurz dazu: In dem Vorbericht tun Sie es. Da wird aus diesem Grund die methodische Qualität der Studien heruntergestuft.

Herr Kaiser: Nur damit das nicht verloren geht; das ist ein ganz neuer Punkt: Ich folge Ihnen, und wir haben auch geschrieben, dass man in einer solchen nicht verblindbaren Situation eben die Double-Dummy-Technik nicht einsetzen kann, wenn es gerade um diese Frage des erlebbaren Vorteils unterschiedlicher Anwendbarkeit geht.

Nichtsdestotrotz muss man natürlich die vorhandene – auch empirische – Evidenz sehen, dass in offenen Studien Ergebnisse möglicherweise gerade für subjektive Endpunkte eher verzerrt sind. Damit kann man aber umgehen. Das heißt, man kann sagen: Wenn der Effekt besonders groß ist, dann ist es möglicherweise nicht wahrscheinlich, dass durch das offene Studiendesign diese Effekte alleine erklärt werden.

So könnte man mit diesem Problem umgehen. Das ist auch keine besondere Vorgehensweise, die wir gehen. Wir kommen gleich bei den nicht schweren Hypoglykämien zu diesem Punkt. Da ist das auch ein Thema gewesen. Auf dieses Problem haben auch andere Behörden und HTA-Institutionen reagiert, indem sie das beschrieben. Denken Sie an den Brief der FDA zum Zulassungsverfahren von Insulin Glargin. Darin ist explizit dargestellt worden, dass gerade nicht eine Überlegenheit bezüglich Hypoglykämien konstatiert werden kann – das steht so explizit in dem Brief der FDA –, weil es sich um offene Studien handelt.

Die Studien an sich haben trotzdem eine für sich genommen hohe Qualität. Wir sagen ja auch, dass sie leichte Mängel haben. Leichte Mängel gibt es praktisch immer. Keine Mängel gibt es praktisch nie. Damit gehen wir grundsätzlich erst einmal nicht davon aus, dass die Studien ein Problem haben. Aber nichtsdestotrotz kann das offene Studiendesign bei diesen Endpunkten zu einer Verzerrung führen. Das ist auch internationaler Konsens.

Herr Frick: Wir haben hier – glaube ich – inhaltlich einen Konflikt, den wir etwas zugedeckt haben. Ich folge Ihnen völlig darin, dass da ein erhöhtes Verzerrungspotenzial ist. Die Frage ist, wie wir mit diesem Verzerrungspotenzial, das nicht abwendbar ist, umgehen.

Moderator Sawicki: Genau.

Herr Frick: Sie gehen in der Weise damit um, dass Sie sagen: Die Ergebnisse, die die Patienten schildern, die ja wissen, was sie bekommen haben, glauben wir nur eingeschränkt.

Moderator Sawicki: Nein.

Herr Frick: Ich halte es für einen methodisch falschen Ansatz. Man muss mit dem leben, was man hat, zumindest mit dem, was theoretisch geht. Das ist Ihr Ansatz, den Sie eigentlich fahren. Hier ist genau das Maximum dessen erreicht, was methodisch geht, und Sie relativieren trotzdem das Ergebnis, das dabei herauskommt. Das halte ich erkenntnistheoretisch für einen falschen Ansatz, weil Sie – Sie nicken, Sie wissen, worauf ich hinauswill – die beste Evidenz, die überhaupt denkbar ist, relativieren. Das verlässt das Gedankengebäude der evidenzbasierten Medizin und führt einen erkenntnistheoretisch auf einen Abweg.

Moderator Sawicki: Das haben wir verstanden, sehen es aber anders.

Herr Sandow: Zu Ihrer Frage, ob die einmalige Gabe ein Vorteil ist oder nicht: Man muss nicht immer das Offensichtliche prüfen. Ich würde Ihnen eine Plain Language Summary als Antwort geben. Es gibt das Instrument der Telefonumfrage. Es hat sich bereits in zahlreichen US-Publikationen über die „Fear of Needles“ gezeigt, dass eine einmalige Injektion pro Tag von Patienten eher akzeptiert wird als eine zweimalige und dass die drei- oder viermalige eher ein Hindernis ist, wenn man zur intensivierten Insulintherapie geht. Das heißt, ich verstehe nicht, warum Sie danach fragen, ob man diesen offensichtlichen Vorteil einer einmaligen täglichen Injektion in irgendeiner Weise noch durch randomisierte Studien zur Häufigkeit der Injektion prüfen soll. Dafür gibt es die vorliegenden Studien, die natürlich über den Rahmen der eng gefassten Rahmen der Studienauswahl hinausgehen, die Sie bisher benutzt haben. Ich kann nicht verstehen, warum man einen so offensichtlichen Vorteil durch randomisierte klinische Studien prüfen sollte.

Moderator Sawicki: Wie groß ist denn der Vorteil, Herr Sandow?

Herr Sandow: Telefonumfragen.

Moderator Sawicki: Es geht mir nicht darum, wie er gemessen wurde, sondern darum, wie groß der Vorteil ist.

Herr Sandow: Die Patienten wählen laut den US-Publikationen lieber die einmalige als die zweimalige Injektion.

Moderator Sawicki: Wie viele Patienten sind es, und wie groß ist der Vorteil?

Herr Sandow: Darf ich Ihnen vielleicht die Publikationen nachreichen?

Moderator Sawicki: Es geht darum, dass Sie es untersuchen müssen. Nur dann können Sie es quantifizieren.

Herr Sandow: Es ist untersucht. Und es ist in den Studien über „Fear of Needles“ publiziert worden. Das ist das Schlagwort dabei.

Moderator Sawicki: Gut. Dann können Sie es nachreichen. – Herr Knollmeyer.

Herr Knollmeyer: Ich möchte gerne den Fokus, was die praktische Anwendung anbelangt, gerne ausweiten. Er bezieht sich nicht nur auf die einmal tägliche Gabe und die Flextime. Ich möchte an der Stelle auch darauf hinweisen, dass Flextime-Studien gemacht worden sind, die bedingt durch die Ein- und Ausschlusskriterien nicht betrachtet worden sind. Zum Typ-2-Diabetes sind das die Studien von Herrn Standl und die Studie 4001 von Herrn Fritsche, die einen Morgen- und einen Abendarm im Vergleich zu NPH hatte. Sie haben den Morgenarm auch entgegen der eigenen Berichtsplanung ausgeschlossen.

Für Typ 1 gibt es eine dreiarmige Studie, die die Gleichwertigkeit der Einstellung unabhängig von der Wahl des Injektionszeitraums im Verlaufe des Tages bewiesen hat.

Wie gesagt: Ich würde dringend anregen, dass diese Studien von Ihnen auch gesichtet und in die Bewertung einbezogen werden.

Ein weiterer wesentlicher Aspekt ist aus unserer Sicht die Adhärenz. Für Adhärenz haben Sie Hinweise über die Frequenz der Studienabbrecher. An den Studienabbrechern und an den Abbruchgründen können Sie deutliche Unterschiede der Insulinarten erkennen. Sie haben sie ansatzweise beschrieben, ohne dieses als patientenrelevanten Vorteil anzuerkennen. Ich möchte dringend anregen, dieses zu tun.

Moderator Sawicki: Danke schön. – Herr Paar.

Herr Paar: Ich möchte noch einmal auf das Thema zurückkommen, dass man nämlich bei einer Einmalgabe nicht per se von einem Vorteil ausgehen kann und dass man weitere Studien fordern muss. Es ist auch vom Patientenvertreter und von der DDG gesagt worden, dass das ein Nutzen per se ist. Ich habe an Herrn Kaiser die Frage, ob er glaubt, dass die von Ihnen

geforderte Studie durchführbar ist und ob sich genügend Patienten bereitfinden, an so einer Studie teilzunehmen?

Moderator Sawicki: Welche Studie meinen Sie?

Herr Paar: Ich meine die, die Herr Kaiser fordert. Er fordert eine Studie, die ich – und das ist es, was die Frage impliziert – für nicht durchführbar halte. Denn ich glaube, dass man die Patienten nicht finden wird. Ich wäre auch sehr dankbar, wenn die Kliniker, die die klinischen Studien durchführen, mal kommentieren würden, ob meine Einschätzung völlig aus der Luft geholt ist.

Herr Gallwitz: Das kann ich nur unterstützen.

Moderator Sawicki: Gut, Herr Gallwitz, das haben Sie jetzt deutlich gesagt, obwohl Sie nicht dran waren. – Herr Fritsche.

Herr Fritsche: Ich kann es auch nur unterstützen …

Moderator Sawicki: Gut.

Herr Fritsche: … und möchte hinzufügen: Herr Frick hat es basierend auf der Erkenntnistheorie dargelegt.

Ich muss es so formulieren: Ich finde es geradezu zynisch, wenn man bestimmte Dinge, die der Patient berichtet – dazu kommen wir nachher noch bei den Hypoglykämien –, oder die Injektionshäufigkeit … Es muss doch jedem in die Augen springen, dass eine Injektion etwas ist, was man nicht gerne macht. Also ist es ein Nutzen per se, wenn man weniger Injektionen durchführen muss. Es sei denn, man kommt aufgrund weniger Injektionen zu Tode oder es gibt andere schwere Nebenwirkungen. Wenn es aber mit gleichem Outcome möglich ist, dann sind weniger häufige Injektionen ein Vorteil an sich, den man nicht prüfen muss. Ich kann mir nicht vorstellen, warum ich das machen sollte oder muss.

Moderator Sawicki: Das kann ich Ihnen schnell kurz beantworten: Weil Sie die anderen Dinge, die Sie genannt haben – die Nachteile dadurch –, nicht erfassen. Deswegen müssen Sie überprüfen, welche Nachteile es hat. Würden Sie sonst allen Patienten mit Typ-1-Diabetes eine Injektion Glargin morgens geben?

Herr Fritsche: Wir reden jetzt über Typ-2-Diabetes.

Moderator Sawicki: Sie haben aber eine allgemeine Stellungnahme zur Philosophie der Forschung abgegeben. Sie haben richtig gesagt, dass Sie dadurch auch Nachteile in Kauf nehmen müssen. Also muss man es evaluieren, damit man weiß, welche Vorteile und welche Nachteile es hat. – Herr Knollmeyer.

Herr Knollmeyer: Herr Prof. Sawicki, ich möchte darauf hinweisen, dass es – glaube ich – zwischen den Therapeuten, den Patienten und auch den Herstellern unbestritten ist, dass der Wunsch existiert, Behandlungen zu haben, bei denen nicht mehr so häufig gespritzt werden muss.

Moderator Sawicki: Richtig.

Herr Knollmeyer: Die Anforderung ist von vielen Kreisen an uns gerichtet worden, und dem haben sich die Insulinhersteller gestellt.

Jetzt kann man die Frage stellen, was wir erreicht haben. Was wir erreicht haben, das können wir mit Daten belegen, dass etwa Typ-2-Diabetiker derzeit beim NPH-Insulin etwa 1,5-mal täglich injizieren müssen und dass beim Lantus-Patient im Regelfall die einmal tägliche Gabe ausreicht.

Es besteht unter uns wohl Einigkeit darüber, dass jede Injektion ein potenzieller Schaden ist. Wenn ich einen Diabetiker oder jemanden, der noch nicht spritzen muss, frage, dann wird er sagen: Die Injektionsvermeidung ist eine Eigenschaft, der ich einen hohen Stellenwert beimesse. – Der Kundige wird das aus vielen Gründen sagen. Zum einen reduziert er den Schmerz, den er sich zufügen muss. Zum anderen reduziert er das Potenzial, lokale Unverträglichkeiten zu haben. Er reduziert das Potenzial, dass sich an der Injektionsstelle irgendwelche Veränderungen bilden können. Wir ersuchen das IQWiG, dieses dringend als Zielgröße, als patientenrelevanten Vorteil auch in der Nutzenbewertung zu beschreiben.

Herr Frick: Ich möchte noch einmal auf Sie eingehen. Sie sagten gerade, dass die einmalige Gabe im Vergleich zu einer mehrmaligen Gabe keinen Nutzen an sich darstellen würde, weil man die möglichen Schadenaspekte, die in diesem Endpunkt verbunden sein könnten, berücksichtigen und messen müsste. Dieser ganze Vorbericht ist wie jeder Vorbericht ein multidimensionales Konstrukt. Das heißt, es werden sowohl Nutzen- als auch Schadensaspekte in einzelnen Endpunkten abgeprüft. Von daher ist es eine Überforderung eines jeden einzelnen Endpunktes, das für diesen noch einmal differenziert zu tun. Sie kommen ja nachher zu dem Schluss, der in diesem Fall zusammengefasst grob heißt: Es gibt keine Evidenz für Zusatznutzen und keine Evidenz für Zusatzschaden; es ist also eine Gleichbewertung.

Dann hat natürlich ein eigener Endpunkt, der einen Vorteil in einem Aspekt definiert, wie zum Beispiel einmalige Gabe oder invasive Maßnahme vs. mehrfach, für sich einen Wert. Denn die anderen Endpunkte prüfen natürlich das – Sie sagten es –, was man in diesem Endpunkt mit abprüfen müsste, nämlich welchen Schadenaspekt man sich damit potenziell einkauft. Die werden dann in anderen Endpunkten nicht geprüft. Das wäre eine Frage der Zusammenrechnung des Fazits dieses Berichts.

Moderator Sawicki: Jetzt haben Sie sich doch nicht an meine Bitte gehalten.

Herr Frick: Ich bin noch nicht ganz fertig. Ich habe noch einen zweiten Punkt, dessentwegen ich mich eigentlich gemeldet hatte. Es ging darum, dass Sie eben bei den telefonisch erhobenen Daten zur Patientenzufriedenheit auf die Stärke des Effektes zu sprechen kamen. Da erinnere ich mich an Herrn Kaiser – Thema ist wohl Verzerrung – und dass Sie sagten, wir können mit einem Verzerrungspotenzial nur in der Art und Weise leben, dass wir sagen, dass das trotzdem Evidenz ist, wenn der Effekt so stark ist, dass er durch die Verzerrung nicht allein erklärbar ist. Da sind wir bei dem Thema dramatische Evidenz.

Moderator Sawicki: Dramatische Effekte.

Herr Frick: Oder dramatischer Effekt, der dann allein als Evidenz akzeptiert wird. Da muss ich sagen: Das ist wieder kein Ausweg. Da sind – jetzt will ich mich zu Ihrer Beruhigung nicht wiederholen – wir genau an dem Punkt, wo wir eben waren. Da kippt man eine Menge Evidenz mit dem Bade aus, weil etwas nicht Erreichbares nicht erreicht wird. Es sind nicht nur dramatische Effekte. Sie haben mit jeder Menge Verzerrungen zu tun, und auch randomisierte Studien können nur Teile dieser Verzerrungen beherrschen. Das heißt, wir – und Sie viel mehr als wir – müssen in unserer täglichen Arbeit dauernd mit Verzerrungen arbeiten, und wenn Sie mit verzerrungsfreien Daten arbeiten wollten, dann könnten Sie dieses Institut heute schließen. Es gibt sie nicht.

Das heißt, Sie müssen mit Verzerrungen irgendwie arbeiten. Und mit dieser Art, zu sagen, nur bei dramatischen Effekten ist die Wahrscheinlichkeit so gering, dass eine Verzerrung den Effekt nicht erklärt, kommen Sie nicht weiter.

Moderator Sawicki: Es besteht bestimmt kein Dissens. Es geht darum, ob die Studie Fehler zulässt, die die Größe des Effektes nicht infrage stellen. Wenn Sie einen kleinen Effekt mit einer großen möglichen Verzerrung haben, dann ist das etwas anders zu interpretieren als ein großer Effekt bei einer kleinen Verzerrung. – Herr Paar.

Herr Paar: Ich wollte noch mal die Frage wiederholen: Halten Sie die Studie für machbar?

Herr Vervölgyi: Ich möchte noch mal auf die DDG und auf die Applikationsfrequenzen von NPH-Insulin zurückkommen. Sie haben in Ihrer Stellungnahme gesagt, dass diese Unterschiede in Studiendesigns immer eine Rolle spielen müssen, dass man ihnen also in einer Art und Weise Rechnung tragen müsse. Wir verstehen es so, dass man einerseits beide Therapien, also Glargin und NPH-Insulin, optimiert einsetzen müsste.

Andererseits sagen Sie, dass man die gleiche Zahl von Injektionen vergleichen müsste. Da sehen wir einen gewissen Widerspruch.

Herr Gallwitz: Warum ich mich gemeldet habe, ist der Vergleich mit den Typ-1-Diabetikern und den mehrfach täglichen Injektionen. Ich glaube nicht, dass man den Vergleich heranziehen kann, weil es vor allem bei Typ-1-Diabetes um die prandiale Insulinsubstitution

bei sehr instabiler Stoffwechsellage geht, durch die der Patient einen unmittelbaren Vorteil hat. Ich denke, man muss da ganz klar trennen.

Moderator Sawicki: Aber Sie haben doch verstanden, was ich meinte, oder?

Herr Gallwitz: Ich habe klar verstanden, was Sie meinten, aber ich denke, es ist trotzdem nicht vergleichbar; andere Stoffwechselsituationen, andere Nutzen-Schaden-Relationen für den Patienten.

Moderator Sawicki: Es reicht nicht, nur die Injektionshäufigkeit anzugucken. Man muss sich auch etwas anderes angucken. Genau. – Jetzt zu dem, was Herr Vervölgyi gesagt hat. Herr Vervölgyi, könnten Sie es noch mal wiederholen?

Herr Vervölgyi: Sie haben einerseits in Ihrer Stellungnahme in Bezug auf die Applikationsfrequenz von NPH-Insulin gesagt, dass das in Studien immer eine Rolle spielt, dass man dem Unterschied praktisch Rechnung tragen muss. Wir verstehen es halt im Sinne eines fairen Vergleichs einer optimierten Therapie von Glargin und NPH-Insulin, sprich NPH-Insulin auch mehrfach täglich zuzulassen.

Andererseits sagen Sie, dass man die gleiche Anzahl von Injektionen vergleichen muss.

Herr Fritsche: Darf ich darauf antworten? – Gerade in dieser 4001-Studie, bei der ich Mitautor bin, wurde das auch im Vorfeld diskutiert. Das ist noch einmal ein Design: NPH-Insulin abends vs. Glargin abends vs. Glargin morgens. Da wurde auch im Vorfeld der Studie diskutiert, ob man NPH-Insulin zweimal geben soll. Das wäre eine Möglichkeit, oder wegen der kürzeren Wirkdauer von NPH-Insulin müssten wir, um eine faire Ausgangssituation zu gewährleisten, zweimal NPH-Insulin geben.

Das ist wiederum vom Design her schwierig, weil man dann eine häufigere Insulininjektion mit einer selteneren vergleichen würde, und dann würde gesagt: Wenn wir Patientenzufriedenheit messen würden, dann würden wir durch dieses Studiendesign zu einer höheren Patientenzufriedenheit kommen, weil die einen weniger Glargin und die anderen häufiger NPH spritzen müssten. Klar entscheiden sie sich dann für Glargin und finden es toller.

Das war mit ein Grund der Fairness halber, einmal gegen einmal in dieses Design zu bringen. Man hätte die gleiche Studie – natürlich mit den doppelten Kosten – auch mit unterschiedlichen Designs machen können. Man hätte also die gleiche Studie mit zweimal NPH-Insulin machen können. Das wäre sicherlich eine interessante Sache gewesen. Das gebe ich zu. Aber ob man es machen muss, ist die andere Frage.

Herr Kaiser: Damit beantworte ich auch die Frage von Herrn Paar: Ja, ich denke, man muss solche Studien machen, weil genau das die Fragestellung ist.

(Herr Fritsche: Aber man kann sie nicht machen!)

Sie liegen vor, aber leider nicht für alle Therapiegebiete. Das Problem ist, dass wir für die basalunterstützte orale Therapie für Insulin Glargin solche Studien nicht haben. Aber es gibt natürlich die Studie 3006, in der NPH, bezogen auf den individuellen Patienten, optimal angewandt und in der individuellen Einstellung geprüft wurde, ob er es einmal oder zweimal braucht, und zwar so, wie es dem Zulassungsstatus und dem normalen Umgehen mit NPH entspricht, also nicht zwanghaft einmal abends, sondern so, wie man es in der Behandlung optimal macht. Also sind solche Studien durchführbar, machbar und natürlich auch zu fordern, um zu prüfen, ob es auch in dieser basalunterstützten Oraltherapie zu einem Vorteil kommt. Gerade da könnte es tatsächlich zu einem deutlicheren Vorteil kommen, weil die Injektion durch die prandiale Insulingabe bei der Gesamtinjektionsfrequenz keine Rolle spielt. Denn dann geht es nur um das Basalinsulin.

Ich möchte noch zwei Missverständnisse ausräumen. – Herr Knollmeyer, Sie haben – glaube ich – gesagt, der Morgenarm der Studie 4001 ist nicht eingeschlossen worden. Das kann ich überhaupt nicht nachvollziehen. Er zieht sich durch alle Tabellen. Er ist in der Landkarte der Beleglage. Er findet sich auch im Fazit wieder. Also, ich glaube, da gibt es ein Problem.

Was nicht drin ist, ist der Vergleich zwischen dem Abend- und dem Morgenarm. Aber es ist natürlich der Vergleich zwischen dem Morgenarm und NPH-Insulin drin, und das ist ja die Fragestellung, die den Entscheider im G-BA interessiert. Das ist hier drin.

Das zweite Missverständnis: Es wird hier so getan, als wäre die basalunterstützte orale Therapie das Nonplusultra und das, was bei Patienten mit Typ-2-Diabetes angewendet würde. Die Versorgungsdaten zeigen eine ganz andere Lage. Wir haben auch die aus unserer Sicht bestverfügbaren Versorgungsdaten dargestellt, die systematisch erhoben worden sind. Das spielt eine untergeordnete Rolle im Vergleich zur konventionellen, aber insbesondere auch zur intensivierten Insulintherapie auch bei Typ-2-Diabetes. Insofern reden wir bei der basalunterstützten oralen Therapie nur von einer Minderheit der Patienten mit Typ-2-Diabetes, aber ganz sicherlich nicht davon, wie Basalinsuline bei Typ-2-Diabetes eingesetzt werden. Uns sind zumindest keine Daten von Ihnen eingereicht worden, die diese Untersuchungen irgendwie infrage gestellt hätten.

Herr Knollmeyer: Herr Kaiser, direkt dazu. Auf Seite 110 des Vorberichts ist im Bereich der Veränderung des HbA1c-Wertes bei der Studie 4001 nur der Abendarm, aber nicht der Morgenarm enthalten. Korrespondierendes gibt es im Bereich der Häufigkeiten der Hypoglykämien. Ich möchte das IQWiG dringend ersuchen, auch den Morgenarm in den beiden Beurteilungsparametern mit darzustellen.

Zweiter Punkt. Herr Sawicki, ich bitte um Ihre Hilfe. Sie haben uns vorhin erklärt, dass die Größe des Effektes invers mit dem Verzerrungspotenzial verknüpft sei. Gibt es darüber systematische Erhebungen? – Falls es sie gibt, würde ich sie gerne kennen.

Moderator Sawicki: Nein, das habe ich nicht gesagt. Ich kann noch mal wiederholen, was ich gesagt habe. Die Richtigkeit der Aussagen, ob Sie also mit einer Aussage recht haben, hängt zum einen davon ab, wie groß das Verzerrungspotenzial in einer solchen Studie ist, und zum anderen davon ab, wie groß der Effekt ist. Ich gebe Ihnen ein Beispiel: Wenn Sie aufstehen und sich neben Herrn Sandow stellen und ich in 50 oder 100 Metern Entfernung stehe, dann werde ich unter Umständen nicht feststellen können, wer größer ist. Wenn Sie aber neben einer Ameise stehen, werde ich sagen: Herr Knollmeyer ist größer.

Haben Sie es verstanden? – Also, die Messunggenauigkeit auf große Entfernung ist bezüglich der Interpretation von dem Effekt, den man dann ermittelt, abhängig. Das meinte ich.

Herr Kaiser: Herr Knollmeyer, es ist völlig richtig, dass sich der Morgenarm nicht in den Meta-Analysen zur Abendbehandlung befindet, aber natürlich ist der Arm eingeschlossen und für sich dargestellt. Sie müssen sich den Bericht nur anschauen. Blättern Sie zwei Seiten weiter. Dort ist der Morgenarm dargestellt. Sie blättern in die Landkarte der Belegkarte. Sie blättern in das Fazit. Dort ist die Morgenbehandlung dargestellt. Ich verstehe die Aussage überhaupt nicht.

Ihre Aussage, die Sie konkret gemacht haben, deutet eher darauf hin, dass Sie sagen: Auch den Morgenarm sollte man in die Meta-Analysen einfließen lassen, also dort subsumieren. – Dann muss ich sagen: Dann sehen Sie zwischen der morgendlichen und abendlichen Gabe von Insulin Glargin keinen Unterschied.

Herr Knollmeyer: Genau. Das hat die Studie gezeigt.

Herr Kaiser: Das hat sie gerade nicht gezeigt, weil es in den Parametern einen Unterschied zwischen Morgenarm und NPH gibt, für Abendarm und NPH aber nicht. Insofern zeigt die Studie das gerade nicht.

Herr Knollmeyer: Die Unterschiede zwischen Morgenarm und Abendarm waren meines Wissens statistisch nicht signifikant. Es gab eine separate Studie von Herrn Standl zum Typ-2-Diabetes und Insulin Glargin, die nur morgendliche und abendliche Gabe untersucht hat. In dieser sind auch nicht signifikante Unterschiede festgestellt worden, was die Behörden dann dazu veranlasst hat, eine Formulierung zuzulassen, dass der Patient den Tageszeitpunkt der Injektion frei wählen kann. Das können Sie auf dem EPAR nachlesen.

Herr Kaiser: Da muss man tatsächlich von Conflicting Evidence reden. Das gibt es ja schon mal. Sie haben in dem Fall der Standl-Studie – das ist völlig richtig – keine signifikanten Unterschiede bei Parametern gesehen. In dieser Studie sehen Sie signifikante Unterschiede im Vergleich zu NPH bei dem Morgenarm, aber nicht beim Abendarm. Insofern ist es erst einmal eine Conflicting Evidence.

Herr Paar: Herr Kaiser, vielleicht habe ich meine Frage eben nicht richtig gestellt; sie ist noch nicht ganz beantwortet. Ich habe Sie so verstanden, dass Sie gesagt haben, dass man die Studie jetzt machen muss, weil es untersucht werden soll. Gemeint habe ich: Glauben Sie, dass man Patienten für so eine Studie findet?

Herr Kaiser: Das habe ich damit beantwortet, dass ich gesagt habe, dass es solche Studien gibt. Ich denke beispielsweise an Ihre neue 5-Jahres-Studie. Da ist NPH doch nicht zwanghaft einmal abends eingesetzt worden.

Herr Paar: Die Frage war gewesen: Wenn ich im Design zweimal vs. einmal hineinschreibe, dann frage ich mich, ob Sie glauben, dass man dafür Patienten findet, die eine Unterschrift unter den Informed Consent setzen, um an dieser Art von Studie teilzunehmen.

Herr Kaiser: Vielleicht ist es nur ein Missverständnis, aber es ist nie eine Behauptung gewesen, dass wir gefordert haben, dass man genau solche Studien mit zwanghaft zweimal durchführt. Die Forderung ist vielmehr – so ist es auch in der Diskussion des Vorberichts dargestellt –, NPH patientenindividuell optimal einzusetzen. Das kann im Einzelfall eine einmal tägliche Gabe sein. Das kann im Einzelfall eine zweimal tägliche Gabe sein, im Einzelfall auch eine dreimal tägliche. Das entspricht dem Zulassungsstatus und auch der tatsächlich angewandten Therapie. Und wenn Sie sich die 3006-Studie angucken, dann sehen Sie, dass es auch so durchgeführt worden ist.

Herr Paar: Sie haben gesagt, eine Einmalgabe per se ist kein Vorteil.

Herr Kaiser: Richtig.

Herr Paar: Wenn ich das untersuchen will, dann muss ich doch eine Studie prospektiv machen, in der ich eine Zweimalgabe gegen eine Einmalgabe randomisiere. Oder verstehe ich das falsch?

Moderator Sawicki: Vielleicht kann ich das mal erklären. – Wir können Medikamente miteinander vergleichen, die auf eine unterschiedliche Art und Weise verwendet werden. Dann muss man sich aber ein davon unabhängiges Outcome-Kriterium aussuchen.

Ich greife mal Asthma auf. Sie haben unterschiedlich anzuwendende Präparate oder Oraltherapie vs. Transdermaltherapie. Das können Sie miteinander vergleichen. Aber Sie können nicht annehmen, dass es per se besser ist, ein transdermales Pflaster aufzukleben vs. eine Spritze oder eine Tablette zu geben. Sie müssen dann vielmehr ein anderes Outcome-Kriterium wählen, zum Beispiel die Patientenzufriedenheit. Sie können nicht annehmen, dass es alleine durch die unterschiedliche Art der Verwendung quasi axiomatisch besser ist. Sie müssen ein anderes Outcome-Kriterium wählen. Das ist das, was Herr Kaiser meinte.

Herr Richter: Ich möchte noch einmal auf das Beschreibungsfeld von Zielgrößen allgemein zurückkommen, hinsichtlich derer ich um Klarifikation bitten würde. Hier wabert ein wenig

der Eindruck rum, als würden Zielgrößen mehr oder minder manipulativ eingesetzt, verwandt oder ausgesucht werden. Soweit ich weiß, ist im Auftrag des G-BA ein gewisser Katalog an Zielgrößen umschrieben. Ich würde auch meinen, dass beim G-BA eine gewisse Fachkompetenz angesiedelt ist. Also kann es gar nicht sein, dass wesentliche Zielgrößen manipulativ von irgendeiner Seite eingesetzt werden. Diesen Widerspruch würde ich gerne aufgeklärt wissen. Es ist schließlich eine Kompetenz eines Organs, das eine entsprechende Vorgabe gibt.

Moderator Sawicki: Herr Sandow. – Bitte achten Sie darauf, dass Sie verschiedene Sachen sagen.

Herr Sandow: Zum einen kann ich nicht nachvollziehen oder prüfen – ich weiß nicht, ob Sie es offenlegen möchten –, welche Zielgröße Sie vom G-BA im Einzelnen vorgegeben bekommen haben.

Zum anderen erscheint es mir heute hinsichtlich der Zielgrößen bei der heutigen Diskussion sinnvoll, zu fragen: War die Zielgröße, die hier beschrieben wurde, im Einzelnen sinnvoll oder nicht? – Ich denke, dazu haben wir in der Stellungnahme bestimmte Auffassungen dargelegt, dass es sicherlich Zielgrößen gab, die sicherlich sinnvoll und richtig waren, und dass es andere gab, die unrealistisch waren, weil es dazu keine Evidenz gibt, die sich innerhalb der zur Verfügung stehenden Zeit ergeben konnte. Und es gab Zielgrößen, die offensichtlich unsinnig waren. Also, die Bewertung heute aus der Sicht des Vorberichts kann man sicherlich durchführen, und ich halte sie auch für legitim.

Herr Frick: Der Auftrag des G-BA datiert – so ist es auch im Internet einsehbar – vom Februar 2005 und ist sehr allgemein gehalten. Da wird aufgefordert – glaube ich –, acht Volkserkrankungen – darunter Diabetes Typ 1 und Typ 2 – zu bewerten. Das ist es. Da stehen keine Zielgröße und nichts drin. Das heißt, der Auftrag ist in keiner Weise so ausgestaltet, dass er irgendeine Vorgabe hinsichtlich der Zielgröße machen würde. Von daher kann man es nicht mit der Autorität des G-BA verknüpfen, wie Sie es, Herr Richter, getan haben. Die Fragestellung war wesentlich allgemeiner.

Herr Knollmeyer: Ich würde den Themenpunkt gerne ergänzen. Für uns mangelt es an Transparenz bei der Festlegung der Zielgrößen. Für uns ist nicht erkennbar, auf welcher empirischen Basis die Zielgrößen festgelegt worden sind. Sie sind im Übrigen im Laufe des Verfahrens mehrfach modifiziert worden, hier insbesondere in puncto Unterzuckerung.

Die aktuellen Entscheidungsgrundlagen des Unterausschusses Arzneimittel für die Bewertung des therapeutischen Nutzens, der medizinischen Notwendigkeit und der Wirtschaftlichkeit von Arzneimitteln geben in § 6 vor, dass maßgeblich die Beeinflussung patientenrelevanter Endpunkte sei. Diese Überschrift hat das Kapitel vorher mal getragen. Das ist dann geändert worden.

Und unter 3. steht:

„… unter vorrangiger Berücksichtigung randomisierter kontrollierter klinischer Studien, insbesondere direkter Vergleichsstudien mit anderen Arzneimitteln oder Behandlungsmöglichkeiten mit patientenrelevanten Endpunkten."

Da sind wir in einem sehr komplexen Feld, was die Komparatoren und die Methodik anbelangt.

Ich bleibe bei der Einschätzung, für die es auch Empirie gibt – Herr Prof. Sandow hat es angesprochen –, dass die Injektionsfrequenz patientenrelevant ist. Darüber brauchen wir überhaupt nicht zu diskutieren. Die Evidenz, die es dazu gibt, ist erdrückend.

Herr Richter: Ich wollte noch mal in Richtung Herrn Frick sagen: Ich habe nichts verknüpft, sondern eindeutig um Klarifikation gebeten. Wenn Sie die Kompetenz haben, diese Klarifikation zu geben, dann ist es gut. Ich habe schon G-BA-Berichte gelesen, in denen durchaus Zielgrößen formuliert waren. Wenn das hier nicht der Fall war, dann ist ja gut. Dann ist meine Frage beantwortet. Das war sie lediglich.

Herr Frick: Das sage ich jetzt nicht als Freund des kleinen Karos, aber Sie haben gerade gesagt, dass die Richtigkeit von dem Verzerrungspotenzial abhängt. Ich will damit eine Aussage machen, die für uns von Relevanz ist. Sie sagten eben – ich habe es mir aufgeschrieben, weil ich Sie nicht falsch zitieren wollte – in dem Gespräch mit Herrn Knollmeyer, die Richtigkeit hänge vom Verzerrungspotenzial ab. Das ist – glaube ich – ein grundsätzlicher Unterschied, den wir hier machen. Die Richtigkeit hängt meiner Meinung nach nicht davon ab, höchstens die Wahrscheinlichkeit, dass etwas richtig ist. Sie können perfekte Studien haben und sie sind trotzdem falsch, wenn man es so dichotom ausdrückt. Sie können grottenschlecht designte Studien haben, die trotzdem richtig sind, wenn man irgendwann ein Maß hat, sie zu beurteilen. Wir sind beim Problem der Vierfeldertafel hier.

Wenn Sie sagen, die Richtigkeit hängt vom Verzerrungspotenzial ab, dann ist Ihr Job leicht. Denn dann können Sie mit Fug und Recht bei gegebenem Verzerrungspotenzial sagen, dass es falsch ist. Das ist aber nicht das Problem. Das Problem ist: Sie haben falsche Studien ohne Verzerrungspotenzial und richtige Studien mit diversem Verzerrungspotenzial. Sie müssen da tatsächlich im Einzelfall prüfen.

Moderator Sawicki: Die Wahrscheinlichkeit der Richtigkeit der Aussage hängt ab …

Herr Frick: Genau.

Moderator Sawicki: Das müssen wir schriftlich etwas ausführlicher formulieren. – Bernd Richter.

Herr Richter: Ich kann methodisch etwas dazu sagen. Es ist doch ganz einfach: Die Nullhypothese können Sie nicht beweisen, weil sie naturgegeben ist. Sie richtet sich nicht danach, wie wir was untersuchen, sondern sie entfaltet sich über die Wirklichkeit der Natur, und nur über die Fehlerwahrscheinlichkeit, dass wir diese verwerfen, kommt das Verzerrungspotenzial hinein, und dann gehen wir eventuell fehl. Das ist eine relativ einfache methodische Sache. Das muss man gar nicht so kompliziert ausdrücken. Und dass die beiden Sachen miteinander verwoben sind, ist leider nun einmal so. Wir können uns der Wahrheit schließlich nur approximativ nähern.

Herr Grundner: Ich möchte noch mal die Einmalgabe ansprechen. Herr Kaiser hat darauf hingewiesen, dass aus seiner Sicht die Kombinationstherapie Basalinsulin plus OAD eine eher untergeordnete Rolle spielt. Nach den uns vorliegenden Daten, die wir Ihnen auch zur Verfügung gestellt haben, ist es mittlerweile doch deutlich anders. Wir sprechen über mehr als 45 % der Patienten, die nach Umstellung auf Basalinsulin plus OAD eingestellt werden. Das sage ich, um die Relevanz solcher Studiendesigns und solcher Therapiestrategien noch einmal zu unterstreichen.

Moderator Sawicki: Danke schön. – Dann haben wir das erschöpfend diskutiert. Wir kommen zu:

Tagesordnungspunkt 3:
Nicht schwere Hypoglykämien

Herr Vervölgyi: Dem Vorbericht haben die nicht schweren Hypoglykämien aufgrund des durchweg offenen Studiendesigns in den Studien und des durchweg patientenberichteten Outcomes ein hohes Verzerrungspotenzial attestiert. Wir haben sie deshalb als nicht fazitrelevant angesehen. Dieser Punkt wurde eigentlich von allen Stellungnehmenden thematisiert.

Nun ist es so, dass in verschiedenen Arbeiten gezeigt werden konnte, dass fehlende Verblindungen in Studien bei subjektiven Endpunkten – ich denke, die nicht schweren Hypoglykämien gehören sicherlich dazu – einen erheblichen Einfluss auf das Ergebnis haben, auch zugunsten der Intervention. Deswegen die Frage an Sie, wie Sie das Biaspotenzial der nicht schweren Hypoglykämien einschätzen und wie man dem Problem des hohen Verzerrungspotenzials begegnen könnte?

Herr Grundner: Wir möchten zuerst einmal sagen, dass diese Grundannahme, dass ein offenes Studiendesign eher eine Tendenz birgt, dass im Sinne von vorteilhaft für die neu zu untersuchende Substanz berichtet wird, aus unserer Sicht nicht ganz zutreffend ist. Wir hatten Clare Bradley in dem Zusammenhang zitiert. Ein Faktor, der eventuell zur Verzerrung führt und im Rahmen der RCTs teilartifizielle Situationen schafft, ist auch im Themenkreis der Resuspendierung, dass eben Patienten im Rahmen des Studienprotokolls bei Einschluss in die Studie sehr explizit darauf hingewiesen werden, wie sie mit dem NPH-Insulin umzugehen

haben. Das ist sicherlich eine Bedingung, die auf Praxisbedingungen so nicht übertragbar ist, was wieder eine Verzerrung zugunsten des NPH-Insulins bergen würde und im Falle der Resuspendierung sicherlich erheblich auch für die Hypoglykämiezahlen zu werten wäre.

Herr Knollmeyer: Ich würde gerne noch zwei Aspekte ergänzen. – Das eine ist: Wir haben neben den symptomatischen Hypoglykämien in vielen Studien die, die durch Laborwerte gesichert sind, angegeben. Ich glaube, hier haben wir das Verzerrungspotenzial beseitigt. Denn sofern der Patient sein Blutzuckermessgerät richtig bedient, gehe ich davon aus, dass wir uns auf die hier erhobenen Werte verlassen können. Wir haben dazu – das haben Sie auch beschrieben – in einigen Studien zusätzlichen Aufwand getrieben, um das Verzerrungspotenzial zu minimieren. Wir sind enttäuscht darüber, dass Sie die Ergebnisse dieser Studien nicht separat im Ergebnis dargestellt haben. Das hätte man tun können. Denn dort ist mit den damals validen Methoden versucht worden, diesem bekannten Phänomen irgendwo entgegenzutreten.

Wir glauben, dass es auch falsch ist, die Hypoglykämie insbesondere auf die schwere zu fokussieren, weil jede Hypoglykämie eine Gefährdung des Patienten darstellt und somit ein erhebliches Schadenspotenzial birgt. Dazu gibt es auch eine Reihe von Untersuchungen, die diese Behauptung mit stützen und die in die Bewertung der nicht schwerwiegenden Hypoglykämien einbezogen werden sollten.

Herr Fritsche: Der Punkt Hypoglykämien ist aus meiner Sicht der zentrale Punkt und der zentrale Vorteil der langwirkenden Insulinanaloga. Dieser zentrale Punkt wurde in der Bewertung mit der Begründung ausgeklammert – das war auch Ihre Frage, Herr Kaiser –, dass sie einer starken Verzerrung unterliegen und schwer erhebbar sind. Das sind relativ weiche Daten.

Allerdings liegt es in der Natur der Hypoglykämien, dass sie zum Beispiel in vielen Studien underreported werden und dass man Schwierigkeiten hat, eine wirklich genaue Erfassung zu machen. Ich denke, dass das für beide Situationen, also für NPH-Insulin und für das Insulinanalogon, zutrifft und dass man deswegen diesen ganz zentralen Punkt nicht ausklammern kann. Es ist ein zentraler Punkt, weil es für die Patienten die größte Gefährdung ist. Sie leiden darunter. Und auch ich als Arzt habe bei jedem Patienten, bei dem ich eine Insulintherapie anfange, großen Respekt davor und denke: Hoffentlich kommt er nicht mit einer schweren Hypoglykämie oder einem Autounfall zu mir, und hoffentlich habe ich das nicht aufgrund der Insulintherapie, die ich angefangen habe, mitverschuldet.

Also, wenn man das einmal miterlebt hat, dann vergisst man das nie mehr. Das ist für den Patienten extrem belastend. Schon allein die Angst davor ist extrem belastend. Es ist selbstredend, und deswegen ist dieses Ausblenden mit der Begründung, das sei ein schwieriger Parameter und man sei mit der Erhebung nicht sicher, aus meiner Sicht völlig unzulässig.

Moderator Sawicki: Danke schön. – Thomas Kaiser.

Herr Kaiser: Als Bemerkung zunächst dazu: Genau deswegen wollen wir das jetzt hier diskutieren. Und es ist auch nicht so, dass wir das grundsätzlich ausgeblendet hätten. Vielmehr sind wir aufgrund des vorhandenen Verzerrungspotenzials in der Situation, dass wir da ein Problem haben, ob wir eine Ergebnissicherheit haben oder nicht. Ich meine, das ist in den Stellungnahmen aufgeworfen worden, und wir verstehen auch, dass es zu Recht aufgeworfen worden ist. Deswegen müssten wir – denke ich – in der gemeinsamen Diskussion dazu kommen, zu überlegen, wie man damit gegebenenfalls umgehen könnte. Uns ist eben unklar geblieben, welche Wege es möglicherweise gibt, die dann auch das Verzerrungspotenzial mit in die Betrachtung einbeziehen. Es wäre – glaube ich – nicht der richtige Weg, das außen vor zu lassen; darüber müsste man dann diskutieren.

Zu Ihrer Aussage, Herr Knollmeyer, dass die Blutzuckermesswerte einer Verzerrung nicht unterliegen, wenn der Patient das Gerät bedienen kann. Da gibt es zwei Probleme: Zum einen ist es so, dass Blutzuckermesswerte auf Basis von Tagebüchern – ich habe die Studien verstanden, dass das in der Regel so gemacht und dann auch von den Ärzten übertragen worden ist – im Vergleich zu Blutzuckermesswerten, die in den Blutzuckergeräten mit einem elektronischen Speicher gespeichert wurden, erheblich diskrepant waren. Das heißt, die Aussagen aus den Tagebüchern waren erheblich diskrepant zu den Aussagen, die das Blutzuckermessgerät aus dem Speicher ergeben hat.

Das heißt, man könnte das Problem angehen, indem man in solchen Studien das Blutzuckermessgerät benutzt. – Das ist der eine Punkt.

Der zweite Punkt, der wesentlich ist und der sich aus den Daten für uns leider nicht einschätzen lässt, ist die Häufigkeit der Messung. Je häufiger Sie messen, desto höher ist die Wahrscheinlichkeit, dass Sie eine Hypoglykämie feststellen; das gilt insbesondere dann, wenn es um asymptomatische Hypoglykämien geht. Das heißt, Sie bräuchten eine Bezugsgröße „Häufigkeit der tatsächlich durchgeführten Messungen“. Diese findet sich in den Studienberichten, die uns vorliegen, so nicht. Insofern können wir das Verzerrungspotenzial, das dort vorhanden ist – ich sage ganz offen, dass ein Verzerrungspotenzial da ist; wir wissen nicht, wie häufig in den beiden Gruppen jeweils gemessen wurde –, nicht einschätzen. Aber dieses Problem ist vorhanden, und das Problem haben Sie an anderen Stellen so nicht, wenn Sie das beispielsweise bei den tatsächlichen Hypoglykämien mit einem tatsächlich stattfindenden Ereignis verknüpfen, das heißt mit einer Behandlung. Also, Sie können nicht verzerren, dass es eine Krankenhausbehandlung gegeben hat. Insofern sind diese Daten einfach ergebnissicherer. Sie haben allerdings das Problem, dass Sie weniger Power haben, weil es weniger Ereignisse gibt; das ist uns bewusst. Aber mit diesem Umstand müssen wir umgehen.

Herr Frick: Ich stelle mit Freude fest, dass Sie dieses Problem auch als solches erkennen; ich komme gleich dazu, an wen man das vielleicht adressieren kann. Ich denke, wenn man von

Ihrer Methodik ausgeht, muss man es auch adressieren. Im Moment ist es ja im Fazit nicht berücksichtigt. Insofern kann man verkürzt sagen: Es ist nicht berücksichtigt. – Allerdings geht es hier zweifelsohne um einen Schadensaspekt. Unter Schadensaspekten sind Sie ausweislich Ihrer Methodik viel bereiter, Verzerrungspotenzial zu akzeptieren. Ich verweise nur darauf, dass Sie da nicht einmal die Randomisierung fordern, sondern bis zu Tierversuchen …

Herr Kaiser: Das ist ein Missverständnis.

Herr Frick: Ja? Wo ist das Missverständnis? – Das würde mich jetzt interessieren. Ich verstehe Hypoglykämien per se als ein Schadensrisiko, und ich verstehe Ihre Methodik so, dass Sie dem Grundsatz – ich glaube, das steht auch so im Methodenpapier – „Primum nihil nocere“ folgend bereit sind, ein höheres Verzerrungspotenzial als akzeptable Evidenz heranzuziehen, sofern es um Schaden geht.

Herr Kaiser: Das Missverständnis, das da auch auf Basis der Formulierung im Methodenpapier passiert, ist das, dass wir solche Evidenz zunächst einmal heranziehen, um zu gucken, wo möglicherweise ein Schaden ist, der einer bestimmten Therapie zuzuordnen ist und der vielleicht auch bei einer Therapieentscheidung eine Rolle spielt, um dann aber in die sicheren Werte hineinzugehen. Und das werden abhängig von der Fragestellung und abhängig von der Größe des Effekts eben auch randomisierte Studien sein. Ich meine, dazu gibt es auch verschiedene Veröffentlichungen und verschiedene Herangehensweisen. Ich denke beispielsweise an die Suizidalität bei Antidepressiva. Letztendlich ist man bei den Zulassungsbehörden dabei geblieben, am Ende zu sagen: Nichts anderes als tatsächlich die Gesamtbetrachtung aller randomisierten Studien hilft uns hier weiter. Die Betrachtung der Beobachtungsdaten hat uns eben nicht weitergeholfen.

Also, das ist das Missverständnis, in dem Zusammenhang die Tierstudien als Beleg für einen Schaden zu nennen.

Herr Frick: Kurze Antwort darauf, woher das Missverständnis vielleicht beruht; das kann vielleicht bei der Formulierung der Version 4 helfen. Es steht bei Schadenspotenzialen in dem Methodenpapier zumindest bis Version 3, dass eine Evidenz für einen Schaden, die nicht auf RCT beruht, so lange als gegeben genommen wird, bis sie durch eine höherwertige Studie, also eine RCT, entwertet würde. Von daher: Solange wir sie hier nicht haben – ich übertrage es jetzt auf diesen Fall –, müssen wir mit einem höheren Verzerrungspotenzial bei dem Risikopotenzial Hypoglykämie natürlich leben und damit umgehen und es nicht aus dem Fazit ausblenden.

Mir ist als Gedanke dazu gekommen: Wäre es möglich – die Daten sind teilweise in Datenbanken –, noch einmal in die Definition zu gehen und die Daten – Stichwort: nächtliche Hypoglykämien, abweichende Definitionen – zu sichten, damit wir noch einmal in einen

Austausch treten, ob es auswertungstechnisch Möglichkeiten gibt, dieses Verzerrungspotenzial aus Ihrer Sicht für Sie handhabbarer zu machen?

Moderator Sawicki: Wir werden uns das merken. – Herr Gallwitz.

Herr Gallwitz: Ich möchte noch eine Anmerkung klinischer Art an Herrn Fritsche machen, und zwar sind leichte Hypoglykämien auch ein Risikofaktor für eine Abnahme der Hypoglykämiewahrnehmungsschwelle an sich im Verlauf der Diabeteserkrankung. Ich denke, das ist ein ganz wichtiger Punkt, der eher dazu führen sollte, auch die leichten Hypoglykämien sehr ernst zu nehmen und mit in die Bewertung eingehen zu lassen.

Moderator Sawicki: Es ist unbestritten, dass leichte Hypoglykämien eine Bedeutung für den Patienten haben. Das haben wir auch nie bestritten. Die Frage ist: Wie geht man damit um? Wie fasst man es so, damit man eine möglichst sichere Aussage dazu macht? – Herr Knollmeyer.

Herr Knollmeyer: Ich möchte in diesem Kontext noch einmal auf das Kapitel 3.1.3 „Ermittlung des Schadenspotenzials medizinischer Interventionen“ rekurrieren. Danach beschreibt das Schadenspotenzial das Risiko für das Auftreten unerwünschter Wirkungen bei der Anwendung einer medizinischen Intervention. Die Hypoglykämie ist die klassische unerwünschte Wirkung der Insulintherapie. Insofern glaube ich, dass hier auch im Rahmen der Beurteilung des Schadenspotenzials nach den dort beschriebenen Kriterien das Schadenspotenzial der verschiedenen Insulinprodukte verglichen werden muss.

Jetzt komme ich zu meinem zweiten Punkt. – Für mich ist es die Frage, ob ich deskriptiv oder konfirmatorisch an die Sache herangehe. Aus meiner Sicht ist ein deskriptiver Ansatz durchaus ausreichend. Ich muss nicht konfirmatorisch an diese Fragestellung herangehen.

Dritter Punkt: Das Vorgehen im Bereich der klinischen Studien ist für uns sehr standardisiert gewesen und gewissen zeitlichen Schwankungen angepasst worden. Standard ist gewesen, dass eine 8-Punkt-Messung verpflichtend für jeden Studienteilnehmer durchgeführt wird und auch tatsächlich durchgeführt wurde. Der Patient war aber frei, wenn er Dinge bemerkt hat, um zusätzliche Blutzuckermessungen durchzuführen. Das ist erst einmal die Stallregie gewesen, die an ihn ergangen ist. Und somit haben wir eigentlich über die Studien ein gleichmäßiges Raster, das über die ganze Zeit so angewendet worden ist.

Nun schlage ich die Brücke zu Herrn Frick. Wir haben in unseren Dokumenten mit aufgeführt, zu welcher Uhrzeit und welchen Wert der Patient gemessen hat. Und wir haben erhoben, zu welcher Uhrzeit die Unterzuckerung stattgefunden hat. Diese Informationen befinden sich in unseren Datenbanken, und über diese Informationen können wir natürlich unterschiedliche Vorgaben, was nächtliche Hypoglykämien usw. anbelangt, nachträglich bereinigen. Es kann durchaus sein, dass das der Grund gewesen ist, warum Sie in den Studienunterlagen im Studienbericht andere Informationen vorfinden als die, die wir Ihnen

heute weitergeben, weil wir die Daten noch einmal hergenommen und im Rahmen der veröffentlichten Mullins-2-Analyse nach diesem Raster nicht bearbeitet haben, um hier eine vergleichbarere Aussage zum Auftreten und zur Schwere der Unterzuckerung machen zu können. Hier kann man im Sinne der Best Available Evidence nur dringend sagen: Diese Untersuchungen, die wir im Nachgang gemacht und veröffentlicht haben, sind, was dieses Kriterium anbelangt, das Beste, was sich derzeit in öffentlichem Besitz befindet.

Moderator Sawicki: Danke schön. – Herr Koenen.

Herr Koenen: Grundsätzlich zu den nicht schweren Hypoglykämien: Ich denke, es ist sehr wichtig, zu realisieren, dass diese Hypoglykämien von Patienten selbst dann, wenn sie nicht mit Blutzuckermessungen substantuiert wurden, empfunden wurden und damit natürlich auch die Lebensqualität der Patienten beeinträchtigen.

Um das noch einmal klarzumachen: Bei den Blutzuckermessungen, die einen Hypoglykämiewert gezeigt haben, gab es zwei unterschiedliche Möglichkeiten. Zum einen waren es Blutzuckermessungen, die im Rahmen von vorgegebenen 8- oder 9-Point-Profilen erhoben wurden. Das heißt, die Frequenz wird da wohl nicht so eine große Rolle spielen. Zum anderen haben wir Blutzuckermessungen, die spontan vom Patienten durchgeführt wurden, weil der Patient glaubte, Hypoglykämie zu empfinden. Wenn das wichtig ist, sind wir bereit, in unsere Datenbanken hineinzugehen und das irgendwie genau zu präzisieren.

Zum anderen haben wir bei der Nichtanerkennung der nicht schweren Hypoglykämien ein Grundsatzproblem. Das Grundsatzproblem ist: Hinsichtlich Hypoglykämien bei Typ-2-Diabetikern ist es so, dass schwerwiegende Hypoglykämien bei Typ-2-Diabetikern nicht so oft vorkommen und dass die Patientenzahl so groß bzw. die Beobachtungsdauer so lang sein müsste, um einen Unterschied darzustellen. Aus dem Grunde denke ich, dass insbesondere bei Typ-2-Diabetikern die nicht schweren Hypoglykämien auf jeden Fall in die Bewertung einbezogen werden müssen. Denn das ist der Vorteil, den eben langwirksame Insulinanaloga bei Typ-2-Diabetikern bieten.

Moderator Sawicki: Die Begründung, die Sie eben gebracht haben, ist natürlich interessant. – Bitte schön, Herr Paar.

Herr Paar: Herr Kaiser, Sie haben am Anfang gesagt, dass man jetzt irgendwie aus diesen Stellungnahmen heraus einen Weg finden könnte bzw. sollte, um das unter Umständen anders zu formulieren. Ich bin schon der Meinung, dass Sie das im Review, den Sie mit den externen Sachverständigen gemacht haben – ich habe ja heute von Herrn Sawicki gelernt, dass es die gleichen sind, die auch den Vorbericht geschrieben haben –, in der Diskussion sehr zielführend diskutiert haben. Sie gehen von dem statistisch signifikanten Vorteil, den Sie finden, aus; das ist eine Feststellung, die Sie bei einigen Endpunkten hier treffen und gefunden haben, und dann diskutieren Sie, was es für einen Bias geben kann. Ich habe unverändert den Eindruck, dass in dem Endbericht die Argumentationskette – ich mag es

falsch verstanden haben – etwas andersherum läuft, dass man eben sagt: Es sind so schlechte Daten, dass ich nicht einmal von diesem statistischen Finding am Ende des Tages ausgehen kann.

Deswegen glaube ich: Wenn Sie da noch einmal hineinschauen, könnte das ein Weg sein, das anders zu diskutieren. Sie sind immerhin Koautor.

Herr Thole: Meine Frage geht in die gleiche Richtung. Wir hatten das bei uns in der Stellungnahme auch angesprochen: die Nichtberücksichtigung der nicht schweren Hypoglykämien. In anderen Verfahren findet sie statt, und ich finde, das ist eine Inkonsistenz zwischen den Verfahren. Denn ich entnehme Ihren Ausführungen, dass das, was Sie als methodischen Grund zur Unsicherheit und nicht Nichtbelastbarkeit der Daten angeführt haben, der hier zur Ablehnung führt, ja immer wirken müsste. Es ist nämlich grundsätzlich ein Problem der Messung der Daten. Das heißt, Sie müssten diesen Effekt mehr oder weniger konstant vorliegen haben, sodass ich das nur als Bestätigung unserer Stellungnahme dazu sehe. Ich sehe, dass es in Bezug auf einen patientenrelevanten Endpunkt, der meiner Ansicht nach zu erheben ist, zu einer Ungleichbehandlung zwischen den Verfahren kommt. Ich denke, es muss zumindest dargestellt werden, warum es speziell in diesem Verfahren zu einer Nichtberücksichtigung kommt. Ich bin gespannt auf die Darstellung im Ergebnis, wie dann verfahren wird. Wir haben es schon lange diskutiert. Ich habe Sie auch so verstanden, dass Sie hier keine umfassende Antwort geben möchten.

Moderator Sawicki: Wenn ich Herrn Vervölgyi richtig verstanden habe, dann wollte er von Ihnen Vorschläge haben. Es ist nicht die Frage, ob man es nicht berücksichtigt. Es ist ganz klar wichtig für die Patienten, wenn sie eine leichte Unterzuckerung haben; keine Frage.

Herr Thole: Wie haben Sie das denn in anderen Verfahren berücksichtigt?

Moderator Sawicki: Unsere Frage an Sie ist, ob Sie uns Vorschläge unterbreiten können, wie man mit einem solchen weichen Parameter, der nicht zuverlässig erhebbar ist, umgehen kann.

Herr Thole: Ich glaube, Herr Knollmeyer hat vorhin schon gesagt, dass es deskriptiv dargestellt werden kann. Das habe ich als einen Lösungsansatz im Sinne der Best Available Evidence verstanden, dass man also die Daten nimmt und sie darstellt und auf das in den Daten möglicherweise innewohnende Unsicherheitspotenzial hinweist. Das ist ein Standardvorgehen, das ich kenne. Das halte ich für extrem pragmatisch.

Die Gegenfrage wäre: Wie sind Sie bislang in den anderen Diabetesverfahren mit diesem Problem umgegangen, das Sie scheinbar – zumindest teilweise – in anderen Verfahren gelöst haben müssen, weil Sie Daten zu nicht schweren Hypoglykämien in anderen Verfahren dargestellt haben? – Das heißt, das Problem der Ergebnisunsicherheit in Bezug auf diesen Endpunkt muss relativ ähnlich sein. Es kann nicht sein, dass das nur in diesem Verfahren hier

auf einmal ein Problem ist und dass es vorher quasi ganz anders war. Das würde ich gerne von Ihnen erfahren.

Moderator Sawicki: Es ist natürlich nicht unsere Frage an Sie, wie wir damit umgegangen sind. Ich habe die Anfrage anders formuliert.

Herr Thole: Die Antwort ist schon gekommen, wie damit umzugehen ist.

(Zuruf: Ich habe auch den Vorschlag gemacht; das können Sie nachlesen!)

Moderator Sawicki: Ja, gut so. – Herr Fritsche.

Herr Fritsche: Noch einmal zur Frage von Herrn Kaiser, wie man das diskutieren kann. Ich habe bereits gesagt, dass diese unsichere Dokumentation der leichten Hypoglykämien auf beide Therapiearme zutreffen muss. Ich sehe da keinen Unterschied.

Man kann sich aber theoretisch überlegen: Wenn Sie Detemir oder Lantus abends geben vs. NPH-Insulin abends, dann ist die höchste Hypoglykämiegefahr durch das Wirkprofil von NPH mitten in der Nacht. Eigentlich – das haben wir auch bei der 4001-Studie diskutiert – wird man da den Hauptvorteil dieser langwirksamen Insulinanaloga gar nicht dokumentieren können, weil der Patient diese leichte Hypoglykämie nachts verschläft. Das kann trotzdem Folgen haben, weil seine Hypoglykämiewahrnehmung angesichts vieler dieser leichten Hypoglykämien nachts immer schlechter wird. Wenn etwas nicht dokumentiert wird, dann geschieht dies zuungunsten der langwirksamen Insulinanaloga. Sie können also ihren Vorteil über die Nacht gar nicht ausspielen, weil wir nachts gar nicht regelmäßig messen können. Also, wir haben versucht, häufiger Nachtmessungen in dieses 4001-Protokoll einzuführen.

Da das aber eine internationale Multicenterstudie war, wurde mir gesagt: Ihr könnt nicht jede Woche fünf nächtliche Hypoglykämietestungen in das Protokoll hineinschreiben. Dann habe ich gesagt: Gut, dann wird man den Vorteil von Lantus vielleicht gar nicht sehen.

Trotzdem gab es einen eindeutigen Vorteil, was die Hypoglykämien angeht. Ich denke, der Vorteil von diesem Analoginsulin wird – wenn überhaupt – noch viel größer sein, als es in den Studien herauskam. – Das war mein Punkt.

Moderator Sawicki: Danke. – Herr Sandow.

Herr Sandow: Noch einmal Plain Language. Wenn ich über subjektiv wahrgenommene Hypoglykämien spreche – zunächst ohne biochemische Validierung, also keine biochemischen Hypoglykämien –, dann sind es Ereignisse, die von Patienten in ihrem Carnet oder irgendwo aufgenommen und protokolliert werden: Ich habe das Gefühl, ich hatte eine Hypoglykämie.

Diese Dinge sollten zumindest nach einem Artikel, der neulich im Ärzteblatt stand, durch die Randomisierung den beiden Armen gleichmäßig zugewiesen werden, wenn denn eine ausreichende Zahl von Patienten in die Studie eingeschlossen wurde. Das heißt, kann man damit nicht zunächst einmal arbeiten? Kann man nicht sagen, dass man hier eine Wahrnehmung von Patienten aufgezeichnet hat? – Dazu gibt es ja Häufigkeiten. Und dann kann man noch sagen, ob es dazu biochemische Verifizierungen gab.

Dass man diese Wahrnehmung so einfach als subjektiv verzerrt abtut, kann ich nicht verstehen. Sie ist immer verzerrt. Es gibt welche mit guter Glykämiewahrnehmung und solche, die meinen, eine Hypoglykämie zu haben, welche sich gar nicht verifizieren lässt. Es sind subjektiv empfundene Ereignisse, die aufgeschrieben wurden. Mehr ist das doch nicht.

Herr Kaiser: Kurz zu den subjektiven Daten. Da werden zwei Sachen durcheinandergeworfen: Das eine ist die Messung des Parameters selber. Es geht also darum, was man bei Darstellung einer subjektiven Befindlichkeitsstörung, also eines Symptoms, misst – das will ich nicht abqualifizieren –, das dann mit Hypoglykämien gleichgesetzt wird. Das können Hypoglykämien sein, müssen es aber nicht sein. In einer randomisierten Studie kann man erst einmal grundsätzlich davon ausgehen, dass sich dann vielleicht der Level der Ereignisse dadurch ändert, aber der Unterschied zwischen den Therapien gegebenenfalls nicht. Es sei denn, es kommt dann irgendwann so weit, dass es bei allen Patienten zu so einem Ceiling-Effekt kommt. Dann hätten Sie das Problem, aber den Unterschied dazwischen nicht. Insofern wäre diese Messung tatsächlich so, wie Sie das beschreiben, nicht unbedingt ein Problem für die Darstellung eines Therapieeffekts.

Das gilt aber nur dann, wenn nicht andere Punkte zu einer Verzerrung zwischen den Gruppen führen können. Und da kommt eben das offene Studiendesign rein. Das ist das Problem. Das heißt, es ist möglich, dass Sie aufgrund des offenen Studiendesigns – und dafür spricht die empirische Evidenz eindeutig – in einer Gruppe häufiger solche Sachen berichten als in einer anderen, obwohl es sich gar nicht um solche Symptome handelt. Das ist das Problem, mit dem man umgehen muss. Das eine ist die Verzerrung der Messung selber. Was misst man genau? – Das andere ist der Effekt zwischen den beiden Gruppen.

Mit diesem Umstand müssen wir umgehen, weil wir durchweg offene Studien haben. Das beantwortet auch die Frage von Herrn Thole. Wenn man beispielsweise auf die Glitazone guckt, dann sind es eben häufig verblindete Studien. Da kann man mit diesem Parameter anders umgehen, weil es dann die Verzerrung zwischen den Gruppen so nicht gibt. Also, das ist ein wesentlicher Punkt, der da eine Rolle spielt.

Herr Schnorpfeil: Ich möchte an der Stelle darauf hinweisen, dass die EMEA das ein bisschen anders sieht, insbesondere was die Berücksichtigung der nächtlichen Hypoglykämien anbelangt. Das hat ja auch bei Detemir dazu geführt, dass der Unterschied zwischen NPH-Insulin und Detemir eindeutig zugunsten von Detemir ausgegangen ist. Also,

die Rate der nächtlichen Hypoglykämien war deutlich niedriger als beim NPH-Insulin. Das ist auch entsprechend in der Fachinformation niedergelegt worden.

Also, die EMEA akzeptiert dieses Kriterium sehr wohl. Sie hat damit nicht die Probleme wie das IQWiG. Es ist aber auch ein Punkt – den möchte ich hier noch einmal formulieren –, dass letztlich das IQWiG und die Nutzenbewertung nicht hinter die Fachinformationen der EMEA zurückfallen können. Insofern denke ich, dass Sie schon daran gebunden sind, letzten Endes auch die Bewertungskriterien der EMEA an der Stelle zu berücksichtigen und die Fachinformationen entsprechend einfließen zu lassen.

Herr Richter: Ich finde es interessant, sich dem Phänomen leichter hypoglykämischer Ereignisse zu nähern, und es ist auch schwierig, weil man sich letzten Endes fragen muss – wenn man es summatorisch sieht –, ob es ein positives oder negatives Ereignis ist. Wenn es so ist, dass repetitive Hypoglykämien vor allem nächtlich auftreten und nicht von den Patienten wahrgenommen werden und den Patienten vielleicht irgendwann in eine schwere Hypoglykämie bringen, weil er sie nicht mehr wahrnimmt, dann könnte es ja sein, dass für ihn eine leichte Hypoglykämie, die wahrgenommen wird, eher ein positives Ereignis ist, was nicht darauf hindeutet, dass repetitiv nicht wahrgenommene aufgetreten sind. Ich bitte zu bedenken, dass man einem Ereignis, das einem primär negativ erscheint, auch einen negativen Touch gibt. Man muss es in der Gesamtheit sehen. Das ist wahrscheinlich nicht so einfach.

Moderator Sawicki: Sehr richtig. Wenn ich es richtig verstehe, könnte man eine leichte Unterzuckerung als ein positives Ereignis sehen, weil sie dann zum Beispiel zu einer Insulindosiskorrektur führt und dadurch einer schweren Unterzuckerung vorgebeugt wird.

Herr Richter: Das könnte theoretisch so sein. Ich weise nur darauf hin, dass es möglich ist.

Herr Knollmeyer: Ich möchte zwei Dinge ansprechen. – Sie, Herr Kaiser, haben gerade die Glitazone angesprochen. Zu Glitazonen gibt es Studien, die im offenen Design durchgeführt worden sind, und mir ist nicht bekannt, dass Sie dort die Ergebnisse entsprechend relativiert hätten.

Der zweite Punkt, auf den ich rekurrieren möchte, ist praktisch eine Verstärkung dessen, was Herr Schnorpfeil gesagt hat. Das Papier des Unterausschusses sagt wieder in meinem berühmten § 6 Satz 3 – diesmal ist es der erste Spiegelstrich:

„Die Bewertung des therapeutischen Nutzens erfolgt grundsätzlich auf der Basis der arzneimittelgesetzlichen Zulassung, der Fachinformation sowie Publikationen von Zulassungsberichten."

Hier stellen wir erhebliche Diskrepanzen fest und bitten Sie, diese Diskrepanzen auszugleichen.

Herr Thole: Ich komme auf den Punkt noch einmal zurück. Soweit ich das Methodenpapier des IQWiG bislang verstanden habe, gehen Sie mit den Evidenzanforderungen in Bezug auf Schadensnachweise anders um. Das heißt, Sie akzeptieren dort auch die Evidenz ab Level 2 abwärts bis runter zum Level 4 oder 5 je nach Schema, um daraus Schadenshinweise abzuleiten. Herr Kaiser schüttelt den Kopf.

Moderator Sawicki: Es kommt drauf an.

Herr Thole: Es ist aber möglich, das zu tun. Ich möchte es an einem Argument spiegeln. Ich finde es interessant, was wir gerade in der Diskussion gehört haben, dass nämlich eine leichte Hypoglykämie eventuell ein positiver Effekt ist. Ich würde es dann so verstehen, dass nur RCT-basierte Daten herangezogen werden, um diese leichten Hypoglykämien zu beschreiben. Das wäre dann ein positiver Nutzen und kein Schaden, den Sie nachweisen würden. Das wäre innerhalb dieser Denkkaskade.

Wenn man das zu Ende denkt und sagt, eine leichte Hypoglykämie ist ein positives Ereignis, dann kommen wir zu interessanten Setzungen für die Medizin. Dann ist möglicherweise die Angina pectoris ein schönes Ereignis für den Warnhinweis auf einen Herzinfarkt.

Entschuldigung, ich verstehe es nicht ganz. Ich muss ehrlich sagen, dass es hier quasi zu Deutungen kommt, die ich interessant finde. Als das gesagt wurde, hatte ich gerade wirklich das Gefühl, dass hier ein Argument im Raum steht, um das, was getan wurde, zu rechtfertigen, obwohl sehr klar im Raum steht, dass es gegen das eigene methodische Vorgehen scheinbar weitestgehend verstößt. Das ist zumindest mein Verständnis. Entschuldigung, wenn das jetzt vielleicht anzüglich war, aber das Beispiel lag mir wirklich auf der Zunge. Ich bin selber Arzt und habe einiges gelernt. Ich finde das gerade interessant. Sie hatten dem auch weitestgehend zugestimmt, und wir können uns heute vielleicht darauf einigen, dass wir die leichten Hypoglykämien aus den Nebenwirkungen streichen können, wenn Sie der Ansicht sind. Das wäre vielleicht auch ein Diskussionspunkt.

Moderator Sawicki: Noch mal ganz kurz: Was fanden Sie anzüglich?

Herr Thole: Meine Bemerkung vielleicht, als ich Ihnen das Beispiel der Angina als Vorläufer des Herzinfarktes als positiven Nutzen …

Moderator Sawicki: Dass die Angina praktisch positiv ist, weil sie den Patienten praktisch davor warnt …

Herr Thole: Dass er einen Herzinfarkt haben kann. Das ist vielleicht auch anzüglich zu verstehen, und das möchte ich auch gleich … Das war hier bewusst karikierend in den Raum gestellt, aber ich glaube, das haben auch alle hier so verstanden.

Moderator Sawicki: Das habe ich jetzt nicht als anzüglich …

Herr Thole: Umso besser. Dann freut mich das.

Moderator Sawicki: Dann möchte ich bitten, dass wir das hier – aber nicht allzu lange – ausdiskutieren. – Herr Sandow, möchten Sie direkt etwas dazu sagen?

Herr Sandow: Zwei Sätze, weil es wichtig ist, das zu unterscheiden.

Moderator Sawicki: Bitte.

Herr Sandow: Wenn wir zwischen subjektiv wahrgenommenen und aufgeschriebenen berichteten Hypoglykämien und dem Begriff der leichten Hypoglykämie unterscheiden würden, dann wäre ich sehr viel glücklicher. Denn die leichte Hypoglykämie als eine Kategorie, die erfasst oder nicht erfasst werden könnte und vielleicht mal nachts oder unbemerkt stattfindet, bringt uns hier nicht weiter. Uns bringen hier nachgewiesene, gezählte Ereignisse weiter. Insofern würde ich den Begriff „subjektive Hypoglykämie“ und dann die Unterkategorie „biochemisch bestätigte Hypoglykämie leichter Art“ akzeptieren. Dann sollten wir die Diskussion darüber, welche Bedeutung eine leichte Hypoglykämie hat – selbst wenn man sie nicht erkennt und sie vielleicht nachts stattfindet –, ausblenden. Wäre das eine Möglichkeit?

Herr Grundner: Wenn ich Herrn Richter richtig verstanden habe, dann hat er die Angst formuliert, dass eine neue Substanz X Hypoglykämien vielleicht weniger gut wahrnehmbar macht, was ich nur unterstreichen kann. Das dürfte sicherlich nicht der Fall sein.

Ich möchte nur darauf hinweisen, dass es im Fall von Insulin Detemir keinerlei Hinweis darauf gibt, dass das so ist. Wir haben durchaus Untersuchungen mit kleinen Fallzahlen, die darauf hinweisen, dass das Gegenteil der Fall ist: Wenn es Unterschiede gibt, dann gibt es sie eher im Sinne einer besseren Wahrnehmung. Das wollte ich hier anführen, weil es ein wichtiger Einwand ist.

Herr Fritsche: Ich musste erst verarbeiten, was Herr Richter gesagt hat, weil ich seine Gedankengänge nicht für möglich gehalten habe. Wenn man sich die Forschung zu den schweren und leichten Hypoglykämien anschaut, so sieht man, dass es beispielsweise aus den USA Daten von einem Herrn Daniel Cox gibt. Er hat den Patienten Computer gegeben, welche randomisiert zu irgendwelchen Zeiten des Tages gepiepst haben, und dann mussten die Patienten ihren Blutzucker messen. Dann hat man diese Daten über lange Zeit ausgewertet und gesehen, dass es allesamt Patienten mit schweren Hypoglykämien in der Vorgeschichte waren. Immer dann, wenn sich die leichten Hypoglykämien gehäuft haben, wenn also an zwei Tagen hintereinander drei oder vier leichte Hypoglykämien auftraten, dann kam die schwere Hypoglykämie. Es gab eine Häufung von leichten Hypoglykämien vor dieser Bewusstlosigkeit.

Es gibt also überhaupt keinen Anlass dafür, die leichten Hypoglykämien als Warnzeichen und als etwas Positives darzustellen. Das halte ich, mit Verlaub gesagt, dem Patienten gegenüber – ich möchte nicht zu harte Worte wählen – für problematisch.

Moderator Sawicki: „Problematisch“ ist in Ordnung.

Herr Fritsche: Als ich Ihre Bemerkung gehört habe, habe ich mich wirklich erschrocken und musste überlegen, ob er es wirklich so gemeint haben kann.

Herr Richter: Also, wenn es Sie erschrocken hat, dann ist es gut. Denn dann regt es Sie zum Nachdenken an. Es gibt schließlich immer wieder neue Gedanken in der Medizin, die man vielleicht verfolgen sollte.

Herr Fritsche: Wir sind jetzt aber nicht zum Theoretisieren da, sondern es geht um Patienten.

Herr Richter: Ich habe nur Ihre eigenen Gedanken, die Sie vorher entwickelt haben, konsekutiv weiterentwickelt. Und weil es eine Kombination aus Hypoglykämiewahrnehmung und leichtem Ereignis gibt – das haben Sie eben selbst in Bezug auf die nächtlichen Ereignisse dargestellt –, kann man sich überlegen, ob ein Zusammenhang besteht. Es ging um nicht mehr. Es baute auf Ihren eigenen Gedanken auf.

Moderator Sawicki: Ich würde die Diskussion jetzt gerne abschließen, weil …

(Herr Fritsche: Nein!)

… sie uns ein bisschen wegführt, und die anzügliche Bemerkung von Herrn Thole aufgreifen. Vielleicht wird daran deutlich, dass es als theoretische Situation durchaus sein könnte, dass es so ist. Wenn wir tatsächlich die Möglichkeit hätten, bei Patienten mit koronarer Herzerkrankung einen Herzinfarkt wirksam zu verhindern, und es gäbe sehr viele Patienten mit einer unerkannten koronaren Herzerkrankung, dann wäre Angina pectoris vielleicht günstig. Es könnte durchaus sein, dass das dann so wäre. Es ist vorstellbar.

Ich kann mir auch nicht vorstellen, dass man derzeit Pectangina als einen positiven patientenrelevanten Endpunkt beschreibt, aber wenn wir eine solche therapeutische Intervention irgendwann hätten – nehmen wir einmal an, es gäbe den Super-Stent, einen Gesamt-Stent aller Koronarien, und dann hätte man nie wieder ein Problem mit seinen Koronarien; sie würden innen ausgekleidet –, dann könnte man sich vorstellen, dass man das als Symptom so aufnimmt. Das könnte so sein, aber ich möchte das jetzt nicht vertiefen.

Herr Thole: Ich möchte eine letzte Antwort darauf geben.

Moderator Sawicki: Ja, aber ich möchte es nicht vertiefen. Es ist ein Gedankenexperiment. Es hat mit unserer Bewertung jetzt nichts zu tun.

Herr Thole: Ich verstehe Ihre Ausführungen, dass Sie sagen, es könnte so sein. Fakt ist, dass es zurzeit eher als schadenbehaftet betrachtet wird, und dementsprechend ist eine Darstellung der entsprechenden Daten über diesen Bericht zu den leichten Hypoglykämien sehr relevant. Sie kann getrennt erfolgen – auch deskriptiv – mit einer klaren Darstellung der Unsicherheit der Daten und dem Entscheider, nämlich dem G-BA, über den Impact dieser Aussagen.

Ich glaube, Herr Vervölgyi hat initial nach einer Lösung für dieses Problem gefragt. Ich stelle hiermit fest, dass das eine ganz zentrale mögliche Lösung ist, um mit dem von Herrn Vervölgyi in den Raum gestellten Problem tatsächlich umzugehen.

Moderator Sawicki: Ja, genau. Aber bitte bleiben Sie doch jetzt beim Thema.

Herr Thole: Ich bin beim Thema geblieben.

Moderator Sawicki: Wir haben in unserem Vorbericht die leichten Unterzuckerungen nicht als positives Ereignis dargestellt. Ganz im Gegenteil: Wir haben betont, dass es ein negatives Ereignis ist, welches möglichst zu vermeiden ist. Es war eben ein Gedankenexperiment.

Herr Fritsche: Ich muss noch mal Schärfe hineinbringen. Darf ich?

Moderator Sawicki: Schärfe?

Herr Fritsche: Ja.

Moderator Sawicki: Ja. Sie bringen die Schärfe hinein, und Herr Richter wird es dann wieder entschärfen. Dann machen wir Schluss mit diesem Thema, weil es uns nicht wirklich – außer einem interessanten Gedankenexperiment – weiterbringt.

Herr Fritsche: Ich habe auch nach Bildern gesucht, mit denen man diese leichten Hypoglykämien vergleichen kann. Man kann es zum Beispiel mit Beinaheunfällen bei Flugzeugen oder bei Testungen von Autos machen. Da kann man auch aus Fehlern und Funktionsstörungen lernen, und dann kann man die Autos für den Straßenverkehr besser machen. Aber wir haben es hier mit Patienten zu tun – darauf möchte ich noch mal hinweisen –, und wir können hier nicht wunderbar theoretisieren über leichte Hypoglykämien und darüber, was diese vielleicht noch an Vorteilen haben können. In meiner täglichen Praxis und in der Praxis der Patienten ist es etwas Furchtbares. Ich kann einer Hypoglykämie aber auch gar nichts Positives abgewinnen.

Moderator Sawicki: Es hat niemand vorgeschlagen, hier die patientenrelevanten Endpunkte als ein positives Ereignis unterzubringen. Den Vorschlag gibt es hier nicht.

Herr Richter: Ich wollte auf Herrn Sandow rekurrieren, weil er meiner Meinung nach einen wichtigen Punkt gebracht hat. Sie haben Daten darüber, wie häufig biochemische Hypoglykämien aufgezeichnet worden sind, und gleichzeitig darüber, wie häufig die

Patienten subjektiv über Hypoglykämien berichtet haben. Daran hätte man es abgleichen können, und man hätte eine Idee – wenn die Daten vorhanden und gut sind –, ob die Hypoglykämiewahrnehmung bei diesem Patienten verändert war. Das wäre doch ein interessanter Punkt.

(Herr Sandow: Gehört aber nicht zum Auftrag hier!)

Moderator Sawicki: Ich würde diese erkenntnistheoretisch-philosophische Diskussion gerne beenden, weil sie uns auf Abwege führt. Es war aber gut, dass man angesprochen hat, dass man negative Ereignisse durchaus auch mal anders sehen kann. – Thomas Kaiser zusammenfassend zum Schluss.

Herr Kaiser: Danke für die Hinweise. Wie gesagt, wir haben es auf Basis der Stellungnahmen als wesentlichen Punkt gesehen. Wir werden diese Informationen mitnehmen. Wir werden auch den Punkt, zusätzliche Informationen zu bereits relevanten und von uns eingeschlossenen Studien anzufragen, mitnehmen, und dann werden wir uns an Sie wenden.

Moderator Sawicki: Noch einmal als Abschluss zu Tagesordnungspunkt 3: Nicht schwere Hypoglykämien sind wichtig. Es ist schwierig, sie in Studien zu evaluieren und sie zu interpretieren.

Wir müssen uns diesem Punkt stellen und haben es auch intern hier im Institut diskutiert. Es ist nicht so, als ob es eine einfache Lösung gäbe. Hier haben wir von Ihnen außer dem Vorschlag, es deskriptiv zu beschreiben – das kann man natürlich immer machen; das ist immer eine Option –, auch nichts bekommen. Es ist durchaus eine interessante Aufgabe, solche patientenberichteten Ereignisse, die schwer zu validieren sind und unter Umständen einem erheblichen Bias unterliegen, so zu evaluieren, dass man daraus belastbare Aussagen ableitet. Das haben wir schon aufgenommen, und wir haben gedacht, dass wir das mit Ihnen kurz diskutieren. Wir sind ein bisschen vom Thema abgekommen, aber das ist auch nicht schlimm, zumal wir sehr viel Zeit für diese Erörterung haben.

Wir kommen damit zu

Tagesordnungspunkt 4:
Gemeinsame Betrachtung von Hypoglykämien und dem HbA1c-Wert

Was wollen wir da wissen?

Herr Vervölgyi: Dieser Punkt war schon Thema in der Berichtsplanerörterung letztes Jahr. Wir haben im Vorbericht einen Vorschlag für die gemeinsame Betrachtung von HbA1c und Hypoglykämien gemacht. Jetzt hat die DGG in ihrer Stellungnahme darauf hingewiesen, dass der Zusammenhang zwischen HbA1c und Hypoglykämien als Hyperbel zu verstehen ist. Sie beziehen sich damit auf die DCCT von 1993.

(Herr Fritsche: Genau!)

Unsere Frage dahin gehend ist: Wie würden Sie konkret an einem Beispiel von Daten ein Produkt aus Hypoglykämien und dem HbA1c-Wert bilden? – Das ist uns nicht ganz klar.

Moderator Sawicki: Haben Sie dazu überhaupt einen Vorschlag?

Herr Fritsche: Ja.

Moderator Sawicki: Dann wären wir dafür dankbar. – Bitte, Herr Fritsche.

Herr Fritsche: Das ist mit statistischen Methoden zu machen, indem man auf die x-Achse den HbA1c-Wert und auf die y-Achse die Häufigkeit der Hypoglykämien aufträgt und dann eine Funktion bildet und diese Funktion jeweils für NPH-Insulin oder Detemir oder Lantus macht. Dann kann man diese beiden Funktionen miteinander vergleichen.

Sie haben es gemacht, indem Sie eine Tabelle mit Pluspunkt und Minuspunkt erstellt und dann diese Plus- und Minuspunkte bewertet haben. Wenn man, wie es in der DCCT-Studie 1993 erstmals gemacht wurde, diese Funktion aufträgt, dann kann man mit mathematischen Methoden diese Funktionen miteinander vergleichen und einen Unterschied feststellen.

Moderator Sawicki: Und wie soll ich die Auswertung verstehen? Würde sie dann so laufen, dass die Bereiche pro Abnahme des HbA1c-Wertes untersucht würden? – Das wird ja keine lineare Abhängigkeit sein.

Herr Gallwitz: Man kann dann die Kurven, die Krümmung in bestimmten Bereichen beschreiben. Man kann es in Deltas beschreiben.

Herr Fritsche: Zum Beispiel würde herauskommen, dass es in einem Bereich HbA1c 13 bis 12 % keinen Unterschied geben wird zwischen den Analoga.

Moderator Sawicki: Es geht um die unteren Bereiche.

Herr Fritsche: Es geht um die unteren Bereiche, also um die Bereiche, wo wir therapeutisch hinwollen. Man kann auch die gesamte Kurve nehmen und sagen: Alles, was mehr in der linkeren unteren Ecke liegt – kurventechnisch beschrieben –, ist das Bessere, also pro HbA1c weniger Hypoglykämien.

Moderator Sawicki: Also pro HbA1c-Reduktion?

Herr Fritsche: Ja. Oder pro absolutem HbA1c.

Man könnte auch das Produkt nehmen, wobei wir dann in den oberen und unteren Kurven Verzerrungen bekämen. Ich bin kein Mathematiker, aber man macht es anschaulich, indem man HbA1c mit Hypoglykämiehäufigkeit multipliziert. Dann hat man einen gemeinsamen

Wert. Aber die Statistiker sagen, dass das in Einzelfällen gefährlich werden kann, oder es kann falsche Ergebnisse ergeben.

Moderator Sawicki: Die Wahrscheinlichkeit von Verzerrungen?

Herr Fritsche: Genau.

Herr Knollmeyer: Wir waren auch enttäuscht über die Einfachheit der Darstellung der Zusammenhänge zwischen dem HbA1c-Wert und der Hypoglykämiehäufigkeit. In der veröffentlichen Meta-Regression von Peter Mullins haben Sie solche Kennlinien für Lantus und NPH für die Häufigkeit von Hypoglykämien in Abhängigkeit vom patientenindividuellen HbA1c-Wert. Wir haben damals in der Diskussion zum Berichtsplan sehr deutlich darauf hingewiesen, dass dieses unserer Auffassung nach die bestmögliche Chance darstellt, hier die Charakteristika der unterschiedlichen Insuline möglichst genau zu beschreiben. Insofern sind wir sehr enttäuscht gewesen, dass just diese Ergebnisse vom IQWiG nicht in die Nutzenbewertung eingeschlossen worden sind.

Wir möchten noch ein weiteres Problem ansprechen, das sich aus der Vorgehensweise letztendlich ergibt. Die Akzeptanz der Analysen auf der Basis von patientenindividuellen Daten ist erst im Nachgang der mündlichen Anhörung zu den Berichtsplänen entstanden. Dann haben Sie als eines der Kriterien benannt, dass es sich um einen vollständigen Studieneinschluss bestimmter Fragestellungen handeln muss, damit solche Analysen vom IQWiG berücksichtigt werden können. Das war uns natürlich nicht bekannt, und insofern bestand gar keine Chance, diese Evidenz herzustellen.

Es ergeben sich weitere praktische Gesichtspunkte, die man in den Entscheidungsgang einbeziehen muss. Wenn man im Jahr 2002 eine solche Analyse macht, dann muss man natürlich auch die veröffentlichte Evidenz des Jahres 2002 als Entscheidungskriterium heranziehen. Dann kann man nicht hergehen und sagen: In 2007 oder 2008 hat es zur Fragestellung diese und jene Studien gegeben. Nicht alle sind eingeschlossen worden, also nehme ich diese im Jahr 2004 veröffentlichte Meta-Analyse nicht in meine Bewertung auf. – Das bedarf der jeweils zeitlichen Beurteilung. Ich glaube, da haben wir noch ein Problem, das wir irgendwo lösen müssen. Wenn man sagt, es sollten systematisch bestimmte Studien eingeschlossen werden, um Verzerrungen auszuschließen, dann sind wir völlig bei Ihnen. Aber die Beurteilung der Einzelstudie muss vermutlich differenzierter auf der Basis weiterer Fragen erfolgen, die Sie dann an uns richten, um zu klären, ob hier ein Selektionsbias vorhanden gewesen ist oder nicht.

Moderator Sawicki: Danke schön. – Den Vorschlag von der DDG haben wir verstanden, und Sie haben es auch in Ihrer Stellungnahme dargestellt.

Herr Knollmeyer: Herr Prof. Sawicki, ich glaube, wir haben noch zwei Punkte. Im Sinne der Best Available Evidence möchten wir das IQWiG dringend ersuchen …

Moderator Sawicki: Zu TOP 4?

Herr Knollmeyer: Zu TOP 4. – Ja, das sind ja zwei Veröffentlichungen von Meta-Analysen auf der Basis von patientenindividuellen Daten. An der einen ist Herr Fritsche beteiligt gewesen; diese ist zu einem früheren Zeitpunkt gemacht worden. Eine andere ist im letzten Jahr veröffentlicht worden. Da ist unsere Frage, ob die Ergebnisse vom IQWiG unter diesen Gesichtspunkten einbezogen werden. Im Sinne der Best Available Evidence halten wir das für dringend geboten.

Herr Kaiser: Damit das nicht verloren geht: Wir werden damit so umgehen, dass wir es uns noch mal genauer anschauen werden. Wir werden ähnlich wie eben in der Diskussion um die nicht schweren – ich nenne es einfach so, um bei dem Term zu bleiben; das Problem ist ja, dass der Term auch in den Studien leider nicht einheitlich benutzt wird – Hypoglykämien mit den Meta-Analysen so umgehen, dass wir abzuschätzen versuchen, wie verlässlich das Ergebnis ist. Dann wird es entsprechende Auswirkungen auf die Bewertung haben.

Moderator Sawicki: Gut. – Noch mal zwecks Klarstellung zur Best Available Evidence: Diese Grundlage, die sich übrigens auf die Behandlung individueller Patienten und nicht systematischer Übersichten bezieht, heißt nicht, dass man die Best Available Evidence immer in der Interpretation akzeptiert. Vielmehr schaut man sie sich bei der Behandlung individueller Patienten an. Nur zur Klarstellung. Das heißt nicht, dass ich etwas Bestimmtes mache, wenn ein Experte sagt, dass ich es zu machen habe. So ist das nicht. Wie gesagt, das bezog sich nicht auf systematische Übersichten, sondern auf die individuellen Patienten.

Herr Frick: Ich möchte noch mal auf Herrn Kaiser zurückkommen, weil es mir vom Verständnis her etwas kryptisch war. Wir haben hier das Problem, dass wir eigentlich eine Methode gefunden haben – da rekurriere ich ganz wesentlich auf die Diskussion zum Berichtsplan, die wir geführt haben –, dass eine Meta-Regressions-Analyse als Sonderform einer Meta-Analyse auf Basis von patientenindividuellen Daten eigentlich die optimale Art ist, um das auszuwerten. Bloß, die Daten, die wir jetzt im Evidenzpool haben – das war der wesentliche Grund, sie nicht zu berücksichtigen –, enthalten nicht alle Studien, die die Ein- und Ausschlusskriterien erfüllen bzw. nicht erfüllen. Das heißt, grundsätzlich ist es das Prinzip, um die Frage zu beantworten, aber im Moment fehlt die Auswertung unter Berücksichtigung aller von Ihnen berücksichtigten Studien.

Das, was Herr Knollmeyer eben sagte, kann ich nur unterstreichen: Letztlich war es erst ab dem Vorliegen des Vorberichts möglich, zu wissen, welche Studien Sie wie klassiert einnehmen. Es ist technisch einfach nicht möglich, in der Zeit eine solche Analyse auf patientenindividuellen Daten zu fahren, um das, was Sie eben zu den Hypoglykämien angedeutet haben, zu hinterfragen, ob man diese Daten in einer Art und Weise auswerten kann, dass sie in die Bewertung sinnvoll einfließen können. Ich denke, diese Frage stellt sich.

Herr Knollmeyer: Hierzu habe ich eine Ergänzungsfrage. In der Bewertung der Studien stratifizieren Sie nach dem gewählten Therapiekonzept, und dieser Stratos müsste sich auch in der Meta-Analyse auf der Basis von patientenindividuellen Daten wiederfinden können. Das heißt, ich muss nicht alle Studien vollständig enthalten haben, sondern ich muss für eine bestimmte Fragestellung die gesamten Daten nach dieser Vorgabe analysieren. Wie gesagt, auch das ist für uns ein Faktum, das erst mit der Veröffentlichung des Vorberichts bekannt geworden ist.

Moderator Sawicki: Gut. – Sind Sie damit einverstanden – oder gibt es einen höheren Diskussionsbedarf –, dass wir für die Unterpunkte 1 bis 3 des Tagesordnungspunktes 5 fünf bis maximal sechs Minuten verwenden? Oder werden wir länger darüber diskutieren? – Gut, dann machen wir jetzt eine Pause.

(Unterbrechung von 12:05 bis 12:30 Uhr)

Moderator Sawicki: Wir fahren nun fort. Ich hoffe, Sie haben sich für Tagesordnungspunkt 5 gestärkt, den ich hiermit aufrufe:

Tagesordnungspunkt 5: Sonstige Aspekte der Studienbewertung

Zunächst kommen wir zu Unterpunkt 1:

Tagesordnungspunkt 5.1 Bewertung fehlender Werte bei Patient-reported Outcomes.

Herr Vervölgyi: Da beziehen wir uns auf die Stellungnahme von den Firmen Sanofi-Aventis und Novo Nordisk. Sie gehen in ihren Stellungnahmen auf die Bewertung des Vorberichts bezüglich fehlender Werte bei patientenberichteten Endpunkten ein und sagen, dass hohe Nichtberücksichtigungsraten bei der Erhebung von PROs nicht unüblich seien. Wenn wir unterstellen, dass es so ist, stellt sich für uns die Frage, ob diese hohen Nichtberücksichtigungsraten die Validität der Ergebnisse weniger infrage stellen als bei anderen Endpunkten. Das wäre unsere Frage dazu.

Herr Frick: Erlauben Sie mir vielleicht die Vorabbemerkung, dass das natürlich eine Frage ist, die die Entwicklerin unseres Instruments – in unserem Fall Frau Bradley – sicherlich noch besser als ich hätte beantworten können. Sie ist leider nicht hier. Ich sehe noch einen freien Stuhl. Es wäre Platz gewesen. Also, ich kann nur die Second-best-Antwort gleich geben, weil sie selber nicht da ist.

Eine Frage, die dahinter steht, ist sicherlich: Warum fehlen die Daten? – Dafür gibt es ja verschiedene Gründe. Ich denke, es gibt Gründe, die relativ klarmachen, sie sind nicht Ursache eines Bias. Und es gibt Gründe, die die Wahrscheinlichkeit eines Bias viel größer machen. Da muss man gucken, warum von den Patienten die Daten nicht erhoben wurden.

Der Punkt, der – denke ich – sehr gering biasgefährdet ist, ist schlichtweg, dass keine validierte Sprachversion des Bogens zur Verfügung stand. Das ist völlig transparent. Das kann man nachvollziehen, und das war auch prospektiv, wenn man so will. Es war schlichtweg nicht möglich, mit diesem Instrument zu dieser Zeit diese Patienten zu evaluieren. Man müsste dann genau aufschlüsseln, wie viele dieser Patienten deshalb herausfallen. Da sehe ich – es geht ja immer um Verzerrungsrisiken – praktisch kein Verzerrungsrisiko. Ich sehe, dass Sie nicken. Bei den anderen ist eher ein Problem der Verzerrungsrisiken gegeben. Das konstatiere ich völlig. Da muss man gucken, inwieweit es Ungleichverteilungen zwischen den Armen, wie groß es wirklich ist und ob es im Einzelfall dafür eventuell auch Gründe gibt. Da muss man in die Daten gucken. Und wie viele sind es überhaupt? Wie viele bleiben noch übrig, bei denen es nicht die fehlende Version des Bogens war?

Das Gleiche gilt übrigens für das Instrument, wenn es Patienten sind, die in der Auswaschphase des Bogens sind. Ich glaube, das sind die ersten vier Wochen beim DTSQ, weil immer vier Wochen rückwirkend gefragt wird. Das heißt, die Studienabbrecher waren in den ersten vier Wochen. Das ist Manual des DTSQ. Sie sollen nicht eingehen, weil sie nicht vier Wochen Recall machen können. Auch das halte ich für das Instrument gemäß und würde denken, dass es auch da gut erklärlich ist, wenn es so ist, dass Patienten deshalb nicht eingegangen sind. Darüber muss man sicherlich auch noch diskutieren.

Also, es gibt verschiedene Gründe. Ich glaube, es gibt Gründe, die mit einem nicht vorhandenen Verzerrungsrisiko oder einem äußerst geringen Verzerrungsrisiko verbunden sind. Es gibt Gründe wie diese vier Wochen, bei denen ich es auch so sehe bzw. bei denen es eigentlich keine andere Möglichkeit gibt, weil das Instrument vier Wochen Recall abfragt. Und diese sind dann nicht da, weil innerhalb dieser vier Wochen abgebrochen wurde. Und es gibt dann noch weitere Gründe, die man sich im Einzelfall angucken muss und bei denen sicherlich ein höheres Verzerrungspotenzial liegt. Ich weiß nicht, ob Ihnen das jetzt als Antwort genügt.

Herr Knollmeyer: Ich würde gerne noch zwei Ergänzungen machen. Für die PRO-Bewertung brauchen wir keine eigene Fallzahlschätzung; das wird für sekundäre Endpunkte üblicherweise nicht gemacht. In der Empirie ist es immer so, dass die Fallzahlen, die wir in die Studien einschließen, weit über denen liegen, die man eigentlich für sichere Aussagen von PRO-Auswertungen nehmen würde. Das heißt, auch wenn wir eine gewisse Rate an Patienten verlustig werden, weil sie ihren Bogen nicht ausfüllen oder nicht dem Doktor zurückgeben, dann haben wir immer noch eine ausreichende Power, um zuverlässige Aussagen zu machen.

Was wir in dem Kontext als schwierig angesehen haben – dazu haben wir uns auch schriftlich geäußert –, ist, dass die Ergebnisse der Fragebögen DTSQc und DTSQs zusammen ausgewertet worden sind. Das muss man trennen. Dazu gibt es auch entsprechende Hinweise

von der Autorin, weswegen wir dringend darum ersuchen, wenn man diese Ergebnisse zusammenführt, dann sind sie auch entsprechend zu beachten.

Moderator Sawicki: Danke schön. – Ist das ausreichend? – Es ist ausreichend.

Dann kommen wir zu:

Tagesordnungspunkt 5.2:
Umgang mit fehlender Evidenz

Das wird, wie ich annehme, sicherlich ein etwas ausführlicherer Punkt sein. Herr Vervölgyi, was möchten wir konkret wissen?

Herr Vervölgyi: Ich habe eine Frage an die DDG, die in ihrer Stellungnahme den Satz aufführt, dass man das Fehlen von Evidenz für einen Substanzvorteil nicht mit dem Vorliegen von Evidenz für Gleichwertigkeit gleichsetzen darf. Das sehen wir prinzipiell auch so. Deswegen unsere Frage: Wo im Bericht meinen Sie, dass Evidenz zur Bewertung des Fehlens eines Nutzens geführt hat?

Herr Gallwitz: Als wir die Stellungnahme geschrieben haben, waren wir eher allgemein von dieser Feststellung ausgegangen. Wir würden das vor allem auch noch mal auf den Punkt leichte Hypoglykämien beziehen, der für Sie methodisch schwierig anzugehen war. Beantwortet das Ihre Frage?

Herr Fritsche: Ich kann es nicht genau zitieren, aber wie wir alle wissen, liegen keine Daten für makrovaskuläre Endpunkte oder Mortalität vor. Das kommt in der Zusammenfassung so, wie wir das formuliert haben, unserer Meinung nach heraus.

Moderator Sawicki: Was kommt heraus?

Herr Fritsche: Für diese langwirksamen Analoga gibt es keinen Hinweis, dass sie die Mortalität oder die makrovaskulären Endpunkte reduzieren. Von daher bieten sie keinen Vorteil.

Moderator Sawicki: So haben Sie es verstanden?

Herr Fritsche: So habe ich es verstanden.

Moderator Sawicki: So war es aber nicht gemeint.

Herr Fritsche: Okay.

Herr Frick: Ich kann auch mit einer Formulierung aushelfen, in der diese sprachliche Schärfe meiner Meinung nach nicht ganz getroffen worden ist. Das ist die Seite 199; ich glaube, das ist in der Zusammenfassung.

„Zusammenfassend gibt es bez. der Behandlungszufriedenheit keinen Unterschied zwischen Insulin Glargin und NPH-Insulin."

Das ist angesichts der Vorgeschichte, die auf den Seiten vorher ausgebreitet wird, ein signifikanter Unterschied, der dann aus diversen Gründen relativiert wird. Daraus dann die positive Aussage, es gebe keinen Unterschied, abzuleiten, stellt die Frage, ob die Abwesenheit von Evidenz, die nicht mal hier gegeben ist, mit dem Nichtvorhandensein eines Unterschiedes gleichzusetzen ist.

Moderator Sawicki: Das ist richtig. Danke schön.

Wir kommen zu

Tagesordnungspunkt 5.3:
Bewertung der morgendlichen Gabe von Insulin Glargin

Herr Vervölgyi: Da haben wir eine Frage an die Firma Hoechst, die in ihrer Stellungnahme darauf eingeht, dass man von der Gleichwertigkeit der Therapieschemata von Insulin Glargin in morgendlicher und abendlicher Gabe ausgehen sollte. Aber in der Studie 4001 zeigt sich ein Unterschied zwischen den beiden Gruppen. Deswegen möchten wir gerne fragen, ob Sie den Standpunkt noch mal kurz vorbringen können.

Herr Sandow: Das ist nicht meine Stellungnahme. Das ist die von Herrn Knollmeyer.

Herr Knollmeyer: Dann ist es aber von Sanofi-Aventis.

Herr Sandow: Ist es auch.

Herr Vervölgyi: Ich meine, dass ich das der Stellungnahme von Hoechst entnommen habe, aber wenn Sie dazu Stellung nehmen …

Herr Knollmeyer: Ich muss zu meiner Schande gestehen, dass ich nicht richtig zugehört habe, da Sie sich an Hoechst gewendet haben. Ich möchte Sie bitten, Ihre Frage noch einmal zu wiederholen.

Herr Vervölgyi: Das kann ich gerne tun. – Es geht darum, dass in der Stellungnahme gesagt wird, dass von der Gleichwertigkeit der Therapieschemata von Insulin Glargin morgens und abends ausgegangen werden soll. In der Studie 4001 zeigen sich aber jeweils Unterschiede zu NPH-Insulin. Das sehen wir ein bisschen als Widerspruch. Deswegen wollte ich nachfragen, ob Sie Ihren Standpunkt noch mal kurz vorbringen könnten.

Herr Knollmeyer: Ich hatte es vorhin schon mal angesprochen. Der Unterschied ist ja nicht auf allen Beurteilungsebenen zwischen Morgen- und Abendarm der Studie 4001 relevant. Es gibt eine zusätzliche Studie, die extra gemacht wurde, um die Gleichwertigkeit zu beweisen.

Das ist die Studie von Herrn Standl, die wir auch im Literaturverzeichnis aufgeführt und Ihnen zur Verfügung gestellt haben. Sie unterstützt dieses noch einmal.

Bei Typ 1 wäre es die Studie von Herrn Hamann, die dreiarmig ist, also morgens, mittags und abends miteinander verglichen hat. Auch da wurde die Gleichwertigkeit der Einstellung unabhängig von der Wahl des Tageszeitpunktes beschrieben. Insofern glauben wir, dass diese Frage doch mit hochwertigen Studien ausreichend belegt worden ist.

Moderator Sawicki: Gut. Danke schön für die Klarstellung. Ich muss noch mal in die Studien hineingucken und das genau anschauen.

Dann haben diese drei Punkte doch weniger an zeitlichem Aufwand benötigt. – Bitte schön, Herr Lilienfeld-Toal.

Herr Lilienfeld-Toal: Ich würde unter TOP 5, „Sonstige Aspekte“, noch einen Gesichtspunkt ansprechen, der nicht extra ausgeführt ist, und zwar handelt es sich um die Bewertung der Ergebnisse der Therapiezufriedenheit der Patienten. Sie haben hier eben darüber diskutiert, dass irgendwelche Daten fehlen. Wir haben aber noch nicht darüber diskutiert, dass Sie der Präsentation der nummerischen Auswertung, also diesem Cohen-Index, eine Bewertung dahin gehend, dass es so niedrig ist, dass es nichts bedeutet, nachschieben. Ich möchte Sie bitten, dass Sie auf diese Bewertung verzichten und nur präsentieren, wie die Datenlage ist, dass eben Patienten eine höhere Behandlungszufriedenheit unter dem Glargin haben. So ist die Datenlage. Was man daraus machen muss, sollte man den dann folgenden Gremien überlassen.

Ich weise auch noch mal darauf hin, dass sich für die Patienten sicher die Konstellation IQWiG / G-BA so darstellt, dass die individuelle Situation des Betroffenen eben nicht ausreichend abgebildet wird, zum Beispiel in diesen Studien. Es gibt durchaus Patienten – das weiß jeder, der Diabetiker betreut; das wissen die Patienten untereinander, die sehr stark von einem Wechsel von dem einen auf Drucksache andere Insulin profitieren, also von NPH auf ein langwirksames Insulin –, die in diesen Studien sozusagen statistisch untergehen. Wenn der Gesamtmittelwert schon zeigt, dass die Patientenzufriedenheit besser wird, dann ist ganz klar, dass sich dahinter eine relativ große Gruppe von Betroffenen befindet, die eben sehr stark davon profitieren, und andere nicht.

Durch dieses System, das wir entwickelt haben – G-BA richtet sich nach den Bewertungen des IQWiG –, geht das unter. Es wird praktisch in der Therapie der Patienten die Möglichkeit, für diesen einzelnen Patienten diese Substanz zu benutzen, beseitigt. Darauf möchte ich noch mal hinweisen. Unterdrücken Sie bitte die Bewertung dieser Unterschiede, damit das dann im G-BA stattfinden kann.

Moderator Sawicki: Es geht schon allein deshalb nicht unter, weil Sie es jetzt gesagt haben. Das ist in die Stellungnahmen aufgenommen, und es wird auch an den G-BA weitergeleitet.

Das heißt, es geht keine Information verloren. Es geht nichts unter. Aber wir sind gehalten, die wissenschaftlichen Daten zu bewerten und eine Empfehlung an den Gemeinsamen Bundesausschuss diesbezüglich abzugeben. Das werden wir tun; das machen wir.

Bitte schön, Herr Frick.

Herr Frick: Frage dazu, ob wir das Thema Patientenzufriedenheit nicht zu einem eigenen Punkt machen wollen.

Moderator Sawicki: Das können wir unter „Verschiedenes“ machen.

Herr Frick: Dieses Thema hat aus meiner Sicht nämlich viele Aspekte. Dann würde ich es zurückstellen.

Herr Kaiser: Das kann man übergangsweise behandeln.

Moderator Sawicki: Ich wollte erst einmal den Punkt hinsichtlich der morgendlichen Gabe beenden.

Herr Kaiser: Das haben wir doch schon beendet, oder nicht?

Moderator Sawicki: Nein, das haben wir noch nicht beendet. Da hat sich Herr Lilienfeld-Toal zu Punkt 5.3 geäußert. Oder waren Sie schon bei „Verschiedenes“?

Herr Lilienfeld-Toal: Nein, ich war noch beim Punkt „Sonstige Aspekte der Studienbewertung“.

Moderator Sawicki: Das war die Überschrift zu Tagesordnungspunkt 5. Das haben wir jetzt aber besprochen. – Herr Sandow.

Herr Sandow: Ich denke, in den Kontext gehört es vielleicht. Ich habe es folgendermaßen formuliert, und es gehört in diese morgendliche oder abendliche Gabe und in den Vergleich mit NPH. Qualifikatoren wie fairer Vergleich sind für eine Nutzenbewertung aufgrund vorliegender Studienergebnisse nicht geeignet. Das war mein Kommentar. Ich hätte gerne, dass Sie sich zu dem Begriff „fairer Vergleich“ noch mal artikulieren.

Moderator Sawicki: Sie meinen, was ein fairer Vergleich ist?

Herr Sandow: Sie verwenden den Begriff „fairer Vergleich“.

Moderator Sawicki: Ja.

Herr Sandow: Der birgt natürlich viele Interpretationsmöglichkeiten. Der Begriff „fair“ stammt ja eigentlich aus dem Sport. Wir haben es hier nicht mit sportlichen Veranstaltungen

zu tun, sondern mit einer Sachbewertung. Entschuldigung, wenn ich das so pointiert formuliere. Deswegen wüsste ich gerne, was Sie unter „fairer Vergleich“ verstehen.

Moderator Sawicki: Das stammt nicht nur aus dem Sport. Vielleicht kann ich darauf antworten, wenn das keiner macht. Oder will das jemand machen?

Jetzt im Rahmen der EM 2008 ist die Fairness der Spieler sicherlich eine Bewertung. Aber es gibt auch faire Gerichtsverfahren und unfaire Gerichtsverfahren. Und auch andere Beurteilungen können fair oder unfair sein.

Wenn man vergleicht, ist das Wesentliche, dass es in diesem Vergleich nicht von vornherein Benachteiligungen für eine bestimmte Gruppe gibt. Wenn sich zum Beispiel – bleiben wir beim Sport – die eine Mannschaft die Stollen aussuchen darf und bei der anderen alles streng nachgemessen und kontrolliert wird, dann wäre das ein unfairer Vergleich. Deswegen muss man eine Strukturgleichheit der beiden Gruppen, die miteinander verglichen werden, herstellen. Die Strukturgleichheit ist ganz wesentlich. Das wird auch beim Fußball so angestrebt. Beantwortet das Ihre Frage?

Herr Sandow: Nicht ganz.

Moderator Sawicki: Was ist denn unklar?

Herr Sandow: Dann setze ich direkt nach und sage: Wenn ich einmal mit einmal vergleiche – das entspricht ja auch der Versorgungssituation …

Moderator Sawicki: Wenn Sie einmal mit einmal vergleichen?

(Zuruf: Die einmal tägliche Gabe!)

Ach so.

Herr Sandow: Ich präzisiere es gleich.

Wenn ich einmal …

Moderator Sawicki: Einmal was?

Herr Sandow: NPH-Insulin.

Moderator Sawicki: Okay. Einmal NPH-Insulin mit was?

Herr Sandow: Der Praxis wohlgemerkt entsprechend, dass man es also vorwiegend abends „at bedtime“ gibt. Wenn ich das mit einmal morgens Glargin vergleiche, dann frage ich mich, was daran unfair ist. Muss ich da zweimal NPH vergleichen, oder ist das nicht auch ein fairer Vergleich?

Moderator Sawicki: Darf ich eine unfaire Gegenfrage stellen?

Herr Sandow: Natürlich.

Moderator Sawicki: Wenn Sie ein Medikament, das zweimal oder dreimal pro Tag gegeben werden muss – ein Antihypertensivum –, mit einem Medikament vergleichen, das nur einmal gegeben werden muss – Nephidipin mit Ramipril –, würden Sie dann sagen, dass man eine Tablette mit einer Tablette pro Tag vergleicht, oder wäre das eher ein unfairer Vergleich? Muss man nicht betrachten, wie die Dosis verteilt ist und wie die Abdeckung der 24-Stunden-Wirksamkeit aussieht?

Herr Sandow: Mein Vorschlag ist – das hat sich aus dem, was ich gesagt habe, schon angedeutet –, von der Versorgungssituation auszugehen. Wir haben ja gesagt, eine einmal tägliche Injektion ist für die praktische Anwendung dem Diabetiker nach bekannten Publikationen die liebere Möglichkeit. Warum ist ein solcher Vergleich unfair? Was ist daran unfair, Herr Kaiser? Können Sie es bitte sagen?

Herr Kaiser: Ein paar Dinge dazu. – Zunächst einmal würde ich die Grundbehauptung infrage stellen, dass es dem Diabetiker das Liebste ist. Die Versorgungssituation ist nämlich tatsächlich eine andere. Also, die häufigste Anwendungsfrequenz von NPH ist nicht einmal – das zeigen die Daten –, sondern zweimal oder mehrmals. Das zunächst zur Versorgungssituation. Davon unbenommen kann es sein, dass es dem Diabetiker am liebsten ist. Aber in der individuellen Behandlung führt es offensichtlich dazu, dass NPH doch häufiger angewandt wird. So sind zumindest die Versorgungsdaten zur Verwendung.

Dann ist es nicht so, dass wir diese Studien grundsätzlich ausgeschlossen haben. Vielmehr können wir lediglich zu solch einem Vergleich zwischen Glargin einmal vs. NPH einmal nur sagen: Das Problem dabei ist nicht, dass es einmal gegeben wurde, sondern dass es gezwungenermaßen nur einmal gegeben werden durfte, im Gegensatz zur Studie 3006, in der es möglich war, individuell auch auf eine andere Frequenz zu gehen. Also, zu diesem gezwungenen Verhaften selbst bei der individuellen Stoffwechselsituation würde ich die Kliniker befragen. Es ist ja nicht der Alltag, dass man sagt: Wenn du Probleme mit nächtlichen Hypoglykämien hast unter NPH, aber unter NPH bleiben möchtest, dann musst du gezwungenermaßen bei NPH einmal abends bleiben. – Dann wird man doch versuchen, eine individuelle Behandlungssituation zu schaffen, die das Risiko der Hypoglykämien reduziert. Das ist das, was an dieser Stelle einen unfairen Vergleich ausmacht.

Ich gehe noch mal auf die Yokoyama-Studie ein. Hier haben Sie ein Beispiel für den unfairen Vergleich. Hier haben Sie eine Kombination aus dem Medikament und einer bestimmten Art von Behandlung, also zum Beispiel tatsächliche Dosisadaptation ja oder nein in bestimmtem Umfang. Aus solchen Studien können Sie nicht schließen, dass sich das Ergebnis allein auf das Medikament bezieht, sondern es ist eine Kombination aus Medikament und Vorgehensweise. Das heißt, der Effekt kann durch das Medikament bedingt sein, er kann aber

auch durch die unterschiedliche Vorgehensweise bedingt sein. Konkretes Beispiel: Nehmen Sie die ACCORD-Studie, die eine sehr intensive Therapie mit einer weniger intensiven Therapie verglichen hat, mit einer Mortalitätssteigerung unter der intensiveren Therapie. Dann kann man unseres Erachtens daraus nicht die Schlussfolgerung ziehen, dass Glargin die Mortalität steigert. Aber es war Bestandteil nur des intensiveren Arms, und das ist unseres Erachtens kein fairer Vergleich für eine Beurteilung von Glargin vs. NPH.

Wenn man aber sagen würde, man müsste im Grunde genommen alles vergleichen, bei dem Glargin und NPH in beiden Armen war, wäre man sofort bei der ACCORD-Studie. Denn das war eine klare Unterscheidung zwischen den beiden Armen. Unseres Erachtens ist die Studie kein Beleg für eine Mortalitätssteigerung unter Glargin.

Herr Sandow: Die Erklärung leuchtet ein. Aber die Begründung des Begriffes „fair“ ist damit für mich nicht gegeben.

Moderator Sawicki: Aus Ihrer Sicht ist das also nicht richtig. Haben Sie einen anderen Vorschlag statt „fair“?

Herr Sandow: Wenn Sie es einfach so erklären, dass es ein erzwungener Vergleich ist, dann bin ich damit einverstanden.

Herr Kaiser: Es steht in dem Fazit so beschrieben, dass es ein nicht optimierter Einsatz von NPH-Insulin ist. Das ist der Hintergrund dieser Studien.

Moderator Sawicki: Herr Knollmeyer.

Herr Knollmeyer: Ich möchte darum ersuchen, dass wir einheitliche Kriterien für alles anwenden. Bei Yokoyama ist es so gewesen, dass es sämtlich Patienten gewesen sind, die auf NPH-Insulin stabil eingestellt gewesen sind, und dann der eine Teil im NPH-Arm geblieben ist. Der andere ist im Umstellungsprozedere Glargin unterzogen worden.

So ist es in sehr vielen Studien, die wir bei Patienten, die nicht insulinnaiv gewesen sind, durchgeführt haben. Die hatten eine relativ lange Vorlaufzeit mit NPH. Und wenn ich diese stabile Therapie weiterführe, dann habe ich natürlich in der Initialphase ein geringeres Komplikationsrisiko als bei Patienten, bei denen ich einen Wechsel auf ein anderes Insulin mit einer anderen Kinetik – bei Zulassungsstudien auch noch ein Insulin, mit dem der Arzt noch keine Vorerfahrung in der Therapie hatte –, vornehme. Insofern: Wenn Sie bei Yokoyama beklagen, am NPH sei nicht geändert worden und damit seien die Stollen ungleich lang gewesen, dann sind in vielen anderen Studien auch in bestimmten Studienphasen die Stollen ungleich lang gewesen, und zwar eher zum Nachteil des Analoginsulins.

Moderator Sawicki: Danke schön. – Direkt dazu?

Herr Kaiser: Direkt dazu. – Sie haben die ideale Begründung dafür geliefert, dass man mit Kurzzeitstudien keine vernünftigen Aussagen treffen kann.

Herr Knollmeyer: Wir haben keine Kurzzeitstudien bei Ihnen in der Bewertung gefunden. Diese Effekte sind vielleicht in den ersten vier Wochen zu beobachten. Dann ist üblicherweise das Findungsphänomen beseitigt, sodass man im Nachgang zu den ersten vier Wochen von den längerfristigen Einstellungsparametern sprechen kann, die dann auch wirklich die unterschiedlichen Eigenschaften des Produktes beschreiben. Aber dafür sind 24 Wochen völlig ausreichend lang, und das Thema Studiendauer müssen wir auch noch mal grundsätzlich besprechen.

Moderator Sawicki: Herr Fritsche.

Herr Fritsche: Ich möchte noch etwas zur 4001-Studie sagen, weil es hier immer zwischen Sanofi-Aventis und Ihnen hin und her geht. Es ist bei dieser Studie so gewesen, dass auch tatsächlich die Mitautoren am Studiendesign beteiligt waren und wir damals sehr intensiv darüber gesprochen haben. Vielleicht kann ich das noch einmal erzählen. Diese Einmalgabe NPH-Insulin war natürlich ein Thema beim Design dieser Studie. Da das aber ein international agierender Pharmakonzern ist, wurden auch andere Leute gefragt, aus Finnland und aus den USA. Und diese Therapie mit dem Bedtime-Insulin kommt aus den USA, und da ist es üblich gewesen, dass man es einmal abends gibt.

Natürlich sagen Sie jetzt, dass Sie Deutschland betrachten. Aber Sie sagen gleichzeitig auch, das sei unsere Versorgungsrealität. Herr Grundner hat gesagt, dass das heute schon gar nicht mehr zutrifft. Man kann aber nicht sagen: Weil es in Deutschland alle so machen, ist es das Richtige. – Man hat damals die Studie so designt, weil es auch andere Realitäten gab und man diese auch abbilden wollte.

Sie haben aber völlig recht: Man hätte das auch mit zweimal NPH machen können. Dann wäre es eine andere Studie gewesen. Dann wären andere Vergleiche möglich gewesen und andere Vergleiche wiederum nicht.

Herr Kaiser: Ich verstehe die Problematik. Nichtsdestotrotz müssen wir dem hiesigen Entscheidungsgremium eine Empfehlung an die Hand geben, und eine Empfehlung, die primär auf einer Art der Versorgung aus einem anderen Land basiert, ist dann nicht brauchbar. Dies gilt insbesondere dann, wenn sich die Versorgung zwischen dem Land und den hiesigen Verhältnissen unterscheidet.

Bei dem, was Herr Grundner eben gesagt hat, ging es nicht um die Frequenz der Insulingabe, sondern es ging um die Häufigkeit der Kombination aus Basalinsulin und oralen Antidiabetika, zu der Daten geliefert wurden. Aber das ist eine andere Situation. Wir sprechen gerade über die Häufigkeit …

Herr Fritsche: Da ging diese Studie 4001 … Entschuldigung.

Herr Kaiser: Aber er hat nicht davon gesprochen, dass jetzt inzwischen in der überwiegenden Zahl nur noch einmal NPH angewandt wird. Er hat davon gesprochen, dass überhaupt die Kombination mit oralen Antidiabetika in ihrer Häufigkeit steigt auf Basis der Daten, die die Firma Novo Nordisk eingereicht hat. Das war eine andere Frage als die: Wenn denn Basalinsulin angewandt wird – in welchem Therapieschema auch immer –, wie häufig wird dann NPH-Insulin angewandt?

Da sind die Daten aus der deutschen Versorgungsrealität eben so – und davon müssen wir ausgehen –, dass die einmal tägliche Gabe eben nicht am häufigsten geschieht, sondern die mehrmals tägliche Gabe.

Herr Fritsche: Darf ich dazu nur einen Satz sagen?

Moderator Sawicki: Ja.

Herr Fritsche: Das ist aber jetzt inzwischen die häufigste Therapie, und das ist nicht die Folge von irgendwelchen Marketingdingen. Das liegt daran, dass sich diese Therapie in der Praxis als gut erwiesen hat bzw. von dem Patienten akzeptiert wird.

Herr Sandow: Ich habe diese Begründung in Bezug auf einmal NPH ausführlich in der Stellungnahme eingereicht.

(Herr Fritsche: Die kenne ich nicht! Deswegen!)

Es bezieht sich selbstverständlich auf die international geplante Studie und nicht auf die deutsche Situation. Insofern war mir der Begriff „fair“ eben eine Klarstellung wert.

Herr Knollmeyer: Noch mal zur Injektionsfrequenz. Sie hatten hier zur Beurteilung der Injektionsfrequenz Daten – ich glaube – aus Ulm konsultiert. Meines Wissens handelt es sich hierbei um ein Register, das ausschließlich Daten des Typ-1-Diabetes betrachtet. Da wir uns hier über den Typ-2-Diabetes austauschen, hatte ich mir erlaubt, in unserer Stellungnahme die Typ-2-Injektionshäufigkeiten so zu benennen, wie sie sich aus uns zugänglichen Marktdaten ergeben.

Was mich auch gewundert hat, ist, dass Sie nur die Quelle, aber nicht das Ergebnis dann im Vorbericht zitiert haben. Die hätte uns natürlich auch interessiert.

Moderator Sawicki: Danke schön. – Dann haben wir TOP 5 jetzt abgeschlossen, und wir kommen zu

Tagesordnungspunkt 6: Verschiedenes

Gibt es etwas, was Ihnen unbedingt auf dem Herzen liegt? – Bitte schön, Herr Thole.

Herr Thole: Ich habe in Ihrem vorliegenden Vorbericht die Probleme aus unserer Stellungnahme heraus nicht beantwortet gefunden, wie Sie mit den nicht berücksichtigten Kommentaren unter anderem von Herrn Freemantle umgehen. Diese sind in der Anhörung zum Berichtsplan nicht dokumentiert. Das empfinde ich als grundsätzliches Problem im Ablauf des Verfahrens. Da sind Lücken offen. Das fehlt mir hier, und ich kann zurzeit anhand der vorliegenden Dokumentation überhaupt nicht nachvollziehen, ob das gegebenenfalls irgendwo passiert ist. Das wäre eine Frage, die wie auch immer geklärt werden kann.

Im Rahmen der Überarbeitung Ihrer Methoden von Version 2.0 zu Version 3.0 taucht in der Präambel oder in den allgemeinen Anmerkungen, also auf der ersten Seite des Dokuments, auf, dass die Methoden Version 3.0 jetzt an die gesetzliche Forderungssituation angepasst wurden. Ich fände es hilfreich, wenn für dieses Verfahren und logischerweise auch für alle anderen Verfahren irgendwo nachvollziehbar gemacht würde, was im Verlaufe des Verfahrens wann und wie an die Forderungen des Gesetzes angepasst wurde. Ich finde es zurzeit schwer nachvollziehbar.

Moderator Sawicki: Bezogen auf die Methoden Version 3.0?

Herr Thole: Ja, von Version 2.0 zu Version 3.0 sagen Sie selber, dass die Methoden gesetzeskonform gemacht wurden. Man kann aber meiner Ansicht nach aus den bisherigen Produkten nicht nachvollziehen, wann in welchem Verfahrensschritt was passiert ist, um das zu tun. Also, ich fände es hilfreich, wenn es irgendwie aufgearbeitet würde und man das von außen irgendwie nachvollziehen könnte, was auch nötig wäre. Ich könnte mir vorstellen, dass es an einigen Stellen – das wurde auch im Vorfeld so gesagt – zu erneuten Anhörungen kommt, um das wieder darzulegen und Dinge zu besprechen, die sich neu ergeben haben. Ich glaube, für das Gesamtverfahren wäre es hilfreich, wenn das einmal zusammenhängend ausgearbeitet würde, damit man nachvollziehen kann, wie das Verfahren gelaufen ist.

In dem Zusammenhang fände ich es extrem hilfreich – das habe ich gerade heute wieder gemerkt –, wenn die Tagesordnung und die Fragen, die gestellt werden, vorab in einem angemessen Zeitraum herausgegeben werden könnten. Ich glaube nämlich, es ist im Interesse aller hier Anwesenden, dass die Fragen suffizient diskutiert werden. Ich glaube, da wird niemand Nein sagen. Und die Tatsache, dass die Fragen ad hoc im Raum gestellt werden und teilweise die Aspekte der Stellungnahme so berühren, dass man sich, obwohl man die Stellungnahme selber geschrieben und abgegeben hat, nicht en detail vorbereiten kann, setzt einen in eine sehr schwierige Situation, die wahrscheinlich auch für Sie nicht immer zufriedenstellend ist. Dann wird gesagt: Werden wir nachreichen. Ist uns jetzt erst klargeworden. – Das haben wir heute erlebt.

In dem Zusammenhang fände ich es bei diesem Auftrag auch hilfreich, wenn die durchgeführten Auftragspräzisierungen – das war heute ein ganz wichtiger Gegenstand der Diskussion – zwischen Ihnen und dem G-BA dargelegt werden könnten. Denn diese Klarstellungen – die sind quasi die Mutter des Verfahrens hier – und das, was da besprochen worden ist, also das Substrat, das dabei herauskommt, führen zu den suchtauglichen Fragen, zur Recherchestrategie und zu allen Annahmen, die in das Verfahren wirken.

Da wir jetzt – das hatten wir heute ganz akut gesagt bekommen – die Möglichkeit haben, Studien in den jeweiligen Anhörungsphasen nachzureichen, fände ich es hilfreich, wenn man wüsste, wo man steht, welche Evidenz Sie haben, welche Kriterien für Ein- und Ausschluss greifen und wie die Ein- und Ausschlussergebnisse aussehen. Ich rede noch nicht über die Datenanalysen, sondern nur darüber, dass man sieht: Diese 20 Studien sind drin, die sind draußen, und die sind deswegen draußen. – Dann kann man viel besser vorgehen und sagen: Wir sehen, dass hier etwas fehlt, und wir können die Evidenz aufbereiten – Sie möchten es doch in Ihren Methoden auch aufbereitet geliefert bekommen –, sodass man es Ihnen wirklich bringen und sagen kann: Das ist die Aufbereitung, und das ist die Begründung.

Ich denke, es wäre diskussionswert, hier umfangreicher tätig zu werden. Ideal wäre es, wenn mehr Hintergrundmaterial aus dem Verfahren da wäre, also zum Beispiel die Bewertung der Gutachter, also die Produkte. Dann könnte man das Verfahren wirklich besser begleiten. Ich glaube, wir merken alle, dass wir sehr konstruktiv sind. Wir bringen sehr viel Input. Sie haben sehr viele Fragen an uns nach dem Input, und wir möchten ihn gerne leisten. Ich sehe hier eine Lücke.

Wenn ich das Ganze für mich werte – ich habe das jetzt ein paar Mal erlebt –, komme ich zu folgender Aussage: Ich finde es immer sehr lehrreich. Ich lerne jedes Mal sehr viel hier. Ich fände es gut, wenn man die ganze Klärung zu Beginn des Verfahrens leisten könnte. Das wird durch die Methoden aus meiner Sicht auch dadurch unterstützt, dass Sie in den Methoden schreiben, dass regelhaft eine Auftragspräzisierung zwischen Ihnen und dem G-BA nötig ist. Wenn es tatsächlich so ist, dass da regelhaft präzisiert werden muss und der Auftrag herausgearbeitet wird, dann empfände ich es als logisch, dass diese Diskussion, die wir heute erleben, an den Anfang des Verfahrens gerückt würde. Von da an könnte man nämlich viel konstruktiver und besser arbeiten. Ich empfinde, hier ist eine Lösung möglich, die für alle Beteiligten das Verfahren hier viel besser durchführbar macht in der Gestalt, dass man viel besser miteinander und mit den richtigen Daten diskutieren kann auf einer ganz anderen Ebene. Das ist mein Gefühl. Ich würde mich freuen, wenn sich etwas ändern würde. Ich denke, wir alle hier möchten wie ich dafür sorgen, dass das alles zu einem guten Ergebnis kommt.

Moderator Sawicki: Sollen wir direkt kurz dazu Stellung nehmen? – Das wäre vielleicht am besten.

Was Sie jetzt sagen, ist natürlich eine ganz grundsätzliche Verfahrenskritik, dass Sie sich am Gemeinsamen Bundesausschuss als Verband Forschender Arzneimittelhersteller – andere Fachgesellschaften haben das auch gefordert – beteiligen. Ja, das kann man diskutieren. Allerdings muss man dann auch in Kauf nehmen, dass man anschließend das Ergebnis trägt. Es geht nicht beides. Man kann sich nicht quasi als Bestandteil des Systems sehen und sagen: Ich mache hier mit, aber wenn mir das Ergebnis nicht passt, gehe ich wieder raus und klage dagegen oder mache wieder etwas anderes.

Also, wenn man von vornherein konstruktiv mitmacht, dann muss man sich auch dem Verfahren und vor allem dem Verfahrensergebnis unterwerfen. Da bin ich nicht sicher, ob Sie diese gesamte Verantwortung auch für das Ergebnis des Verfahrens so voll tragen. Wenn das so ist, dann müsste man das auf der Ebene des Gemeinsamen Bundesausschusses und gegebenenfalls auch mit dem Ministerium besprechen. Dann kann man darüber reden. Ich bin hier nicht in der Position, das zu entscheiden. Wir können hier nicht für den Gemeinsamen Bundesausschuss entscheiden.

Der andere Punkt ist, dass Sie nicht nachvollziehen können, warum Studien ein- oder ausgeschlossen werden. Das erschließt sich mir nicht. Nach meinem Verständnis ist es doch genau dargestellt im Vorbericht. Es ist doch einer der Inhalte des Vorberichts, dass man genau sehen kann, welche Studie eingeschlossen oder nicht eingeschlossen worden ist.

Herr Thole: Entschuldigen Sie, dass ich Sie unterbreche.

Moderator Sawicki: Bitte.

Herr Thole: Meine Empfindung ist, dass wir heute tatsächlich eine Diskussion darüber hatten, dass die Kriterien für den faktischen Ausschluss, der erfolgt ist, der Gruppe hier – hier sitzen viele Fachleute, die das sehr gut nachvollziehen – tatsächlich nicht klar waren. Mein Gefühl ist, dass der Vorbericht genau den exakten Grund oder die Konsistenz des Grundes nicht transportiert und nicht darlegt. Das ist mein Gefühl. Sonst hätten wir heute nicht eine Stunde lang über Studien diskutiert.

Moderator Sawicki: Das ist ja der Grund, warum wir hier miteinander sprechen. Es kann ja auch sein. Sonst könnten wir uns ja solche Runden sparen, wenn sie uns nichts bringen würden. Wenn Sie uns sagen, dass Sie es nicht nachvollziehen konnten, dann nehmen wir es auf. Sie wissen ja, dass es unser Ziel ist, dass es nachvollziehbar ist. Warum Studien ein- oder ausgeschlossen werden, muss im Bericht transparent dargestellt werden. Wenn Sie jetzt hier sagen, dass das so für Sie nicht der Fall ist, dann nehmen wir das ernst und werden es überprüfen. Sollte Ihre Kritik tatsächlich zutreffend sein, dann werden wir das korrigieren.

Zu den anderen Transparenzaspekten, die Sie genannt haben. Das beruht zum Teil auf einem Missverständnis des Verfahrens hier. Wir haben keinen Bericht – so wie NICE – von externen Sachverständigen. Es ist nicht so, dass ein externer Gutachter einen Bericht erstellt, uns

diesen hier zur Verfügung stellt und wir auf dieser Basis den IQWiG-Bericht erstellen. Es ist vielmehr ein fortwährender Prozess, der auch dynamisch von jeweiligen Akteuren mit beeinflusst wird. Also, es gibt nicht den abgeschlossenen Bericht externer Sachverständiger.

Darüber hinaus haben wir übrigens so etwas wie eine Art Abstimmung mit allen, also nicht nur mit den Stellungnehmenden oder den Stellungnahmeberechtigten. Das sind die Erörterung bzw. die Anhörung und die Stellungnahmen zum Berichtsplan. Da gibt es die Möglichkeit, zu allen methodischen Aspekten, wie wir vorhaben, den Bericht zu erstellen, zu diskutieren.

Herr Thole: Kurze Anmerkung zur Methode. Das sehe ich ganz konkret in diesem Verfahren nicht gegeben; Stichwort: Freemantle. Und wenn ich mich richtig erinnere, dann wurde in der Einladung zur heutigen Erörterung gesagt, dass nur zwei pro Stellungnahmeantrag anreisen dürfen. Das heißt, es ist explizit in der Anhörung des Berichtsplans so gewesen, dass die Kommentare von Herrn Freemantle explizit nicht protokolliert wurden und damit aus meiner Sicht nicht in das Verfahren Eingang finden konnten.

Und das Zweite ist, ich schaue mir die Parteien hier an: Es fehlen mindestens drei oder vier Leute von denen, die sich angemeldet haben. Und vorhin war zu hören, dass Stellunggeber gar nicht eingeladen worden sind. Das war eine Frau Bradley. Sie ist wie Herr Freemantle englischsprachig. Ich weiß nicht, wo die Kriterien liegen. Das kann ich von außen nicht nachvollziehen. Ich sehe nur, dass wir hinsichtlich der Teilnehmerzahl heute eingeschränkt waren. In Bezug auf die vollständige Wertung der Anhörung teile ich Ihre Ansicht nicht, weil – Stichwort: Freemantle – es vom Berichtsplan zum Berichtsplan so nicht war.

Herr Kaiser: Ich glaube nicht, dass wir lange über die allgemeine Methodik diskutieren sollten. Das hier ist nämlich nicht der passende Ort. Ich würde mir wünschen, dass jeder, der einen Zettel unterschreibt, diesen auch liest und dem, was er unterschreibt, auch folgt. Mit der Unterschrift unter das Formblatt beschreiben Sie, dass Sie den Leitfaden zur Kenntnis genommen haben. Im Leitfaden sind die Kriterien des Stellungnahmeverfahrens einschließlich der Erörterung beschrieben, einschließlich auch des Umstandes, dass es hier deutschsprachig durchgeführt wird. Natürlich sind die Kommentare von Herrn Freemantle berücksichtigt worden. Die schriftliche Stellungnahme, die er eingereicht hat, ist im Berichtsplan auch abgebildet worden. Was nicht berücksichtigt worden ist – das ist vorab aber auch so beschrieben worden –, sind seine mündlichen Einlassungen, die er auf Englisch getätigt hat.

Noch mal: Ich würde mir wünschen, dass man auch liest, was man unterschreibt, und dem auch folgt. Ich denke, das sollte für alle gelten.

Zum Transparenzgebot, das Sie angemahnt haben: Ich würde mir sehr wünschen, dass zu allen Studien alle Studienberichte öffentlich gemacht würden, damit für alle die Bewertung nachvollziehbar wäre. Das ist hier nicht der Fall. Wir haben das große Glück, dass wir Studienberichte bekommen haben. Wir konnten die Daten extrahieren. Das ist sicherlich

etwas, wovon auch der Bericht profitiert. Davon unbenommen ist natürlich nichtsdestotrotz, dass es für die Öffentlichkeit so ist, dass sie sich auf unsere Extraktionsarbeit verlassen muss, und auch der Gemeinsame Bundesausschuss muss sich darauf verlassen. Ich denke, Transparenz kann nicht einseitig sein. Wenn Sie Transparenz fordern, sollte man auch fordern, dass alle Studiendaten – und zwar vollständig – offengelegt werden. Wir haben im Stellungnahmeverfahren den Hinweis auf eine längst fertiggestellte Studie bekommen, die wir eben nicht vorab bekommen haben. Diese konnten wir nicht in den Vorbericht einarbeiten. Diese konnte jetzt nicht diskutiert werden. Diese können wir nur nachträglich einarbeiten. Das ist für mich nicht Transparenz, sondern zunächst einmal aus welchen Gründen auch immer das Zurückhalten von Daten. Also, die Transparenzdiskussion darf man nicht einseitig führen.

Herr Knollmeyer: Direkt dazu. Sie fordern Transparenz ein. Sie kennen unsere Gründe, die dazu geführt haben, dass eine primäre Lieferung unmöglich gewesen ist. Dann hätte man vielleicht auch den individuellen Bedürfnissen der Firmen etwas mehr entgegenkommen müssen. Ich glaube, diese Diskussion müssen wir noch mal führen. Ich weiß auch, dass unser Unternehmen hier nicht alleine steht, sondern dass sich diese Schwierigkeit bei vielen Firmen ergibt.

Wir haben versucht, uns so compliant und förderlich, wie es uns zugelassen worden ist, zu verhalten. Wir müssen sehen, dass es andere Vorbilder gibt, mit denen es uns etwas leichter fällt, vertrauens- und schutzwürdige Informationen auszutauschen. Dem ist so. Und wie gesagt: Ich will den Punkt an der Stelle gar nicht weiter vertiefen.

Für uns hat sich aus der Intransparenz der Agenda und Ihrer Fragen die Schwierigkeit ergeben, welche Personen aus dem Unternehmen am ehesten geeignet sind, als Diskussions- und Informationspartner für Ihre offenen Fragen zur Verfügung zu stehen. Wenn ich die Agenda und die Fragen, die Herr Vervölgyi sehr spezifisch gestellt hat, vorher gewusst hätte, dann hätten wir heute eine Antwort für Sie parat gehabt. Auch das muss man letztendlich sehen. Ich glaube, damit haben wir letztendlich eine Chance vertan, auf einem ganz anderen Niveau zu diskutieren. Es ist in Unternehmen eben so, dass Kompetenz ganz unterschiedlicher Art auf sehr vielen verschiedenen Schultern ruht.

Unsere Chance wäre, wirklich den bestmöglichen Gesprächspartner für die jeweils von Ihnen zu diskutierenden Themen zugänglich zu machen. Wir würden das gerne noch weiter verbessern, aber auf der Basis der uns verfügbaren Infos können wir das derzeit nicht. Da bitten wir darum, darüber noch einmal nachzudenken, ob Sie nicht unserem schon öfter geäußerten Wunsch entsprechen und uns die Agenda im Vorhinein, die Stellungnahmen der anderen Stellungnehmer und vielleicht auch Ihre Fragen, die Sie dezidiert haben, zur Kenntnis geben. Ich glaube, damit könnte eine solche Anhörung sehr deutlich an Qualität gewinnen.

Moderator Sawicki: Das werden wir uns überlegen. Das ist zum Teil ein konstruktiver Vorschlag. – Herr Frick.

Herr Frick: Ich darf nur ganz kurz ergänzen. Ich glaube, speziell dieses Thema englische Sprache als Lingua franca der Wissenschaftlichkeit … Das waren ja Regeln, die Sie erzählt haben, warum Sie das nicht akzeptieren können, die nicht gottgegeben waren, sondern es waren interne Regeln, die Sie sich gegeben haben. Von daher mein Appell, diese Regeln vielleicht noch mal zu überdenken und zumindest die englische Sprache durchaus auch als Anhörungssprache zu akzeptieren.

Herr Kaiser: Wir stehen international nicht alleine da, dass wir die lokale Sprache als die Diskussionssprache sehen.

Herr Frick: Ja, das sieht das NICE auch so.

(Heiterkeit)

Herr Kaiser: Bei NICE ist es aber Englisch.

Herr Frick: Trotzdem denke ich, es ist erwägenswert, wenn es darum geht, alle Informationen von den bestinformierten Leuten zu bekommen.

Herr Kaiser: Das ist ja schriftlich möglich.

Herr Frick: Ich meine, es gibt einen Mehrwert dieser Anhörung. Wir erleben gerade, dass er hier eindeutig verloren geht.

Moderator Sawicki: Ja, für andere Teilnehmer der Anhörung ist es aber kein Mehrwert, sondern eine Minderung ihrer Teilnahmemöglichkeit und auch ihres Verständnisses. Und wenn wir jetzt anfangen, Simultanübersetzungen für Englisch, Französisch oder für welche Sprachen auch immer zu machen, dann geht dieser Charakter in Form dieses Gesprächs – und deshalb auch die Begrenzung der Teilnehmeranzahl – verloren. Sonst machen wir eine andere Veranstaltung. Dann machen wir die Anhörung so, wie Sie es im Bundestag wahrscheinlich schon gesehen haben. Das ist doch eine andere Struktur. Dann sitzen Sie da, Herr Knollmeyer. Dann kommen Sie unter Umständen gar nicht dran, wie es meistens ist. Dann haben Sie sich zwei Wochen vorbereitet, aber Sie werden gar nicht gefragt. Dann können Sie sich nicht so äußern, wie Sie es jetzt hier können, und es kommt auch nicht zu einem Gespräch. Dann wird jeder eingeladen, der stellungnahmeberechtigt ist, und es gibt keine Beschränkung.

Herr Kaiser: Das können wir uns natürlich überlegen.

Moderator Sawicki: Nein, das wollen wir nicht. Wir wollen ein Gespräch. Wir wollen einen Austausch. Wir wollen verstehen, warum bestimmte Stellungnehmer es so und nicht anders gesagt haben. Sonst hat diese Veranstaltung hier keinen wirklichen Sinn. Das ist der Hintergrund. – Herr Paar meldet sich schon ganz lange.

Herr Paar: Erst einmal glaube ich, dass Sie ein Beispiel im eigenen Haus haben, an dem Sie sehr gut die Mischung zwischen heute und dem Bundestag haben. Das war nämlich die Hochdruck-Anhörung. Da haben wir in einer sehr guten – ich habe auch mit sehr vielen Kollegen gesprochen – Mischung eine sehr konstruktive Diskussion auch im großen Kreis geführt; das ist aber meine Interpretation.

Das Zweite ist: Ich würde es begrüßen, wenn Sie die Fragen vorher schicken würden. Ich finde es gut, dass Sie darüber nachdenken. Sie haben uns gerade erklärt, was fair ist. Ich würde das auch unter dem Begriff „fair“ einordnen, wenn man es vorher hätte. Wenn man den Weg dieser Vorabschickung von Fragen geht, dann kann man auch wunderbar Frau Prof. Bradley hier unterbringen. Ich denke, man kann auch mal eine Frage an sie auf Englisch formulieren, und dann geht das auch nicht unter. Also, wenn Sie darüber nachdenken, dann finde ich das klasse.

Moderator Sawicki: Ja, wir werden darüber nachdenken. Wenn wir konkrete Fragen an bestimmte Stellungnehmer haben, dann können wir sie natürlich auch außerhalb einer solchen Erörterung stellen.

Herr Paar: Bestens.

Moderator Sawicki: Dafür muss man sich hier nicht treffen.

Herr Paar: Das ist gut.

Moderator Sawicki: Dann können wir Sie anschreiben und sagen: Wir haben nicht verstanden, warum Sie es so und so machen.

Herr Paar: Das wäre wunderbar.

Herr Kaiser: Dann verzichten wir auf die Erörterung.

Moderator Sawicki: Dann ist die Erörterung in diesem Fall nicht von Bedeutung. Es geht hier um die Diskussion bestimmter Aspekte.

Herr Paar: Aber, Herr Sawicki, das muss doch nicht immer so schwarz-weiß sein. Die Fragen waren heute gut vorbereitet. Und Sie merken doch selber: Wenn Sie zu Herrn Knollmeyer sagen: Sagen Sie mir doch zu Tabelle 18 …

Moderator Sawicki: Dann sagt Herr Knollmeyer, er wird uns das schicken.

Herr Paar: Das kann man aber auch vorher eruieren.

Moderator Sawicki: Das kann man auch vorher machen.

Herr Paar: Die Zeitspanne ist die gleiche.

Moderator Sawicki: Herr Paar, das war heute an zwei oder drei Stellen so, aber das waren nicht einmal 10 % der Zeit, die hier dafür verwendet wurde, dass die Stellungnehmer gesagt haben: Dazu kann ich jetzt nichts sagen. Da muss ich zu Hause in die Daten reinschauen. Ich werde Ihnen antworten. – Das war jetzt nicht der Hauptaspekt.

Herr Paar, wir werden darüber nachdenken. Es kann sein, dass es hilfreich ist, Ihnen diese Themen mitzuteilen, damit Sie sich vorbereiten können und damit die Diskussion noch konstruktiver und noch besser vorbereitet ist. Ich habe gesagt, wir werden darüber nachdenken, ob wir das so tun.

Das Problem sind die Zeit und auch die Vorbereitung für Sie. Wir haben meistens ein paar Tage vor diesem Termin die Agenda fertig. Wenn Sie das natürlich auch vorbereiten wollen, dann verzögert das die ganze Beurteilung noch weiter. Wir sind jetzt bei anderthalb Jahren in der Beurteilungsdauer. Wir können das nicht noch weiter verzögern. Das geht nicht. Das ist nicht mehr akzeptabel. – Herr Sandow.

Herr Sandow: Direkt zu dem Punkt. Wir haben jetzt Folgendes gehört, und ich finde das alles sehr richtig: Es muss zu bestimmten Punkten eine Klarstellung geben. Nach meiner Kenntnis ist das bisher nie erfolgt in der Form, dass zu den konkreten Stellungnahmen einer bestimmten Institution eine konkrete Rückfrage kam. Ich halte das für eine sehr praktische … Klären Sie mich auf, wenn es anders ist.

Moderator Sawicki: Heute.

Herr Sandow: Ich habe mit Herrn Knollmeyer darüber gesprochen.

Herr Kaiser: Heute?

Herr Sandow: Nein. Wenn ich beispielsweise eine Stellungnahme erhalte, in der mir ein Punkt nicht klar ist, dann kann ich an den Stellungnehmenden ohne Weiteres zurückschreiben und sagen: Auf Seite 4, Zeile 15, ist mir der und der Zusammenhang nicht klar, ich brauche dazu noch zusätzliche Informationen. – Nach meiner Kenntnis wäre das eine günstige Möglichkeit, schon im Vorfeld der Erörterung Fragen zu klären, und das würde vielleicht für die Erörterung einen Gewinn bedeuten. Denn dann ist die Präzisierung der wesentlichen Fragen leichter.

Moderator Sawicki: Das tun wir. Wir fragen konkret.

Herr Sandow: Gibt es so etwas?

Moderator Sawicki: Natürlich.

Herr Sandow: Okay.

Herr Gallwitz: Ich habe noch eine Anmerkung zum vorherigen Versenden der Agenda. Ich denke, dass die Erörterung ein ganz wichtiger und wertvoller Punkt im gesamten Berichtsverfahren darstellt und dass es für die gute und suffiziente Vorbereitung unabdingbar ist, dass man vorher die Agenda hat, um sich gut vorbereiten zu können. Das gilt gerade für unsere Seite, zumal wir noch andere Jobs nebenbei haben.

Moderator Sawicki: Meinen Sie die Tagesordnung?

Herr Gallwitz: Ja, genau.

Moderator Sawicki: Okay.

Herr Gallwitz: Und das sage ich, damit ein bisschen signalisiert wird, dass man sich hier auf einer Augenhöhe und nicht auf unterschiedlicher Augenhöhe befindet. – Vielen Dank.

Moderator Sawicki: Dann darf ich mich bei Ihnen für Ihre …

(Herr Knollmeyer: Wir haben noch den Punkt „Verschiedenes“!)

Das war „Verschiedenes“.

(Herr Knollmeyer: Dann möchte ich mich noch mal melden!)

Okay.

Herr Knollmeyer: Es gibt noch ein paar Punkte, die wir gerne mit Ihnen besprochen hätten. Im Vorhinein würde ich anmerken wollen: Für mich ist der mündliche wissenschaftliche Diskurs nicht unidirektional, sondern immer bidirektional. Auch wir haben natürlich Fragen an Sie, die wir in diesem Rahmen diskutieren möchten. Ich glaube, der mündliche Diskurs ergänzt immer den schriftlichen Diskurs. Das ist zumindest nach meinem persönlichen Erleben sehr sinnvoll.

Ich möchte einige Punkte ansprechen, hinsichtlich derer es unsererseits Fragen an Sie gibt.

Moderator Sawicki: Haben Sie uns die vorher mitgeteilt?

Herr Knollmeyer: Ja, mit der Stellungnahme.

Moderator Sawicki: Okay.

Herr Knollmeyer: Das eine ist das Thema MID als Maß für die Therapiezufriedenheit. Das ist aus unserer Sicht ein veraltetes und nicht allgemein übliches Konzept, um hier festzulegen, welches Maß letztendlich als relevant eingestuft werden kann oder nicht.

Ein weiteres Thema ist die Frage der Studiendauer im Cross-over-Design. Auch hier sind wir der Auffassung, dass die zweimal 24 Wochen zu hoch festgelegt sind. Die Begründung dazu kann man auch aus den Papieren der EMEA ableiten.

Das nächste Thema, das wir auch schriftlich angesprochen haben, ist die Nonkonformität der Ergebnisse im Vergleich zu anderen HTAs und Bewertungen der Wirkstoffe, die teilweise auch mit der gleichen Aktualität entstanden sind. Da erkennen wir Unterschiede, was die eingeschlossenen Studien anbelangt. Yokoyama ist nur das einzelne Beispiel. Da gibt es weitere Unterschiede, und auch was die Unterschiede der Zielgrößen anbelangt, kommen wir wieder auf das Thema der Hypoglykämiebewertung und auf das Thema der praktischen Anwendung zurück.

Und als allerletztes Thema nenne ich die Beurteilung der Studienqualität. Auch hier haben wir einen Großteil im Verlauf schon diskutiert, aber ich glaube, wir haben noch nicht alle Gesichtspunkte angesprochen.

Moderator Sawicki: Ja, gut, aber das sind doch auch Inhalte Ihrer Stellungnahme.

Herr Knollmeyer: Ja.

Moderator Sawicki: Was ist jetzt die Frage an uns?

Herr Knollmeyer: Wie schätzen Sie das ein?

Moderator Sawicki: Das werden wir dann in der Würdigung aller Stellungnahmen wie immer als Anlage zum Abschlussbericht publizieren.

Herr Knollmeyer: Das würde ich als sehr spät empfinden.

Moderator Sawicki: Warum?

Herr Knollmeyer: Das hieße, dass wir kein weiteres diskursives Verfahren über Ihre Bewertung hätten. Da hätten wir schon ganz gerne im Rahmen dieser Anhörung von Ihnen die Rückmeldung, wie Sie das qualitativ bewertet haben, was wir letztendlich hier an Vorschlägen gemacht haben bzw. an Einwänden vorgetragen haben.

Moderator Sawicki: Ja, das können wir dann zu dem Zeitpunkt leisten, wenn wir alle Stellungnahmen gewürdigt haben. Das werden wir dann auch dem Transparenzgebot entsprechend publizieren. Das können Sie dann auch nachlesen. Und Sie können sehr wohl dazu Stellung nehmen, wenn Sie damit zum Beispiel nicht einverstanden sind – das tun Sie auch regelhaft –, gegenüber dem Gemeinsamen Bundesausschuss. Ich glaube nicht, dass es weltweit ein System gibt, das so viele Möglichkeiten der Stellungnahmen, Diskussionen, Erörterungen und Gesprächen zulässt wie dieses, das wir jetzt hier haben.

Sie haben an drei Punkten zweimal die Möglichkeit – also im Grunde sechsmal die Möglichkeit –, Ihre Meinung zu sagen und immer wieder zu sehen, welches Ergebnis es gibt, um dann noch mal Ihre Meinung zu sagen. Wenn Sie das dann nicht gesagt haben, dann kann man es auch mit einer weiteren Erörterung nicht heilen.

Herr Knollmeyer: Herr Sawicki, darf ich direkt etwas dazu sagen?

Moderator Sawicki: Sicher.

Herr Knollmeyer: Was wir leider dann in den tragenden Gründen des Gemeinsamen Bundesausschusses wiederfinden, ist, dass der Gemeinsame Bundesausschuss den Einblick in bestimmte Sachthemen verweigert, indem er praktisch apodiktisch feststellt: Das hat das IQWiG so bewertet, und das übernehmen wir so.

Moderator Sawicki: Nein.

Herr Knollmeyer: Das finden Sie sogar in den Akten, die dann im Rahmen der gerichtlichen Auseinandersetzungen entstehen. Die Schriftsätze beinhalten immer wieder diesen Passus: Das IQWiG hat festgestellt, und wir haben keinen Anlass, an dieser Feststellung zu zweifeln.

Insofern muss ich schon um Verständnis dafür werben, dass es für uns wichtig ist, bestimmte Dinge so früh wie möglich im Entstehungsprozess einer Erstattungsentscheidung in allen Facetten beleuchten zu können. Denn zu einem bestimmten Zeitpunkt ist die Bereitschaft des G-BA, hier noch Änderungen vorzunehmen, fast null.

Moderator Sawicki: Das verstehe ich. Sie haben beim IQWiG die Möglichkeit, es zweimal zu tun: einmal zum Zeitpunkt der vorläufigen Methodenerstellung, also beim Berichtsplanentwurf, und einmal – da sind wir jetzt mitten im Prozess – zum Zeitpunkt der Vorberichtsdiskussion über Stellungnahmen. Dazu können Sie alles sagen. Wir hören uns das an. Wir interpretieren das mit Ihnen zusammen hier, aber auch darüber hinaus. Wir haben darüber hinaus auch die Möglichkeit, konkret mit Ihnen Nachfragen nachzugehen. Und dann würdigen wir das im Rahmen einer Stellungnahme von uns; das kann man auch als eine Stellungnahme zu den Stellungnahmen werten. Es ist also eine Würdigung der Stellungnahmen. Das ist dann transparent. Wenn Sie damit nicht einverstanden sind, können Sie immer noch auf einer anderen Ebene sagen, dass Sie damit nicht einverstanden sind.

Sie müssen es dem Gemeinsamen Bundesausschuss gegenüber plausibel darstellen. Sie können sagen, dass wir eine Studie von Ihnen nicht berücksichtigt haben. Sie können sagen: Wir haben eine Studie eingereicht. Die ist doch so und so. Sie entspricht doch den Kriterien, aber das IQWiG hat sie nicht berücksichtigt. – Dann wird der Gemeinsame Bundesausschuss mitnichten sagen, dass das IQWiG das so entschieden hat und er uns blind folgt.

Herr Hess und auch alle anderen sagen, dass sie uns mitnichten blind folgen. Sie gucken sich das an, und wenn Ihre Einwände Hand und Fuß haben und auch gut begründet sind, dann wird

diesen selbstverständlich gefolgt. Natürlich. Dazu ist der Gemeinsame Bundesausschuss auch verpflichtet. Sie können aber nicht darauf hoffen, dass Ihre Einwände, Stellungnahmen und Vorschläge einfach so übernommen werden.

Sie können auch nicht auf einen Konsensprozess hier hoffen. Eine Beteiligung an der letztendlichen Formulierung der Empfehlung an den Gemeinsamen Bundesausschuss wird es nicht geben. Außer der Gesetzgeber macht es anders und sagt, dass die Industrie daran beteiligt werden soll. Dann ja. Dann wird sich auch vieles andere ändern. Aber erst einmal ist es so, dass wir dies tun.

Sie haben aber ganz häufig die Möglichkeit, alles zu sagen, was Sie möchten, und es wird alles gelesen, alles gehört und alles berücksichtigt. – Herr Frick.

Herr Frick: Wir sind ja jetzt bei „Verschiedenes“, und ich habe einen größeren und einen kleineren Punkt, welche ich gerne auf die Tagesordnung setzen würde. Den größeren Punkt – das ist eben schon ein bisschen angesprochen worden – würde ich Patientenzufriedenheit nennen, der ein paar Aspekte hat. Und das kleinere Thema – das ist eine Sache von zwei oder drei Minuten – wären die Meta-Analysen.

Moderator Sawicki: Gut. Dann fangen wir mit dem ersten Punkt an.

Herr Frick: Das Thema Patientenzufriedenheit hat aus meiner Sicht zwei Aspekte. Das eine ist ein ordnungspolitischer Aspekt, und der andere ist ein methodischer Aspekt.

Kurz zum ordnungspolitischen Aspekt; das war eben schon kurz angesprochen worden. Wir haben hier Empfehlungsfunktionen für den G-BA. Und das ist wieder diese Trennung zwischen Assessment und Appraisal. Das heißt, wenn Sie am Beispiel der Patientenzufriedenheit sagen, dass da zwar ein signifikanter Effekt ist, aber dass Sie aus diversen Gründen sagen – ich zitiere diesen Satz, den Sie ändern wollen: „Zusammenfassend gibt es bez. der Behandlungszufriedenheit keinen Unterschied zwischen Insulin Glargin und NPH-Insulin“, und wenn der G-BA nicht die 300 Seiten, sondern eventuell nur diesen einen Satz liest, dann mag das misleading sein.

Ich möchte einfach nur darauf hinweisen – das ist jetzt der ordnungspolitische Teil –, dass die Entscheidung natürlich eine Werteentscheidung, Ethical Judgement, ist, ob ein nachgewiesener, ein statistisch signifikanter Effekt qua seines Ausmaßes oder aus irgendeinem anderen Grund nicht relevant sei. Ich denke, dass das keine wissenschaftliche Entscheidung ist, sondern eine gesellschaftlich-mandatierte Entscheidung ist.

Moderator Sawicki: Eine wissenschaftliche Werteentscheidung.

Herr Frick: Eine gesellschaftlich-mandatierte Entscheidung, denke ich. Dafür hat man diese Trennung zwischen Assessment und Appraisal. Wenn Sie – vielleicht habe ich jetzt eine Brücke für Sie; jetzt komme ich aufs Methodische – jetzt eine methodisch richtig gute

Methode hätten – internationaler Standard oder irgendetwas, was es gut ermöglichen würde, jenseits der statistischen Signifikanz mit klinischer Relevanz umzugehen –, dann wäre das natürlich schon einmal was, aber die Situation ist nicht so. Das schreiben Sie auch in Ihrem Methodenpapier, und ich glaube, Sie finden überhaupt niemanden, der widersprechen würde. Natürlich ist es viel schwieriger. Da fehlen uns ein bisschen die Instrumente, um klare und eindeutig positivistische Aussagen zu klinischer Relevanz machen zu können.

Dann wird zum Beispiel MID diskutiert. Das diskutieren Sie, dass das Bradley-Instrument, das nicht verwendet hätte und dass es nicht selber Minimal Important Difference definiert hätte, wozu Frau Bradley eigentlich sagt: Das ist eigentlich eingepreist im Ding. Unterschiede, die wir messen, sind relevant, und letztlich kann es nur der Patient entscheiden. Von daher macht es für das Instrument keinen Sinn.

Sie gehen einen anderen Weg und gehen über Cohen's d, und bei dem Zitat von 1988, das Sie zitieren, geht es um Fallzahlkalkulation. Genau das ist es, was Cohen damals gemacht hat. Er hat eine Dicke-Daumen-Regel gemacht für: Wenn ich die Effect Size nicht genau weiß, wie kann ich dann die Fallzahl herausbekommen? – Zynisch kann man sagen: Wenn p immer 0,05 und Power immer 80 % sind, dann weiß ich nach Cohen: kleiner Effekt 400 Patienten, mittlerer Effekt 64 Patienten, großer Effekt 26 Patienten.

Viel mehr sagt einem das nicht. Dass diese Gleichung von Sozialwissenschaftlern gerne mal umgedreht und daraus eine normative Wertung abgeleitet wird – also, wenn diese dimensionslose Größe 0,2 ist, dann heißt es kleiner Effekt, wenn sie um die 0,5 ist, dann heißt es mittlerer Effekt, und um 0,8 ist großer Effekt –, versetzt Cohen – jetzt lebt er nicht mehr – vermutlich in Rotation. Er hat dem früher immer widersprochen, und das und mehrere, weil man es umdreht und etwas, was eine Dicke-Daumen-Regel für eine Fallzahlkalkulation, in eine normative Entscheidungsregel umsetzt. Das ist – glaube ich – kein optimaler Weg, um der Politik die Entscheidung abzunehmen. Das ist eine echte Schwäche von Cohen. Von daher denke ich, dass Sie einfach keine gute methodische Antwort haben, damit umzugehen.

Dann kommen wir jenseits der klinischen Relevanz …

Herr Kaiser: Können wir kurz darüber diskutieren? – Denn sonst wird es …

Herr Frick: Okay, wenn Sie mich nachher noch mal drankommen lassen. Denn ich kann nichts für die Komplexität. Sie haben nämlich ganz viele Punkte gemacht, mit denen Sie das signifikante Ergebnis Patientenzufriedenheit relativieren, und ich versuche, diese Themen … Und wenn Sie sagen, klinische Relevanz ist ein Thema, dann bin ich an dieser Stelle fertig.

Herr Kaiser: Sie haben auf ein Dilemma hingewiesen, dass es nämlich für die Beurteilung der klinischen Relevanz für die patientenberichteten Endpunkte im Grunde genommen keine international konsentierte Herangehensweise gibt. Allerdings gibt es einen mehr oder weniger vorhandenen internationalen Konsens, so auch in dem FDA-Papier, in dem Draft-Dokument,

dargestellt. Aber je nachdem, in welchem Zusammenhang man es mal sehen möchte, wird es – auch in den Stellungnahmen, die wir bekommen haben – als Draft-Dokument betitelt, aber dann ist es doch das FDA-Dokument, das einen großen Stellenwert hat. Also, jedenfalls ist es nicht von zwei Personen zusammengestellt, die mit dem Thema nichts zu tun haben. Es ist vielmehr aus einer langjährigen Diskussion auch von Leuten entstanden, die sich mit diesem Thema beschäftigen. Frau Bradley ist ja nicht die Einzige auf der Welt, die sich damit beschäftigt, sondern sie ist auch eine, die eine öffentliche Kontraposition dazu eingenommen hat.

Aber das FDA-Papier beschreibt ja das Ziel, zunächst einmal ein solches Instrument als eine klinische Relevanzbeurteilung bei solchen patientenberichteten Endpunkten zu nehmen. Es beschreibt aber auch, dass die verschiedenen Herangehensweisen zum Teil Schwächen haben und dass es zum Teil leider eben nicht definiert ist für die einzelnen Instrumente. Nichts anderes haben wir zunächst einmal für die Patientenzufriedenheit beschrieben. Es ist auch weiterhin unsere Sichtweise, dass es ein Problem ist.

Jetzt ist die Frage, wie man damit umgeht. Sie haben völlig recht: Das Verfahren von Cohen ist eigentlich für etwas anderes gedacht worden. Es ist aber nicht richtig, dass es ein Sonderweg ist, den wir hier gehen. Es ist etwas, was viele in der Erkenntnis machen, dass es eine Ermangelung von instrumentenspezifischen Definitionen für klinische Relevanz gibt. Also, wir sind da nicht alleine. Das findet nicht nur bei Sozialwissenschaftlern, sondern auch in normativen Prozessen statt.

Das ist das Problem, mit dem wir hier umzugehen haben. Jetzt kann man natürlich an dem konkreten Beispiel noch mal diskutieren, ob wir eine falsche Einschätzung an der Stelle vorgenommen haben. Und sicherlich ist richtig – deswegen sind wir auch dankbar für den Hinweis –, dass die Schlussfolgerung, die in dem einen Satz steckt, sich so nicht aus den Daten ergibt. Also, es ist tatsächlich ein Fehler in dem Bericht, den wir korrigieren werden. Das ist also eine falsche Schlussfolgerung, die wir gezogen haben. Aber grundsätzlich denke ich, dass das Thema klinische Relevanz bei patientenberichteten Endpunkten ein international diskutiertes Thema und kein Einzelfall ist, den wir hier vornehmen.

Herr Frick: Aber kein standardisiertes Thema.

Herr Kaiser: Ja, das ist das Problem. Jetzt könnte man auch sagen: Das mit den Hypoglykämien ist so nicht standardisiert, und deswegen bewertet man es nicht so oder so. – Aber wir müssen mit dem Problem umgehen. Auch da haben wir das Problem, dass in verschiedenen Studien verschiedene Operationalisierungen vorgenommen wurden. Ich erinnere auch an – ich weiß gar nicht, welcher Stellungnehmer das eingereicht hat; ich glaube, es war die Firma Novo Nordisk – die Guideline der ADA zu Hypoglykämien. Auch dort wird beklagt, dass es keine Standardisierung der Erfassung innerhalb der Studien gibt, aber man muss auch mit dem Problem umgehen. Das heißt ja nicht, dass man dann sagt: Na gut, dann ist uns alles egal, und wir nehmen die Daten, wie sie eben sind. – Man muss vielmehr

irgendwie versuchen, sich dem zu nähern, und die Stellungnahmen – jetzt komme ich auf den Punkt zurück, was wir damit machen – werden wir uns dahin gehend genau anschauen. Also, haben wir an einer Stelle eine Interpretation oder eine Herangehensweise vorgenommen, die sich im Grunde genommen so nicht halten lässt? Oder könnte man es anders machen? – Dann muss man möglicherweise zwei Wege darstellen.

Herr Frick: Ich möchte es nicht auswalzen, aber ich denke, es gäbe auch eine Möglichkeit, etwas demütiger methodisch heranzugehen und einfach zu sagen: Das sind die Daten. Das geben wir euch als G-BA an Informationen. Das ist der Effektschätzer. Das ist die Steuerung. – Es geht aber nicht um diese Werteentscheidung, weil ich den G-BA eher für mandatiert halte. Aber darüber könnten wir lange diskutieren.

Moderator Sawicki: Es ist nicht gesetzeskonform, was Sie vorschlagen.

Herr Frick: Okay.

Moderator Sawicki: Es wäre wissenschaftlich in Ordnung, geht aber gesetzlich nicht. Das Institut teilt seine Ergebnisse dem G-BA als Empfehlung mit. Und anderes machen wir …

Herr Frick: Es mangelt an meiner juristischen Bildung. Entschuldigung, dass ich den Link jetzt nicht ziehen kann. Ich will es auch nicht ausführen. Ich habe es einfach nur kommentiert, und Sie haben das letzte Wort dazu. Ich glaube, der Punkt ist rübergekommen.

Ein technisches Manko – wenn ich es so nennen darf – noch dazu: Sie haben bei Cohen einen Irrelevanzbereich von unter 0,2 gefunden. Cohen ist anders. Cohen um 0,2 ist klein, Cohen um 0,5 ist mittel, Cohen um 0,8 ist groß. Es gibt keinen Irrelevanzbereich bei Cohen. Es ist aber nicht unter 0,2 irrelevant. Das haben Sie aber wörtlich daraus gemacht. Das heißt, da haben Sie das Instrument – selbst wenn Sie akzeptieren, das anzuwenden – noch überstrapaziert. Das können Sie nicht machen. Das können Sie aus dem Ganzen nicht ableiten. Sie können es in den Bereich der kleinen Effekte nehmen, wenn Sie es machen möchten. Sie können es aber nicht in den Bereich der Irrelevanz nehmen. Denn den gibt es nicht.

Moderator Sawicki: Danke schön. Das waren jetzt beide Aspekte?

Herr Frick: Das waren jetzt der ordnungspolitische und der methodische Aspekt zum Thema klinische Relevanz. Die gute Nachricht: Zu den anderen Punkten gibt es keine ordnungspolitischen Aspekte, sondern nur rein methodische.

Es gibt leider einige andere Punkte. Das kommt daher, dass Sie sagen: Ich habe hier einen signifikanten Effekt. – Dann kommen ganz viele Argumente, die das relativieren, aber am Ende sagen Sie: Der ist aber nicht relevant.

Das wesentliche Argument war die fehlende klinische Relevanz, aber es sind noch andere. Das ist einmal, dass Sie „Verblindung“ sagen. Das führe ich jetzt nicht aus, weil es genau die Diskussion ist, die wir vorhin geführt haben: Nichtverblindbarkeit der Zuordnung bei Nichtverblindbarkeit von PROs der Datenerhebung schlichtweg, weil die Patienten natürlich wissen, was sie bekommen haben und was sie sagen. Es ist die Nichterreichung eines unerreichbaren Ziels als qualitatives Manko zu sehen. Das halte ich auch in diesem Fall – das ist wie mit den Hypoglykämien – für ein Problem.

Das Nächste ist, dass Sie sagen, es habe keine Fallzahlkalkulation und keine Adjustierung gegeben. Das kann man für jeden Punkt sagen. Das ist beinahe konstitutiv für HTAs oder fürs IQWiG. Also, wenn, dann muss Fairness überall oder nirgendwo gelten.

Zum Thema Meta-Analysen: Da haben Sie dieses Instrument DTSQs und DTSQc; es gibt ja zwei. Da ist dieses DTSQc, dieses Veränderungsinstrument, entwickelt worden. Es ist eine Antwort darauf gewesen, dass die Entwickler gesehen haben, dass dieses s-Instrument belastet ist durch einen Ceiling-Effekt. Das heißt, sie sind meistens so hoch, dass sie keinen room for improvement haben und dass es keine Sensitivität zur Verbesserung ist. Das ist ein großer Teil der Patienten, um die es hier geht.

Sie haben aber diese beiden Instrumente, die schlichtweg etwas Unterschiedliches messen, in eine Meta-Analyse gepackt. Was Sie daraus machen, ist, dass folgende Summe herauskommt: DTSQs wird etwas nach oben gezogen, und der andere, der DTSQc, wird nach unten gezogen. – Das heißt – das ist eigentlich aberwitzig –, dass zwei unterschiedliche Instrumente mit unterschiedlicher Eignung in eine gemeinsame Auswertung getan werden.

Herr Kaiser: Vielleicht dazu direkt: Ich denke, das ist ein berechtigtes Argument. Ich meine, man muss auch irgendwie mit den vorhandenen Daten umgehen, die erhoben worden sind.

(Herr Frick: Richtig!)

Gegebenenfalls kann man das so machen, dass man sich auch im Rahmen der Sensitivitätsanalyse die Dinge einzeln anschaut, dass man also zum Beispiel nur DTSQc betrachtet.

Herr Frick: Das Letzte zur Meta-Analyse ist, dass Sie umgeschwenkt haben auf die reine Auswertung nach statistisch zufälligen Effekten nach Methodenpapier 2, was ich für das Einschlägige halte, bei beiden Effekten parallel. Es sei denn, es gibt eine Begründung; diese habe ich nicht gefunden. Das heißt, auch hier hätte man erwartet, dass die Meta-Analysen sowohl nach zufälligen als auch nach festen Effekten vorhanden wären und auch dargestellt würden. Das ist hier nicht erfolgt. Hier ist nur eine konservative Auswertung nach zufälligen Effekten erfolgt.

Moderator Sawicki: Das gucken wir uns an.

Herr Frick: Damit bin ich auch schon fertig.

Moderator Sawicki: Danke schön. – Die letzte Wortmeldung des heutigen Tages kommt von Herrn Fritsche.

Herr Fritsche: Gut, dann habe ich das Schlusswort.

Moderator Sawicki: Nein, ich habe das Schlusswort.

(Heiterkeit)

Herr Fritsche: Ich habe zwei Dinge. – Das eine ist eine Frage an Sanofi-Aventis, weil es heute in der Erörterung nicht klargeworden ist. Ich habe beim amerikanischen Diabeteskongress diese Langzeitstudie zu Lantus und Retinopathie und Krebserkrankungen gesehen. Ist das nachgereicht worden? Geht es noch in die Bewertung mit ein? – Das halte ich nämlich für relevant.

Die andere Bemerkung, die ich machen wollte: Ich bin das erste und vielleicht auch das letzte Mal bei so etwas hier. Da mir die Tagesordnung nicht vorlag, hatte ich eine andere Vorstellung. Ich habe jetzt sehr viel über Cohen und so gelernt, was mir in meiner – dafür bin ich Ihnen dankbar – täglichen Arbeit nützt. Ich bin allerdings von meinen Patienten beauftragt worden, hier als ihr Fürsprecher einzutreten. Jetzt bin ich mir nicht mehr sicher, ob das hier der Ort oder die Plattform ist, weil es hier eher um die technische Durchführung der wissenschaftlichen Bewertung geht; es geht hier sehr viel um Wissenschaft. Ich möchte darum bitten, den Patienten dahinter nicht zu vergessen. Das wurde mir so aufgetragen, und nur weil ich eine Studie für eine Pharmafirma gemacht habe oder daran beteiligt war, spreche ich nicht für dieses Insulin oder für diese Insuline. Vielmehr geht es mir um die Patienten, und aufgrund der Geschichte der Insulintherapie weiß ich, dass es ein Vorteil für die Patienten ist. Hier ist zwar die Plattform für RCTs, aber es ist mir ein Bedürfnis, es als Fußnote mitzuteilen, dass ich es als einen wesentlichen Vorteil für die Versorgung der Patienten halte. Das gilt vielleicht nicht für die Patienten in Afrika, aber für die Patienten in unserem Gesundheitswesen.

Moderator Sawicki: Das ist somit festgehalten.

Herr Fritsche: Danke. – Da habe ich auch für mich gesprochen und nicht für die DDG. Da kann ich nicht für die DDG sprechen, oder?

(Heiterkeit)

Herr Gallwitz: Ich bin zwar im Vorstand der DDG, aber wir sind gemeinsam aus der DDG: Wir kommen aus der gleichen Klinik, und ich denke, dass es ganz wichtig ist, dass diese Patienteneinstellung, diese Patientensicht hier noch einmal festgehalten wird. Wenn wir auf

einem Weg zu einer Entwicklung von Diabetes.de oder Diabetes Deutschland sind, bei der die Patientenvereinigung ganz wichtig sein wird, dann kann ich das nur unterstützen.

Moderator Sawicki: Danke schön. – Dann bedanke ich mich bei allen für die engagierte Diskussion, für die Stellungnahmen, für die Kommentare. Ich wünsche allen einen guten Heimweg bzw. meinen Mitarbeitern einen schönen Resttag.

Anhang: Dokumentation der Stellungnahmen

Inhaltsverzeichnis

Seite

A 1 Stellungnahmen von Organisationen, Institutionen und Firmen ..... A 2

A 1.1 Bundesverband der Pharmazeutischen Industrie e. V. ..... A 2

A 1.2 Bundesverband der niedergelassenen Diabetologen ..... A 9

A 1.3 Deutsche Diabetes Gesellschaft ..... A 12

A 1.4 Deutscher Diabetiker Bund e. V. ..... A 17

A 1.5 Hoechst GmbH ..... A 29

A 1.6 Novo Nordisk Pharma GmbH ..... A 40

A 1.6.1 Analysen auf Basis von individuellen Patientendaten (IPD) ..... A 55

A 1.7 Sanofi-Aventis Deutschland GmbH ..... A 401

A 1.7.1 Analysen auf Basis von individuellen Patientendaten (IPD) ..... A 418

A 1.8 Verband forschender Arzneimittelhersteller e. V. ..... A 539

A 2 Stellungnahmen von Privatpersonen ..... A 545

A 2.1 Bradley Clare, Prof., PhD ..... A 545

A 2.2 Bramlage, Peter, PD Dr. ..... A 552

A 2.3 Bretzel, Reinhard, Prof. Dr. med. ..... A 561

A 2.4 Chantelau, Ernst, Prof. Dr. med. ..... A 566

A 2.5 Debusmann, Edeltraud und Horst ..... A 569

A 1 Stellungnahmen von Organisationen, Institutionen und Firmen

A 1.1 Bundesverband der Pharmazeutischen Industrie e. V.

Autoren:
Sickmüller, Barbara, Prof. Dr.
Paar, Dieter, PD Dr.
Schnorpfeil, Willi, Dr.

Adresse:
Prof. Dr. Barbara Sickmüller
Bundesverband der Pharmazeutischen Industrie e. V.
Friedrichstraße 148
10117 Berlin

Stellungnahme

Bundesverband der Pharmazeutischen Industrie e.V.
BPI
Leben ist Vielfalt

Friedrichstraße 148
10117 Berlin
www.bpi.de

Medizin und Pharmazie

Telefon: (0 30) 2 79 09-1 70
Telefax: (0 30) 2 79 09-3 70
E-Mail: bsickmueller@bpi.de

Stellungnahme zum Vorbericht „Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2“, Bewertungsvorhaben A05-03

Der Bundesverband der Pharmazeutischen Industrie e.V. (BPI) nimmt zum IQWiG-Vorbericht „Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2“ (Bewertungsvorhaben A05-03) wie folgt Stellung:

1. Auftragsbezogene Erstellung von Methoden und Kriterien, § 35 b Abs.1 S. 6 SGB V

Gemäß § 35b Abs. 1 S. 6 SGB V hat das IQWiG seine Nutzenbewertungen *auftragsbezogen* zu erstellen.

Grundlage (auch) dieses Bewertungsvorhabens ist der Auftrag des Gemeinsamen Bundesausschusses (G-BA) vom 18. Januar 2005. Dieser lautet: „*Behandlungsfelder: Diabetes mellitus Typ 1, Diabetes mellitus Typ 2*, Bluthochdruck, Asthma bronchiale, Chronisch obstruktive Lungenerkrankungen, Demenz, Depression“[1]. Daher ist die Frage nach dem Abschluss und der auftragsgemäßen Durchführung von Bewertungsaufträgen an dem Umfang des Auftrags des G-BA vom 18. Januar 2005 zu messen.

Da der vorgelegte Vorbericht A05-03 jedoch lediglich verschiedene Insulintherapien vergleicht, wird er in Ermangelung der beauftragten Gesamtbewertung des bezeichneten Behandlungsfeldes dem Auftrag des G-BA vom 18.Januar 2005 nicht gerecht. Das Bewertungsvorhaben A05-03 steht somit auch nicht in Übereinstimmung zu den gesetzlichen Anforderungen, wonach „*die Bewertung durch Vergleich mit anderen Arzneimitteln und Behandlungsformen*“ zu erfolgen hat, § 35 b Abs. 1 S. 3 SGB V.

2. Sachgerechtigkeit der Trennung der Indikationsgebiete

Es ist zudem zweifelhaft, ob die getrennte Bewertung der Indikationsgebiete Diabetes mellitus Typ 1 und 2 in der Umsetzung durch den Gemeinsamen Bundesausschuss (G-BA) im Rahmen der Arzneimittelrichtlinien (z.B. in Form von Festbetragsgruppen/Festbeträgen) oder auch im Rahmen der Festsetzung eines Erstattungshöchstbetrages zu sachgerechten Ergebnissen führen kann. Denn mit einer indikationsspezifischen Bewertung ist insbesondere eine umfassende Bewertung der Kosten-Nutzen-Relation eines Arzneimittels nicht möglich. So trennt z.B. der rapid report zu inhalativem Insulin nicht zwischen Diabetes mellitus Typ 1 und Typ 2.

[1] Abrufbar unter http://www.g-ba.de/downloads/39-261-383/2005-01-18-Auftraege-IQWiG.pdf

Zudem wird dadurch auch systematisch Evidenz ausgeblendet. Es gibt Belege, dass Patienten von einer Behandlung mit Insulinanaloga (kurz- und langwirksam) in intensivierten Therapieregimes (Basis-Bolus-Therapie) gegenüber einer Behandlung mit Humaninsulin (kurz- und langwirksam) profitieren[2]. Diese Fallkonstellation wird aber durch die Trennung sowohl in Indikationen als auch in kurz- und langwirksame Insulin/Insulinanaloga nicht analysiert. Gerade bei der intensivierten Insulintherapie spielt die Interaktion und Dosierung der eingesetzten Präparate eine entscheidende Rolle. So kann die Situation entstehen, dass ein auf der Kombination modernen Insulins basierendes Therapieschema bessere Ergebnisse und damit einen zusätzlichen Nutzen gegenüber Humaninsulin liefert, vom IQWiG aber überhaupt nicht untersucht wird. Aufgrund dieser selektiven Vorgehensweise bei der Bewertung der Insulinanaloga besteht die Gefahr, dass gängige Therapieregime und die wirksamste Therapie nicht in die Bewertung eingehen und auf der Basis dieser Bewertung faktisch von der Versorgung ausgeschlossen werden.

Hier ist ein umfassendes Vorgehen, welches verschiedene Therapieschemata berücksichtigt, erforderlich.

3. Ungleichbehandlung der Wirkstoffe

Für die einzelnen Bewertungsvorhaben zur Diabetestherapie werden zudem unterschiedliche Komparatoren herangezogen: Zum Teil wird ein Wirkstoff (z.B. Glitazone bei Diabetes mellitus Typ 2, Bewertungsvorhaben A05-05-A) mit allen in Deutschland zugelassenen und verfügbaren Blutzucker senkenden Verfahren verglichen, was auch - wie bereits ausgeführt - weitestgehend den Anforderungen für die Arzneimittelbewertung gemäß § 35 b Abs. 1 S. 3 SGB V entspricht. Im Gegensatz dazu beschränkt sich das vorliegende Bewertungsvorhaben lediglich auf die vergleichende Nutzenbewertung der Insulintherapie. Neben einem Verstoß gegen den gesetzlich vorgegeben Grundsatz der umfassenden Arzneimittelbewertung führt dies zudem zur Ungleichbehandlung der zu bewertenden Therapien, zumal eine sachgerechte Begründung für die unterschiedliche Herangehensweise bei der Ausgestaltung der Fragestellung für die einzelnen Bewertungsvorhaben fehlt.

Die Ausgestaltung des vorliegenden Vorberichts ist daher in höchstem Maße inkonsistent und beinhaltet für pharmazeutische Unternehmen aufgrund der durch die Heranziehung unterschiedlicher Komparatoren erzielten Ergebnisse wettbewerbsrechtliche Implikationen.

4. Bindungswirkung der Zulassung

---

[2] 1) Rašlová K, Bogoev M, Raz I, Leth G, Gall MA, Hâncu N. Insulin detemir and insulin aspart: a promising basal-bolus regimen for type 2 diabetes. Diabetes Research and Clinical Practice 2004; 66:193-201

2) Rašlová K, Bogoev M, Raz I, Leth G, Gall MA, Hâncu N. Corrigendum to “Insulin detemir and insulin aspart: A promising basal-bolus regimen for type 2 diabetes” [Diabetes Research and Clinical Practice 66 (2004) 193-201]. Diabetes Research and Clinical Practice 2006;72:112

Soweit nutzen- (oder auch schaden)bezogene Feststellungen der bewerteten Arzneimittel bereits in der Fachinformation niedergelegt sind, ist das IQWiG an diese Feststellungen gebunden. Eine nochmalige Überprüfung der Feststellungen der Zulassungsbehörden ist dem G-BA als erstattungsrechtlich entscheidungsbefugter Institution und damit dem IQWiG als von dieser beauftragtem Gutachter entzogen:
„*Im Rahmen der Bewertung von Arzneimitteln anhand der in der GKV geltenden Maßstäbe ist allerdings - im Unterschied zur Methodenanerkennung - die arzneimittelrechtliche Zulassung eines Wirkstoffes zu beachten, bei der gemäß § 21 Abs.2 AMG Qualität, Wirksamkeit und medizinische Unbedenklichkeit des Wirkstoffes für die vorgesehene Indikation geprüft und abschließend bewertet werden (...). Diese Kriterien darf der GBA unter dem Aspekt des "medizinischen Nutzens" eines Arzneimittels oder Wirkstoffes nicht abweichend von der Beurteilung der für die Zulassung nach dem AMG zuständigen Behörde bewerten*." (BSG, vom 31.05.2006, B 6 KA 13/05 R -, Rn. 68ff., 71).

Diesen Grundsatz berücksichtigt auch das SGB V. Gemäß § 35 Absatz 1b Satz 4 SGB V erfolgt nämlich der Nachweis einer therapeutischen Verbesserung auch aufgrund der Fachinformation.

Zu monieren ist beim vorliegenden Vorbericht in diesem Zusammenhang vor allem die Nichtbeachtung des in den jeweiligen Fachinformationen festgestellten signifikant geringeren Risikos aller und aller nächtlicher Hypoglykämien bei der Behandlung mit Insulinanaloga im Vergleich zu einer Behandlung mit NPH-Insulin. Obwohl die Zulassungsbehörden die Risikoverringerung als relevant und hinreichend belegt angesehen haben, bewertet das IQWiG diese Aspekte nochmals selbst und kommt im Unterschied zu den Zulassungsbehörden zu dem Schluss, dass die nicht schwerwiegenden Hypoglykämien aufgrund der offenen Studiendesigns „als nicht ausreichend ergebnissicher“ anzusehen seien. Daher hat das IQWiG dies auch nicht im Fazit der Nutzenbewertung berücksichtigt.

Die in den jeweiligen Fachinformationen ausgewiesenen Vorteile der langwirksamen Insulinanaloga in Bezug auf nicht schwerwiegende Hypoglykämien sind daher aufgrund der Drittbindungswirkung der Feststellungen der Zulassungsbehörden zwingend in der Nutzenbewertung zu berücksichtigen.

5. Sperrwirkung des § 31 Absatz 2a Satz 8 SGB V

Das Anwendungsgebiet von Insulindetemir wurde von der europäischen Zulassungsbehörde EMEA im März 2007 um die Kombination von Insulindetemir mit oralen Antidiabetika zugelassen. Zumindest für diese Anwendungserweiterung sperrt § 31 Absatz 2a Satz 8 SGB V eine Berücksichtigung im vorliegenden Bewertungsvorhaben. Denn gemäß § 31 Absatz 2a Satz 8 SGB V kann eine Bewertung erst erstellt werden, „*wenn hinreichende Erkenntnisse über die Wirksamkeit des Arzneimittels nach den Grundsätzen der evidenzbasierten Medizin vorliegen können.*“ Ausweislich der Gesetzesbegründung soll dem Hersteller hier nach der Zulassung genügend Zeit gegeben werden, aussagekräftige Studien zu erstellen.

Diese Regelung, die sich im Wortlaut zum einen auf die Kosten-Nutzen-Bewertung, zum anderen wiederum auch auf die Nutzenbewertung bezieht, ist mit ihrer Verortung in § 31 Absatz 2 a SGB V systematisch verfehlt. Dass der Gesetzgeber hier eben gerade nicht nur die Kosten-Nutzen-

Bewertung im Blick hatte, wird dadurch deutlich, dass die Regelung auf die „Wirksamkeit“ eines Arzneimittels (im Sinne eines Zusatznutzens) und *nicht* auf die Kosteneffektivität bzw. auf Daten zum Kosten-Nutzen-Verhältnis abstellt. Tatsächlich ist bereits die Bewertung des Langzeitnutzens eines Arzneimittels im Versorgungsalltag in der Regel direkt nach der Zulassung nicht möglich, da hier nur die Zulassungsdaten zur Verfügung stehen, die Qualität, Wirksamkeit, Sicherheit und das Nutzen-Risiko-Verhältnis (§ 1 i.V.m. § 4 Absatz 28 AMG) belegen. Genau dies ist jedoch nicht Gegenstand einer Nutzenbewertung durch das IQWiG (siehe Ausführungen zu Nr. 4). Nach Wortlaut, Sinn und Zweck verbietet § 31 Absatz 2 a Satz 8 SGB V daher bereits regelmäßig eine Nutzenbewertung direkt nach der Zulassung.

Da die relevante Erweiterung des Anwendungsbereiches von Insulindetemir in Kombination mit oralen Antidiabetika erst vor einem Jahr erfolgte, können Daten zum Langzeitnutzen noch nicht vorliegen. Mithin ist dieses Anwendungsgebiet wegen der Sperrwirkung von § 31 Absatz 2a Satz 8 SGB V von der Nutzenbewertung auszunehmen.

6. Studienauswahl

Auch für diesen Berichtsplan ist erneut zu monieren, dass für die Bewertung lediglich Studien der Evidenzklasse 1b herangezogen werden sollen. Als Grund hierfür wird ausgeführt, dass randomisierte klinische Studien (RCTs) möglich und praktisch durchführbar sind. Eine subsidiäre Heranziehung von Studien niedrigerer Evidenz für den Fall, dass RCTs zur Beantwortung der konkreten Fragestellung nicht vorliegen, ist nicht vorgesehen.

Die sogenannten *Randomised Controlled Trials* (RCTs) gelten gemeinhin als studientechnischer Goldstandard innerhalb der Ermittlung bestmöglicher Evidenz bei medizinischen Entscheidungen. Bei entsprechender Durchführungsqualität werden diese Studien mit der Evidenzklasse 1b versehen, Metaanalysen auf Grundlage dieser RCTs erhalten die Evidenzklasse 1a und somit die höchste Wertigkeit innerhalb dieser Klassifizierung gemäß Agency for Health Care Policy and Research/AHCPR und Scottish Intercollegeate Guidelines Network/ SIGN, die jedoch explizit sowohl andere Studientypen wie auch Expertenwissen als Quelle wissenschaftlicher Evidenz nennt (wie auch im Evidenzbegriff nach Sackett[3] ausdrücklich enthalten). Bei der Ausgestaltung der Einschluss- und Suchkriterien ist daher darauf zu achten, dass Metaanalysen bei der Literaturrecherche auch aufgefunden werden.

Stellen diese Studien bei entsprechender Qualität der Durchführung und Dokumentation auch eine Quelle gut kontrollierter und valider Daten dar, so bleiben naturgemäß aufgrund des Charakters und Designs dieser Studien viele Fragestellungen unbeantwortet. Hierbei besteht die grundsätzliche Frage, ob diese Verfahren allein in der Lage sind, die Therapiewirklichkeit ausreichend abzubilden. Nicht zuletzt aufgrund der Tatsache, dass RCTs oftmals multizentrisch weltweit durchgeführt werden, können sich im Einzelnen z. T. eklatante Diskrepanzen zur Therapiewirklichkeit in der täglichen Praxis ergeben. Auch Aspekte der Lebensqualität wie überhaupt die Frage nach der Erreichung individueller Therapieziele bleiben in einem Design, das primär den Nachweis von Wirksamkeit und Sicherheit zum Ziel hat, weitgehend unberücksichtigt.

[3] David Sackett, „Evidence based medicine:what it is and what it isn`t“, BMJ 1996;312:71-72

Eine strikte Beschränkung auf die höchste Evidenzstufe steht zudem im Gegensatz zu dem nunmehr anzulegenden Maßstab, wonach die auftragsbezogenen Methoden und Kriterien dem in den Fachkreisen anerkannten internationalen Standard der evidenzbasierten Medizin zu entsprechen haben, § 35 b Abs. 1 S. 5 SGB V.

Denn danach ist jeweils die für die konkrete Fragestellung *best verfügbare Evidenz* heranzuziehen[4].

Auch Jönsson[5] stellt in einer vergleichenden Bewertung der Vorgehensweise des IQWiG im Vergleich zu den Verfahren des englischen NICE und der schwedischen Erstattungsagentur LFN fest, dass die genannten Behörden zur Nutzenbewertung von Arzneimitteln alle vorhandenen Daten nutzen, da der Nutzen für den Patienten, z. B. hinsichtlich der Compliance, nicht nur in randomisierten klinischen Studien (RCTs) ermittelt werden kann.

Daher sollten auch Daten, z. B. aus nicht-interventionellen Studien, Registern oder epidemiologischen Studien bei der Nutzenbewertung mit berücksichtigt werden.

Gleiches gilt für valide intermediäre Endpunkte,.. Internationale Institute wie NICE und LIF sehen auch Modellierungen als gut akzeptierte und regelhaft notwendige Vorgehensweisen an. Sie werden auf jeden Fall auch beim Vorliegen von harten klinischen Endpunkten, z. B. Tod oder Herzinfarkt benötigt, um daraus z. B die Überlebensdauer abzuleiten.

Zur Frage der Validität der Ergebnisse von Anwendungsbeobachtungen im Vergleich zu klinischen Studien gibt es darüber hinaus eine Untersuchung, die zu dem Ergebnis gelangt, dass die Studien trotz unterschiedlichem Design zu gleichen Erkenntnissen führen[6]. Damit werden die Vorbehalte hinsichtlich der Ergebnisunsicherheit von nicht randomisierten Studien widerlegt. Im Rahmen der Untersuchung wurden 136 Meta-Analysen von in Medline und Cochrane gelisteten Publikationen für den Zeitraum 1985 bis 1998 überprüft.

7. Zusammenfassung

- Der Vorbericht erfüllt mangels umfassender vergleichender Bewertung mit anderen Arzneimitteln und Behandlungsformen insgesamt nicht den Auftrag des G-BA zur Bewertung des Behandlungsfeldes Diabetes mellitus Typ 2 und verstößt gleichzeitig auch gegen den Bewertungsgrundsatz gemäß § 35 b Abs. 1 S. 3 SGB V.

- Die Aufspaltung des Bewertungsauftrags in die Indikationen Diabetes mellitus Typ 1 und 2 und in kurz- und langwirksame Insulinanaloga führt dazu, dass bestimmte Fallkonstellationen und die in der Versorgungswirklichkeit tatsächlichen relevanten

[4] s. Fn. 3

[5] Jönson Bengt, Comments on Methods IQWiG; Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, VFA Gutachten Januar 2006

[6] Benson, K. et al., N. Engl. J. Med. 342, 1878 (2000); Concato, J. et al., N. Engl. J. Med. 342, 1887 (2000)

Bundesverband der Pharmazeutischen Industrie e.V.

BPI

Leben ist Vielfalt

Therapieregime ausgeblendet werden und so die nutzbringendste Therapie womöglich nicht identifiziert wird. Zudem ist der Nutzen einer derart selektiver Bewertung für den G-BA beschränkt.

- Durch die unterschiedliche Ausgestaltung der Fragestellungen/ Kriterien der Bewertungsvorhaben zur Diabetestherapie verstößt das IQWiG wegen Fehlens sachgerechter Begründungen zum einen gegen das Transparenzgebot und zum anderen gegen das Gebot der Gleichbehandlung, dem das IQWiG als privater Verwaltungshelfer mittelbar unterliegt. Dieses Vorgehen hat wettbewerbsrechtlich relevante Konsequenzen.

- Das IQWiG muss den in den jeweiligen Fachinformationen dokumentierten Vorteil der Insulinanaloga in Bezug auf nicht schwerwiegende Hypoglykämien gegenüber NPH-Insulin aufgrund der Drittbindungswirkung der Feststellungen der Zulassungsbehörde in der Nutzenbewertung berücksichtigen. Eine nochmalige Überprüfung ist dem G-BA als entscheidungsbefugter Institution und damit dem IQWiG als von dieser beauftragtem Gutachter entzogen.

- Aufgrund der Sperrwirkung von § 31 Absatz 2a Satz 8 SGB V ist die Bewertung von Insulindetemir in Kombination mit oralen Antidiabetika von der Bewertung auszunehmen, da die Erweiterung des Anwendungsbereiches von Insulindetemir erst vor einem Jahr erfolgte, so dass Daten zum Langzeitnutzen noch nicht vorliegen können.

- Die ausnahmslose Beschränkung der Studienauswahl auf RCTs widerspricht international anerkannten Standards der evidenzbasierten Medizin, § 35 b Abs. 1 S. 5 SGB V. Es ist die für die konkrete Fragestellung best verfügbare Evidenz in die Bewertung einzubeziehen.

Berlin, den 25. April 2008

Prof. Dr. Barbara Sickmüller | PD Dr. Dieter Paar | Dr. Willi Schnorpfeil

A 1.2 Bundesverband der niedergelassenen Diabetologen

Autoren:
Kusterer, Klaus, Prof. Dr. med. Dr. rer. nat.

Adresse:
Prof. Dr. med. Dr. rer. nat. Klaus Kusterer
P 7, 24
68161 Mannheim

Prof. Dr. med. Dr. rer. nat. Klaus Kusterer

Internistische Gemeinschaftspraxis
- Endokrinologie, Diabetes u. Stoffwechsel -

Kurfürsten-Passage P 7, 24
68161 Mannheim
Telefon: 0621 / 156 8008
Fax: 0621 / 156 8010
kusterer@endokrima.de

Mannheim, den 26.11.2008

Stellungnahme des BVND zum Vorbericht A05-03
Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2

Der BVND gratuliert dem IQWiG zu der sorgfältig durchgeführten Nutzenbewertung der langwirksamen Insulinanloga.

Die Nutzenbewertung ergab bezüglich der Zielgrößen HbA1c und Hypoglykämie für beide Insuline eine Tendenz zur Überlegenheit gegenüber NPH Insulin. Der Beleg für einen klinisch relevanten Zusatznutzen nach der Methodik ihrer Bewertung fehlt. Daraus folgt aber auch, dass es bisher keinen Hinweis auf eine Gefährdung der Patienten durch Insulinanaloga gibt. Dies ist für die niedergelassenen Diabetologen, die diese Insuline einsetzen eine wichtige Information.
Wir stimmen mit der Analyse überein, dass die überwiegende Zahl der bisher durchgeführten und im Vorbericht A05-03 bewerteten Studien für Glargin nur einen geringen Teil der in Deutschland durchgeführten Insulintherapie abbildet. Die erfahrenen Diabetologen des IQWiG wissen, welche Vorbehalte bei den Patienten bzgl. der Insulintherapie bestehen. Die BOT war bei ihrer Einführung ein innovatives Konzept, die Arzt und Patient den Einstieg in die Insulintherapie erleichtert. Der früher insbesondere auch von Hochschullehrern gelehrte Dogmatismus des „entweder oder“ d.h. orale Medikation oder Insulintherapie, hatte dazu geführt, dass viele Patienten zu spät auf Insulin eingestellt wurden. Mit der BOT ist der Einstieg erleichtert worden und der Beweis, dass dieses Konzept zumindest für einen gewissen Zeitraum funktioniert, ist erbracht. Natürlich können die durchgeführten Studien dies nicht Abbilden. Versorgungsforschung oder „real life“ Studien, die den tatsächlichen Nutzen oder Schaden einer Therapie evaluieren, haben in Deutschland keinen Stellenwert.

Auch der Vorbericht A05-03 ist für die praktische Tätigkeit der Diabetologen nur von eingeschränkten Wert in der Beurteilung ob ein Insulinanalogon einen Zusatznutzen im Vergleich zu NPH Insulin hat, da er einen methodischen Mangel aufweist. Ein wesentliches Problem in der Einstellung ist der Nüchternblutzucker. Wir denken, es ist hier nicht notwendig, dass wir die Bedeutung guter Nüchternblutzuckerwerte mit Studien belegen. NPH Insuline haben den Nachteil, dass sie zu kurz wirksam sind und ein ausgeprägtes Maximun ungefähr 4 Stunden nach der Applikation haben. Mit der konventionellen Therapie sind selten gute Nüchternblutzuckerwerte erreichbar, da die Applikation zu Abendessen zu früh in Bezug auf die Wirkdauer des NPH Insulins ist. Ein Meilenstein in der Insulintherapie war der Vorschlag das Basalinsulin vor dem Schlafengehen zu verabreichen. Erstmals war es möglich vernünftige Nüchternblutzuckerwerte bei einem Teil der Patienten zu erreichen. Trotzdem

war bei vielen Patienten eine gute Einstellung wegen schlechter Nüchternblutzuckerwerte nicht möglich. Es war die Forderung der Diabetologen „wir brauchen bessere Basalinsuline“ an die Insulinhersteller, die dazu geführt hat, dass Insulinanaloga entwickelt wurden. Bei der BOT oder ICT werden in der diabetologischen Praxis die Patienten zunächst mit NPH Insulin behandelt. Das NPH Insulin wird titriert bis zum Erreichen des individuell vorgegeben Nüchternzielwerts. Bei Nichterreichen des Zielwerts z.B. durch nächtliche Hyoglykämien oder zu weites Absinken der nächtlichen Blutzuckerwerte mit dadurch induzierter Gegenregulation wird ein Insulinanalogon eingesetzt. In den Zielgrößen der Nutzenbewertung (4.4.4 Seite 17-18) des Vorberichts vermissen wir den Nüchternblutzucker als Zielgröße und die Anzahl der Patienten mit denen sich der Nüchternblutzuckerzielwert mit NPH im Vergleich zum Analogon erreichen ließ. Diese, für den praktisch tätigen Diabetologen entscheidenden Fragen zur Beurteilung, ob ein Analog sinnvoll ist, wurden nicht untersucht. In allen in den Vorbericht einbezogenen Studien wurden die Nüchternblutzuckerwerte gemessen und in den Studien in denen auf einen vorgegebenen Nüchternblutzuckerwert titriert wurde, kann die Anzahl der Patienten, die diesen Wert nicht erreichen, ermittelt werden.

Für Langzeitnutzen und –schaden fehlen die Daten. Es macht jedoch keinen Sinn, wenn das IQWiG dauernd nicht realisierbare Studien einfordert. In der HOPE Studie konnte noch ein Nutzen von Ramipril belegt werden, in der ONTARGET Studie konnte keine Überlegenheit von Telmisartan bzgl kardiovaskulären Ereignissen gezeigt werden. In der HOPE Studie hatten z.B. 30% der Patienten ein Statin bei ONTARGET 70%. Ähnlich Unterschiede bestanden auch bzgl. Thrombozytenaggregationshemmern. Studien ohne diese begleitende Medikation sind heute ethisch nicht vertretbar. Um einen Zusatznutzen zu belegen, ist eine so große Studiendauer und Patientenanzahl erforderlich, dass diese bei dem verlangten bürokratischen Aufwand nicht mehr finanzierbar sind. Dem niedergelassenen Diabetologen nützen unrealistische Forderungen für die Praxis nichts.

Für den BVND

Prof. Dr. med. Dr. rer. nat. Klaus Kusterer
Internist, Endokrinologe, Diabetologe DDG

A 1.3 Deutsche Diabetes Gesellschaft

Autoren:
Fritsche, Andreas, Prof. Dr. med.
Gallwitz, Baptist, Prof. Dr. med.
Klein, Harald, Prof. Dr. med.
Tschritter, Otto, Dr. med.

Adresse:
Prof. Dr. med. Andreas Fritsche
Medizinische Klinik IV
Ottfried Müller Straße 10
72076 Tübingen

DEUTSCHE
DIABETES
GESELLSCHAFT

Geschäftsstelle DDG, Bürkle-de-la-Camp-Platz 1, 44789 Bochum

Ausschuss Pharmakotherapie
Prof. Dr. Harald Klein / Vorsitzender
Berufsgenossenschaftliches Universitätsklinikum
Bergmannsheil
- Klinikum der Ruhr-Universität Bochum -
Bürkle-de-la-Camp-Platz 1
44789 Bochum
Tel.: 0234/302-6400 od. -6401
Fax: 0234/302-6403
Internet: www.ddg.info

Stellungnahme vom 14. 04.08 der DDG zum Vorbericht des IQWG

„Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2“ *

Auftrag A05-03 Version 1.0

(* im Vorbericht nicht berücksichtigte Studien werden im Kontext zitiert und sind im Literaturverzeichnis **fett** gedruckt; sie liegen der Stellungnahme wie gewünscht bei)

Generelle Anmerkungen:

1.) Es ist nicht nachvollziehbar, warum Stellungnahmen in einem sehr kurzen Zeitraum erstellt werden müssen, der in keinem Verhältnis zum Umfang des Vorberichtes steht (307 Seiten). Dieses Vorgehen schließt eine vollständige wissenschaftliche Bearbeitung durch andere Parteien neben ihrer sonstigen Tätigkeit von vornherein aus.

2.) Allgemein ist anzumerken, dass in anderen Ländern ähnliche Berichte von vergleichbaren Institutionen zu anderen Bewertungen von Analoginsulinen kommen (1). Diese Berichte sollten dringend zur Kenntnis genommen werden. Dadurch wird deutlich, dass der vorliegende Bericht des IQWIG eine Außenseiterposition vertritt.

3.) Die Autoren dieser Stellungnahme widersprechen ausdrücklich der Einschätzung des IQWIG vom fehlenden Vorteil der langwirkenden Analoginsuline (Insulin Glargin und Insulin Detemir) gegenüber NPH-Insulin bei Typ 2 Diabetespatienten. Insbesondere die geringere Hypoglykämierate macht diese Insuline sicherer in der Therapie des Typ 2-Diabetes mellitus.

4.) Vorteile in der praktischen Anwendung von langwirkenden Insulinanaloga gegenüber NPH-Insulin werden nicht erwähnt und gehen somit nicht in die Bewertung ein. Hierzu gehören:

   a. Langwirkende Insulinanaloga liegen als klare Lösung vor und müssen nicht gemischt werden. Deshalb kommt es gerade bei alten oder stärker behinderten Patienten nicht zu Fehldosierungen, die in Hypo- oder Hyperglykämien enden (2).

   b. Langwirkende Insulinanaloga können sowohl morgens als auch abends gespritzt werden. Dies ermöglicht eine wesentlich patientenfreundlichere Therapie (Injektion zum Zeitpunkt, den der Patient wünscht, Injektion wird seltener vergessen, flexiblere Handhabung für zum Beispiel Pflegedienste). Solche Vorteile sind in RCTs nicht geprüft oder können schwer geprüft werden, sie sind aber deswegen nicht weniger wichtig und verbessern die Versorgung der Patienten.

Kritikpunkte im Einzelnen:

1.) Bewertung von Hypoglykämiehäufigkeit:

Das IQWIG schließt die Betrachtung von nicht schwerwiegenden Hypoglykämien aus mit der Begründung, diese seien nur mit der Komponente einer „objektiven Fremdbeurteilung ergebnissicher“. Diese Aussage missachtet den Patienten und ist zynisch. Auch eine nicht in der Bewusstlosigkeit endende Hypoglykämie ist ein schwerwiegendes negatives Ereignis für den

Deutsche Diabetes-Gesellschaft
Vorstand 2007/2008: Prof. Dr. Th. Haak (Präsident), Prof. Dr. Th. Danne (Vizepräsident), Prof. Dr. B. Gallwitz (Schatzmeister), PD Dr. O. Kordonouri (Schriftführerin), PD Dr. R. Lundershausen (Pressesprecher), Prof. Dr. H. Hauner (Tagungspräsident 2008), Prof. Dr. D. Müller-Wieland (Tagungspräsident 2009), Prof. Dr. W. Kerner, Prof. Dr. M. Tiedge
Geschäftsführer: Assessor G.-P. Buyken, **Stellv. Geschäftsführer:** R. Weichbrodt
Vereinsregister: AG Düsseldorf VR 4546, **Finanzamt:** Bochum-Mitte, **St.-Nr.:** 306/5792/0410
Bankverbindung: National-Bank AG, BLZ 360 200 30, Konto-Nr.: 64 64 777,
IBAN: DE39 3602 0030 0006 4647 77, BIC: NBAG DE 3 E

Patienten und und häufig auch für Partner/Angehörige und beeinträchtigt ihn und seine Umgebung schwer.

Es sind ferner noch die folgenden Forschungsergebnisse zu beachten:
- Die Zahl der Hypoglykämien unter einer bestimmten Therapie sind ein Hauptfaktor für die Lebensqualität und Lebenserwartung von Diabetespatienten. (3)
- Bereits eine einzige schwere Hypoglykämie kann zum Tode führen (4)
- Ein Haupt-Risikofaktor für schwere Hypoglykämien ist die Häufigkeit von leichten Hypoglykämien (3,4). Je weniger leichte Hypoglykämien, desto besser die Hypoglykämiewahrnehmung, desto höher die Wahrscheinlichkeit, eine Hypoglykämie rechtzeitig zu bemerken und Gegenmaßnahmen durchführen zu können. Häufige leichte Hypoglykämien dagegen sind ein Risikofaktor für Hypoglykämiewahrnehmungsstörung und schwere Hypoglykämien (5).

Das Abheben auf „Rate schwerer Hypoglykämien" als besonders wichtiges Kriterium durch das IQWIG beinhaltet, dass die sogenannten leichten Hypoglykämien als weniger wichtig erachtet werden. Leichte Hypoglykämien führen außerdem zu einer starken Beeinträchtigung der kognitiven und motorischen Leistungsfähigkeit (relevant zum Beispiel am Arbeitsplatz) und der Lebensqualität von Typ 2 Diabetespatienten und sind an sich ein wichtiges Kriterium.
Das Vorgehen des IQWIG ist also nicht adäquat. Insbesondere wurde schon in der Stellungnahme der Deutschen Diabetes Gesellschaft zum Berichtsplan und zum Amendment 1 zum Berichtsplan „Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ2" (A05-03) angemahnt, dass auch sogenannte „leichte" Hypoglykämien zu berücksichtigen seien. Dies wurde durch das IQWIG aus nicht nachvollziehbaren Gründen ignoriert.

2.) Gleichzeitige Betrachtung von Hypoglykämie und HbA1c-Absenkung

Hypoglykämien und der HbA1c-Wert stehen in einem engen Zusammenhang. Ein niedriger HbA1c-Wert führt zu einer erhöhten Rate von Hypoglykämien (6). Man kann die patientenrelevanten Zielgrößen „HbA1c" und „Hypoglykämierate" also nicht getrennt voneinander betrachten. Sie müssen bei dieser gegenseitigen Abhängigkeit zusammen evaluiert werden und die gegenseitigen Abhängigkeiten berücksichtigt werden. So kann beispielsweise ein geringer (auch nicht signifikanter) Rückgang des HbA1c-Wertes bei gleichzeitigem geringem (auch nicht signifikantem) Rückgang der Hypoglykämierate einen patientenrelevanten Vorteil erbringen. Da der Zusammenhang von HbA1c-Wert und Hypoglykämiehäufigkeit durch eine Hyperbel beschrieben wird (6), ist der Effekt von HbA1c und Hypoglykämiehäufigkeit als Produkt zu berechnen und nicht als bloße Summe. Daraus ergibt sich, dass kleine Verbesserungen gleichzeitig in HbA1c und Hypoglykämiehäufigkeit einen großen Vorteil für den Patienten in der Gesamtbetrachtung ergeben. Das IQWIG führt eine solche Betrachtung anhand von Tabelle 43, 44 und 45 (Seite 151-155) mit einer subjektiven „Plus" oder „Minus" Bewertung vor. Dieses Vorgehen ist wissenschaftlich nicht haltbar. Es sei hier auf die Metanalyse von Mullins et al. verwiesen (7).

3.) Umgang mit Übersichtsarbeiten und früheren Metaanalysen zu diesem Thema

Das IQWIG führt in seiner Suchstrategie auch die Suche nach Übersichtsarbeiten und früheren Metaanalysen an. Allerdings werden diese Arbeiten nur verwendet, um möglicherweise noch nicht entdeckte relevante RCTs zu identifizieren. Es wird offensichtlich darauf verzichtet, sich mit den Schlussfolgerungen, der Expertenmeinung und der inhaltlichen Diskussion auseinanderzusetzen. Insbesondere verzichtet das IQWIG vollständig auf einen Vergleich des Ergebnisses seiner eigenen mit den Ergebnissen anderer Metaanalysen und begründet in keiner Weise, wie es dazu kommen konnte, dass es hier gravierende Abweichungen gibt. Besonders fällt auf, dass die erwähnten Metaanalysen einen Vorteil der Analoginsuline beim Hypoglykämierisiko zeigen, den das IQWIG nicht findet.
In einem Cochrane Report (8) wird eine Metaanalyse der zur Verfügung stehenden Daten durchgeführt. Der Cochrane Report wurde vor ca. 1 Jahr veröffentlicht, wird aber vom IQWIG nicht einmal erwähnt. Es ist vollkommen unverständlich, dass das IQWIG trotz ausgefeilter Suchstrategie diesen Report übersehen hat, zumal andere international angesehene Institutionen (NICE) sich bei ihren Nutzenbewertungen langwirksamer Insulinanaloga auf diesen Report

beziehen. In dem Cochrane-Report kommen die Autoren zu anderen Ergebnissen als das IQWIG. So sieht der Cochrane-Report einen eindeutigen Vorteil beider langwirkenden Insulinanaloga beim Hypoglykämierisiko, insbesondere bei nächtlichen Hypoglykämien. Demnach kommen auch das NICE (1) und das CADTH zu anderen Schlussfolgerungen und Empfehlungen als das IQWIG.
Auch die Studie von Mullins et al. (7), die eine qualitativ hochwertige Metaanalyse zum Vergleich NPH-Insulin und Insulin Glargin durchführt, ebenso wie die Metaanalyse von Rosenstock et al. (9) und Tschritter et al. (10) (letztere zum Vergleich NPH-Insulin und Insulin Detemir) werden nicht berücksichtigt.

4.) Endpunkte: Werden die zur Verfügung stehenden Daten, die Ziele und die Auswahlkriterien aufeinander abgestimmt?

Das IQWIG bezieht sich unter anderem auf „patientenrelevanten Nutzen", es wird beklagt, dass über Endpunkte wie Gesamtmortalität keine ausreichenden Studien und Daten vorliegen! Die Bewertung hierzu müsste also lauten: **„Es gibt noch keine abgeschlossenen Studien, die eine Beurteilung zulassen. Diese müssen durchgeführt werden. In diesem Punkt muss das Ergebnis der Nutzenbewertung offen bleiben. Insbesondere kann hier noch keine negative Nutzenbewertung ausgesprochen werden."**

Das Fehlen von Evidenz für einen Substanzvorteil darf also nicht gleichgesetzt werden mit dem Vorliegen von Evidenz für Gleichwertigkeit! Das gilt besonders dann, wenn das Fehlen von Evidenz auf dem vollständigen Fehlen von Daten beruht, die für die Evidenz maßgeblich sind! Bei einer ausgewogenen Betrachtungsweise ist noch zuzufügen, dass auch für NPH-Insulin keine Daten zu entsprechenden Endpunkten vorliegen.

Bei der Bewertung weiterer Endpunkte wie Lebensqualität, Behandlungszufriedenheit und Hypoglykämien müssen die Ein- und Ausschlusskriterien für die Studien entsprechend angepasst werden. Ggf. ist die Auswahl der Studien für die Endpunkte unterschiedlich vorzunehmen. Während für die harten Endpunkte eine Laufzeit von einem oder mehreren Jahren wünschenswert wäre, kann man dieses Kriterium nicht einfach für alle anderen Endpunkte übernehmen. Man kann für harte Endpunkte die Kompromisslösung mind. 6 Monate (bzw. 24 Wochen) wählen, wenn es zu wenige Studien von mindestens 1 Jahr Laufzeit gibt, aber für die Bewertung anderer Endpunkte sind auch kürzere Studien als 24 Wochen relevant. Durch die Nichtbeachtung solcher Studien bei diesen Endpunkten wird willkürlich Information aus Studien mit hoher Evidenz (RCTs) ignoriert (z.B. 11).

Damit ist die Bewertung von Nüchternblutzucker, HbA1c, Hypoglykämiehäufigkeit und Lebensqualität nach den gleichen Kriterien, die man für die Mortalität und Morbidität festgelegt hat, nicht nur willkürlich sondern auch wissenschaftlich falsch.

5.) Praxis der Insulintherapie

Das IQWIG fasst mittel- und langwirksame Insuline zusammen zu Verzögerungsinsulinen und begründet dies damit, dass diese ja in der Praxis zum selben Zweck eingesetzt werden (S. 1). Nun ist aber NPH-Insulin mittellang wirksam, Insulin Detemir und Insulin Glargin sind langwirksam. Solche Unterschiede bringen automatisch mit sich, dass die Möglichkeit unterschiedlicher Zeitpunkte und Frequenzen der Applikation besteht, und dass dies bei den Studiendesigns und bei der Bewertung der Ergebnisse immer wieder eine Rolle spielen wird. So handelt es sich hier nicht nur um den Vergleich verschiedener Insulinpräparationen, die immer völlig identisch angewandt werden können.

Das IQWIG beanstandet, dass in den RCTs manchmal NPH-Insulin nur einmal gespritzt wurde, obwohl bei NPH-Insulin laut Fachinfo empfohlen wird, dass die Applikationsfrequenz an die individuellen Bedürfnisse angepasst werden soll. Im Vergleich von einmal gespritztem NPH mit Insulin Glargin ergibt sich ein Vorteil für Insulin Glargin, der aber vom IQWIG nicht anerkannt wird. Man sollte aber, um gleiche Ausgangesbedingungen zu haben, auch die gleiche Anzahl von Injektionen vergleichen.

Zudem schließt das IQWIG einen Studienarm aus, in dem Insulin Glargin morgens verabreicht wird, mit der Begründung, die Vergleichbarkeit mit dem NPH-Insulin, das seinen Eigenschaften entsprechend abends gegeben wurde, sei nicht gegeben. Wenn die Grundlage dieses Berichtes

ist, dass verschiedene Therapien verglichen werden, muss es auch möglich sein, die erweiterten Möglichkeiten der neuen Therapie zu nutzen.

Aus unserer Sicht ist es also notwendig, bei den RCTs verschiedene Designs zuzulassen. Letztendlich wird damit hohe Evidenz für den Vergleich von Therapieformen unter verschiedenen Bedingungen gewonnen.

6.) Fehlende Verblindung

Die Kritik des IQWIG an der fehlenden Verblindung ist nicht angebracht. Eine Verblindung ist durch das unterschiedliche Aussehen der Insuline (trüb versus klar) schlecht möglich. Weitherhin ist zur korrekten Anwendung der unterschiedlichen Insuline auch das Wissen des Patienten über dessen Wirkdauer und Wirkstärke nötig. Da diese bei den langwirkenden Analoginsulinen und dem NPH-Insulin unterschiedlich ist, ist eine Verblindung nicht möglich, ohne den Patienten als wesentlichen Teil der Therapie auszuschließen, zu entmündigen und ggf. sogar zu gefährden.

Literatur:

(1) **http://www.nice.org.uk/nicemedia/pdf/Insulin_Analogues.pdf**

(2) Jehle PM, Micheler C, Jehle DR, Breitig D, Boehm BO. Inadequate suspension of neutral protamine Hagendorn (NPH) insulin in pens. Lancet. 1999 Nov 6;354(9190):1604-7.

(3) Fanelli CG, Porcellati F, Pampanelli S, Bolli GB.Insulin therapy and hypoglycaemia: the size of the problem. Diabetes Metab Res Rev. 2004 Nov-Dec;20 Suppl 2:S32-42. Review.

(4) Fritsche A, Stefan N, Haring H, Gerich J, Stumvoll M. Avoidance of hypoglycemia restores hypoglycemia awareness by increasing beta-adrenergic sensitivity in type 1 diabetes. Ann Intern Med. 2001 May 1;134(9 Pt 1):729-36.

(5) Mokan M, Mitrakou A, Veneman T, Ryan C, Korytkowski M, Cryer P, Gerich J. Hypoglycemia unawareness in IDDM. Diabetes Care. 1994 Dec;17(12):1397-403.

(6) DCCT Research Group: The effect of intensive diabetes treatment on the development and progression of long-term complications in insulin-dependent diabetes mellitus: the Diabetes Control and Complications Trial . N Engl J Med. 1993;329:977–986

(7) **Mullins P, Sharplin P, Yki-Jarvinen H, Riddle MC, Haring HU. Negative binomial meta-regression analysis of combined glycosylated hemoglobin and hypoglycemia outcomes across eleven Phase III and IV studies of insulin glargine compared with neutral protamine Hagedorn insulin in type 1 and type 2 diabetes mellitus. Clin Ther. 2007 Aug;29(8):1607-19.**

(8) **Horvath K, Jeitler K, Berghold A, Ebrahim SH, Gratzer TW, Plank J, Kaiser T, Pieber TR, Siebenhofer A. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Database Syst Rev. 2007 Apr 18;(2):CD005613. Review**.

(9) **Rosenstock J, Dailey G, Massi-Benedetti M, Fritsche A, Lin Z, Salzman A. Reduced hypoglycemia risk with insulin glargine: a meta-analysis comparing insulin glargine with human NPH insulin in type 2 diabetes. Diabetes Care. 2005 Apr;28(4):950-5.**

(10) **Tschritter O, Fritsche A, Gallwitz B, Haäring H-U. langwirkende Insulinanaloga in der Therapie des Diabetes mellitus Typ 1 und Typ2. Eine Zusammenfassung der klinischen Studien und Metaanalysen. Diabetes und Stoffwechsel 2005, 14:375-82. Review**

(11) **Raslová K, Tamer SC, Clauson P, Karl D. Insulin detemir results in less weight gain than NPH insulin when used in basal-bolus therapy for type 2 diabetes mellitus, and this advantage increases with baseline body mass index. Clin Drug Investig. 2007;27(4):279-85.**

Für den Pharmakotherapieausschuss der Deutschen Diabetes Gesellschaft:

Prof. Dr. med. Andreas Fritsche, Med. Klinik IV, Universität Tübingen
Prof. Dr. med. Baptist Gallwitz, Med. Klinik IV, Universität Tübingen
Dr. med. Otto Tschritter, Med. Klinik IV, Universität Tübingen

Prof. Dr. med. Harald Klein, Berufsgenossenschaftliches Universitätsklinikum Bergmannsheil, Ruhr Universität Bochum, Vorsitzender des Ausschusses

A 1.4 Deutscher Diabetiker Bund e. V.

Autoren:
Möhler, Dieter
von Lilienfeld-Toal, Herrmann, Prof. Dr. med.
Windisch, Heinz

Adresse:
Heinz Windisch
Käthe-Kollwitz-Str. 28
01477 Arnsdorf

DEUTSCHER DIABETIKER BUND

Der Bundesvorstand

Sachbearbeiter:
RA Dieter Möhler
stellvertretender Bundesvorsitzender
Tel.: 05 61 / 70 34 77 - 0
0175 4659860
Fax: 05 61/70 34 77 - 1
E-Mail: moehler@diabetikerbund.de

Institut für Qualität und Wirtschaftlichkeit
im Gesundheitswesen
Stellungnahme zum Vorbericht A05/03
Prof. Dr. med. Peter T. Sawicki
Dillenburger Straße 27

51105 Köln

Meiningen, den 24.04.2008

Betreff: Stellungnahme des Deutschen Diabetiker Bund e.V. zum Vorbericht langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2 Auftrag des IQWiG A05/03 Version 1.0 vom 18.03.20008

Sehr geehrte Damen und Herren,

innerhalb der aus § 139 a SGB V resultierenden Beteiligungsrechte der Selbsthilfeorganisationen in Bezug auf chronisch kranke und behinderte Menschen gibt unsere Organisation nachfolgende Stellungnahme sowie in der Anlage beigefügt die Stellungnahme des hier gegründeten und satzungsgemäß bestellten wissenschaftlichen Beirates in unserem Namen ab, beigefügt als Anlage 1 im derzeitigen Verfahrensstadium ab.

1.

Zunächst ist festzustellen, dass der gemeinsame Bundesausschuss gemäß § 139 a Abs. 1 SGB V ein fachlich unabhängiges Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen gegründet sehen wollte.

Die Beteiligungsrechte der betroffenen Patienten sind garantiert über Abs. 5 der genannten Vorschrift.

Die Stellungnahmen sind entgegen der Auffassung der Verfasser des Vorberichts in die Entscheidung zwangsnotwendig einzubeziehen, so dass die Ausführungen des Instituts, wonach relevante Stellungnahmen berücksichtigt werden *können*, nicht der geltenden Rechtslage entspricht.

Betreffend die Wahrnehmung der rechtlich, garantierten Interessen der Patienten gehört in diesem Zusammenhang in jedem Bereich auch das einzuräumende rechtliche Gehör betreffend ein formell richtig gestaltetes Verfahren.

Hierzu werden aus Sicht des DDB e.V. folgende Einwände erhoben:

Es findet sich an keiner Stelle der vorformulierte Auftrag des gemeinsamen Bundesausschusses.

Insoweit ist eine Überprüfung nicht möglich, ob sich der erteilte Auftrag auf die gesetzgeberisch zugewiesenen Aufgabengebiete beschränkt oder insbesondere auch hinterher nachträglich Abänderungen, gegebenenfalls ergebnisorientiert aus Absprachen zwischen dem Institut selbst und Vertretern des gemeinsamen Bundesausschusses stattgefunden haben. Der Auftrag ist dem entsprechend in seiner Gesamtheit konkret wörtlich zu nennen und auch gegebenenfalls als Anhang dem Vorbericht beizufügen.

Der Vorbericht spricht selbst von einem ursprünglichen Auftragsschreiben, welches den vorliegend im Anhörungsverfahren Beteiligten nicht zugänglich gemacht ist ebenso wenig wie die erwähnte Auftragskonkretisierung.

Insbesondere kann nicht nachvollzogen werden, inwieweit vom Auftraggeber bereits konkrete Vorgaben zur Bewertung des Nutzens oder konkrete Vorgaben hinsichtlich der Zugrundelegung von Anschlusstatsachen gemacht wurden. Diese Informationen müssen allesamt zugänglich gemacht werden, um den Beteiligten den Anspruch auf rechtliches Gehör und Abgabe einer umfassenden Stellungnahme innerhalb des vorgesehenen Verfahrens zu ermöglichen.

2.

Gemäß § 139 a Abs. 5 SGB V hat das Institut in allen wichtigen Abschnitten des Bewertungsverfahrens, insbesondere nach Inanspruchnahme von Sachverständigen, deren Stellungnahmen in die Entscheidung einfließen, auch preiszugeben, wer diese Sachverständigen waren.

Nur durch entsprechende Bekanntgabe der an dem Vorbericht beteiligten Sachverständigen ist auch eine Wahrung der Rechte der weiteren Beteiligten möglich insbesondere was eine Überprüfung derselben auf das Vorliegen von Fachkunde und einer fachlichen Unabhängigkeit anbelangt. Es sind sämtliche Namen preiszugeben. Die abgegebenen Stellungnahmen sind insoweit kenntlich zu machen, als diese von diesen als Hilfspersonen stammen und im Vorbericht Eingang gefunden haben.

Dies liegt ersichtlich nicht vor, was ausdrücklich gerügt wird.

3.

Des weiteren wird die konkrete Ausgestaltung des Verfahrens als Verstoß gegen das gesetzlich verankerte rechtliche Gehör vorliegend gerügt.

Eine Beschränkung der Stellungnahmemöglichkeit auf eine bestimmte Seitenanzahl, wie in dem Verfahren vorgeschlagen, verbietet sich.

Aufgrund der Formulierung des § 139 a Abs. 5 S. 2 SGB V steht fest, dass die Stellungnahmen in dem förmlich ausgestalteten Verfahren in die Entscheidung einzubeziehen **sind.**

Allein aus der gesetzgeberischen Formulierung ergibt sich, dass eine Beschränkung nicht stattfinden darf.

4.

Der Vorbericht nimmt keine wissenschaftliche Abgrenzung dahingehend vor, wann von mittellangen und wann von langwirksamen Insulinen gesprochen werden kann und welche Voraussetzungen dafür vorliegen müssen, ein zu beurteilendes Insulin in eine bestimmte Gruppe einzuordnen.

Es ist nicht erkennbar, inwieweit diese Unterscheidung, die offensichtlich vorgenommen wird, berichtsrelevant geworden ist, was ausdrücklich gerügt wird. Die Anschlusstatsachen, die dem Bericht insoweit zu Grunde gelegt sind, werden damit nicht öffentlich preisgegeben.

Nachdem vorliegend der Auftrag und dessen Abänderungen nicht vollumfänglich bekannt gegeben sind, werden rein vorsorglich folgende Einwendungen in Bezug auf die Nutzenbewertung gemacht:

So weit der Bericht ausführt, es sei nicht gesichert, ob die Senkung des Blutzuckers in jedem Fall Folgeschäden verhindere, wird darauf verwiesen, dass sich eine solche Überlegung von vornherein im Rahmen einer ordnungsgemäßen Nutzenbewertung verbietet.

Die Bewertung hätte in diesem Zusammenhang, sollte eine solche Überlegung darüber angestellt werden dürfen, nachprüfbar darzulegen, wie die betreffenden Kenntnisse erlangt wurden, die andere Beurteilungen zuließen.

Anerkannt wissenschaftlicher Kenntnisstand ist zum gegenwärtigen Zeitpunkt, daß niedrige Blutzuckerwerte weniger Folgeschäden nach sich ziehen.

Es besteht eine einhellige Auffassung, dass die Blutzuckersenkung daher vorrangiges Ziel der Diabetes-Therapie und der erwünschten Folgen in Form der Vermeidung von Folgekomplikationen ist.

Die diskutierten Studien umfassen zu kurze Zeiträume und sind daher von den Anschlusstatsachen her betrachtet schon gar nicht geeignet, die fast einhellig übereinstimmend in der Fachwelt vertretene Auffassung zu erschüttern, dass sich ein Nachweis nicht führen ließe, ein niedriger Blutzuckerwert gehe mit weniger Folgeschäden einher.

Im übrigen hätte gerade aufgrund dieser einhellig vertretenen Auffassung der Vorbericht im einzelnen darzulegen, welche Befundtatsachen sich für das Institut ergeben, von diesem medizinischen Wissensstand abzuweichen.

In dem Bericht müsste nachprüfbar dargelegt sein, wie die betreffenden Behauptungen, die für das Institut die Geltung einer Kenntnis haben sollen, erlangt wurden.

Dies gilt um so mehr, als nach dem derzeitigen wissenschaftlichen Ansatz von einem Erfahrungssatz auszugehen ist. Es ist eben gerade typischer Geschehensablauf, daß niedrige Blutzuckerwerte in der Regel weniger Folgeschäden bedeuten.

In Bezug auf den Auftragsumfang einer Nutzenbewertung verbieten sich Ausführungen zur Kanzerogenität der Insulinanaloga.

Wenn es keinen Hinweis darauf gibt, dass eine erhöhte Kanzerogenität besteht, verbietet sich die Aufnahme einer Überlegung in eine Nutzenbewertung als Anschlusstatsache. Dies gilt selbstverständlich auch für Hinweise, dass eine erhöhte Kanzerogenität nicht besteht aber dies nicht abschließend geklärt sei.

Soweit gerade entgegenstehende Anschlusstatsachen durch Auftreten erhöhter Kanzerogenität nicht zu sehen sind, ist dieser Punkt im Rahmen einer vorzunehmenden Nutzenbewertung irrelevant.

Mangels eindeutig nachvollziehbarer Anschlusstatsachenfeststellung lässt sich eine konkrete Sachverhaltsfeststellung nicht ableiten.

Dem HbA1C-Wert der grundsätzlich nutzenrelevant ist, wird, dies bezogen auf den Vergleich der Insuline zueinander, eine Bedeutung beigemessen, die es in dieser Form nicht gibt.

Der HbA1C-Wert, lässt den Rückschluss zu, dass die Insulintherapie, flankiert von einem entsprechenden Aufwand in der Blutzuckermessung und hieraus resultierend einer Insulindosisanpassung im Behandlungsergebnis messbar wird und damit unmittelbare Ausformungen der Lebensqualität einhergehen.

Die Nutzenbewertung hat sich an den gesetzgeberischen Erfordernissen auszurichten.

Gemäß § 2 a Abs. 5 SGB V ist den Belangen der Behinderten und chronisch Kranken besondere Rechnung zu tragen.

Durch gerade diese Ausformulierung hat der Gesetzgeber beabsichtigt, der Überlegung Geltung zu verschaffen, dass bei chronisch Kranken ein normnaher Zustand der Einstellung an sich durch unterschiedliche Behandlungen erreicht werden kann. Dem entsprechend sind Umstände wie die durch die jeweilige Behandlung eintretende Lebensqualität und Behandlungszufriedenheit in die Ziele miteinbezogen, eben besonders zu berücksichtigen.

Der Vorbericht lässt einfache Erfahrungssätze außer Acht, die zum Inhalt haben, dass bei Nichtvorliegen einer niedrigeren Hypoglykämiefrequenz als auch hyperglykämischer Erscheinungen negative Auswirkungen im Berufsleben gehäuft auftreten.

Die Nutzenbewertung ist, wie im Vorbericht wahrgenommen, nicht nachvollziehbar.

Der konkrete Nutzen für den einzelnen Patienten zeigt sich in der Durchführung der Therapie vor Ort, nicht in einer generalisierenden Betrachtung aller Betroffener, wie vom Institut vorgenommen.

Allein aus der Vielfalt unterschiedlicher Patienten und damit unterschiedlicher Stoffwechselsituationen, beeinflusst von äußeren Lebensumständen lassen sich generalisierende Aussagen im Sinne eines nicht vorliegenden Nutzens nicht treffen.

Als gesichert kann gelten der einfache Erfahrungssatz, wonach in Bezug auf Hypoglykämien, bei Einsatz eines Insulinanalogons ein erhöhter Nutzen für den betreffenden Patienten besteht.

Die Zugrundelegung von Datenmaterial, wonach die Behandlungszufriedenheit zwar höher unter dem Insulinanalogon sei, aber hierbei in Bezug auf eine Nutzenbewertung eine klinische Relevanz nicht gegeben sei, ist unzulässig.

Die Nutzenbewertung ist nicht an einer klinischen Relevanz, sondern an einem gesetzgeberische formulierten Ziel auszurichten.

Es ist ebenfalls einfacher Erfahrungssatz, dass insbesondere bei unzureichender Diabeteseinstellung und damit einhergehender Behandlungsunzufriedenheit des Patienten dessen psychische Belastung zunimmt und sich in eben solchen Folgeerkrankungen auswirkt.

Hier reichen schon höhere postprandiale Blutzuckeranstiege und ein verlängerter Zeitraum zur Erreichung des Zielwertes nach einer Dosisanpassung aus, den Patienten psychisch zu belasten.

Auch in diesem Zusammenhang ist wieder das gesetzgeberische Ziel, formuliert in § 2 des SGB V deutlich hervorzuheben, wonach es staatliche Aufgabe ist, eine Verschlimmerung der Krankheit zu verhüten.

Das heißt natürlich auch, dass Folgeerkrankungen z.B. im psychischen Bereich als aus einer Behandlungsunzufriedenheit resultierend kraft gesetzgeberischer Verpflichtung zu verhüten sind.

Das IQWiG lässt im Vorbericht die Anschlusstatsachen hierzu außer Acht.

Jede Nutzenbewertung hat die gesetzgeberischen Vorgaben zu beachten, was vorliegend in Bezug auf die Behandlungszufriedenheit und die erhöhte Gefährdung betreffend den Eintritt psychischer Erkrankungen nicht berücksichtigt ist. Insoweit besteht in der Nutzenbewertung ein Defizit, weil Anschlusstatsachen nicht eingestellt sind.

Wenn weiter statistische Maßnahmen sich nicht in eine konkrete Bedeutung für den jeweiligen Patienten übersetzen lassen, zeigt dies die Unfähigkeit, vorliegend überhaupt eine auf wissenschaftlichen Tatsachen beruhende Nutzenbewertung abzugeben.

Der Nutzen ergibt sich aus dem konkreten Nutzen für den einzelnen Patienten resultierend aus dessen Lebensumständen und persönlicher Gesundheitssituation.

Jede generalisierende Betrachtung verbietet sich.

Auch ist eine geringere Applikationsfrequenz in Bezug auf die zu beurteilenden Insulinanaloga vorliegend nutzenrelevant, jedoch von den Verfassern des Vorberichts bestritten.

Es steht ohne weiteres fest, dass der Rückgang der Anzahl von Injektionen ein Nutzen für den betroffenen Patienten ist, insbesondere unter Beachtung der Tatsache, dass jeder Gesunde eben keine entsprechende Insulinapplikation braucht.

Der Unterschied zum Gesunden besteht darin, auf Injektionen angewiesen zu sein. Im Zustand chronischen Krankseins ist eine gehäufte Anzahl notwendiger Injektionen belastender, als eine geringere. Also liegt schon denkgesetzlich ein Nutzen vor durch die geringere Injektionsanzahl.

Der Vorbericht ist damit in Bezug auf das gesetzlich vorgesehene Verfahren formell und darüber hinausgehend auch materiell insbesondere durch Außerachtlassung von Denkgesetzen und Erfahrungsgrundsätzen grob fehlerbehaftet.

Das IQWiG lässt in seinem Vorbericht die wissenschaftliche anerkannten Grenzen außer Acht, wonach bei unzureichenden Studienmaterial (fehlendes oder fehlerhaftes durch Nichtberücksichtigung von Anschlusstatsachen) die Nutzenbewertung an Denkgesetzen und Erfahrungsätzen auszurichten hat.

Die Ausschlusstatsachen werden nicht ausreichend berücksichtigt, eine Abwägung der feststehenden Nutzen im Verhältnis zueinander wird nicht vorgenommen und einzelne Nutzen ausgeklammert.

Damit ist die insgesamt vorgenommene Abwägung in ihrem Ergebnis, dem Insulinanalogen käme kein entsprechender Nutzen zu, offensichtlich fehlerhaft und nicht verwendbar.

Auf die beigefügte Stellungnahme des wissenschaftlichen Beirates nehmen wir ergänzend Bezug und machen dieselbe vollinhaltlich zum Gegenstand unserer Darlegungen.

Mit freundlichen Grüßen

i.V. Dieter Möhler
stellvertretender Bundesvorsitzender

Heinz Windisch
Bundesvorsitzender

Anlage 1 Stellungnahme des wissenschaftlichen Beirates

Anlage

DEUTSCHER DIABETIKER BUND

Der Bundesvorstand

Deutscher Diabetiker Bund · Goethestraße 27 · 34119 Kassel

Institut für Qualität und Wirtschaftlichkeit
im Gesundheitswesen
Stellungnahme zum Vorbericht A05/03
Prof. Dr. med. Peter Sawicki
Dillenburger Str. 27

51105 Köln

Ihr Ansprechpartner:
Prof. Dr. med Hermann von
Lilienfeld-Toal
Vors. des wissenschaftlichen
Beirats des DDB
Tel.: 06051/872293
Fax: 06051/872292
E-Mail: hermann.
lilienfeld@mkkliniken.de

Kassel, 24.04.2008

Stellungnahme des Deutschen Diabetiker Bundes e. V. zum Vorbericht langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2
Auftrag des IQWiG A05/03 Version 1.0 vom 18.03.2008

Sehr geehrte Damen und Herren,

für den wissenschaftlichen Beirat des DDB gebe ich folgende Stellungnahme ab:

Wissenschaftliche Untersuchungen haben Autoren:
Auf Seite II wird verschwiegen, wer die externen Sachverständigen dieses Vorberichtes sind.

Seite 1: Es wird von „mittellang" wirksamen Insulinen gesprochen. Welche Insuline sind diese? Welche Rolle spielen sie für den Bericht?

Seite 4: Es wird eine Überlegung darüber angestellt, ob die Senkung des Blutzuckers in jedem Fall Folgeschäden verhindert. Auch wenn es einzelne Studien gibt, die die direkte Gleichung: „niedrige Blutzucker bedeutet weniger Folgeschäden" in Frage stellen, wie sie beispielsweise zitiert wurden, besteht eine große Einigkeit in der Fachwelt, dass Blutzuckersenkung ein unbestrittenes Therapieziel ist. Da in den folgenden diskutierten Studien jeweils Zeiträume nicht über 1 Jahr behandelt werden, ist die hier vorgestellte Diskussion unsinnig und irreführend.

Seite 5: Zur Kanzerogenität der Insulinanaloga: Unter Tabelle 1, Seite 6: Nachdem ablesbar an dieser Tabelle sowohl die EMEA und auch FDA keine Bedenken gegen den Einsatz dieser Medikation wegen einer erhöhten Kanzerogenität hat, sollte dieser Aspekt auch entsprechend im Text behandelt werden: Offensichtlich gibt es zur Zeit keinen Hinweis darauf, dass eine erhöhte Kanzerogenität besteht. Ob dieses abschließend geklärt ist oder nicht, wie im Text zitiert, hilft zur Zeit nicht weiter. Der augenblickliche Stand des Wissens ist, dass kein Hinweis auf eine erhöhte Kanzerogenität besteht. Wenn von Seiten des IQWiG neue Literatur hierzu beigetragen werden kann, wäre es interessant. Dieses ist aber offensichtlich nicht der Fall.
Wegen der weiten Verbreitung der Insulinanaloga halten wir im DDB die immer wieder neu von Seiten des IQWiG vorgebrachten Bedenken wegen einer möglichen Kanzerogenität dieser Präparate für unverantwortlich. Wenn man Sorgen induziert, muss dieser Vorgang auf einer realen Basis bestehen, und diese existiert hier offensichtlich nicht.

Seite 22: Letzter Absatz „in dieser Studie kam es jedoch zu einer unterschiedlichen Art der Anwendung von Insulin Glargin und NPH Insulin.....".
Natürlich müssen Glargin und NPH Insulin, die verschiedene Wirkprofile haben, unterschiedlich verwendet werden. Gerade die unterschiedliche Handhabung ist ja Teil des Vergleiches. Aus diesem Grunde ist der Ausschluss dieser Studie unverständlich. Es müssen doch Therapieschemata verglichen werden und nicht Dosis gegen Dosis.

Seite 86: Wir haben Angst, dass das Ergebnis des Vergleiches zwischen Analoga und NPH in dieser Untersuchung über den Nutzen der Insulinanaloga bereits vorher feststand. Diese Angst ist nicht unbegründet, wie ein genaues Lesen des Textes offen legen:

Im letzten Absatz der Seite 86 wird erst ein Befund zuungunsten des NPH-Insulins zitiert: „In der Fundusfotographie wurde ein Fortschreiten der Retinopathie.....numerisch, aber statistisch nicht signifikant häufiger in der Gruppe mit NPH-Insulin diagnostiziert." Dann folgt ein Befund zuungunsten des Insulinanalogons Glargin: „.....in der Fundusfotographie jedoch ergaben sich unter Insulin Glargin auffällig mehr Patienten, die ein Makulaödem entwickelten." Also: Bei NPH-Insulin: statistisch nicht signifikant häufiger, bei Glargin: auffällig mehr Patienten. Wie der folgende Text zeigt, war dieser Befund der Glargin Gruppe übrigens auch nicht statistisch signifikant.

Deutscher Diabetiker Bund e.V.
Vorstand: Bundesvorsitzender: Heinz Windisch
Stellv. Bundesvorsitzende: Rosmarie Johannes · RA Dieter Möhler
Bundesschatzmeister: Jens Pursche · Bundesjugendreferent: NN
Bundesgeschäftsstelle: Goethestr. 27 · 34119 Kassel · Telefon 05 61 – 7 03 47 70 · Telefax 05 61 – 7 03 47 71
E-Mail: info@diabetikerbund.de · Internet: www.diabetikerbund.de
Bankverbindung: Bank für Sozialwirtschaft, Essen · Kto.-Nr.: 7 219 500 · BLZ 370 205 00

Seite 104: Zum Ausmaß der Blutzuckersenkung: Es sei an dieser Stelle noch einmal darauf hingewiesen, was schon mehrfach dem IQWiG zur Kenntnis gegeben wurde, dass ein Vergleich zwischen blutzuckersenkenden Maßnahmen im Grunde die Fähigkeit des Betroffenen, mit dem eingesetzten Medikament und, bei Insulin, seinem Wirkprofil umzugehen, widerspiegelt. Daher sind Unterschiede zwischen verschiedenen Insulintherapieschemata im Grunde nur anhand des Aufwandes und damit der Behandlungs- und Lebensqualität messbar. Bei entsprechendem Aufwand wird jeder Diabetiker ein gutes HBA1c erreichen. Die Bedeutung des Vergleiches von HBA1c ist daher gar nicht so groß, während die Bedeutung von Lebensqualität und Behandlungszufriedenheit zur Frage des Nutzens sehr groß ist.

Seite 110 bis 118: Unter Glargin-Medikation durchgehend bessere Hba1c-Werte: Warum wird dieses nicht entsprechend gewürdigt? Das IQWiG bewertet diesen Unterschied als irrelevant. Das IQWiG berücksichtigt hierbei nicht, dass der Einzelne unter Umständen sehr verschieden auf dieses Präparat, wie übrigens auf andere Insuline auch, reagiert. Es wird durch die große Zahl der Patienten, die an den Studien teilnehmen, verdeckt, dass Einzelne durch das eine oder andere Präparat stark profitieren. Mit anderen Worten, die Frage, ob für den einzelnen Betroffenen das Präparat geeignet ist oder nicht – so wie sie sich im Alltag stellt – wird nur ganz entfernt durch derartige Studien abgebildet. Daher kann man erwarten, dass bei selbst einem kleinen Vorteil, was die Hba1c-Senkung betrifft, unter Umständen einzelne Patienten einen dramatischen Effekt erleben. Für den Betroffenen ist auch der kleinste Vorteil wichtig, da es sich ja um eine jahrelange Therapie handelt.

Seite 134: Eine ähnliche Überlegung muss über die Häufigkeit von Hypoglykämien angestellt werden. Wie in den hier diskutierten Studien belegt ist, ist sehr wahrscheinlich, dass unter Glargin eine geringere Inzidenz einer Hypoglykämie besteht.

Seite 195: Es werden vom IQWiG Daten zitiert, wonach Behandlungszufriedenheit höher unter Glargin als NPH Insulin ist. Der Grund weswegen dieser Befund aus Sicht des IQWiG angeblich ohne klinische Relevanz sei, ist nicht nachvollziehbar. Warum die Einführung einer Minimal Important Difference (MID) in die Diskussion einen durchgehenden Effekt in den 4 zitierten Studien (Abb 21,S196) aufhebt, ist nicht verständlich.
Nach unserer Kenntnis ist es außerordentlich schwierig, derartige Kriterien wie Behandlungszufriedenheit objektiv zu quantifizieren. Wenn die hierzu entwickel

Deutscher Diabetiker Bund e.V.
Vorstand: Bundesvorsitzender: Heinz Windisch
Stellv. Bundesvorsitzende: Rosmarie Johannes · RA Dieter Möhler
Bundesschatzmeister: Jens Pursche · Bundesjugendreferent: NN
Bundesgeschäftsstelle: Goethestr. 27 · 34119 Kassel · Telefon 05 61 – 7 03 47 70 · Telefax 05 61 – 7 03 47 71
E-Mail: info@diabetikerbund.de · Internet: www.diabetikerbund.de
Bankverbindung: Bank für Sozialwirtschaft, Essen · Kto.-Nr.: 7 219 500 · BLZ 370 205 00

ten Instrumente nun durchgehend in Richtung Vorteil für die Insulinanaloga Ergebnisse zeigen, darf dieses nicht durch einen methodischen Trick entwertet werden. Dieser besteht darin, einen zusätzlichen Parameter (MID) einzufordern, der offensichtlich von den Studienautoren selbst als nicht relevant eingestuft wurde und daher nicht angegeben wurde. Tatsache ist, dass ein Unterschied besteht. Im Sinne der EBM (verfügbare Evidenz!) ist es ein methodischer Fehler, diese Ergebnisse über die Behandlungszufriedenheit auszublenden. Dies ist besonders unzulässig angesichts der Tatsache, dass es sich mit diesen Parametern wahrscheinlich um die wichtigsten Indikatoren des Nutzens handelt. Für die Wissenschaftler und Interessenten an dieser Diskussion, die nicht an die Zusammenarbeit mit Diabetikern gewöhnt sind, möchten wir an dieser Stelle unterstreichen, dass es sich für den Diabetikern nicht allein um eine „Lebensqualität „ im Sinne des Wohlfühlens handelt, sondern dass aus der Lebensqualität und der Behandlungszufriedenheit heraus die Fähigkeit wächst, seine Stoffwechselstörung selbst zu behandeln, wie es nötig ist: „Empowerment".

Seite 196, letzter Satz vor Abbildung 21.: Dieser Satz ist falsch. Mit den vorliegenden Daten besteht ein Hinweis, auch wenn er vom IQWiG offensichtlich niedrig bewertet wird.

Seite 245: Es besteht eine statistisch signifikante geringere Gewichtszunahme unter Detemir. Für diesen Befund gilt dieselbe Überlegung wie oben für das gesenkte Hba1c unter Glargin: Unter diesem Mittelwert verbergen sich Betroffene, bei denen zum Teil erhebliche Gewichtsreduktionen beobachtet werden. Warum soll wegen dieses nur geringen mittleren Effektes die Möglichkeit, dass der Einzelne hiervon profitieren kann, gestrichen werden?

Seite 251, 1. Absatz, letzter Satz: „Die Übersetzung dieser statistischen Maße in eine konkrete Bedeutung für den Patienten, ist jedoch kaum möglich".
Das IQWiG hat als Grundlage seiner gesamten Überlegung die Frage gestellt, ob ein Unterschied bei möglichen Parametern eines Therapieeffektes zwischen therapeutisch unterschiedlich behandelter Gruppen besteht. Mit einem Satz, wie dem jetzt zitierten, wird darauf verzichtet, auf den beobachteten Effekten basierend Bewertungen vorzunehmen. Es wird mit einer solchen Überlegung jeder Willkür Tür und Tor geöffnet, denn wer wird „eine konkrete Bedeutung für Patienten" wie begründen können? Das ist im konkreten vorliegenden Fall der langwirksamen Insulinanaloga bei Typ 2 Diabetes, in diesem Vorbericht jedenfalls nicht durchgeführt. Daher wird das IQWiG aufgefordert, auf derartige, zu einer völlig willkürlichen Bewertung führende Gedanken zu verzichten. Hier wird durch die Willkürlichkeit das Prinzip der Wissenschaftlichkeit eklatant verletzt.

Deutscher Diabetiker Bund e.V.
Vorstand: Bundesvorsitzender: Heinz Windisch
Stellv. Bundesvorsitzende: Rosmarie Johannes · RA Dieter Möhler
Bundesschatzmeister: Jens Pursche · Bundesjugendreferent: NN
Bundesgeschäftsstelle: Goethestr. 27 · 34119 Kassel · Telefon 05 61 – 7 03 47 70 · Telefax 05 61 – 7 03 47 71
E-Mail: info@diabetikerbund.de · Internet: www.diabetikerbund.de
Bankverbindung: Bank für Sozialwirtschaft, Essen · Kto.-Nr.: 7 219 500 · BLZ 370 205 00

Seite 253, letzter Absatz: „Es fanden sich jedoch für keine dieser Studien Angaben darüber, ob Patienten eine spezielle Insulintherapieschulung erfahren hatten". Unser Kommentar: Endlich findet das IQWiG einen Blickwinkel, wie die Frage bearbeitet werden muss. Die ganz klare Konsequenz aus dieser Feststellung: Es kann zu der Frage, ob langwirksame Insulinanaloga für Typ 2 Diabetiker geeignet sind, nur unter dem Blickwinkel etwas gesagt werden, wenn man den Aufwand der Betroffenen und ihre individuelle Situation berücksichtigt. Das ist zum Beispiel mit der Behandlungszufriedenheit geschehen. Eine richtige Bewertung dieses Umstandes fehlt aber völlig in dem vorliegenden Vorbericht.

Seite 254: Überschrift „Applikationsfrequenz des NPH-Insulins"
In dem folgenden Abschnitt werden Überlegungen angestellt, wie oft NPH-Insulin im Vergleich zum Glargin und Detemir appliziert wurde und appliziert werden müsste.
Es scheint als hätten die Autoren dieses Abschnittes übersehen, wie die Wirkprofile sind. Es ist ganz klar, dass bei etwa anhaltendem Wirkprofil des Glargins über 24 Stunden ein vergleichbarer Insulinspiegel allenfalls durch 2 Injektionen NPH-Insulin oder Detemir erreicht wird. Allein die Einsparung von Injektionen (selbst wenn es sonst keinen Vorteil von Gargin gegenüber NPH Insulin gäbe) stellt einen erheblichen Nutzen für den Diabetiker zugunsten Glargin gegenüber NPH dar. Wo ist dieses berücksichtigt in dem Vorbericht?

Wir hoffen, dass wir durch unsere Anmerkungen und Einwände das IQWiG dazu bewegen können, offensichtlich Fehler in der Bewertung, wie sie im vorliegenden Vorbericht existieren, zu beseitigen, so, dass die Bewertung real ist und Präparate Diabetikern, die sie brauchen, weiterhin zur Verfügung stehen.

gez. Prof. Dr. med. H. v. Lilienfeld-Toal
Vorsitzender des wiss. Beirates des DDB

H. Windisch

Heinz Windisch
Bundesvorsitzender

Deutscher Diabetiker Bund e.V.
Vorstand: Bundesvorsitzender: Heinz Windisch
Stellv. Bundesvorsitzende: Rosmarie Johannes · RA Dieter Möhler
Bundesschatzmeister: Jens Pursche · Bundesjugendreferent: NN
Bundesgeschäftsstelle: Goethestr. 27 · 34119 Kassel · Telefon 05 61 – 7 03 47 70 · Telefax 05 61 – 7 03 47 71
E-Mail: info@diabetikerbund.de · Internet: www.diabetikerbund.de
Bankverbindung: Bank für Sozialwirtschaft, Essen · Kto.-Nr.: 7 219 500 · BLZ 370 205 00

A 1.5 Hoechst GmbH

Autoren:
Kurrle, Roland, Dr.
Landgraf, Wolfgang, Dr.
Sandow, Jürgen, Prof. Dr. med.

Adresse:
Prof. Dr. med. Jürgen Sandow
An der Linde 2
61479 Glashütten

Stellungnahme der Hoechst GmbH zum Vorbericht A 05-03,
Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus-Typ 2
J. Sandow, W. Landgraf, R. Kurrle

Allgemeine Bemerkungen

Der Vorbericht "Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus-Typ 2“ wurde nach mehreren Änderungen des Berichtsplans von 2005 durch ein Amendment in 2006 erstellt. Dabei wurden eine Reihe von konstruktiven Änderungen in der Berichtsmethodik vorgenommen. Ein Beispiel ist der Bezug auf die Versorgungssituation in Deutschland. Hier werden die praktischen Bedingungen der klinischen Anwendung von oralen Antidiabetika (OAD), Humaninsulin (Normalinsulin), NPH-Insulin und den zwei zu bewertenden Basalinsulinen Glargin und Detemir im Rahmen der gebräuchlichen Therapieschemata angesprochen. Der Vorbericht wird damit den tatsächlichen Bedingungen einer Behandlung des Diabetes Typ 2 besser gerecht als in der ersten Fassung des Berichtsplans angelegt. Es ergeben sich unserer Auffassung nach immer noch deutliche Lücken, weil wesentliche Formen der Behandlung wie das Therapieschema konventionelle Insulintherapie mit Mischinsulinen (ohne oder mit OAD) nicht berücksichtigt sind und auch das Therapieschema "Supplementäre Insulintherapie" nicht zum Vergleich herangezogen wird. Wir sehen deshalb dringend den Bedarf, den Vergleich mit der tatsächlichen Versorgungssituation in Deutschland konsequent herzustellen durch Verwendung der vorhandenen Evidenz aus Studien, die für den Vergleich der Therapieschemata geeignet und erforderlich sind. Weiter kommt es durch die restriktive und ungewöhnliche Bewertung der Hypoglykämien zu einem Verlust von Erkenntnis, der sich nicht hinlänglich durch die Hypothese von der "subjektiven Verzerrung“ begründen läßt. Die Häufigkeit de Auftretens schwerer Hypoglykämien beim Typ 2 Diabetes ist umfassend dokumentiert [14, 15]. Eine sachgerechte Bewertung muß aufgrund der vorliegenden Ergebnisse durchgeführt werden. Es besteht weitgehend Konsens, daß die Hypoglykämie-Wahrnehmung des Typ 2 Diabetikers von der Seite des Patienten her eine klare Aussage ermöglicht, wobei Unterschiede durch das Studiendesign RCT-24 weitgehend ausgeglichen werden. Das Vorgehen des IQWIG in der Bewertung von Hypoglykämien ist ungewöhnlich, zumindest innerhalb der Studien kann eine klare Aussage erkannt und verwendet werden.

Bei der Auswahl von Studien, die in die Bewertung eingehen (restriktive Einschlusskriterien) sind eine Reihe von Studien nicht berücksichtigt, die für die Bewertung der Versorgungssituation in Deutschland wesentlich sind. Die ausschließliche Verwendung von Studien mit NPH-Insulin als Komparator hat sich jedoch als ungeeignet erwiesen für die Darstellung der Versorgungssituation in Deutschland, weil der Vergleich mit Mischinsulinen (Kombinations-Insulinen) wesentlich ist für das Therapieschema "konventionelle Insulintherapie" [1]. Zur Definition der konventionellen Insulintherapie und der stadiengerechten Behandlung des Diabetes mellitus Typ 2 sind die aktuellen Leitlinien im Literaturverzeichnis (267) zu nennen [1, 2] und für die Beschreibung der Therapieschemata zu verwenden.

Darstellung der Grundlagen im Vorbericht (Seite 1 bis 9)

Eine Aktualisierung der zitierten Leitlinien ist geboten [1,2]. Die konventionelle Insulintherapie (CIT) wird üblicherweise mit Misch-Insulinen durchgeführt (Kombinations-Insuline). In die Bewertung des IQWIG sollen nach unserer Auffassung Studien aufgenommen werden, in denen eine CIT unter Verwendung von Mischinsulinen mit einer Basal-Therapie durch Insulinanaloga (mit oder ohne OAD) verglichen wird.

Mischinsuline (Kombinations-Insuline) haben einen wesentlichen Anteil an den Verordnungen in Deutschland, wie in Tabelle 70 ersichtlich (252) und auch unter Abschnitt "Versorgungssituation" diskutiert. In klinischen Studien ist ein Vergleich von Mischinsulinen mit den Therapieschemata Basalinsulin allein, Basalinsulin plus OAD (BOT) sowie mit der intensivierten Insulintherapie (IIT) unter Verwendung eines Basal-Insulins möglich. Wie im Vorbericht ausgeführt werden NPH-Insulin und Basalinsuline (Glargin und Detemir) nicht allein (Mono-Therapie), sondern in Therapieschemata zusammen mit kurzwirksamen Insulinen oder oralen Antidiabetika (OAD) verwendet. Die vergleichende Bewertung der Basalinsuline ist unter dem Gesichtspunkt der Versorgungssituation zu erweitern durch den Einschluss von Studien, in denen Mischinsuline (konventionelle Insulintherapie) [3] oder kurzwirksame Insulin (Supplementäre Insulintherapie) [4] als Komparator verwendet wurden. Dies entspricht dem Auftrag unter Berücksichtigung der tatsächlichen Versorgungssituation in Deutschland (Erweiterung der Einschlusskriterien).

Die Einleitung (Seite 4) ist nicht fachgerecht hinsichtlich der Hypothese von "substanz-spezifischen nützlichen und schädliche Effekten" bei mikrovaskulären und makrovaskulären Komplikationen. Die die genannte Evidenz bezieht sich ausschließlich auf Humaninsulin, bei dessen klinischer Anwendung in Langzeitstudien sehr unterschiedliche Ergebnisse vorliegen. Die angeführten Studien 13 bis 15 sind sämtlich mit Humaninsulin durchgeführt. Damit ist die unterschiedliche klinische Ausgangssituation und das Studiendesign ebenso wie die sich entwickelnde Erkenntnis über die Bedingungen der Anwendung von Humaninsulin zu nennen, es liegen keine substanz-spezifischen Effekte vor.

Eine Diskussion der mitogenen Wirkung von Humaninsulin oder Insulinanaloga ist nicht Gegenstand der Bewertung im Auftrag A05-03. Der Hinweis auf aktuelle Übersichtsarbeiten durch Einfügen im Literaturverzeichnis ist zweckmässig zur Klarstellung [5, 6]. Im Vorbericht werden die Begriffe Mitogenität und Kanzerogenität gleichwertig verwendet. Dies ist fachlich nicht richtig. Die Wachstumswirkung von Insulin ist bekannt, der Begriff Mitogenität bezeichnet die Zellproliferation, der Begriff Kanzerogenität bezeichnet eine tumorerzeugende Wirkung (Tumor-Entstehung).

In den aktuellen Leitlinien und internationalen Guidelines gibt es keine Hinweise auf eine klinische Relevanz. Damit ist auch der Hinweis auf ein "potentiell erhöhtes Mitogenitäts-Risiko" (Seite 257) missverständlich und irreführend im Hinblick auf ein hypothetisches Schadenspotential. Eine in Zusammenarbeit mit der FDA durchgeführte Fünfjahres-Sicherheitsstudie (Studie 4016) bei über 1000 Typ 2 Diabetikern hat für Insulin Glargin im Vergleich zu NPH-Insulin keinerlei Hinweise auf

unterschiedliche Schadenpotentiale und auch keine Hinweise entsprechend einer Mitogenitäts-Diskussion ergeben. Die Ergebisse liegen den Zulassungsbehörden zur Bewertung vor (EMEA und FDA). Ebenso zeigen die laufenden Berichte zur Arzneimittelsicherheit (Pharmakovigilanz) keine Signale hinsichtlich des Auftretens von Tumoren. Hinsichtlich einer fehlenden Wirkung auf das Wachstums bestehender Tumoren liegt eine neue Publikation vor [12].. Auch die wissenschaftliche Evaluierung der experimentellen Studien zu den Wachstumswirkungen von Humaninsulin ist nicht Gegenstand der Bewertung. Es ist nicht zulässig, die Literaturstelle [IQWIG 77] im Vorbericht als Evidenz ohne wissenschaftliche Begründung zu verwenden während aktuelle fachlich begründete Übersichtsarbeiten zur Verfügung stehen [5, 6] und im Bericht genannt werden können.

Die Abschnitte 2 und 3 auf Seite 5 sowie Tabelle 1 müssen nach unserer Auffassung ersatzlos gestrichen werden, weil die Hinweise nicht dem wissenschaftlichen Erkenntnisstand entsprechen und die Bewertung durch die Zulassungsbehörden (EMEA und FDA) in Tabelle 1 auszugsweise und missverständlich dargestellt ist.
Die tatsächliche Feststellungen der EMEA lauten [9]

(Seite 8-9)
"The company emphasised that the long term rodent studies did not show any increase in mammary tumours as would be expected had the compound a mitogenic potential. The clarification provided by the company was acceptable and it was concluded that the carcinogenic potential of insulin glargine was small ".

(Seite 19)
"Based on the CPMP review of data on quality, safety and efficacy, the CPMP considered by consensus that the benefit/risk profile of Lantus was favourable in the treatment of diabetes mellitus, where treatment with insulin is required.
The Committee for Proprietary Medicinal Products recommends the granting of a marketing authorisation for Lantus, subject to the chemical, pharmaceutical and biological as well as clinical follow-up measures undertaken by the company.
Based on the CPMP review of available data of a paediatric variation application, the CPMP considered that the benefit/risk profile of Lantus was favourable in the treatment of children of 6 years or above and recommended therefore the extension of indication."

Für das Fehlen einer klinischen Relevanz der bei Ratten beobachteten Veränderungen an der Injektionsstelle sind im Vorbericht zahlreiche Beobachtungen aus den klinischen Studien benannt (siehe Studienberichte). Auch aus diesem Grund muss Tabelle 1 gestrichen werden, weil die verkürzte Darstellung irreführend ist.

<u>Zielgrößen zur Bewertung von Nutzen und Schaden</u>

Die Überarbeitung des Berichtsplans (Amendment) hat wesentliche, neue Gesichtspunkte ergeben, die im Vorbericht (Seite 8) nicht vollständig berücksichtigt worden sind. Dies betrifft die Bewertung des Nutzens ebenso wie die Erfassung eines hypothetischen Schadens-Potentials durch klinische Studien. Bereits aus den Studienberichten ergibt sich ein Bild über die allgemeine und lokale Verträglichkeit. Es liegen jetzt auch die Ergebnisse der 5-Jahres-Sicherheitstudie 4016 vor. Das Ergebnis dieser Bewertung kann im Abschlußbericht für die Einschätzung des Schadens-Potentials

verwendet werden. Es ergeben sich für diese Studie keine Unterschiede zwischen Insulin Glargin und NPH-Insulin im Hinblick auf Retinopathie- Progression.
Die Einschlusskriterien unter "Intervention und Vergleichsbehandlung" (Seite 9) sind restriktiv und sie führen durch den Ausschluss von Mischinsulinen als Komparator zum Verlust wichtiger klinischer Information (siehe Versorgungssituation Seiten 251-253 im Vorbericht). Bei angemessener Berücksichtigung der Ausführungen im Vorbericht zu den gebräuchlichen Therapieschemata bei der Behandlung des Typ 2 Diabetes sind weitere Studien in die Bewertung aufzunehmen. Dieser Umstand wurde im überarbeiteten Berichtsplan nicht angemessen berücksichtigt, obwohl entsprechende Stellungnahmen eingegangen sind.
Es werden Zielgrößen (Seiten 9- 10) verwendet, die durch die üblichen klinischen Studien vom RCT-24 Typ nicht erfaßt und beschrieben werden können, weil die Entwicklung des Typ 2 Diabetes durch eine langsame Progression des Schweregrads gekennzeichnet ist. Die aufgeführten Therapieziele sind nur dann realistisch, wenn in frühen Stadien behandelt wird, sodaß vor allem die makrovaskulären Komplikationen verringert oder vermieden werden können (ticking clock hypothesis). Es muß also die Insulinbehandlung frühzeitig beginnen und über lange Zeiträume durchgeführt werden. Die Forderung nach *"hochwertigen Langzeitstudien, die primär den Nachweis eines patientenrelevanten Nutzens zum Ziel haben"* (250) ist richtig und verständlich, sie kann aber bei realistischer Betrachtung der Progression des Typ 2 Diabetes mit den vorhandenen RCT-24 Studien nicht erfüllt werden und bleibt damit für den Vorbericht ohne Bedeutung. Entsprechende Studien mit Insulin Glargin wurden frühzeitig begonnen, stehen jedoch wegen der erforderlichen langen Studienzeit (5 Jahre) für die gegenwärtige Bewertung noch nicht zur Verfügung (ORIGIN-Studie) [11]. So ist im Vorbericht die Erfassung der Gesamtmortalität als Therapieziel bei den vorliegenden Studien wenig realistisch, auch wegen der bestehenden Comorbiditäten. Langzeitstudien sind beim Typ 2 Diabetes sinnvoll, wenn die Behandlung bereits in frühen Stadien beginnt und über lange Zeiträume fortgesetzt wird. Die Zielgrößen im Vorbericht leiten sich ab von der raschen Progression des Typ 1-Diabetes. Diese Zielgrössen lassen sich nicht in gleicher Weise auf den Typ 2 übertragen (siehe Stellungnahme zum Amendment). Weil die Prävalenz der bewerteten Zielgrößen durch die langsame Progression bei Typ 2 niedrig ist, ist der im Vorbericht gewählte Weg (clinical outcome data) für eine Bewertung des Nutzens ungeeignet.

Dies gilt auch für die Zielgröße "Symptomatik bedingt durch chronische Hyperglykämie", die bei der durchgehend erreichten guten Blutzuckerkontrolle gänzlich ungeeignet ist. Anders verhält es sich bei der Zielgröße "Erblindung sowie deren Vorstufen". Hier ist das Ergebnis der abgeschlossenen 5-Jahres-Sicherheitstudie 4016 zum Vergleich NPH-Insulin vs. Insulin Glargin relevant und zu berücksichtigen. Dies gilt auch für die vergleichende Erfassung unerwünschter Arzneimittel-wirkungen in dieser 5-Jahres-Studie.
Die Feststellung *"für alle unter 4.1.3 genannten Therapieziele ist eine Evaluation im Rahmen von RCTs möglich und praktisch durchführbar"* verkennt die Prinzipien der evidenzbasierten Medizin (Bewertung der besten verfügbaren Evidenz). Die langsame Progression des Typ 2 Diabetes erfordert eine frühzeitige und langfristige Behandlung, um die erst spät eintretenden

Komplikationen zu verhindern (viz. macrovascular complications).

Berichterstattung über unerwünschte Ereignisse (256)

Zur Bewertung des Schadens-Potentials sind die Periodic Safety Update Reports (PSUR) geeignet, die den Zulassungsbehörden vorliegen und regelmäßig aktualisiert werden. Das Schadens-Potential von Insulin Glargin kann beurteilt werden, weil hier die Meldungen über beobachtete UAW zusammengefaßt sind. Die Studienberichte sind ebenfalls geeignet, während Publikationen nicht darauf angelegt sind, UAW vollständig zu erfassen und zu bewerten. Bei einer Auswertung der PSUR-Berichte zeigt sich, daß während der Behandlung mit Glargin keine Tumoren aufgetreten sind, die als behandlungs-bedingt eingestuft wurden.

Hinsichtlich der Veränderungen des Augen-Hintergrunds unter Therapie liegt das Ergebnis der 5-Jahres-Sicherheitsstudie zur Behandlung mit NPH-Insulin und Insulin Glargin den Zulassungsbehörden zur Bewertung vor (EMEA und FDA). Primäre Endpunkte war die Häufigkeit der Drei-Stufen-Progression (3-Step-Progression), es zeigte sich nach einer Behandlung über fünf Jahre zwischen NPH-Insulin und Insulin Glargin kein signifikanter Unterschied.

Zu der allgemeinen Beurteilung der Mitogenität von Insulinanaloga und möglichen Auswirkungen auf das Wachstum bestehender Tumoren liegt eine Reihe von aktuellen Übersichtsarbeiten vor [5,6], so daß frühere Hinweise der Arzneimittelkommission (AKDÄ) keine Bestätigung hinsichtlich klinischer Relevanz für Basal-Insulinanaloga gefunden haben. Die genannten Übersichtsarbeiten nehmen zur Frage einer klinischen Relevanz der biochemischen Befunde eindeutig und unter Bezug auf die wissenschaftliche Literatur Stellung. Das Instrument zur Erkennung von Signalen über ein mögliches Schaden-Potential sind die bereits genannten PSUR-Berichte, weil sie alle behandelten Patienten betreffen und nicht nur eine begrenzte Anzahl aus randomisierten klinischen Studien erfassen.

Intensivierte Insulintherapie (IIT) oder konventionelle Insulintherapie (257)

Es hängt vom Stadium der Entwicklung eines Typ 2 Diabetes ab, welche Behandlung gewählt wird und ab wann eine Insulintherapie erforderlich ist [2]. Hier ergeben sich aus der klinischen Erfahrung eine Reihe von Möglichkeiten, die eine Erfassung über RCT-24 Studien sinnvoll machen (Therapieschemata zum Beispiel Basaltherapie mit OAD, statt OAD kurz wirksames Insulin zur prandialen Substitution, supplementäre Insulintherapie). In jedem Falle ist die Wirksamkeit der Behandlung über eine Blutzuckermessung (ggf. Selbstmessung) zu kontrollieren. Die Bedingungen für eine frühzeitige, wirksame und nachhaltige Behandlung sind bei der konventionellen Insulintherapie nicht gegeben, weil die Anpassung durch Titration der Dosis entsprechend dem Zielwert für die klinische erforderliche Blutzuckereinstellung nicht durchgeführt wird. Die konventionelle Insulintherapie mit einem Mischinsulin [1] ist im Fall von Glargin und Detemir nicht möglich, weil die Basalinsuline nicht mit kurzwirksamen Insulinen gemischt werden können (feste Kombinationen). Bei dem Vergleich mit einer konventionelle Insulintherapie in der

Versorgungsrealität ist darauf zu achten, daß bei langwirksamen Insulinen nur die basale Substitution gegeben ist und zu den Mahlzeiten (prandiale Insulintherapie) ein kurzwirksames Insulin gegeben wird (intensivierte Therapie IIT). Die Diskussion im Vorbericht soll um die Beschreibung der stadiengerechten Behandlung erweitert werden, wobei die Mischinsuline der CIT zugeordnet werden.

Applikationsfrequenz von NPH-Insulin vs. Insulin Glargin (254)

Die Injektionshäufigkeit von NPH-Insulin richtet sich nach der erzielten Blutzuckereinstellung [1], für den Vergleich von Glargin und NPH-Insulin wurde aus Gründen der klinischen Relevanz in erster Linie die einmalige Injektion verglichen. Zum Zeitpunkt der Studienplanung war die wissenschaftliche Grundlage für eine Anwendung der basalen Insulin-Substitution die nächtlichen Hemmung der hepatischen Glykogenolyse. Durch das intermediär-wirksame NPH-Insulin wird diese Hemmung am besten durch die Injektion abends (bedtime) erreicht, deshalb ist die Wahl der Injektion einmal abends für einen Vergleich unter klinisch relevanten Bedingungen geeignet. Qualifikatoren wie "fairer Vergleich" sind für eine Nutzenbewertung aufgrund vorliegender Studien-Ergebnisse nicht geeignet. In früheren Studien mit NPH-Insulin wurden zwei oder mehr Injektionen pro Tag gegeben, weil die klinische Wirksamkeit unter diesen Bedingungen besser war. Ein wesentlicher Vorteil von Insulin Glargin ist die einmal tägliche Injektion morgens oder abends [10], für die vergleichende Langzeitstudien vorliegen. Für die Bewertung wichtig ist die Frage, ob eine vergleichbare Wirkung auf die mittlere Blutzuckerkonzentration (HbA1c) erreicht wird im Vergleich mit der einmal abendlichen NPH-Insulin Gabe. Im jeweiligen Therapieschema wird nicht allein mit Glargin einmal täglich oder NPH-Insulin 1x oder 2x tgl. behandelt, sondern zusätzlich mit einem OAD oder kurzwirksamen Insulin. Der entsprechende Vergleich der Applikation einmal täglich ist praxisrelevant unter Berücksichtigung des jeweiligen Therapieschema.

Injektionszeitpunkt bei einmal täglicher Gabe von Insulin Glargin morgens vs. abends (255)

Für eine Beurteilung des Patienten-relevanten Nutzens sind diese Vergleiche von wesentlicher Bedeutung. Die Praxis der Anwendung hat gezeigt, daß die Glargin-Injektion morgens die Bedingungen für eine nachhaltige Einstellung der Blutzuckerwerte erfüllt [10] und die gebräuchliche Injektion von NPH-Insulin (bedtime) ersetzen kann. Diese Erkenntnisse müssen transparent und sachgerecht in die Nutzenbewertung aufgenommen werden.
In den Studien zum Vergleich von Insulin Glargin vs. NPH-Insulin war eine Anpassung der Injektionshäufigkeit von NPH-Insulin und auch eine Wahl des Zeitpunkts morgens für NPH-Insulin nicht vorgesehen, weil die Hemmung der nächtlichen Glykogenolyse die Zielwirkung des NPH-Insulin darstellt. Das bekannte Zeit-Wirkungsprofil von NPH-Insulin führte zur Festlegung auf die Injektion abends (bedtime injection).
Die Untersuchung des Injektionszeitpunkts morgens für Insulin Glargin erfolgte auch aufgrund klinischer Beobachtungen, daß bei einzelnen Patienten die morgendliche Injektion mit geringerer Häufigkeit von hypoglykämischer Wahrnehmungen verbunden ist (intra-individueller Vergleich).

Eine Begründung für Glargin ergibt sich auch aus dem Freisetzungsprofil (Heise und Heinemann 2001). Es ist uns nicht ersichtlich, warum *"von einer Gleichwertigkeit dieser Therapieschemata nicht ausgegangen werden kann"* angesichts der Tatsache, daß die Injektion morgens und abends in kontrollierten klinischen Studien verglichen wurde [10]. Diese Option wurde von der EMEA 2003 in die EPAR-Bewertung aufgenommen (flexible timing of injection).
Es wird im Vorbericht Bezug genommen auf eine Bewertung durch die FDA aus dem Jahr 2000 [76]. Später wurde die Zulassung in USA erweitert durch flexible Wahl des Injektionszeitpunkts (flextime). Wenn im Vorbericht festgestellt wird, dass die abendliche Injektion von NPH-Insulin das Risiko schwerwiegender Hypoglykämien erhöht, so ist damit das für den Vergleich gewählte Studien-Design hinlänglich begründet, es entspricht der praktischen Anwendung von NPH-Insulin (injection at bedtime). Eine Teilung der Dosis auf zwei Injektionen morgens und abends entspricht nicht der üblichen Anwendung des NPH-Insulin [1].
Es ist nicht richtig, wenn Kombinationsstudien von Insulin Glargin + OAD mit einer Injektion morgens oder abends in der Bewertung zusammengefaßt werden. Die Wahl des Injektionszeitpunkts (morgens oder abends) ist in der Bewertung durch die EMEA beschrieben als flexible Wahl. Soweit sich für die Injektion morgens Vorteile ergeben sind diese in die Nutzenbewertung aufzunehmen.

Klinische Wirksamkeit

Die klinische Wirksamkeit von Glargin als Teil eines Therapieschemas ist nachgewiesen. Bei einer wirksamen Einstellung des Typ 2 auf Zielwerte der Glykämie sind die gleichen Wirkungen auf klinische Endpunkte zu erwarten wie bei einer vergleichbaren Therapie mit NPH-Insulin. Für eine sachgerechte Bewertung ist es wichtig, auch RCT-24 Studien mit Mischinsulinen oder supplementärer Insulintherapie [9] als Komparator zu bewerten. Die Studien unter [3] , [4]und [9] erfüllen die Bedingungen (Seite 34) daß die in der Studie eingesetzten Wirkstoffe in Deutschland für das Anwendungsgebiet zugelassen sind. Die genannten Studien sind unter dem Gesichtspunkt der Versorgungssituation in Deutschland zu bewerten und sollen in den Bericht A05-03 aufgenommen werden.

Nutzen der Insulintherapie bei verschiedenen Therapieschemata

Die nachhaltige blutzucker-senkende Behandlung beim Diabetes Typ 2 führt nach dem vorliegenden Erkenntnisstand zu einer Reduktion mikrovaskulärer Komplikationen und mit hoher Wahrscheinlichkeit bei frühzeitiger Behandlung zu einer Verringerung makrovaskulärer Komplikationen. Bei einer Therapie, die zu frühzeitiger, wirksamer und nachhaltigen Blutzuckersenkung führt werden sich die klinischen Endpunkte wahrscheinlich in gleicher Weise wie bei den für Humaninsulin dokumentierten Studien verändern (im Vorbericht wurden genannt [13 – 15], weitere Studien sind publiziert bzw. weit fortgeschritten). Es ist nicht begründet, solche Untersuchungen substanzspezifisch für verschiedene Insulin-Analoga zu fordern, um einen Zusatznutzen zu begründen. Die Therapie des Typ 2 erfolgt stadien-gerecht durch die

gemeinsame Anwendung von Insulin und anderen Wirkstoffen (z. B. OAD) zur Senkung der prä-prandialen und post-prandialen Blutzuckerwerte. Die Anwendung von Verzögerungsinsulin (NPH-Insulin) und vergleichbaren Basalinsulinen (Insulinanaloga) ist immer Teil eines Therapieschemas, wie auch im Vorbericht dargestellt (Seite 251 Diskussion).

Eine Insulin-Substitution ist für die Behandlung des Typ 2 Diabetes im Stadium der eingeschränkten Funktion der Betazellen zweckmäßig und erforderlich (Erschöpfung der Betazellen). In diesem Stadium kann Glargin verglichen werden mit

- NPH-Insulin (allein oder mit OAD)
- konventioneller Insulintherapie (Komb-Insuline, Premix-Insuline)
  (auch hier: allein oder mit OAD, [3])
- Supplementärer Insulintherapie (Apollo Studie [9]).

Die Einschränkung unter 4.1.2 ".. *muß die zusätzliche Blutzucker senkende Behandlung sowohl Bestandteil der Vergleichsbehandlung.. als auch in Deutschland zugelassen und verfügbar sein..*“ (Seite 9) führt zu einem Verlust von Information aus Studien, die für die Beschreibung der Versorgungssituation wichtig sind.

Zusatznutzen

Der Zusatznutzen ergibt sich aus der einfachen Behandlung (einmal täglich) mit Wahl der Behandlung morgens oder abends. Die Erweiterung der Behandlung (OAD oder zusätzliche prandiale Insulininjektion entsprechend der intensivierten Insulintherapie) hängt ab vom Stadium der Erkrankung (fortschreitende Erschöpfung der Betazellen).
Die Auswahl von 15 relevanten Studien als Grundlage des Vorberichts entspricht einer restriktiven Festlegung auf Einschlusskriterien, mit denen die Versorgungssituation nicht zutreffend dargestellt wird. Es erscheint wichtig, unter diesem Gesichtspunkt weitere Studien in die Bewertung einzuschließen, soweit sie für diese Frage von Bedeutung sind.

Fazit

Es gibt eine Anzahl von Alternativen in der Auswahl der Therapie-Schemata. Dabei ist zu fragen, welche Behandlung der Versorgungssituation nahekommt. Der Vorbericht stellt diese Frage, gibt aber keine hilfreichen Antworten, weil wesentliche Therapie-Schemata im Berichtsplan des Auftrags A05-03 nicht berücksichtigt worden. Die Vereinfachung auf den Vergleich NPH-Insuline vs. Basal-Insulin ist in der vorgelegten Form nicht ausreichend aussagefähig hinsichtlich der Versorgungssituation. Es sind nach unserer Auffassung hier weitere Studien vorhanden, deren Berücksichtigung Vorteile zeigt hinsichtlich der Vergleiche mit der konventionellen Insulintherapie und der supplementären Insulintherapie.

Ein zusätzlicher Nutzen ist durch das IQWIG außerhalb der Untersuchung klinischer Endpunkte darzustellen: er liegt in der praktischen Anwendung durch den Patienten. Für Glargin ist nachgewiesen, daß die einmal tägliche Injektion mit Wahl des Zeitpunkts morgens oder abends für den Patienten selbst und für Pflegekräfte einen wesentlichen Fortschritt in der Anwendung darstellt. Die Sicherheit hinsichtlich Vermeidung hypoglykämischer Komplikationen ist erhöht, bekannte Fehlerquellen durch unzureichende Mischung von NPH-Insulin Zubereitungen (auch wenn als Pen verwendet) sind ausgeschaltet.
Für die Beurteilung der Langzeitsicherheit geeignet ist nach unserer Auffassung die Berücksichtigung der jetzt vorliegenden Langzeit-Befunde aus einer Fünf-Jahres-Sicherheitsstudie (4016) mit Insulin Glargin im Vergleich zu NPH-Insulin. Hinsichtlich möglicher Veränderungen an der Injektionsstelle ergibt die im Vorbericht vorgelegte Auswertung aller Studien mit Insulin Glargin und Insulin Detemir keine Besonderheiten im Vergleich zu NPH-Insulin und im direkten Vergleich der beiden Basalinsuline.

Literatur

[1] Evidenzbasierte Leitlinie - Antihyperglykämische Therapie des Diabetes mellitus Typ 2 Herausgeber: W. A. Scherbaum, R. Landgraf (2003) http://www.awmf-leitlinien.de/ (Zugriff am 20.04.2008)

[2] Leitlinie – Behandlung des Diabetes mellitus Typ 2 S. Matthaei, H. U. Haering (2008) Diabetologie; 3 Suppl 2: S157–S161

[3] Janka HU, Plewe G, Riddle MC, Kliebe-Frisch C, Schweitzer MA, Yki-Järvinen H. (2005) Comparison of basal insulin added to oral agents versus twice-daily premixed insulin as initial insulin therapy for type 2 diabetes. Diabetes Care. 28(2):254-9.

[4] Bretzel RG, Nuber U, Landgraf W, Owens DR, Bradley C, Linn T. (2008) Once-daily basal insulin glargine versus thrice-daily prandial insulin lispro in people with type 2 diabetes on oral hypoglycaemic agents (APOLLO): an open randomised controlled trial. Lancet. 29;371(9618):1073-84. RCT-44

[5] Zib I, Raskin P. (2006) Novel insulin analogues and its mitogenic potential. Diabetes Obes Metab. 8(6):611-20. Review

[6] Le Roith D. (2007) Insulin glargine and receptor-mediated signalling: clinical implications in treating type 2 diabetes. Diabetes Metab Res Rev. 23(8):593-9. Review.

[7] National Institute for Clinical Excellence. (2002) Guidance on the use of long acting insulin analogues for the treatment of diabetes insulin glargine. United Kingdom: London:National Institute for Clinical Excellence (NICE),. Guidance TAG No. 53 accessed 24-04-08 http://www.nice.org.uk/nicemedia/pdf/53_Insulin_analogues_full_guidance.pdf

[8] Horvath K, Jeitler K, Berghold A, Ebrahim SH, Gratzer TW, Plank J, Kaiser T, Pieber TR, Siebenhofer (2007) A Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus (comments p. 18 ff Rosenstock, Fritsche and Riddle, Zugriff am 22.04.08)

[9] EMEA EPAR Scientific Discussion accessed 2005, CPMP-615-00

The early version of 2000 was updated in 2003 (accessed 24-04-08, see p. 8-9)

[10] Standl E, Maxeiner S, Raptis S, Karimi-Anderesi Z, Schweitzer MA; HOE901/4009 Study Group. (2005) Good glycemic control with flexibility in timing of basal insulin supply: a 24-week comparison of insulin glargine given once daily in the morning or at bedtime in combination with morning glimepiride. Diabetes Care. 28(2):419-20. RCT-24

[11] Origin Trial Investigators , Gerstein H, Yusuf S, Riddle MC, Ryden L, Bosch J. (2008) Rationale, design, and baseline characteristics for a large international trial of cardiovascular disease prevention in people with dysglycemia: the ORIGIN Trial(Outcome Reduction with an Initial Glargine Intervention). Am Heart J. Jan;155(1):26-32

[12] Erbel S, Reers C, Eckstein VW, Kleeff J, Büchler MW, Nawroth PP, Ritzel RA. (2008) Proliferation of Colo-357 pancreatic carcinoma cells and survival of patients with pancreatic carcinoma is not altered by insulin glargine. Diabetes Care. 2008 Mar 3 [Epub ahead of print]

[13] Henderson JN, Allen KV, Deary IJ, Frier BM. (2003) Hypoglycaemia in insulin-treated Type 2 diabetes: frequency, symptoms and impaired awareness. Diabet Med. 20(12):1016-21

[14] Akram K, Pedersen-Bjergaard U, Carstensen B, Borch-Johnsen K, Thorsteinsson B.(2006) Frequency and risk factors of severe hypoglycaemia in insulin-treated Type 2 diabetes: a cross-sectional survey. Diabet Med. 2006 Jul;23(7):750-6.

[15] Akram K, Pedersen-Bjergaard U, Borch-Johnsen K, Thorsteinsson B. (2006) Frequency and risk factors of severe hypoglycemia in insulin-treated type 2 diabetes: a literature survey. J Diabetes Complications. 20(6):402-8.

A 1.6 Novo Nordisk Pharma GmbH

Autoren:
El-Haschimi, Karim, Dr. med.
Grundner, Martin, Dr. med.
Koenen, Christoph, Dr. med.

Adresse:
Dr. med. Karim El-Haschimi
Novo Nordisk Pharma GmbH
Brucknerstraße 1
55127 Mainz

Stellungnahme der Novo Nordisk Pharma GmbH zum Vorbericht des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) „Auftrag A05-03; Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2" Version 1.0 vom 18.03.2008

Vorbemerkung

Wir hatten bereits an früherer Stelle – siehe unsere Schreiben vom 02.04.2007 und 05.07.2007 sowie unsere Stellungnahme zum Berichtsplan vom 26.07.2007 – darauf hingewiesen, dass die Nutzenbewertung aufgrund der positiven Beurteilung von Insulindetemir gegenüber Humaninsulin durch die EMEA als Verfahren nicht sachgerecht ist bzw. nicht der aktuellen Gesetzeslage entspricht[1]. Insbesondere sehen wir auch die Nutzenbewertung durch das IQWiG hinsichtlich der Kombinationstherapie mit oralen Antidiabetika als verfrüht an, da erst 2007 die Erweiterung der Anwendung in die Fachinformation aufgenommen wurde. Langzeitstudien, insbesondere solche, die den Versorgungsalltag abbilden, können hier naturgemäß nicht vorliegen. Das SGB V sieht zudem auch für neue Arzneimittel in SGB V §31 Abs.2a die Möglichkeit vor, neue Studien vorzulegen, ehe eine Bewertung vorgenommen wird. Vor allem, wenn die Nutzenbewertung als Vorstufe einer Kosten-Nutzenbewertung angesehen wird, gilt diese Bestimmung, so dass die Vorgehensweise des IQWiG hier grundsätzlich in Frage zu stellen ist.

Einleitung

Grundsätzlich bemängeln wir an der vorliegenden Nutzenbewertung des IQWiG, dass die Bewertungskriterien, die durch das Institut gewählt und festgelegt worden sind, nicht den üblichen Standards entsprechen. Ein Scoping Workshop unter Einbindung der Fachgesellschaften und Experten hätte hier sicher ein anderes Ergebnis gebracht. Dies betrifft insbesondere die Bewertung und Definition der Hypoglykämien als maßgeblichen Endpunkt für die Beurteilung einer adäquaten Diabetes-Therapie. Der vollständige Ausschluss der nicht schweren/ schwerwiegenden Hypoglykämien durch das IQWiG führt dazu, dass die signifikanten Unterschiede zugunsten von Insulindetemir gegenüber Humaninsulin nicht gewertet werden. Damit wurde ein Hauptparameter, der die Vorteile der langwirksamen Insulinanaloga beschreibt, per definitionem eliminiert. Diese Darstellung der Ergebnisse steht in eklatantem Widerspruch zu der anerkannten Bewertung durch die Zulassungsbehörde, wie sie sich in der aktuellen Fachinformation zu Insulindetemir widerspiegelt[2]. Der durch die Zulassungsbehörden festgestellte Zusatznutzen gegenüber Humaninsulin in der Fachinformation zu Insulindetemir ist dem BSG folgend, als allgemein anerkannter, medizinischer Wissensstand zu werten (siehe auch Stellungnahme zum Berichtsplan)[1,2,3]. Zugleich muss die Fachinformation in die Bewertung des patientenrelevanten Zusatznutzens gemäß SGB V § 35 Abs. 1b einfließen. Die kürzlich durch das Europäische Patentamt (EPA) vorgenommene Nominierung von Insulindetemir als eines von zwölf Produkten zum „Inventor of the Year 2008" mit der Begründung *„…für die entscheidende Verbesserung der Sicherheit und Lebensqualität von Diabetespatienten durch ein neues Insulinderivat, das einmal täglich verabreicht wird*" bekräftigt den Innovationscharakter von Insulindetemir[4].

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien:

In RCTs konnten die Vorteile für Insulindetemir eindeutig belegt werden. Wie auch vom Institut dargelegt, zeigen diese RCTs eine gegenüber Humaninsulin reduzierte Hypoglykämierate für Patienten, die mit Insulindetemir behandelt werden, besonders bei den nächtlichen Hypoglykämien. Des Weiteren wird im Vorbericht gezeigt, dass es unter Behandlung mit Insulindetemir zu einer signifikant geringeren Gewichtszunahme kommt. Ebenso wurden Vorteile für die gesundheitsbezogene Lebensqualität im Bereich Mental Health und für die Behandlungszufriedenheit gezeigt. Diese Daten werden unterstützt durch die Praxisbeobachtung PREDICTIVE™. In dieser international durchgeführten Untersuchung konnten

nach Ein- oder Umstellung auf Insulindetemir u.a. niedrigere Hypoglykämieraten und eine günstige Gewichtsentwicklung in unterschiedlichen Therapieschemata gezeigt werden[5,6,7,8]. Diese Daten stellen somit eine im täglichen Versorgungsalltag bei insgesamt über zehntausend Patienten dokumentierte Bestätigung der Ergebnisse der klinischen Studien dar.

Auch zwei klinische Studien, die mit einer Laufzeit von 20 und 22 Wochen bisher nicht in die Bewertung einbezogen wurden, zeigen deutliche Vorteile gegenüber NPH-Insulin[9,10,11]. So konnte von Philis-Tsimikas et al. eine deutliche Reduktion der Gesamt-Hypoglykämierate gezeigt werden, in beiden Studien bestanden gegenüber NPH-Insulin signifikante Verbesserungen bezüglich der Rate nächtlicher Hypoglykämien, wie auch bezüglich der Gewichtsentwicklung. Diese geringfügig unterhalb der vom IQWiG definierten Mindeststudiendauer liegenden RCTs unterstützen die vorhandene Evidenz zu Insulindetemir.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien:

Hypoglykämien: Auf S. 156 des Vorberichts wird darauf hingewiesen, dass eine Darstellung der hypoglykämischen Gesamtereignisse wie auch der nächtlichen Ereignisse aufgrund einer sehr hohen Ergebnisunsicherheit *„lediglich der Vollständigkeit halber"* erfolge, allerdings keine weitere Berücksichtigung dieser (unter Insulindetemir gegenüber NPH-Insulin z.T. deutlich seltener auftretenden) Ereignisse erfolge. Eine Einbeziehung gerade auch dieser Episoden ist jedoch unbedingt zu fordern. Es wird vom IQWiG anerkannt, dass die Rate nicht schwerwiegender hypoglykämischer Ereignisse unter Insulindetemir signifikant niedriger ist als unter NPH-Insulin. In Bezug auf die im Folgenden gemachten Ausführungen ist es nicht statthaft, diese Erkenntnisse aus der Nutzenbewertung auszuschließen.

Auf S.121 des Vorberichtes wird ausgeführt, dass *„alle Angaben zu Hypoglykämien (…) auf Angaben von Patienten"* beruhen und somit anfällig für Ergebnisverzerrungen seien. In der weiteren Argumentation werden die nicht schwerwiegenden Hypoglykämien als nicht ausreichend ergebnissicher bezeichnet und aus der Nutzenbewertung ausgeschlossen. Es ist unbestritten, dass *„Angaben zu Hypoglykämien (…) auf Angaben von Patienten"* beruhen, dass hier die subjektive Wahrnehmung eine entscheidende Rolle spielt. Weiterhin ist unbestritten, dass ein offenes Studiendesign hinsichtlich dieses Parameters eine grössere Ergebnisunsicherheit beinhaltet als ein (doppel)blindes Design. Der Ausschluss der nicht schwerwiegenden Ereignisse aus der Bewertung ist jedoch hierdurch in keiner Weise gerechtfertigt und führt dazu, dass ein außerordentlich wichtiger Parameter in der Beurteilung der Therapie mit Insulindetemir, mit einer mittlerweile sehr breiten Datenbasis, letztlich ignoriert wird. Eine umfassende und praxisrelevante Beurteilung verschiedener Präparate hinsichtlich ihrer Eignung zur Behandlung des Typ 2 Diabetes ist somit praktisch nicht mehr möglich.

Die Tatsache, dass die Wahrnehmung von Hypoglykämien im subjektiven Erleben des Patienten liegt und dass diese Wahrnehmung durchaus vom Wissen um die konkrete Zuteilung zu einem bestimmten Therapiearm beeinflusst werden kann, ist nicht per se mit einer inakzeptablen Ergebnisunsicherheit gleichzusetzen. Die Annahme, dass das Wissen um den Erhalt der „neuen" Substanz innerhalb einer klinischen Studie eine eher günstigere Wahrnehmung von hypoglykämischen Episoden (im Sinne einer verringerten Hypoglykämierate) bedingt, entbehrt der nötigen Evidenz und wird seitens des IQWiG nicht mit entsprechenden Quellen belegt. Dies wäre jedoch unbedingt zu fordern. In ihrem Kommentar zur *„draft guidance"* der FDA zu *Patient Reported Outcomes*[12] (PRO: gemäß Definition im angeführten Dokument "*a measurement of any aspect of a patient´s health status that comes directly from the patient*") stellt Clare Bradley, eine der international renommiertesten Expertinnen auf diesem Gebiet, fest[13]:

> *„New treatments may carry risks and possible unwanted effects as well as benefits and it is not appropriate to assume that patients will always be more positive about a new treatment than about an old treatment*".

Auch einer generellen Forderung nach Verblindung im Rahmen klinischer Studien kann sie sich nicht anschließen:

> *„Blinding should not be assumed to be universally desirable and in itself can distort study results*".

Auch die mündliche und schriftliche Aufklärung von Patienten vor der Teilnahme an einer klinischen Studie fokussiert aus ethisch-rechtlichen Gründen dann auch auf die möglichen Risiken des neuen Produktes im Falle einer Zuteilung auf dieses neue Therapieprinzip bzw. Präparat und lässt eher darauf schließen, dass Patienten das neuere Präparat kritischer betrachten.

Die Einbeziehung von Daten über subjektiv wahrgenommene Hypoglykämien in die Nutzenbewertung ist allein deshalb sinnvoll, weil gerade auch das subjektive Erleben relevante Entscheidungen triggern kann (Essverhalten, Dosisentscheidungen, Festlegung individueller Zielwerte), die einen erheblichen Einfluss auf das Gesamtbefinden des Patienten haben. Auf diese Aspekte geht die American Diabetes Association (ADA) in ihrem Workgroup Report *„Defining and Reporting Hypoglycemia in Diabetes*" ein[14]:

> *„The rational fear of hypoglycemia can lead to worsening of metabolic control as well as tension with, and restriction of personal freedoms (…) by, anxious and overprotective loved ones, colleagues, or employers"*.

Gerade auch hinsichtlich der Schwierigkeiten, allgemeingültige Grenzwerte für die Definition einer Hypoglykämie festzulegen, führt das Gremium eine praxisorientierte Betrachtungsweise an:

> "*Thus, symptoms of a low plasma glucose, it can be argued, may be included in a definition of hypoglycemia*".

Auch in der UK Prospective Diabetes Study (UKPDS), die grundlegende Erkenntnisse bezüglich der Therapie des Typ 2 Diabetes erbracht hat, wurden *nicht schwere* Hypoglykämien als wichtiger Parameter in die Beurteilung der unterschiedlichen Therapieregimes integriert[15]. So wurden im Rahmen der UKPDS Hypoglykämien auf Grundlage der bei den Visiten gemachten Patientenangaben erfasst und kategorisiert. Neben schweren Ereignissen (Fremdhilfe erforderlich) wurden auch nicht schwere Ereignisse erfasst und berichtet. In der Diskussion führen die Autoren aus:

> *„The risks of hypoglycaemia (…) are perceived by patients as difficulties that limit their ability to achieve improved glucose control*".

Die European Agency for the Evaluation of Medicinal Products (EMEA) gibt in ihrer "Note for guidance on clinical investigation of medicinal products in the treatment of diabetes mellitus" eine Beschreibung der zur Zulassung von Präparaten zur Behandlung des Diabetes Mellitus notwendigen Studien[16]. Sie würdigt die besondere Bedeutung von Hypoglykämien in der Beurteilung eines neuen Verfahrens/ Präparates und schließt neben der Erfassung von schweren („*major*") Hypoglykämien und durch Blutglukosemessung bestätigter Ereignisse (Blutglukose unter 3 mmol/l) explizit allein auf der subjektiven Wahrnehmung des Patienten basierende Hypoglykämienennungen in ihren Katalog ein:

> *„episodes suggestive of hypoglycaemia, where blood glucose measurement were not available*."

In der Fachinformation zu Levemir® wird entsprechend festgestellt[2]:

> *„In Studien zur Kombinationstherapie von oralen Antidiabetika mit Insulin führte die Therapie mit Levemir verglichen mit NPH-Insulin zu einem 61-65% niedrigeren Risiko für leichte nächtliche Hypoglykämien*."

Die Aufnahme eines entsprechenden Hinweises in die Fachinformation lässt zweifelsohne die Annahme zu, dass der geschilderte Sachverhalt für die Zulassungsbehörde einen therapierelevanten Aspekt darstellt.

In einer 2007 veröffentlichten Meta-Analyse berichten Garber et al. über verringerte Hypoglykämieraten unter Insulindetemir vs. NPH-Insulin bei älteren (≥ 65 Jahre) Patienten. Gerade hinsichtlich der oft vorhandenen Komorbiditäten und der besonderen Lebensumstände spielt auch hier die Angst vor Hypoglykämien aus Sicht der Autoren eine besondere Rolle als mögliche „*barrier to adherence*"[17].

In der Zusammenfassung wird auf S. 241 des Vorberichts darauf hingewiesen, dass für „*den Vergleich Insulin Detemir vs. NPH-Insulin, jeweils zweimal täglich morgens und abends und in Kombination mit OAD (…) statistisch signifikant weniger schwerwiegende Hypoglykämien unter Insulin Detemir*" aufgetreten sind. Allerdings wird dieser Feststellung u.a. aufgrund des offenen Studiendesigns lediglich hinweisender Charakter bezüglich einer Überlegenheit von Insulindetemir zugebilligt. Nicht zuletzt unter Berücksichtigung der unter „*5.3.4.2 Definition einer Hypoglykämie*" getroffenen Aussagen (S. 121:"*Diese im Folgenden als „schwerwiegende Hypoglykämien" bezeichneten Ereignisse können (…) als verzerrungsarm angesehen werden*") ist diese Bewertung inhaltlich und methodisch nicht schlüssig.

In Bezug auf die oben gemachten Ausführungen ist es nicht nachvollziehbar, dass die vom IQWiG bestätigten Erkenntnisse bezüglich einer signifikanten Reduktion der nicht schwerwiegenden hypoglykämischen Ereignisse unter Insulindetemir im Vergleich zu NPH-Insulin, aus der Nutzenbewertung ausgeschlossen werden.

Gewicht: In Heft 16 der „*Gesundheitsberichterstattung des Bundes*" gibt das Robert Koch-Institut einen Überblick über das Vorliegen von Übergewicht und Adipositas in Deutschland[18]. Dabei wird auf die stetig zunehmende Zahl übergewichtiger Menschen hingewiesen, vor allem aber werden die damit verbundenen gesundheitlichen Probleme („*Das Risiko, Folge- und Begleiterkrankungen zu entwickeln, steigt mit zunehmendem BMI*"), wie auch die Kostenproblematik dargestellt:

> „Die *entstehenden direkten und indirekten Kosten sind hoch und werden weiter zunehmen, wenn Maßnahmen der Prävention von Übergewicht und Adipositas nicht wirksam umgesetzt werden*."

Es liegt mittlerweile eine Vielzahl von Arbeiten unterschiedlichen Designs zu den gesundheitlichen Auswirkungen von Übergewicht und Adipositas vor. So konnte eine prospektive Kohortenstudie im Rahmen der Framingham Heart Study eine Assoziation zwischen Übergewicht bzw. Adipositas und einer eingeschränkten Lebenserwartung zeigen[19]. Adams et al. konnten in ihrer großen prospektiven, insgesamt 527.265 Teilnehmer umfassenden Untersuchung zu BMI und Mortalität die starke Assoziation zwischen Adipositas und allgemeinem Mortalitätsrisiko zeigen[20]:

> …„*obesity was strongly associated with the risk of death in both men and women in all racial and ethnic groups and at all ages.*"

Die Deutsche Adipositas-Gesellschaft weist insbesondere auch auf Komorbiditäten und Komplikationen, wie auch die Assoziation mit dem Metabolischen Syndrom hin[21]:

> „*Übergewicht/Adipositas gilt als der wichtigste Promotor des Metabolischen Syndroms, welches mit einem hohen Atheroskleroserisiko assoziiert ist*."

Besteht bereits ein Typ 2 Diabetes, so kommt der Gewichtskontrolle eine besondere Bedeutung zu. In ihrem Review zu den Effekten einer Gewichtsabnahme bezüglich kardiovaskulärer Risikofaktoren bei übergewichtigen Patienten mit Typ 2 Diabetes kommen Anderson et al. zu dem Schluss[22]:

> "*Weight management may be the most important therapeutic task for most obese Type 2 diabetic individuals.*"

Daousi et al. analysierten die Daten von 2721 Patienten mit Typ 2 Diabetes und stellten fest, dass adipöse Patienten mit Typ 2 Diabetes im Durchschnitt eine

schlechtere Blutzucker-Einstellung, höheren Blutdruck und schlechtere Lipid-Profile aufwiesen als nichtadipöse Patienten[23]. Sie kommen zu dem Schluss:

> "*Obesity is associated with significantly worse cardiovascular risk factors in this patient group*".

Die DDG stellt in ihren Leitlinien *„Behandlung des Diabetes mellitus Typ 2* " fest[24]:

> *„Die überwiegende Mehrheit der Patienten ist (…) übergewichtig.*"

Dementsprechend wird (bei Übergewicht) eine Gewichtsreduktion als Therapieziel genannt.

In der Zusammenfassung der Beleglage im Vorbericht wird angegeben, dass sich in allen vom IQWiG berücksichtigten Studien zum Vergleich von Insulindetemir vs. NPH-Insulin unter NPH-Insulin eine signifikant höhere Gewichtszunahme dargestellt hat (S. 259). Es wird seitens des IQWiG jedoch angeführt, dass die *„im Allgemeinen beobachtete Gewichtszunahme nach Einleitung oder Intensivierung einer Insulintherapie (…) zum großen Teil als ein Wiedererreichen des vor der Therapie infolge ungenügender Blutzuckereinstellung verlorenen Gewichts einzuschätzen*" sei. Die im Vorbericht zu dieser Thematik zitierte Quelle[25] verweist jedoch auf eine als „Letter" veröffentlichte Arbeit der Review-Autorin selber[26], die methodisch wie inhaltlich kaum als generelle Referenz zu dieser Thematik dienen kann. Das Design (retrospektive Kohortenstudie) erfüllt kaum die vom IQWiG selbst formulierten Mindestanforderungen an belastbare klinische Studien. Weiterhin erfolgte ein Einschluss der Patienten nur, wenn mit Einleitung einer Insulintherapie die oralen Antidiabetika abgesetzt wurden, dies entspricht nicht generell den alltäglichen Gegebenheiten in der Praxis und muss hinsichtlich der Beurteilung des weiteren Gewichtsverhaltens kritisch gesehen werden. Zudem handelt es sich bei den Patienten um ausgesprochen schlecht eingestellte Patienten mit einem mittleren HbA_{1c} von 10,9 % (bei Einleitung der Insulintherapie), die sich alleine in diesem Parameter deutlich von den Behandlungsgruppen der vorgelegten Studien zu Insulindetemir unterschieden. Gerade hinsichtlich der Abgrenzung einer etwaigen wünschenswerten Gewichtszunahme im Rahmen einer Stoffwechselverbesserung gegenüber einer nicht erwünschten Zunahme muss die Übertragbarkeit der Ergebnisse auf eine im Durchschnitt deutlich besser eingestellte Studienpopulation stark angezweifelt werden. Weder Design noch Inhalt der o.a. Quelle sind in der Lage, die damit begründeten Aussagen ausreichend zu stützen.

Verschiedene Mechanismen werden im Zusammenhang mit einer Gewichtszunahme nach Einleitung bzw. Intensivierung einer Insulintherapie diskutiert. Insbesondere der Kalorienverlust infolge einer Glukosurie bei unzureichend eingestellter Stoffwechsellage muss hier angeführt werden, ist doch das Unterschreiten der Nierenschwelle nach Einleitung einer entsprechenden Therapie Ausdruck der angestrebten Stoffwechselverbesserung und sollte nicht als physiologische Angleichung an ein vor Stoffwechseldekompensation bestehendes individuelles Zielgewicht interpretiert werden.

Aber auch eine veränderte Nahrungsaufnahme in Reaktion auf erlebte bzw. befürchtete Unterzuckerungen, sowie Störungen der zentralnervösen Regulation von Hunger und Nahrungsaufnahme sind bedeutsam, führen diese Effekte doch zu einer primär unerwünschten Gewichtszunahme[27,28,29]. Bedeutsam ist weiterhin die Tatsache, dass Patienten einer notwendigen Insulintherapie oftmals ablehnend gegenüberstehen, da sie eine entsprechende Gewichtszunahme erwarten[27,30]. Die antizipierte Angst vor einer weiteren Zunahme des zumeist ohnehin zu hohen Körpergewichtes bei Einleitung einer eigentlich gebotenen Therapie kann hierbei zu einer relevanten Therapiebarriere mit der Folge einer weiterhin unzureichenden Stoffwechseleinstellung führen.

> *„Weight gain is often a major concern for many patients and (…) issues of weight gain as a result of insulin therapy must be addressed*." [30].

Der beobachtete Gewichtseffekt unter Insulindetemir stellt sich konsistent in allen betrachteten Studien dar und realisiert sich bei unterschiedlichen HbA_{1c}-

Ausgangsniveaus und z.T. deutlich unterschiedlichen HbA_{1c}-Senkungen im Studienverlauf. Weiterhin zeigt er eine Korrelation zu den BMI-Daten bei Studienbeginn im Sinne einer geringeren Gewichtszunahme mit zunehmendem Ausgangs-BMI (Kombination mit oralen Antidiabetika[31]) bzw. zunehmender Gewichtsunterschiede zwischen den Behandlungsgruppen (Basis-Bolus-Therapie[32]). Zudem liegen mittlerweile Daten bei Typ 1 Diabetes vor, die über die vom IQWiG *„für eine realistische Abschätzung der klinischen Relevanz*" geforderte Behandlungsdauer von 2 Jahren eine unter Insulindetemir gegenüber NPH-Insulin signifikant geringere Gewichtszunahme zeigen[33].

Vor dem Hintergrund der referierten Daten sind die dokumentierten Gewichtseffekte als Zusatznutzen einer Therapie mit Insulindetemir anzuerkennen.

Aktueller Behandlungsstandard (DIG-Studie): Betrachtet man aktuelle Marktdaten, so ist für die basale Insulintherapie das Humaninsulin als Therapiestandard durch das Analoginsulin abgelöst worden. Die im Vorbericht zitierten Daten der DIG Studie[34] (S. 252 ff.) zu den Behandlungsoptionen und Therapieschemata sind nicht aktuell. Diese Daten sind zu einem Zeitpunkt erhoben worden, zu dem die Therapie mit langwirksamen Insulinanaloga sich gerade etablierte. Insulindetemir war zu dem Zeitpunkt der Erhebung in keiner Indikation auf dem Deutschen Markt zugelassen. Aktuellere Marktzahlen zeigen, dass in den letzten Jahren eine Verschiebung der Therapieschemata stattgefunden hat[35]. 2007 wurden ca. 45% der OAD-Therapie-Versager bei einer Neueinstellung auf ein Basalinsulin eingestellt und nur noch 20%, respektive 17%, erhalten eine konventionelle oder eine Basis-Bolus Therapie. Daten des IMS Disease Analyzer zeigen, dass die langwirksamen Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2, gerade für die reine Basaltherapie, inzwischen zur Standardtherapie geworden sind[35]. Inzwischen werden 67% der Patienten, die eine reine Basaltherapie erhalten, mit Analoginsulin behandelt. Demzufolge wird hier eine moderne etablierte Therapie mit einem veralteten Therapieschema verglichen.

Verringertes mitogenes Potenzial für Insulindetemir: Wenn man davon ausgeht, dass die Bewertung eines hypothetischen Schadenspotenzials relevant ist, dann müssen sich die vorhandenen Informationen in letzter Konsequenz als positives Nutzenverhältnis in der Bewertung wieder finden. EMEA und FDA treffen Zulassungsentscheidungen auf wissenschaftlicher Basis, wesentlicher Bestandteil ihrer Aufgabe ist die Bewertung des Schadenspotenzials. Obwohl die zitierten Aussagen der Zulassungsbehörden in Tabelle 1 S.6 des Vorberichts deutlich darstellen, dass für Insulindetemir von einem geringeren mitogenen Potenzial und daher von einem geringeren Schadenspotenzial im Vergleich zu Humaninsulin ausgegangen werden muss, spiegelt sich das positive Nutzenverhältnis nicht in der abschließenden Bewertung wider.

> *„Es wurde gezeigt, dass Insulin Detemir gegenüber Humaninsulin eine geringere Rezeptor-Bindungsaffinität und […] geringeres mitogenes Potenzial besitzt.*"(FDA, 2003; Vorbericht S. 6, Tab 1).
>
> *„Das mitogene Potenzial von Insulin Detemir scheint gegenüber Humaninsulin reduziert zu sein.*" (EMEA, 2004; Vorbericht S. 6, Tab 1).

Die Bewertung der Zulassungsbehörden findet sich auch in der Formulierung der Fachinformation zu Insulindetemir wieder:

> *„Präklinische Daten, (...) lassen keine besonderen Gefahren für den Menschen erkennen. Rezeptoraffinitätsdaten und in-vitro-Mitogenitätstests erbrachten keinen Nachweis für ein erhöhtes mitogenes Potenzial im Vergleich zu Humaninsulin*." (Fachinformation Levemir®, Juni 2007[2]).

Der anerkannte medizinische Nutzen, wie auch der Schaden ist nicht anders darzustellen als abschließend durch die arzneimittelrechtliche Zulassung vorgegeben (siehe Zitat Grunds. Anmerkungen Punkt 1; *BSG, vom 31.05.2006)*[3]. Demzufolge muss die Bewertung der Zulassungsbehörden Berücksichtigung finden und darf nicht im Vorbericht grundsätzlich in Frage gestellt werden. Insofern muss der

Hinweis auf ein verringertes Schadenspotenzial zu einer Bewertung zu Gunsten von Insulindetemir führen.

Minimal Important Difference (MID): Im Vorbericht (S. 171/ S.187) wird bemängelt, dass keine der Studien eine *Minimal Important Difference* (MID) für einige der verwendeten Messinstrumente angebe. Es ist richtig, dass es für verschiedene Messinstrumente im Bereich der *Patient Reported Outcomes* (PRO) einen validierten MID-Wert gibt. Richtig ist aber auch, dass dies nicht für alle Instrumente gefordert ist. Eine statistisch signifikante Änderung bzw. Verbesserung der Ergebnisse in einem PRO-Instrument, das gezielt Bereiche erfasst und abfragt, die explizit als wichtig und einflussreich von Patienten definiert wurden, ist per definitionem als relevant zu bewerten. Die hier verwendeten Instrumente wurden nach international anerkannten Standards (z.B.: FDA draft guideline für PRO[12]) in Zusammenarbeit mit Patienten entwickelt. Clare Bradley, Entwicklerin des DTSQ Questionnaire stellt dazu eindeutig klar[13]:

> *„I agree that it can be helpful to consider the MID for clinical measures which are intermediate outcomes that may not be important outcomes for the patient in themselves but only in so far as they are predictors of other outcomes that are important (e.g. HbA1c measures of blood glucose in diabetes). With some PRO measures that ask about symptoms, health status or visual functioning without asking about the importance of the issue in question for the patient, it may also be useful to determine MID. However, a statistically significant difference on measures of treatment satisfaction that have been designed explicitly to measure issues of importance to patients (e.g. DTSQ) will necessarily be an important difference. So too will be a statistically significant difference on an individualised measure of the impact of a condition on quality of life, where the importance of an aspect of life for an individual's quality of life is part of the assessment…."*

Demzufolge sind die im Vorbericht beschriebenen statistisch signifikanten Ergebnisse als für den Patienten relevant zu bewerten und der Zusatznutzen für diese Dimensionen ist anzuerkennen.

Hypoglykämiewahrnehmung: Auch für die Dimension der Hypoglykämiewahrnehmung (S. 207) konnten statistisch signifikant bessere Ergebnisse für beide Studien zu Insulindetemir gezeigt werden. Entsprechend der Aussage von Clare Bradley[13] (siehe Punkt MID) ist diese Veränderung als relevant und klinisch signifikant zu betrachten. Die verbesserte Wahrnehmung der Hypoglykämien stellt eine der entscheidenden Dimensionen der Therapiezufriedenheit dar und untermauert die Ergebnisse der klinischen Daten. Die Tatsache, dass es für Patienten mit Problemen in der Wahrnehmung von Hypoglykämien strukturierte Schulungs- und Behandlungsprogramme mit entsprechenden Zertifizierungen bzw. Anerkennungen (z.B. durch die Deutsche Diabetes-Gesellschaft oder das Bundesversicherungsamt) gibt, dokumentiert, dass eine auf der folgenden Aussage basierende Wertung praxisfremd ist und relevante Patientenbedürfnisse und Sicherheitsaspekte ignoriert:

> "*Da es sich nur um den Aspekt der Hypoglykämiewahrnehmung handelt, kann das Ergebnis nicht als ein Hinweis auf die Überlegenheit von Insulin Detemir gegenüber NPH-Insulin in Bezug auf die Therapiezufriedenheit angesehen werden.*" (Vorbericht S. 207).

Nichtberücksichtigungsrate: Die Kriterien, die üblicherweise auf die klinische Datenerhebung angewandt werden, lassen sich nicht in jedem Fall auf die Datenerhebung mit *Patient Reported Outcomes* (*PRO*)-Instrumenten übertragen, dies zeigt sich ganz deutlich in den Bestrebungen sowohl der FDA als auch der EMEA speziell für diesen Bereich geeignete Guidelines zu entwickeln. Es ist ein Problem, dass gerade das Ausfüllen von Fragebögen für chronisch kranke Patienten eine Belastung darstellt und häufig zu einer erhöhten Anzahl fehlender Daten führt.

Eine Nichtberücksichtigungsrate von 17 % ist durchaus nicht unüblich und wird für den Bereich der PRO als geringe Nichtberücksichtigungsrate eher positiv gewertet. Nichtberücksichtigungsraten von über 25% sind im Bereich der PRO Instrumente keine Seltenheit:

> *"For example, the SAP can specify that a domain will be treated as missing if more than 25 percent of the items are missing; if less than 25 percent of the items are missing, the domain score can be taken to be the average of the nonmissing items."* (FDA draft guideline für PRO[12]).

Die hier im Vorbericht vorgebrachte Begründung kann deshalb nicht ausschlaggebend sein, ein statistisch signifikantes Ergebnis zu negieren. Grundsätzlich ist zu beachten, dass die Nichtberücksichtigungsraten zwischen den Gruppen vergleichbar sind.

Korrekturen zum Vorbericht:

S. 261, Zeile 2: Hier wird im Abschnitt zu Insulindetemir fälschlicherweise von Insulinglargin gesprochen.

S. 72, Absatz 2, Zeile 1: „Nur in Studie 1737 wurden"…, hier handelt es sich um die Studie *1373*.

S. 113, Absatz 3, Zeile 6: …"einer einmal täglichen Gabe von NPH-Insulin, jeweils"… Hier muss es heißen: …„einer zweimal täglichen Gabe von NPH-Insulin, jeweils"… Sinngemäß muss es auf S. 136, Absatz 3, Zeile 2 heißen: …"in zweimal täglicher Gabe, jeweils"…

Fazit:

Die vorliegenden Studien zu Insulindetemir in der Indikation Diabetes mellitus Typ 2 zeigen den Zusatznutzen des langwirksamen Insulinanalogons gegenüber Humaninsulin. Neben dem verbesserten pharmakodynamischen und –kinetischen Wirkprofil zeigen sich die verbesserten Produkteigenschaften vor allem in einer niedrigeren Rate von Hypoglykämien.

Zusätzlich wird mit Insulindetemir ein deutlich verbessertes Gewichtsmanagement ermöglicht, welches gerade in der Therapie des Typ 2 Diabetes eine wichtige Rolle spielt. Diese Eigenschaften führen in der Konsequenz dazu, dass auch Parameter zur Lebensqualität bessere Werte aufweisen als bei einer Therapie mit NPH-Insulin.

Obwohl das Verfahren der Nutzenbewertung insgesamt aus rechtlicher Sicht sehr kritisch zu bewerten ist, weil die gesetzlichen Änderungen seit dem GKV Wettbewerbsstärkungsgesetz keine Berücksichtigung gefunden haben, zeigen die vorliegenden Daten den Zusatznutzen von Insulindetemir. Dies wird durch die Fachinformation zu Insulindetemir bestätigt. Wenn die Kriterien der EMEA verwendet werden, die zur Bewertung im Sinne der Fachinformation geführt haben, muss das IQWiG die Überlegenheit der langwirksamen Insulinanaloga anerkennen.

---

[1] Stellungnahme der Novo Nordisk Pharma GmbH; 26.07.2007 zum Berichtsplan „Auftrag A05-03; Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2" Version 1.0 vom 27.06.2007

[2] Fachinformation Levemir®, Juni 2007

[3] Bundessozialgericht, Urteil vom 31.05.2006, B 6 KA 13/05 R

[4] Pressemitteilung – EPA – Brüssel/München (pts/10.04.2008/15:35) - "European Inventor of the Year 2008" URL: http://ddp-direkt.de/portal/details.php?id=35418 ; www.epo.org/topics/news/2008/20080410.html (Internet-Zugriff: 23.04.2008)

[5] Dornhorst A, Lüddeke HJ, Sreenan S on behalf of the PREDICTIVE Study Group. Insulin detemir improves glycaemic control without weight gain in insulin-naïve patients with type 2 diabetes: subgroup analysis from the PREDICTIVE™ study. Int J Clin Pract 2008; 62;4:659-665

[6] Dornhorst A, Lüddeke HJ, Sreenan S et al. on behalf of the PREDICTIVE Study Group. Safety and efficacy of insulin detemir in clinical practice: 14-week follow-up data from type 1 and type 2 diabetes patients in the PREDICTIVE™ European cohort. Int J Clin Pract 2007;61:523-528

[7] Dornhorst A, Lüddeke HJ, Koenen C et al. on behalf of the PREDICTIVE Study Group. Transferring to insulin detemir from NPH insulin or insulin glargine in type 2 diabetes patients on basal-only therapy with oral andidiabetic drugs improves glycaemic control and reduces weight gain and risk of hypoglycaemia: 14-week follow-up data from PREDICTIVE™. Diab Obes Metab 2008;10:75-81

[8] Meneghini LF, Rosenberg KH, Koenen C et al. Insulin detemir improves glycaemic control with less hypoglycaemia and no weight gain in patients with type 2 diabetes who were insulin naïve or treated with NPH or insulin glargine: clinical practice experience from a German subgroup of the PREDICTIVE study. Diab Obes Metab 2007;9:418-427

[9] Philis-Tsimikas A, Charpentier G, Clauson P et al. Comparison of once-daily insulin detemir with NPH insulin added to a regimen of oral antidiabetic drugs in poorly controlled type 2 diabetes. Clin Ther 2006;28:1569-1581

[10] Rašlová K, Bogoev M, Raz I et al. Insulin detemir and insulin aspart: a promising basal-bolus regimen for type 2 diabetes. Diab Res Clin Pract 2004;66:193-201

[11] Rašlová K, Bogoev M, Raz I et al. Corrigendum to "Insulin detemir and insulin aspart: A promising basal-bolus regimen for type 2 diabetes" [Diabetes Research and Clinical Practice 66 (2004) 193-201]. Diab Res Clin Pract 2006;72:112

[12] Food and Drug Administration. Guidance for industry: Patient-reported outcome measures: use in medical product development to support labeling claims. Draft guidance. February 2006. URL: www.fda.gov/CDER/GUIDANCE/5460dft.pdf (Internet-Zugriff: 23.04.2008)

[13] Bradley C. Feedback on the FDA's February 2006 draft guidance on Patient Reported Outcome (PRO) measures from a developer of PRO measures. Health Qual Life Outcomes 2006;4:78

[14] American Diabetes Association Workgroup on Hypoglycemia. Defining and reporting hypoglycemia in diabetes: A report from the American Diabetes Association Workgroup on Hypoglycemia. Diab Care 2005;28:1245-1249

[15] UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 1998;352:837-53

[16] European Medicines Agency. Note for guidance on clinical investigation of medicinal products in the treatment of diabetes mellitus [Online]. 30.05.2002; URL: www.emea.europa.eu/pdfs/human/ewp/108000en.pdf (Internet-Zugriff: 23.04.2008)

[17] Garber AJ, Clauson P, Pedersen CB et al. Lower risk of hypoglycemia with insulin detemir than with neutral protamine Hagedorn insulin in older persons with type 2 diabetes: A pooled analysis of phase III trials. J Am Geriatr Soc 2007;55:1735-1740

[18] Robert Koch-Institut (Hrsg.). Gesundheitsberichterstattung des Bundes, Heft 16. Übergewicht und Adipositas. Gesundheitsberichterstattung des Bundes. August 2003, geänderte Auflage 2005. ISBN 3-89606-144-5, ISSN 1437-5478

[19] Peeters A, Barendregt JJ, Willekens F et al. for NEDCOM, the Netherlands Epidemiology and Demography Compression of Morbidity Research Group. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Ann Intern Med 2003;138:24-32

[20] Adams KF, Schatzkin A, Harris T et al. Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. N Engl J Med 2006;355:763-778

[21] Evidenzbasierte Leitlinie "Prävention und Therapie der Adipositas"; Stand: 25.05.2007; Hrsg.: Deutsche Adipositas-Gesellschaft, Deutsche Diabetes-Gesellschaft, Deutsche Gesellschaft für Ernährung, Deutsche Gesellschaft für Ernährungsmedizin
URL: www.dge.de/pdf/ll/Adipositas-Leitlinie-2007.pdf (Internet-Zugriff: 23.04.2008)

[22] Anderson JW, Kendall CWC, Jenkins DJA. Importance of Weight Management in Type 2 Diabetes: Review with Meta-analysis of Clinical Studies. J Am Coll Nutr 2003; 22:331-339

[23] Daousi C, Casson IF, Gill GV et al. Prevalence of obesity in type 2 diabetes in secondary care: association with cardiovascular risk factors. Postgrad Med J 2006;82:280-284

[24] Matthaei S, Häring HU. Behandlung des Diabetes mellitus Typ 2. DDG Praxisleitlinie. Diabetologie 2008;3 (Suppl 2):S157-S161

[25] Larger E. Weight gain and insulin treatment. Diab Metab 2005;31:4S51-4S56

[26] Larger E, Rufat P, Dubois-Laforgue D et al. Insulin therapy does not itself induce weight gain in patients with type 2 diabetes. Diab Care 2001;24:1849-1850

[27] Meece J. Dispelling myths and removing barriers about insulin in type 2 diabetes. Diab Educator 2006;32:9S-18S

[28] Tschritter O, Hennige AM, Preissl H et al. Cerebrocortical beta activity in overweight humans responds to insulin detemir. PloS ONE 2007;2:e1196

[29] Russell-Jones D, Khan R. Insulin-associated weight gain in diabetes - causes, effects and coping strategies. Diab Obes Metab 2007;9:799-812

[30] Korytkowski M. When oral agents fail: practical barriers to starting insulin. Int J Obesity 2002;26 (Suppl 3):S18-S24

[31] Hermansen K, Davies M, Derezinski T et al. A 26-week, randomized, parallel, treat-to-target trial comparing insulin detemir with NPH insulin as add-on therapy to oral glucose-lowering drugs in insulin-naïve people with type 2 diabetes. Diab Care 2006;29:1269-1274

[32] Răslová K, Tamer SC, Clauson P et al. Insulin detemir results in less weight gain than NPH insulin when used in basal-bolus therapy for type 2 diabetes mellitus, and this advantage increases with baseline body mass index. Clin Drug Invest 2007;27:279-285

[33] Bartley PC, Bogoev M, Larsen J et al. Long-term efficacy and safety of insulin detemir compared to Neutral Protamine Hagedorn insulin in patients with Type 1 diabetes using a treat-to-target basal-bolus regimen with insulin aspart at meals: a 2-year, randomized, controlled trial. Diab Med 2008;25:442-449

[34] Ott P, Benke I, Köhler C et al. Praxis der Insulintherapie bei Typ 2 Diabetikern im Rahmen der DIG- (Diabetes in Germany) Studie. Diabetol Stoffw 2006;1:366-373

[35] IMS Disease Analyzer; Interne Daten-Analyse; Complete Database Dataset September 2007; autorisiert durch IMS Health Fankfurt im April 2008

Langwirksame Insulinanaloga - Diabetes mellitus 2

- Auswertung basierend auf Daten von IMS Health, Disease Analyzer 2006
- Autorisiert durch IMS Health, Frankfurt Office, April 2008

"COMPLETE" DATABASE (mainly general practitioners)(2006/ Q3 2007 data)

- Daten zur Verordnung von langwirksamen Insulinanaloga in 2006 Neueinstellungen - OAD-Therapie-Versager
- Auswertung basierend auf IMS Health Disease Analyzer Daten 2006 und 2007

"COMPLETE" DATABASE (mainly general practitioners)(2006/ Q3 2007 data)

Methodology

Patient groups:

- Type 2 patients: Patients with any diabetes diagnosis (E10-E14) or treated with anti-diabetic (A10), who are not Type 1.
- Type 1: Diagnosis of Type 1 diabetes and no other diagnosis of diabetes 10 years prior to Type 1 diagnosis, no OAD treatment 10 years prior to or after diagnosis.
- Insulin naïve patients: No previous insulin usage (24mth), but active record in database (6 to 24mth).

In the present analyses requirement is made of OAD usage the year prior to initiation.

Insulin types:

- Basal insulin (A10C2, A10C4, A10C5)
- Bolus insulin (A10C1)
- Premix insulin (A10C3)

Insulin treatment regimens:

- Basal & Bolus: Both Basal and Bolus insulin is prescribed within the studied period*. No overlap in treatment is required.
- Basal-only: Basal insulin and no other types of insulin is prescribed within the studied period*.
- Premix-only: Premix insulin and no other types of insulin is prescribed within the studied period*.
- Bolus-only: Bolus insulin and no other types of insulin is prescribed within the studied period*.
- Other: Premix and other type insulin is prescribed within the studied period*.

*Studied period: year 2006 and 2007 in the present analyses.

Patient numbers in current dataset (Dataset September 2007):

Total database:	451782
A10 patients:	206237
E10-E14 patients:	436914
E10-E14 and/or A10 p	451782
Type 2 patients:	423364

Report of brands designed as below:

Matrix Report

"COMPLETE" DATABASE (mainly general practitioners) (2006/ Q3 2007 data)

	Type 2 patients initiated on insulin after OAD failure		Type 2 patients initiated on insulin after OAD failure	
Initiation regimen	Initiated in 2006	Initiated in 2007 (until Q3)	Initiated in 2006	Initiated in 2007 (until Q3)
Basal only	955	812	40%	45%
Premix only	563	371	23%	20%
Bolus only	309	228	13%	13%
Basal & Bolus	406	304	17%	17%
Other	167	105	7%	6%
Total patients	**2400**	**1820**	100%	100%
Practices	**701**	**647**		

	Analogue penetration in Type 2 patients initiated on insulin after OAD	
Initiation regimen	Initiated in 2006	Initiated in 2007 (until Q3)
Basal only	61%	67%
Premix only	9%	7%
Bolus only	19%	15%
Basal & Bolus		
Basal in Basal & Bolus	37%	42%
Bolus in Basal & Bolus	21%	22%
Other		

A 1.6.1 Analysen auf Basis von individuellen Patientendaten (IPD)

Nachfolgend werden die von der Firma Novo Nordisk Pharma GmbH im Rahmen des Stellungnahmeverfahrens zur Verfügung gestellten IPD-Meta-Analysen zur gemeinsamen Betrachtung von Hypoglykämien und der langfristigen Blutzuckersenkung dargestellt.

Statistical Documentation

List of Contents

Page

1: Description of analyses 2
2: Overview of endpoints by study(-pools) 3
3: Overview of endpoints by eot HbA1c categories and by study(-pools) 8
4: HbA1c summary table : End of Trial change from baseline 18
5: HbA1c summary table : End of Trial 20
6: Summary of odds ratio estimates from logistic regression controlling for change in HbA1c 22
7: Summary of odds ratio estimates from logistic regression controlling for End of Trial HbA1c 25

1: Description of analyses

Logistic regression adjusting for End of Trial (EOT) HbA1c or change from basline to EOT HbA1c was used to obtain the odds ratios and 95% confidence intervals. Last observation carried forward was used to obtain the HbA1c EOT values. (Note that EOT varies for the different studies, e.g. some studies are 26 weeks durations and others 52 weeks duration.) Hypoglicaemic episodes with a missing or partial date were treated as treatment emergent. Studies have been pooled where appropriate and the analyses were run by study/study pool and treatment group. Analyses are grouped for each endpoint:

A01: Minor or symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
A02: Minor or symptom, daytime (not nocturnal), BG<3.9 mmol/L (70 mg/dl)
A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
A05: Major, daytime (not nocturnal)
A06: Major, nocturnal
A11: Minor or symptom, BG<2.0 mmol/L (36 mg/dl)
A12: Minor or symptom, BG<3.9 mmol/L (70 mg/dl)
A13: Minor
A15: Major

2: Overview of endpoints by study(-pools)

Overview of endpoints by study(-pools)

	A01				A02			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	21	6.2	11	6.7	152	44.6	81	49.4
II (1337, BOT)	6	1.9	7	4.4	37	12.0	33	20.9
III (1530, BOT)	4	1.7	13	5.5	123	51.9	170	71.4
IV (1659, BB)	5	4.0	14	9.6	40	32.0	70	47.9
V (1373, BOT)	13	4.5	11	3.8	141	48.5	163	56.0
VI (1431,2175, B	15	3.2	8	3.4	296	63.2	156	66.1
BB (1336,1659)	26	5.6	25	8.1	192	41.2	151	48.7
BOT (1337,1530)	10	1.8	20	5.1	160	29.3	203	51.3

Overview of endpoints by study(-pools)

	A03				A04			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	5	1.5	9	5.5	55	16.1	44	26.8
II (1337, BOT)	.	.	3	1.9	21	6.8	25	15.8
III (1530, BOT)	2	0.8	6	2.5	62	26.2	105	44.1
IV (1659, BB)	4	3.2	10	6.8	20	16.0	44	30.1
V (1373, BOT)	4	1.4	5	1.7	87	29.9	86	29.6
VI (1431,2175, B	16	3.4	9	3.8	196	41.9	102	43.2
BB (1336,1659)	9	1.9	19	6.1	75	16.1	88	28.4
BOT (1337,1530)	2	0.4	9	2.3	83	15.2	130	32.8

Overview of endpoints by study(-pools)

	A05				A06			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	5	1.5	3	1.8	2	0.6	1	0.6
II (1337, BOT)	.	.	1	0.6	.	.	1	0.6
III (1530, BOT)	1	0.4	6	2.5	.	.	1	0.4
IV (1659, BB)	.	.	2	1.4	.	.	1	0.7
V (1373, BOT)	3	1.0	4	1.4	3	1.0	4	1.4
VI (1431,2175, B	13	2.8	7	3.0	9	1.9	5	2.1
BB (1336,1659)	5	1.1	5	1.6	2	0.4	2	0.6
BOT (1337,1530)	1	0.2	7	1.8	.	.	2	0.5

Overview of endpoints by study(-pools)

	A11				A12			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	24	7.0	18	11.0	162	47.5	88	53.7
II (1337, BOT)	6	1.9	9	5.7	48	15.5	47	29.7
III (1530, BOT)	5	2.1	15	6.3	135	57.0	186	78.2
IV (1659, BB)	7	5.6	19	13.0	45	36.0	76	52.1
V (1373, BOT)	15	5.2	15	5.2	164	56.4	182	62.5
VI (1431,2175, B	26	5.6	15	6.4	326	69.7	169	71.6
BB (1336,1659)	31	6.7	37	11.9	207	44.4	164	52.9
BOT (1337,1530)	11	2.0	24	6.1	183	33.5	233	58.8

Overview of endpoints by study(-pools)

	A13				A15			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	113	33.1	69	42.1	6	1.8	3	1.8
II (1337, BOT)	24	7.8	24	15.2	.	.	1	0.6
III (1530, BOT)	96	40.5	153	64.3	1	0.4	6	2.5
IV (1659, BB)	39	31.2	63	43.2	.	.	2	1.4
V (1373, BOT)	135	46.4	151	51.9	5	1.7	8	2.7
VI (1431,2175, B	249	53.2	140	59.3	20	4.3	11	4.7
BB (1336,1659)	152	32.6	132	42.6	6	1.3	5	1.6
BOT (1337,1530)	120	22.0	177	44.7	1	0.2	7	1.8

3: Overview of endpoints by eot HbA1c categories and by study(-pools)

Overview of endpoints by EOT HbA1c and by study(-pools)

		A01				A02			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	2	6.5	2	16.7	7	22.6	5	41.7
	6-<7	9	12.2	2	4.5	39	52.7	24	54.5
	7-<8	7	5.2	5	6.9	66	49.3	41	56.9
	9-<10	.	.	.	.	5	17.9	2	22.2
	10-<11	.	.	.	.	1	25.0	.	.
	>=11	.	.	.	.	1	50.0	1	100.0
II (1337, BOT)	<6	.	.	.	.	3	9.1	1	5.0
	6-<7	.	.	.	.	1	6.7	5	33.3
	7-<8	2	2.9	2	4.7	9	13.0	9	20.9
	9-<10	.	.	2	9.1	10	15.4	4	18.2
	10-<11	3	11.5	.	.	6	23.1	.	.
III (1530, BOT)	<6	.	.	2	6.7	9	39.1	15	50.0
	6-<7	2	2.0	5	3.9	50	50.0	103	81.1
	7-<8	1	1.1	5	7.1	59	62.8	45	64.3
	9-<10	.	.	.	.	1	25.0	.	.
IV (1659, BB)	<6	.	.	1	16.7	1	20.0	3	50.0
	6-<7	.	.	1	3.8	2	14.3	12	46.2
	7-<8	1	2.0	5	8.9	17	33.3	26	46.4
	9-<10	1	6.3	1	6.3	5	31.3	6	37.5
	10-<11	1	20.0	1	33.3	2	40.0	2	66.7
V (1373, BOT)	<6	1	2.2	2	6.9	15	33.3	12	41.4
	6-<7	4	3.7	1	0.8	60	55.6	68	57.6
	7-<8	6	7.2	5	5.2	42	50.6	55	56.7
	9-<10	1	7.1	.	.	8	57.1	4	66.7
	10-<11	.	.	.	.	2	66.7	.	.
VI (1431,2175, B	<6	1	3.1	.	.	18	56.3	8	40.0
	6-<7	4	2.9	3	3.7	96	69.6	61	75.3

Overview of endpoints by EOT HbA1c and by study(-pools)

		A01				A02			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
VI (1431,2175, B	7-<8	5	3.1	2	2.6	109	67.3	49	63.6
	9-<10	1	3.2	.	.	18	58.1	11	68.8
	10-<11	.	.	1	25.0	3	30.0	2	50.0
	>=11	1	16.7	.	.	1	16.7	.	.
BB (1336,1659)	<6	2	5.6	3	16.7	8	22.2	8	44.4
	6-<7	9	10.2	3	4.3	41	46.6	36	51.4
	7-<8	8	4.3	10	7.8	83	44.9	67	52.3
	9-<10	1	2.3	1	4.0	10	22.7	8	32.0
	10-<11	1	11.1	1	25.0	3	33.3	2	50.0
	>=11	.	.	.	.	1	50.0	1	100.0
BOT (1337,1530)	<6	.	.	2	4.0	12	21.4	16	32.0
	6-<7	2	1.7	5	3.5	51	44.3	108	76.1
	7-<8	3	1.8	7	6.2	68	41.7	54	47.8
	9-<10	.	.	2	9.1	11	15.9	4	18.2
	10-<11	3	11.5	.	.	6	23.1	.	.

Overview of endpoints by EOT HbA1c and by study(-pools)

Study(-pool)	EOT HbA1c	A03 Detemir Freq	A03 Detemir Perc	A03 Comparator Freq	A03 Comparator Perc	A04 Detemir Freq	A04 Detemir Perc	A04 Comparator Freq	A04 Comparator Perc
I (1336, BB)	<6	1	3.2	2	16.7	2	6.5	4	33.3
	6-<7	.	.	4	9.1	10	13.5	8	18.2
	7-<8	1	0.7	1	1.4	24	17.9	25	34.7
	9-<10	2	7.1	.	.	5	17.9	.	.
	10-<11	.	.	.	.	1	25.0	.	.
	>=11	.	.	.	.	1	50.0	.	.
II (1337, BOT)	<6	.	.	.	.	2	6.1	.	.
	6-<7	.	.	.	.	2	13.3	1	6.7
	7-<8	.	.	.	.	5	7.2	6	14.0
	9-<10	.	.	2	9.1	5	7.7	5	22.7
	10-<11	.	.	.	.	2	7.7	1	16.7
III (1530, BOT)	<6	.	.	2	6.7	5	21.7	11	36.7
	6-<7	.	.	.	.	22	22.0	54	42.5
	7-<8	1	1.1	4	5.7	31	33.0	34	48.6
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	.	.	.	.	1	20.0	2	33.3
	6-<7	.	.	1	3.8	1	7.1	9	34.6
	7-<8	2	3.9	5	8.9	8	15.7	16	28.6
	9-<10	1	6.3	1	6.3	2	12.5	4	25.0
	10-<11	.	.	.	.	.	.	1	33.3
V (1373, BOT)	<6	3	6.7	1	3.4	16	35.6	8	27.6
	6-<7	.	.	2	1.7	28	25.9	34	28.8
	7-<8	1	1.2	1	1.0	27	32.5	34	35.1
	9-<10	.	.	.	.	3	21.4	1	16.7
	10-<11	.	.	.	.	2	66.7	.	.
VI (1431,2175, B	<6	.	.	2	10.0	11	34.4	6	30.0
	6-<7	4	2.9	3	3.7	68	49.3	37	45.7

Overview of endpoints by EOT HbA1c and by study(-pools)

		A03				A04			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
VI (1431,2175, B	7-<8	7	4.3	2	2.6	73	45.1	36	46.8
	9-<10	1	3.2	.	.	12	38.7	4	25.0
	10-<11	.	.	1	25.0	.	.	1	25.0
	>=11	.	.	.	.	1	16.7	.	.
BB (1336,1659)	<6	1	2.8	2	11.1	3	8.3	6	33.3
	6-<7	.	.	5	7.1	11	12.5	17	24.3
	7-<8	3	1.6	6	4.7	32	17.3	41	32.0
	9-<10	3	6.8	1	4.0	7	15.9	4	16.0
	10-<11	.	.	.	.	1	11.1	1	25.0
	>=11	.	.	.	.	1	50.0	.	.
BOT (1337,1530)	<6	.	.	2	4.0	7	12.5	11	22.0
	6-<7	.	.	.	.	24	20.9	55	38.7
	7-<8	1	0.6	4	3.5	36	22.1	40	35.4
	9-<10	.	.	2	9.1	5	7.2	5	22.7
	10-<11	.	.	.	.	2	7.7	1	14.3

Overview of endpoints by EOT HbA1c and by study(-pools)

Study(-pool)	EOT HbA1c	A05 Detemir Freq	A05 Detemir Perc	A05 Comparator Freq	A05 Comparator Perc	A06 Detemir Freq	A06 Detemir Perc	A06 Comparator Freq	A06 Comparator Perc
I (1336, BB)	<6	.	.	.	.	.	.	.	.
	6-<7	2	2.7	.	.	.	.	.	.
	7-<8	1	0.7	2	2.8	1	0.7	1	1.4
	9-<10	.	.	1	11.1	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
II (1337, BOT)	<6	.	.	.	.	.	.	.	.
	6-<7	.	.	.	.	.	.	.	.
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
III (1530, BOT)	<6	1	4.3	1	3.3	.	.	.	.
	6-<7	.	.	5	3.9	.	.	1	0.8
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	.	.	.	.	.	.	.	.
	6-<7	.	.	.	.	.	.	.	.
	7-<8	.	.	1	1.8	.	.	1	1.8
	9-<10	.	.	1	6.3	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
V (1373, BOT)	<6	.	.	1	3.4	.	.	1	3.4
	6-<7	2	1.9	.	.	1	0.9	2	1.7
	7-<8	1	1.2	.	.	1	1.2	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
VI (1431,2175, B	<6	.	.	.	.	.	.	.	.
	6-<7	4	2.9	5	6.2	3	2.2	3	3.7

Overview of endpoints by EOT HbA1c and by study(-pools)

		A05				A06			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
VI (1431,2175, B	7-<8	5	3.1	1	1.3	1	0.6	1	1.3
	9-<10	4	12.9	.	.	1	3.2	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
BB (1336,1659)	<6	.	.	.	.	.	.	.	.
	6-<7	2	2.3	.	.	.	.	.	.
	7-<8	1	0.5	3	2.3	1	0.5	2	1.6
	9-<10	.	.	2	8.0	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
BOT (1337,1530)	<6	1	1.8	1	2.0	.	.	.	.
	6-<7	.	.	5	3.5	.	.	1	0.7
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.

Overview of endpoints by EOT HbA1c and by study(-pools)

		A11				A12			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	2	6.5	2	16.7	7	22.6	6	50.0
	6-<7	9	12.2	6	13.6	41	55.4	24	54.5
	7-<8	7	5.2	6	8.3	70	52.2	45	62.5
	9-<10	2	7.1	.	.	7	25.0	2	22.2
	10-<11	.	.	.	.	1	25.0	.	.
	>=11	.	.	.	.	1	50.0	1	100.0
II (1337, BOT)	<6	.	.	.	.	4	12.1	1	5.0
	6-<7	.	.	.	.	3	20.0	6	40.0
	7-<8	2	2.9	2	4.7	12	17.4	14	32.6
	9-<10	.	.	3	13.6	12	18.5	7	31.8
	10-<11	3	11.5	.	.	6	23.1	1	16.7
III (1530, BOT)	<6	.	.	2	6.7	9	39.1	16	53.3
	6-<7	2	2.0	5	3.9	55	55.0	111	87.4
	7-<8	1	1.1	7	10.0	65	69.1	51	72.9
	9-<10	.	.	.	.	1	25.0	.	.
IV (1659, BB)	<6	.	.	1	16.7	1	20.0	3	50.0
	6-<7	.	.	1	3.8	3	21.4	16	61.5
	7-<8	2	3.9	8	14.3	19	37.3	27	48.2
	9-<10	2	12.5	2	12.5	5	31.3	7	43.8
	10-<11	1	20.0	1	33.3	2	40.0	2	66.7
V (1373, BOT)	<6	3	6.7	2	6.9	22	48.9	14	48.3
	6-<7	4	3.7	3	2.5	64	59.3	74	62.7
	7-<8	6	7.2	6	6.2	47	56.6	65	67.0
	9-<10	1	7.1	.	.	9	64.3	4	66.7
	10-<11	.	.	.	.	2	66.7	.	.
VI (1431,2175, B	<6	1	3.1	2	10.0	21	65.6	11	55.0
	6-<7	7	5.1	5	6.2	107	77.5	62	76.5

Overview of endpoints by EOT HbA1c and by study(-pools)

		A11				A12			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
VI (1431,2175, B	7-<8	9	5.6	4	5.2	121	74.7	54	70.1
	9-<10	2	6.5	.	.	19	61.3	11	68.8
	10-<11	.	.	1	25.0	3	30.0	2	50.0
	>=11	1	16.7	.	.	1	16.7	.	.
BB (1336,1659)	<6	2	5.6	3	16.7	8	22.2	9	50.0
	6-<7	9	10.2	7	10.0	44	50.0	40	57.1
	7-<8	9	4.9	14	10.9	89	48.1	72	56.3
	9-<10	4	9.1	2	8.0	12	27.3	9	36.0
	10-<11	1	11.1	1	25.0	3	33.3	2	50.0
	>=11	.	.	.	.	1	50.0	1	100.0
BOT (1337,1530)	<6	.	.	2	4.0	13	23.2	17	34.0
	6-<7	2	1.7	5	3.5	58	50.4	117	82.4
	7-<8	3	1.8	9	8.0	77	47.2	65	57.5
	9-<10	.	.	3	13.6	13	18.8	7	31.8
	10-<11	3	11.5	.	.	6	23.1	1	14.3

Overview of endpoints by EOT HbA1c and by study(-pools)

Study(-pool)	EOT HbA1c	A13 Detemir Freq	A13 Detemir Perc	A13 Comparator Freq	A13 Comparator Perc	A15 Detemir Freq	A15 Detemir Perc	A15 Comparator Freq	A15 Comparator Perc
I (1336, BB)	<6	5	16.1	4	33.3	.	.	.	.
	6-<7	26	35.1	17	38.6	2	2.7	.	.
	7-<8	50	37.3	36	50.0	2	1.5	2	2.8
	9-<10	6	21.4	2	22.2	.	.	1	11.1
	10-<11	1	25.0	.	.	.	.	.	.
	>=11	1	50.0	.	.	.	.	.	.
II (1337, BOT)	<6	2	6.1	.	.	.	.	.	.
	6-<7	.	.	4	26.7	.	.	.	.
	7-<8	8	11.6	4	9.3	.	.	.	.
	9-<10	2	3.1	5	22.7	.	.	.	.
	10-<11	5	19.2	1	16.7	.	.	.	.
III (1530, BOT)	<6	9	39.1	13	43.3	1	4.3	1	3.3
	6-<7	40	40.0	90	70.9	.	.	5	3.9
	7-<8	43	45.7	43	61.4	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	1	20.0	2	33.3	.	.	.	.
	6-<7	3	21.4	12	46.2	.	.	.	.
	7-<8	15	29.4	21	37.5	.	.	1	1.8
	9-<10	4	25.0	6	37.5	.	.	1	6.3
	10-<11	2	40.0	2	66.7	.	.	.	.
V (1373, BOT)	<6	18	40.0	11	37.9	.	.	2	6.9
	6-<7	49	45.4	63	53.4	2	1.9	2	1.7
	7-<8	45	54.2	52	53.6	2	2.4	.	.
	9-<10	7	50.0	3	50.0	.	.	.	.
	10-<11	2	66.7	.	.	.	.	.	.
VI (1431,2175, B	<6	16	50.0	8	40.0	.	.	.	.
	6-<7	87	63.0	54	66.7	6	4.3	7	8.6

Overview of endpoints by EOT HbA1c and by study(-pools)

Study(-pool)	EOT HbA1c	A13 Detemir Freq	A13 Detemir Perc	A13 Comparator Freq	A13 Comparator Perc	A15 Detemir Freq	A15 Detemir Perc	A15 Comparator Freq	A15 Comparator Perc
VI (1431,2175, B	7-<8	86	53.1	42	54.5	5	3.1	2	2.6
	9-<10	15	48.4	10	62.5	5	16.1	.	.
	10-<11	3	30.0	2	50.0	.	.	.	.
	>=11	1	16.7	.	.	.	.	.	.
BB (1336,1659)	<6	6	16.7	6	33.3	.	.	.	.
	6-<7	29	33.0	29	41.4	2	2.3	.	.
	7-<8	65	35.1	57	44.5	2	1.1	3	2.3
	9-<10	10	22.7	8	32.0	.	.	2	8.0
	10-<11	3	33.3	2	50.0	.	.	.	.
	>=11	1	50.0	.	.	.	.	.	.
BOT (1337,1530)	<6	11	19.6	13	26.0	1	1.8	1	2.0
	6-<7	40	34.8	94	66.2	.	.	5	3.5
	7-<8	51	31.3	47	41.6	.	.	.	.
	9-<10	2	2.9	5	22.7	.	.	.	.
	10-<11	5	19.2	1	14.3	.	.	.	.

4: HbA1c summary table : End of Trial change from baseline

HbA1c summary table: End of Trial change from baseline

StudyPool	Statistic	Detemir	Comparator
I (1336, BB)	N	320	157
	No. missing	21	7
	Mean	-0.28	-0.39
	Standard deviation	1.04	1.02
	Minimum	-4.40	-3.50
	Maximum	2.50	2.00
II (1337, BOT)	N	277	139
	No. missing	32	19
	Mean	-0.78	-1.23
	Standard deviation	1.23	1.16
	Minimum	-4.70	-4.70
	Maximum	2.30	1.90
III (1530, BOT)	N	230	231
	No. missing	7	7
	Mean	-1.63	-1.74
	Standard deviation	0.92	0.91
	Minimum	-4.10	-4.00
	Maximum	1.70	3.10
IV (1659, BB)	N	123	142
	No. missing	2	4
	Mean	-0.96	-0.99
	Standard deviation	1.13	1.01
	Minimum	-7.20	-3.60
	Maximum	1.90	2.80
V (1373, BOT)	N	268	275
	No. missing	23	16
	Mean	-1.42	-1.47
	Standard deviation	1.14	1.03
	Minimum	-4.70	-3.90
	Maximum	2.40	2.20

HbA1c summary table: End of Trial change from baseline

StudyPool	Statistic	Detemir	Comparator
VI (1431,2175, B	N	456	228
	No. missing	12	8
	Mean	-1.00	-1.29
	Standard deviation	1.15	1.10
	Minimum	-4.70	-4.50
	Maximum	4.70	2.10
BB (1336,1659)	N	443	299
	No. missing	23	11
	Mean	-0.47	-0.67
	Standard deviation	1.11	1.05
	Minimum	-7.20	-3.60
	Maximum	2.50	2.80
BOT (1337,1530)	N	507	370
	No. missing	39	26
	Mean	-1.17	-1.55
	Standard deviation	1.18	1.04
	Minimum	-4.70	-4.70
	Maximum	2.30	3.10

5: HbA1c summary table : End of Trial

HbA1c summary table: HbA1c at End of Trial

StudyPool	Statistic	Detemir	Comparator
I (1336, BB)	N	320	157
	No. missing	21	7
	Mean	7.60	7.38
	Standard deviation	1.07	0.93
	Minimum	5.00	5.10
	Maximum	11.60	11.10
II (1337, BOT)	N	277	139
	No. missing	32	19
	Mean	8.66	8.13
	Standard deviation	1.23	1.05
	Minimum	5.90	5.60
	Maximum	12.30	11.60
III (1530, BOT)	N	230	231
	No. missing	7	7
	Mean	6.96	6.78
	Standard deviation	0.72	0.70
	Minimum	5.20	5.20
	Maximum	9.30	10.90
IV (1659, BB)	N	123	142
	No. missing	2	4
	Mean	7.91	7.83
	Standard deviation	1.06	0.95
	Minimum	3.30	5.80
	Maximum	10.50	10.30
V (1373, BOT)	N	268	275
	No. missing	23	16
	Mean	7.19	7.14
	Standard deviation	1.04	0.91
	Minimum	5.00	5.20
	Maximum	11.10	11.30

HbA1c summary table: HbA1c at End of Trial

StudyPool	Statistic	Detemir	Comparator
VI (1431,2175, B	N	456	228
	No. missing	12	8
	Mean	7.52	7.32
	Standard deviation	1.12	1.03
	Minimum	5.30	5.50
	Maximum	12.00	10.90
BB (1336,1659)	N	443	299
	No. missing	23	11
	Mean	7.68	7.59
	Standard deviation	1.07	0.97
	Minimum	3.30	5.10
	Maximum	11.60	11.10
BOT (1337,1530)	N	507	370
	No. missing	39	26
	Mean	7.89	7.28
	Standard deviation	1.34	1.07
	Minimum	5.20	5.20
	Maximum	12.30	11.60

6: Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A01			A02			A03		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.98	0.45	2.16	0.89	0.61	1.31	0.22	0.06	0.74
II (1337, BOT)	0.41	0.13	1.27	0.47	0.28	0.81	0.00	0.00	6E134
III (1530, BOT)	0.31	0.10	0.99	0.42	0.28	0.61	0.39	0.07	2.03
IV (1659, BB)	0.39	0.13	1.16	0.51	0.31	0.84	0.44	0.13	1.44
V (1373, BOT)	1.35	0.58	3.12	0.81	0.58	1.14	0.81	0.22	3.08
VI (1431,2175, B	0.77	0.31	1.88	0.91	0.65	1.28	0.89	0.39	2.07
BB (1336,1659)	0.68	0.38	1.24	0.77	0.57	1.03	0.27	0.12	0.64
BOT (1337,1530)	0.35	0.16	0.77	0.40	0.30	0.53	0.17	0.03	0.81

Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A04			A05			A06		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.57	0.36	0.91	0.83	0.19	3.51	1.02	0.09	11.38
II (1337, BOT)	0.32	0.17	0.61	0.00	0.00	1E150	0.00	0.00	1E150
III (1530, BOT)	0.45	0.31	0.67	0.20	0.02	1.72	0.00	0.00	3E157
IV (1659, BB)	0.45	0.25	0.81	0.00	0.00	3E126	0.00	0.00	1E147
V (1373, BOT)	1.05	0.73	1.51	1.01	0.20	5.04	1.03	0.21	5.15
VI (1431,2175, B	0.98	0.71	1.36	0.94	0.37	2.41	0.85	0.28	2.61
BB (1336,1659)	0.52	0.36	0.74	0.67	0.19	2.34	0.68	0.09	4.87
BOT (1337,1530)	0.38	0.27	0.52	0.13	0.02	1.08	0.00	0.00	8E186

Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A11			A12			A13			A15		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.63	0.32	1.21	0.86	0.59	1.27	0.71	0.48	1.05	0.99	0.24	4.02
II (1337, BOT)	0.29	0.10	0.87	0.38	0.23	0.61	0.40	0.22	0.76	0.00	0.00	1E150
III (1530, BOT)	0.34	0.12	0.95	0.35	0.23	0.53	0.37	0.26	0.54	0.20	0.02	1.72
IV (1659, BB)	0.40	0.16	0.99	0.51	0.31	0.84	0.60	0.36	0.99	0.00	0.00	3E126
V (1373, BOT)	1.10	0.52	2.33	0.85	0.60	1.20	0.88	0.62	1.23	0.85	0.25	2.81
VI (1431,2175, B	0.80	0.41	1.55	0.97	0.68	1.40	0.80	0.58	1.11	0.90	0.42	1.92
BB (1336,1659)	0.52	0.31	0.86	0.75	0.56	1.01	0.66	0.49	0.90	0.80	0.24	2.66
BOT (1337,1530)	0.31	0.15	0.66	0.35	0.26	0.46	0.36	0.27	0.48	0.13	0.02	1.08

7: Summary of odds ratio estimates from logistic regression controlling for End of Trial HbA1c

Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A01			A02			A03		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	1.07	0.48	2.35	0.93	0.63	1.36	0.23	0.07	0.78
II (1337, BOT)	0.36	0.11	1.14	0.47	0.27	0.81	0.00	0.00	2E177
III (1530, BOT)	0.30	0.09	0.95	0.43	0.29	0.64	0.36	0.07	1.90
IV (1659, BB)	0.39	0.13	1.15	0.51	0.31	0.84	0.45	0.14	1.46
V (1373, BOT)	1.34	0.58	3.11	0.81	0.58	1.14	0.77	0.20	2.95
VI (1431,2175, B	0.90	0.37	2.16	0.93	0.66	1.31	0.89	0.39	2.06
BB (1336,1659)	0.73	0.40	1.30	0.78	0.58	1.05	0.28	0.12	0.66
BOT (1337,1530)	0.32	0.14	0.73	0.48	0.36	0.64	0.15	0.03	0.74

Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A04			A05			A06		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.56	0.35	0.89	0.78	0.18	3.34	0.92	0.08	10.37
II (1337, BOT)	0.35	0.18	0.66	0.00	0.00	7E142	0.00	0.00	7E142
III (1530, BOT)	0.44	0.30	0.65	0.22	0.03	1.93	0.00	0.00	7E167
IV (1659, BB)	0.45	0.25	0.82	0.00	0.00	1E123	0.00	0.00	3E136
V (1373, BOT)	1.04	0.72	1.50	0.97	0.19	4.91	1.03	0.21	5.14
VI (1431,2175, B	0.99	0.72	1.37	0.94	0.37	2.39	0.89	0.29	2.70
BB (1336,1659)	0.51	0.36	0.73	0.66	0.19	2.31	0.67	0.09	4.82
BOT (1337,1530)	0.42	0.30	0.59	0.16	0.02	1.33	0.00	0.00	6E185

Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A11			A12			A13			A15		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.67	0.35	1.30	0.90	0.61	1.32	0.72	0.49	1.07	0.95	0.23	3.89
II (1337, BOT)	0.26	0.09	0.79	0.40	0.25	0.66	0.41	0.21	0.77	0.00	0.00	7E142
III (1530, BOT)	0.31	0.11	0.89	0.36	0.24	0.55	0.38	0.26	0.56	0.22	0.03	1.93
IV (1659, BB)	0.39	0.16	0.99	0.52	0.31	0.84	0.59	0.36	0.98	0.00	0.00	1E123
V (1373, BOT)	1.11	0.52	2.34	0.85	0.60	1.21	0.87	0.62	1.23	0.83	0.25	2.75
VI (1431,2175, B	0.85	0.44	1.64	0.99	0.69	1.42	0.82	0.59	1.13	0.90	0.42	1.92
BB (1336,1659)	0.55	0.33	0.92	0.76	0.57	1.02	0.68	0.50	0.91	0.80	0.24	2.65
BOT (1337,1530)	0.28	0.13	0.61	0.41	0.31	0.55	0.42	0.31	0.57	0.16	0.02	1.33

Statistical Documentation

List of Contents

Page

1: Logistic regression for endpoint A01event_any controlling for hba1c_change 4
2: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 22
3: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 23
4: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 24
5: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 25
6: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 26
7: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change 27
8: Logistic regression for endpoint A01event_any controlling for hba1c_eot 28
9: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 46
10: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 47
11: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 48
12: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 49
13: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 50
14: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot 51
15: Logistic regression for endpoint A02event_any controlling for hba1c_change 52
16: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 70
17: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 71
18: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 72
19: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 73
20: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 74
21: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change 75
22: Logistic regression for endpoint A02event_any controlling for hba1c_eot 76
23: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
24: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
25: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
26: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
27: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
28: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot 78
29: Logistic regression for endpoint A03event_any controlling for hba1c_change 78
30: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78
31: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78
32: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78
33: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78
34: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78

35: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change 78
36: Logistic regression for endpoint A03event_any controlling for hba1c_eot 78
37: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
38: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
39: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
40: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
41: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
42: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot 78
43: Logistic regression for endpoint A04event_any controlling for hba1c_change 78
44: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
45: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
46: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
47: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
48: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
49: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change 78
50: Logistic regression for endpoint A04event_any controlling for hba1c_eot 78
51: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
52: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
53: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
54: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
55: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
56: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot 78
57: Logistic regression for endpoint A05event_any controlling for hba1c_change 78
58: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
59: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
60: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
61: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
62: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
63: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change 78
64: Logistic regression for endpoint A05event_any controlling for hba1c_eot 78
65: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
66: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
67: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
68: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
69: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
70: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot 78
71: Logistic regression for endpoint A06event_any controlling for hba1c_change 78
72: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78
73: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78

74: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78
75: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78
76: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78
77: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change 78
78: Logistic regression for endpoint A06event_any controlling for hba1c_eot 78
79: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78
80: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78
81: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78
82: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78
83: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78
84: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot 78

1: Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A01event_any	A01: Minor or symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read 505
Number of Observations Used 477

Response Profile

Ordered Value	A01event_any	Total Frequency
1	0	447
2	1	30

Probability modeled is A01event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	226.052	230.023
SC	230.219	242.526
-2 Log L	224.052	224.023

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0282	2	0.9860
Score	0.0284	2	0.9859
Wald	0.0284	2	0.9859

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0018	0.9665
hba1c_change	1	0.0258	0.8723

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.7081	0.2105	165.5107	<.0001
TRTSEQ_LB	DETEMIR	1	-0.00843	0.2004	0.0018	0.9665
hba1c_change		1	-0.0292	0.1814	0.0258	0.8723

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.983	0.448	2.157
hba1c_change	0.971	0.681	1.386

Association of Predicted Probabilities and Observed Responses

Percent Concordant	36.3	Somers' D	0.044
Percent Discordant	31.9	Gamma	0.065
Percent Tied	31.7	Tau-a	0.005
Pairs	13410	c	0.522

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A01event_any	A01: Minor or symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read 467
Number of Observations Used 416

Response Profile

Ordered Value	A01event_any	Total Frequency
1	0	403
2	1	13

Probability modeled is A01event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	117.699	119.307
SC	121.729	131.399
-2 Log L	115.699	113.307

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.3919	2	0.3024
Score	2.5478	2	0.2797
Wald	2.4078	2	0.3000

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	2.3928	0.1219
hba1c_change	1	0.0306	0.8611

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.3323	0.3645	83.5759	<.0001
TRTSEQ_LB	DETEMIR	1	-0.4463	0.2885	2.3928	0.1219
hba1c_change		1	0.0421	0.2406	0.0306	0.8611

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.410	0.132	1.269
hba1c_change		1.043	0.651	1.671

Association of Predicted Probabilities and Observed Responses

Percent Concordant	52.8	Somers' D	0.220
Percent Discordant	30.8	Gamma	0.264
Percent Tied	16.5	Tau-a	0.013
Pairs	5239	c	0.610

```
Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    445
       2     1                     16

Probability modeled is A01event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	140.984	139.259
SC	145.117	151.660
-2 Log L	138.984	133.259

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.7247	2	0.0571
Score	5.5797	2	0.0614
Wald	5.2218	2	0.0735

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	3.9414	0.0471
hba1c_change	1	1.5100	0.2191

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.9703	0.4736	39.3297	<.0001
TRTSEQ_LB	DETEMIR	1	-0.5822	0.2933	3.9414	0.0471
hba1c_change		1	0.3201	0.2605	1.5100	0.2191

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.312	0.099	0.985
hba1c_change		1.377	0.827	2.295

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.8	Somers' D	0.367
Percent Discordant	30.1	Gamma	0.379
Percent Tied	3.1	Tau-a	0.025
Pairs	7120	c	0.684

```
Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    247
       2     1                     18

Probability modeled is A01event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	133.566	127.260
SC	137.145	137.999
-2 Log L	131.566	121.260

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	10.3056	2	0.0058
Score	9.5028	2	0.0086
Wald	9.3981	2	0.0091

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	2.8515	0.0913
hba1c_change	1	7.1124	0.0077

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.2730	0.2937	59.9075	<.0001
TRTSEQ_LB	DETEMIR	1	-0.4647	0.2752	2.8515	0.0913
hba1c_change		1	0.6729	0.2523	7.1124	0.0077

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.395	0.134	1.161
hba1c_change		1.960	1.195	3.213

Association of Predicted Probabilities and Observed Responses

Percent Concordant	74.1	Somers' D	0.502
Percent Discordant	23.9	Gamma	0.512
Percent Tied	2.0	Tau-a	0.064
Pairs	4446	c	0.751

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A01event_any	A01: Minor or symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read 582
Number of Observations Used 543

Response Profile

Ordered Value	A01event_any	Total Frequency
1	0	520
2	1	23

Probability modeled is A01event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	192.446	195.790
SC	196.743	208.682
-2 Log L	190.446	189.790

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.6555	2	0.7205
Score	0.6583	2	0.7195
Wald	0.6548	2	0.7208

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.4757	0.4904
hba1c_change	1	0.1627	0.6867

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.0177	0.3389	79.3076	<.0001
TRTSEQ_LB	DETEMIR	1	0.1483	0.2149	0.4757	0.4904
hba1c_change		1	0.0777	0.1926	0.1627	0.6867

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.345	0.579	3.124
hba1c_change	1.081	0.741	1.576

Association of Predicted Probabilities and Observed Responses

Percent Concordant	52.3	Somers' D	0.126
Percent Discordant	39.6	Gamma	0.137
Percent Tied	8.1	Tau-a	0.010
Pairs	11960	c	0.563

```
Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    661
       2     1                     23

Probability modeled is A01event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	203.271	197.791
SC	207.799	211.375
-2 Log L	201.271	191.791

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.4803	2	0.0087
Score	9.6917	2	0.0079
Wald	9.8007	2	0.0074

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.3348	0.5628
hba1c_change	1	9.7962	0.0017

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.8884	0.2424	142.0373	<.0001
TRTSEQ_LB	DETEMIR	1	-0.1318	0.2277	0.3348	0.5628
hba1c_change		1	0.5592	0.1787	9.7962	0.0017

Logistic regression for endpoint A01event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

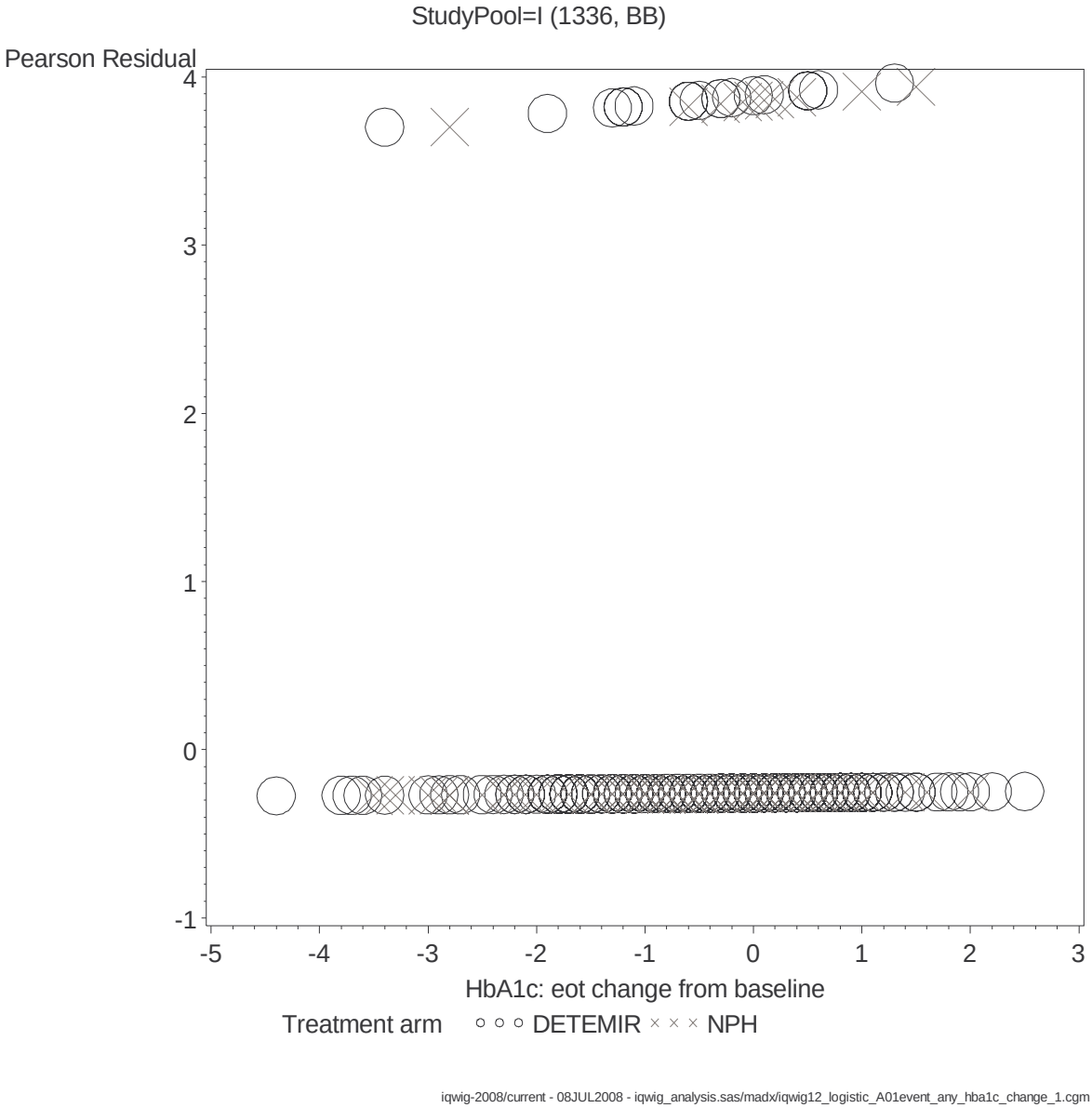
The LOGISTIC Procedure

Odds Ratio Estimates

Effect			Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR	vs GLARGINE	0.768	0.315	1.876
hba1c_change			1.749	1.232	2.483

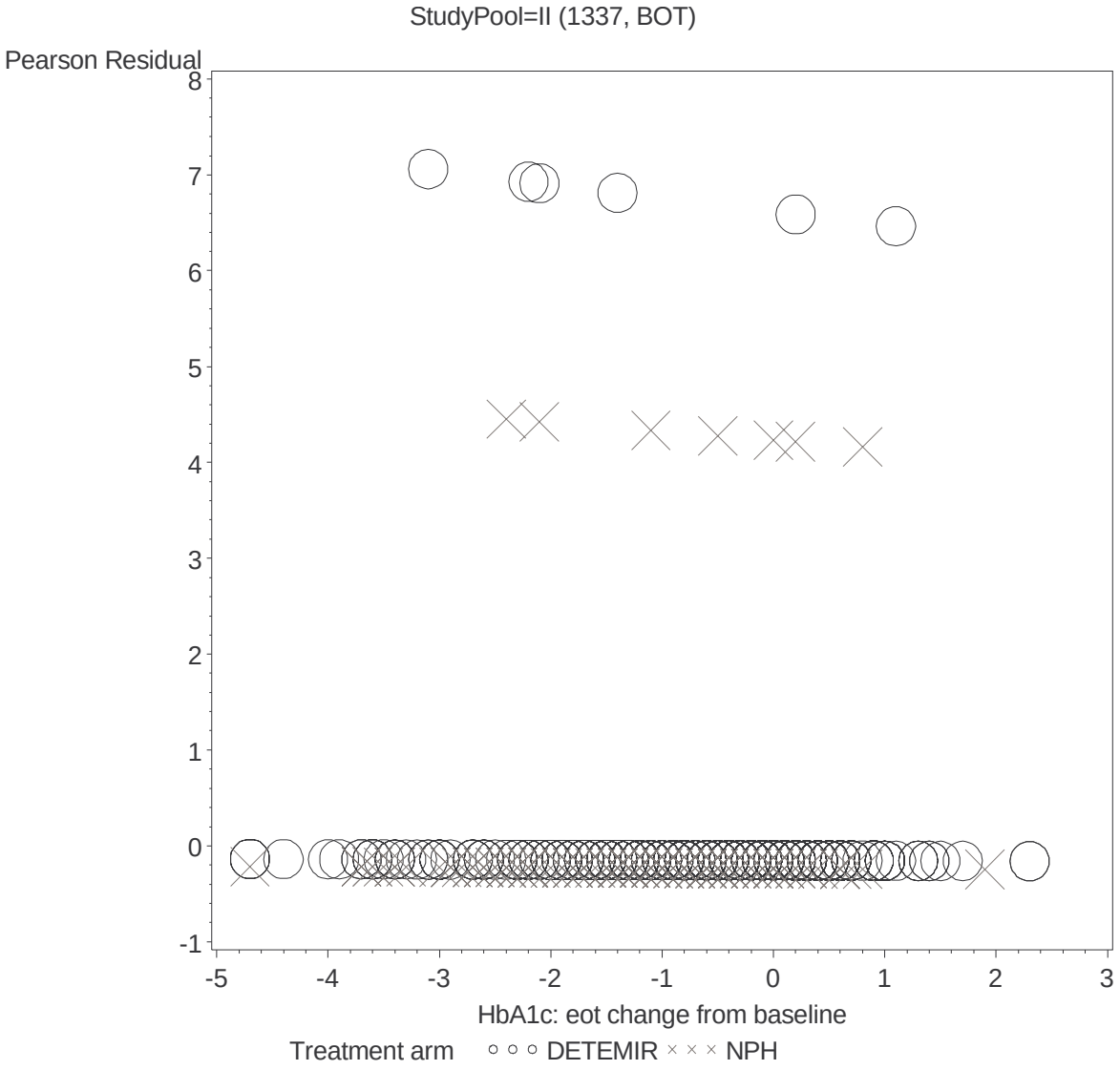
Association of Predicted Probabilities and Observed Responses

Percent Concordant	65.1	Somers' D	0.335
Percent Discordant	31.6	Gamma	0.346
Percent Tied	3.2	Tau-a	0.022
Pairs	15203	c	0.668



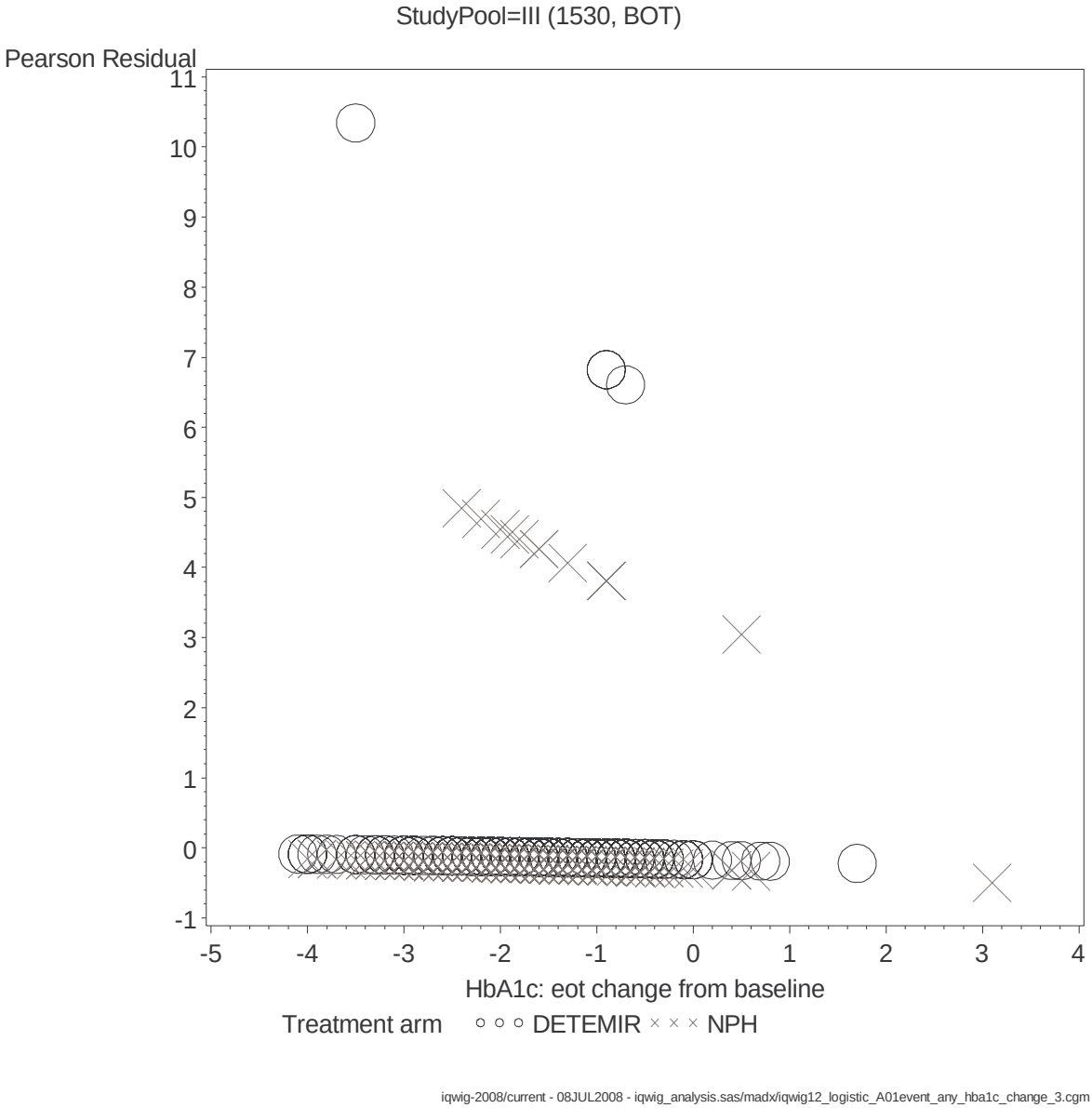
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_change_1.cgm

2: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change

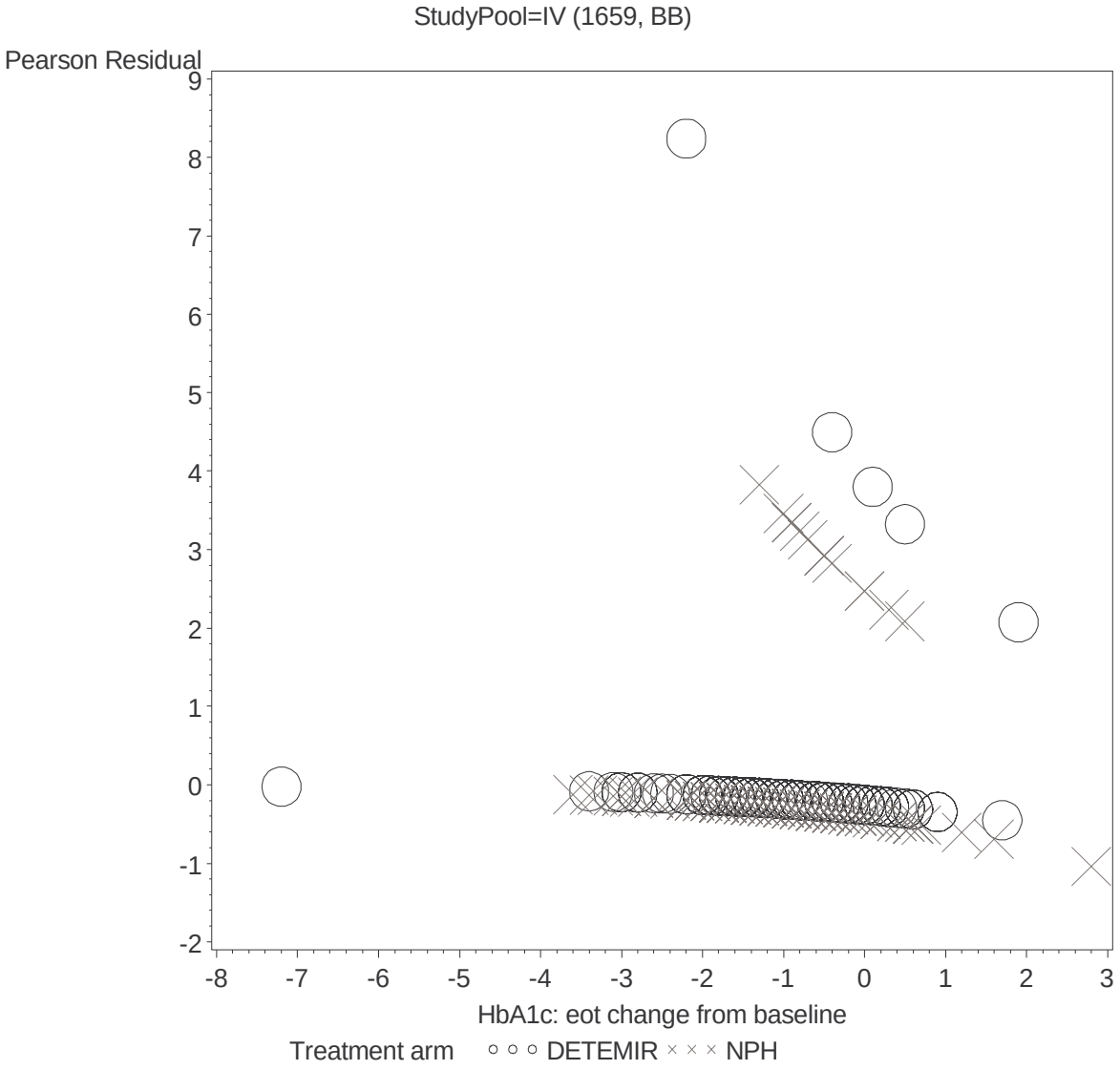


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_change_2.cgm

3: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change

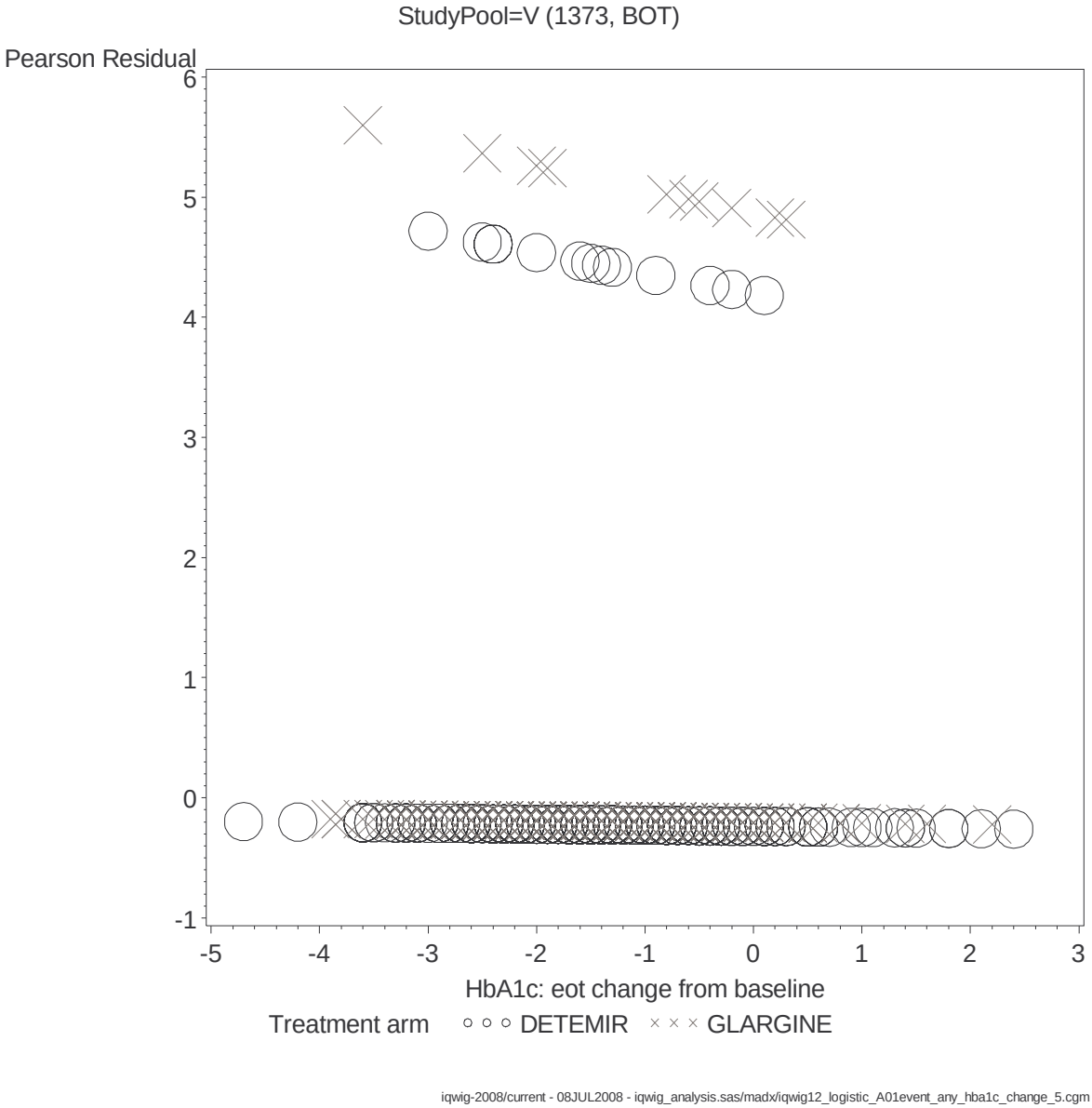


4: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change

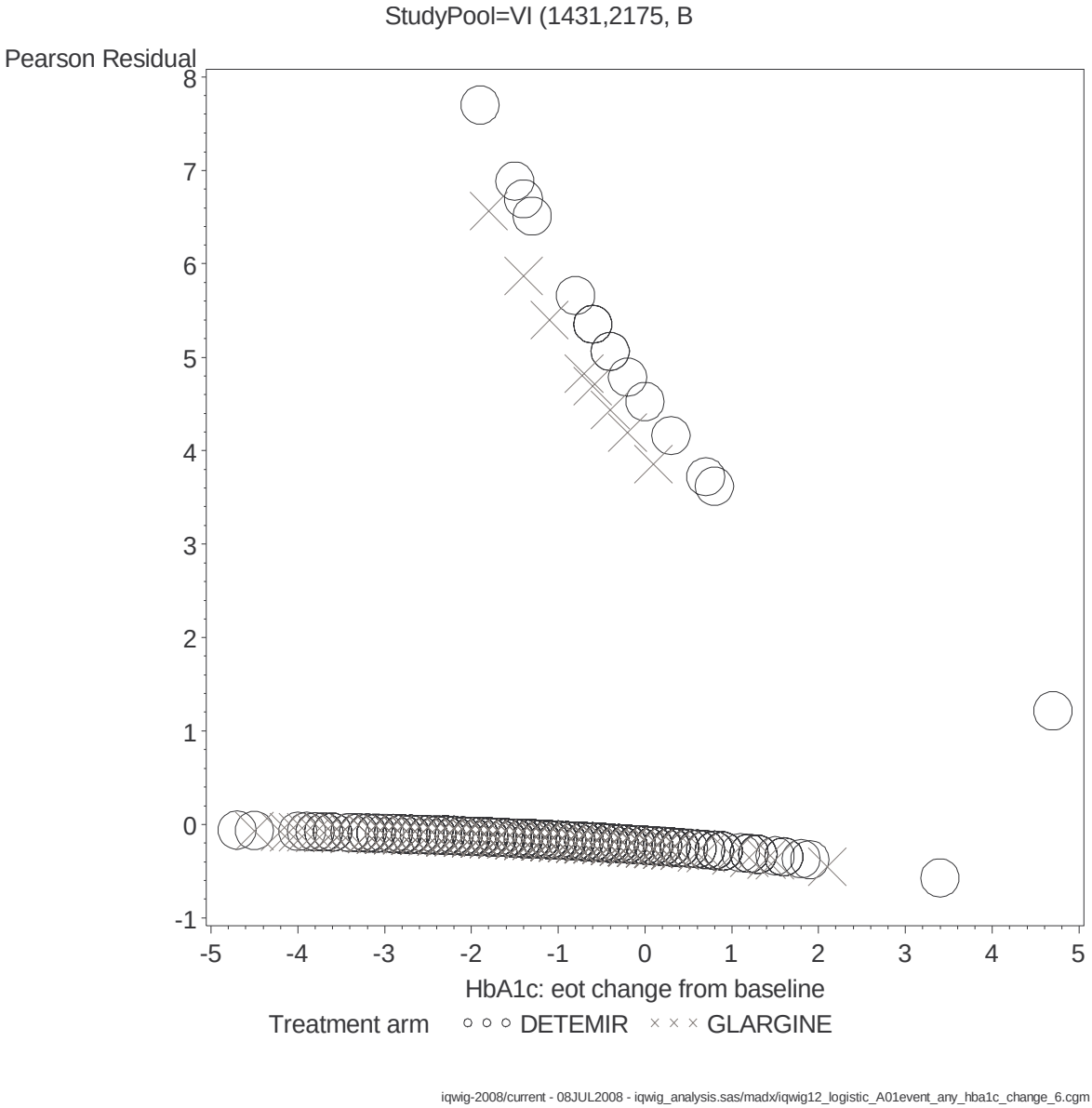


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_change_4.cgm

5: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change



6: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change



7: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_change

8: Logistic regression for endpoint A01event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    447
       2     1                     30

Probability modeled is A01event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	226.052	222.769
SC	230.219	235.272
-2 Log L	224.052	216.769

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.2825	2	0.0262
Score	6.6667	2	0.0357
Wald	6.6641	2	0.0357

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0247	0.8750
hba1c_eot	1	6.6632	0.0098

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.2507	1.5018	0.6936	0.4049
TRTSEQ_LB DETEMIR	1	0.0317	0.2018	0.0247	0.8750
hba1c_eot	1	-0.5431	0.2104	6.6632	0.0098

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	1.066	0.483	2.350
hba1c_eot	0.581	0.385	0.877

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.0	Somers' D	0.266
Percent Discordant	35.4	Gamma	0.273
Percent Tied	2.6	Tau-a	0.031
Pairs	13410	c	0.633

```
Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                                  Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    403
       2     1                     13

Probability modeled is A01event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	117.699	118.220
SC	121.729	130.312
-2 Log L	115.699	112.220

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.4787	2	0.1756
Score	3.5994	2	0.1653
Wald	3.4282	2	0.1801

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	3.0334	0.0816
hba1c_eot	1	1.1471	0.2842

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.5495	2.0899	7.0511	0.0079
TRTSEQ_LB DETEMIR	1	-0.5126	0.2943	3.0334	0.0816
hba1c_eot	1	0.2543	0.2375	1.1471	0.2842

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.359	0.113	1.137
hba1c_eot	1.290	0.810	2.054

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.2	Somers' D	0.300
Percent Discordant	33.2	Gamma	0.311
Percent Tied	3.6	Tau-a	0.018
Pairs	5239	c	0.650

```
Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1             0            445
       2             1             16

Probability modeled is A01event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	140.984	139.182
SC	145.117	151.582
-2 Log L	138.984	133.182

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.8021	2	0.0550
Score	5.7300	2	0.0570
Wald	5.3591	2	0.0686

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	4.1923	0.0406
hba1c_eot	1	1.6575	0.1979

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-6.2813	2.2371	7.8834	0.0050
TRTSEQ_LB DETEMIR	1	-0.6030	0.2945	4.1923	0.0406
hba1c_eot	1	0.4036	0.3135	1.6575	0.1979

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.299	0.094	0.950
hba1c_eot	1.497	0.810	2.768

Association of Predicted Probabilities and Observed Responses

Percent Concordant	65.6	Somers' D	0.344
Percent Discordant	31.2	Gamma	0.355
Percent Tied	3.2	Tau-a	0.023
Pairs	7120	c	0.672

```
Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A01event_          Total
   Value     any            Frequency

       1     0                    247
       2     1                     18

Probability modeled is A01event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	133.566	131.367
SC	137.145	142.107
-2 Log L	131.566	125.367

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.1981	2	0.0451
Score	5.9525	2	0.0510
Wald	5.7484	2	0.0565

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	2.9037	0.0884
hba1c_eot	1	3.3735	0.0663

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-6.4269	2.0809	9.5387	0.0020
TRTSEQ_LB DETEMIR	1	-0.4661	0.2735	2.9037	0.0884
hba1c_eot	1	0.4584	0.2496	3.3735	0.0663

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.394	0.135	1.150
hba1c_eot	1.582	0.970	2.579

Association of Predicted Probabilities and Observed Responses

Percent Concordant	65.7	Somers' D	0.331
Percent Discordant	32.6	Gamma	0.336
Percent Tied	1.7	Tau-a	0.042
Pairs	4446	c	0.665

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A01event_any	A01: Minor or symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read 582
Number of Observations Used 543

Response Profile

Ordered Value	A01event_any	Total Frequency
1	0	520
2	1	23

Probability modeled is A01event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	192.446	195.535
SC	196.743	208.426
-2 Log L	190.446	189.535

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.9115	2	0.6340
Score	0.9335	2	0.6270
Wald	0.9282	2	0.6287

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.4606	0.4973
hba1c_eot	1	0.4319	0.5111

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.1026	1.5131	7.3511	0.0067
TRTSEQ_LB DETEMIR	1	0.1460	0.2151	0.4606	0.4973
hba1c_eot	1	0.1351	0.2055	0.4319	0.5111

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.339	0.576	3.111
hba1c_eot	1.145	0.765	1.712

Association of Predicted Probabilities and Observed Responses

Percent Concordant	53.6	Somers' D	0.141
Percent Discordant	39.5	Gamma	0.151
Percent Tied	6.9	Tau-a	0.011
Pairs	11960	c	0.570

```
Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                                  Model Information

Data Set                        WORK.ANALYSIS_IQWIG
Response Variable               A01event_any            A01: Minor or symptom, daytime (not
nocturnal), BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels       2
Model                           binary logit
Optimization Technique          Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A01event_         Total
   Value     any           Frequency

       1           0             661
       2           1              23

Probability modeled is A01event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              GLARGINE     -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	203.271	206.162
SC	207.799	219.746
-2 Log L	201.271	200.162

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.1091	2	0.5743
Score	1.1764	2	0.5553
Wald	1.1707	2	0.5569

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0599	0.8067
hba1c_eot	1	1.1513	0.2833

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.7897	1.3824	12.0052	0.0005
TRTSEQ_LB DETEMIR	1	-0.0548	0.2240	0.0599	0.8067
hba1c_eot	1	0.1915	0.1785	1.1513	0.2833

Logistic regression for endpoint A01event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

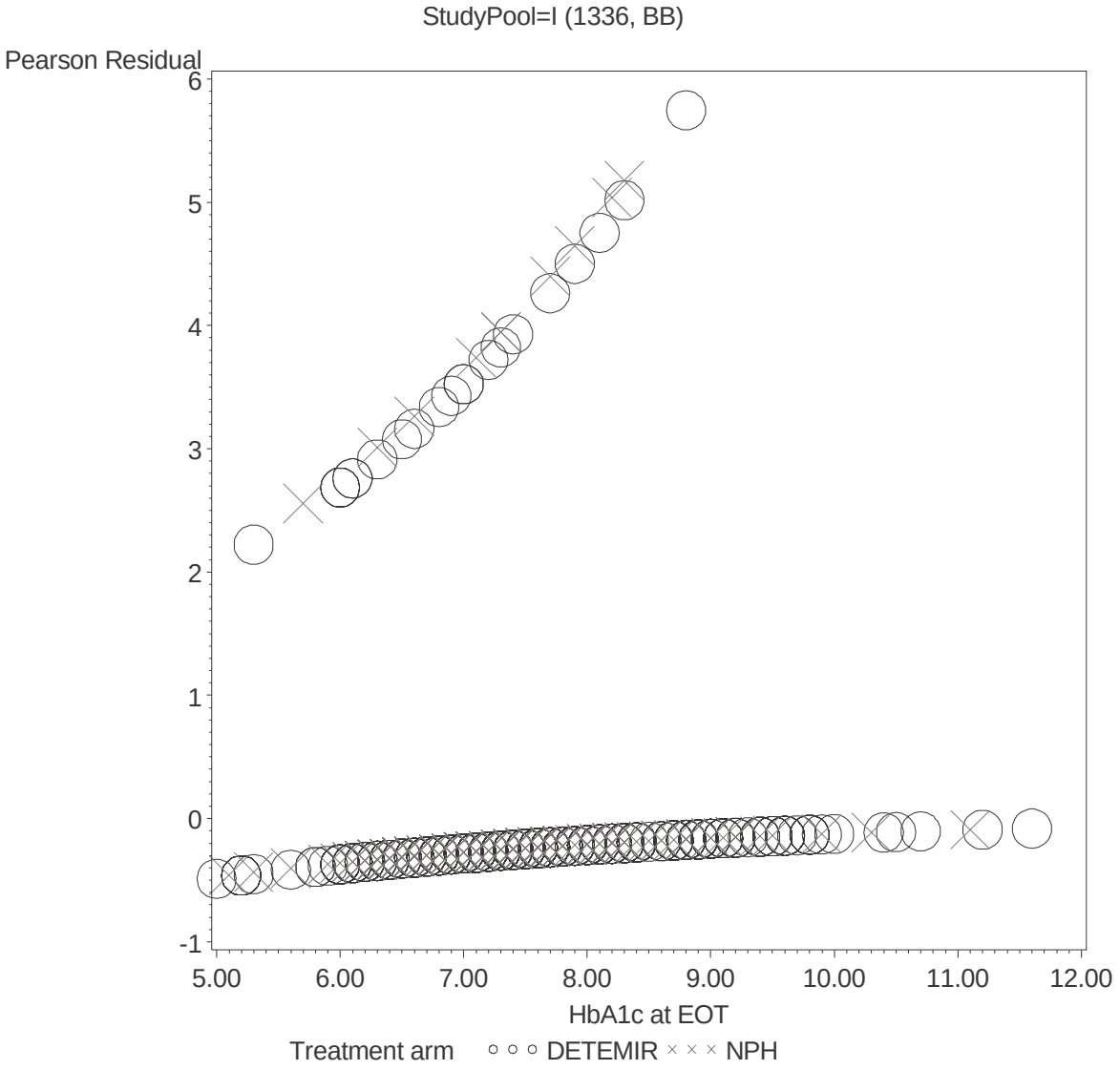
The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.896	0.372	2.157
hba1c_eot	1.211	0.854	1.718

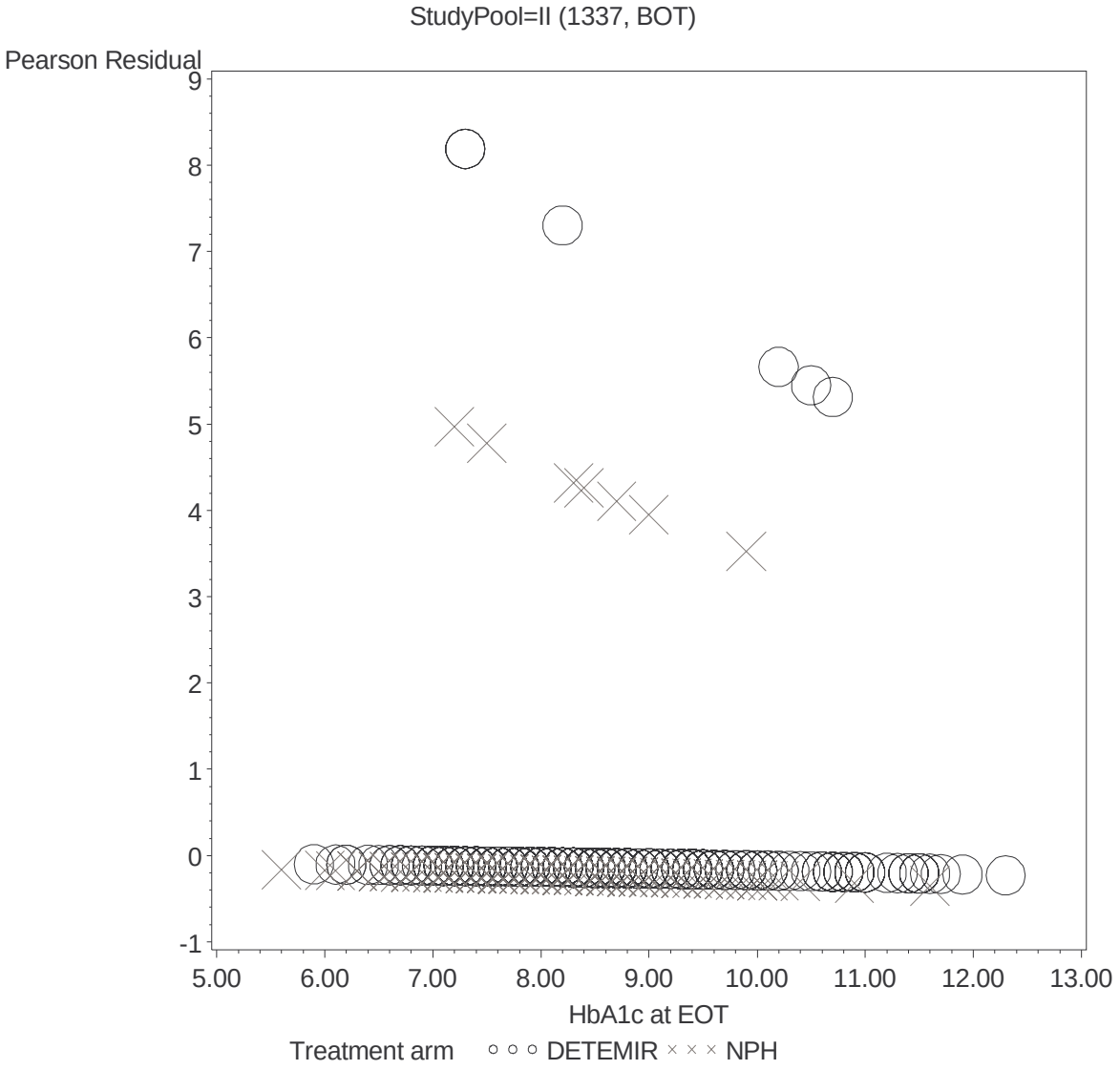
Association of Predicted Probabilities and Observed Responses

Percent Concordant	50.5	Somers' D	0.099
Percent Discordant	40.6	Gamma	0.109
Percent Tied	8.9	Tau-a	0.006
Pairs	15203	c	0.549



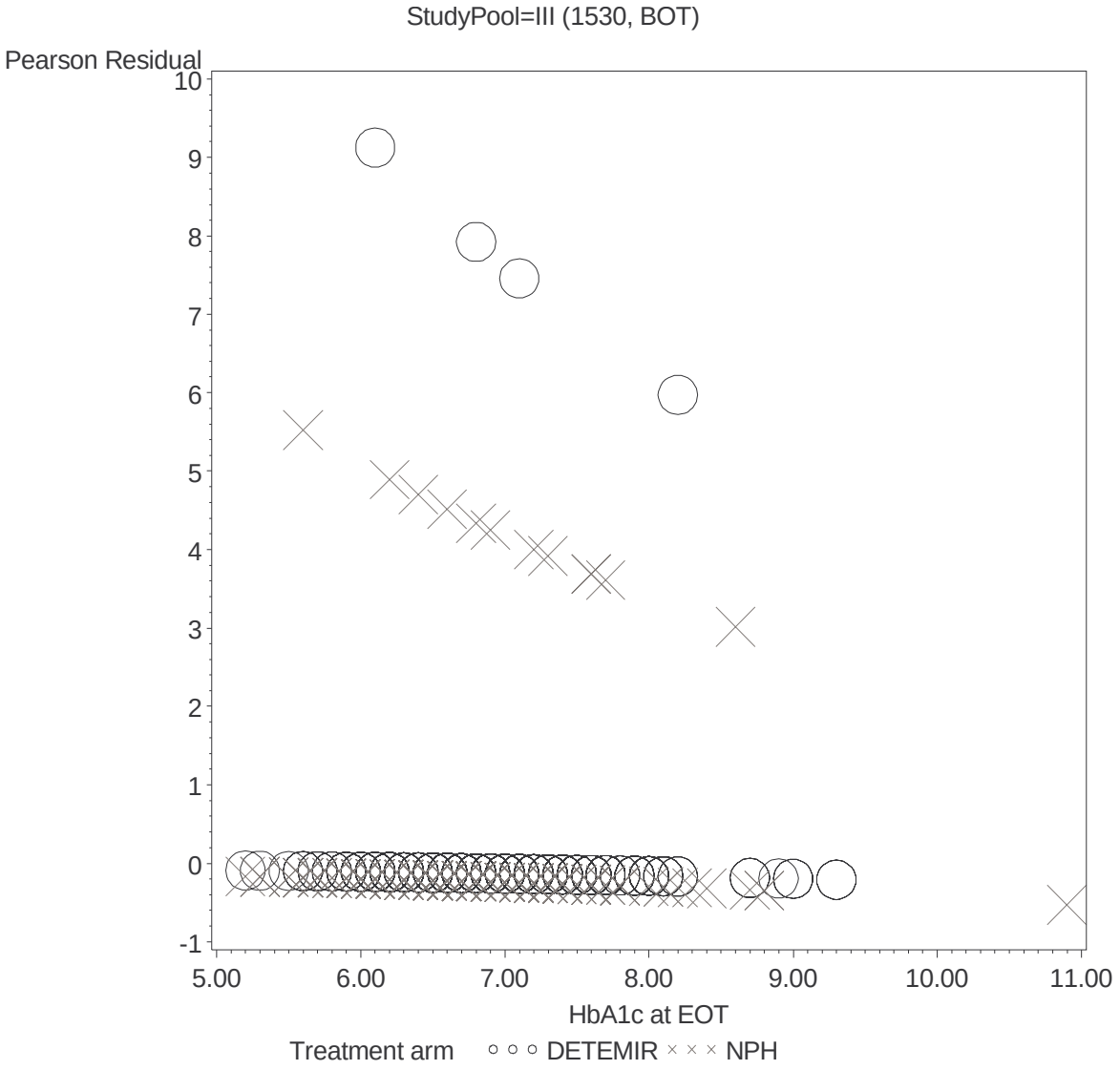
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_1.cgm

9: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot



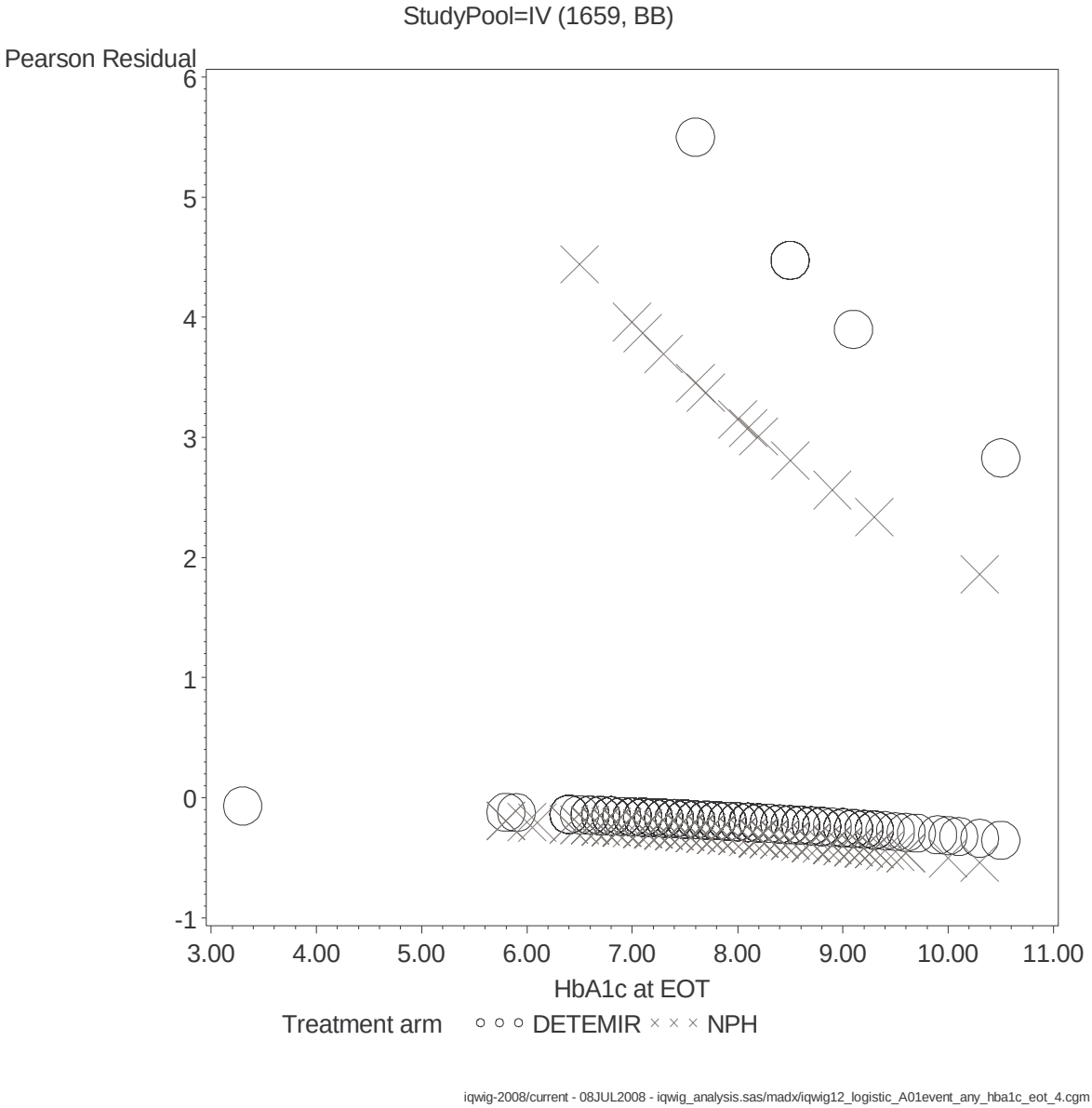
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_2.cgm

10: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot



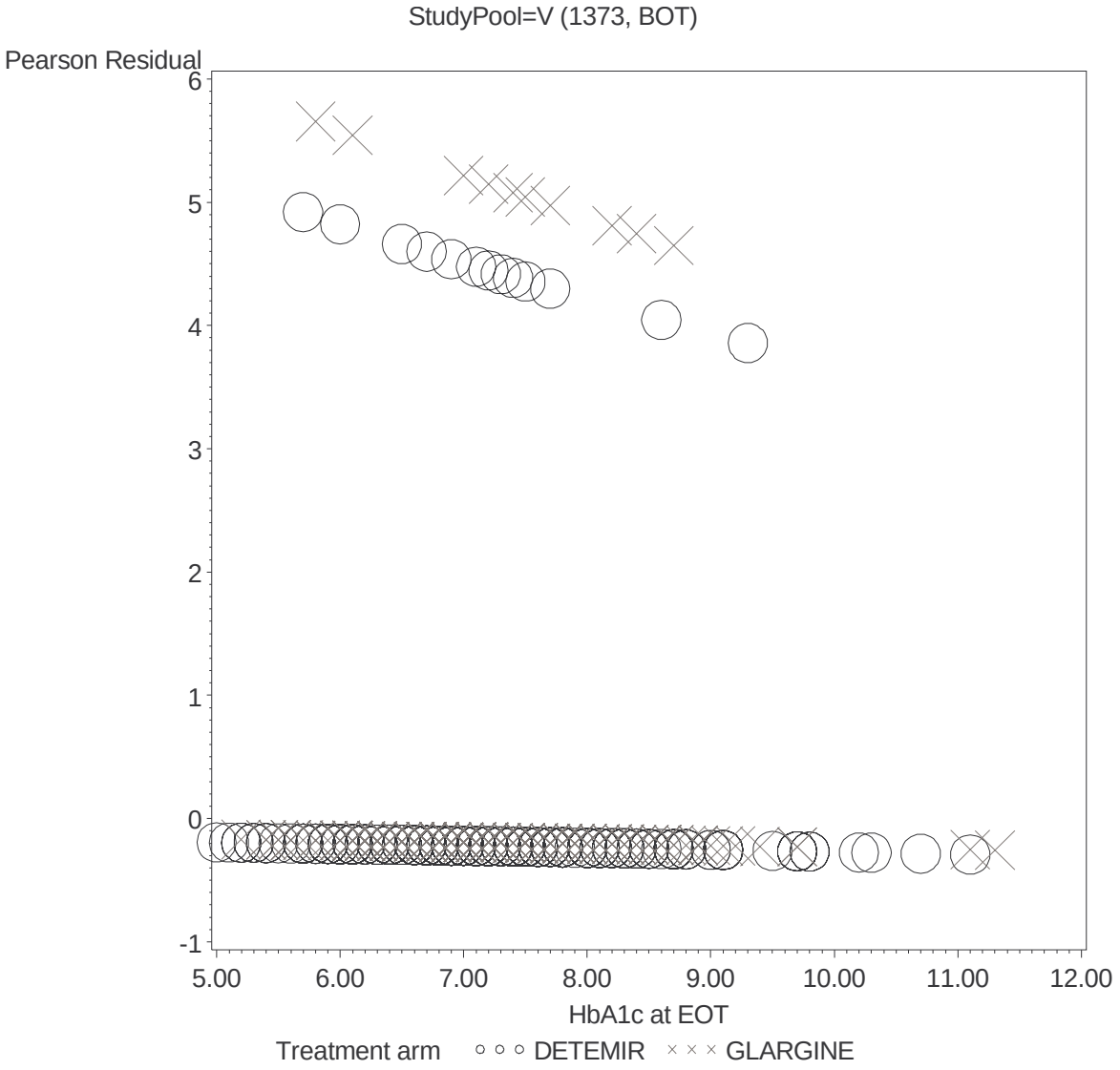
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_3.cgm

11: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot



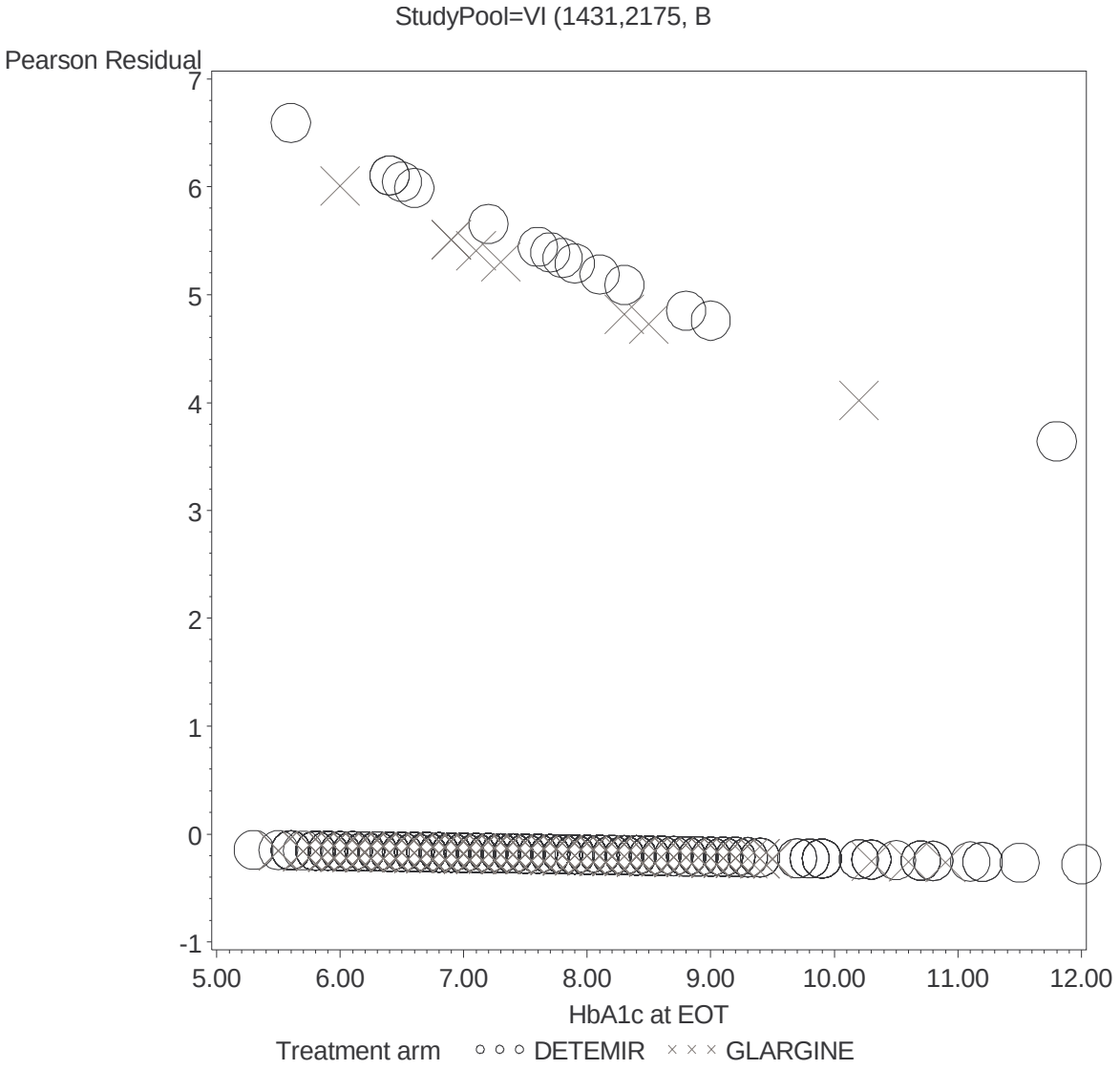
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_4.cgm

12: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_5.cgm

13: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A01event_any_hba1c_eot_6.cgm

14: Logistic regression: residuals for endpoint A01event_any controlling for hba1c_eot

15: Logistic regression for endpoint A02event_any controlling for hba1c_change

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A02event_any          A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A02event_          Total
   Value     any            Frequency

       1     0                    250
       2     1                    227

Probability modeled is A02event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	662.153	664.329
SC	666.320	676.832
-2 Log L	660.153	658.329

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.8235	2	0.4018
Score	1.8215	2	0.4022
Wald	1.8131	2	0.4039

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.3344	0.5631
hba1c_change	1	1.4060	0.2357

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.1108	0.1022	1.1752	0.2783
TRTSEQ_LB	DETEMIR	1	-0.0565	0.0978	0.3344	0.5631
hba1c_change		1	-0.1062	0.0895	1.4060	0.2357

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.893	0.609	1.310
hba1c_change		0.899	0.755	1.072

Association of Predicted Probabilities and Observed Responses

Percent Concordant	51.7	Somers' D	0.057
Percent Discordant	46.0	Gamma	0.058
Percent Tied	2.2	Tau-a	0.028
Pairs	56750	c	0.529

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             349
       2           1              67

Probability modeled is A02event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              NPH          -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	369.261	363.554
SC	373.292	375.646
-2 Log L	367.261	357.554

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.7069	2	0.0078
Score	10.0789	2	0.0065
Wald	9.7449	2	0.0077

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	7.4798	0.0062
hba1c_change	1	1.1102	0.2920

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.6905	0.1833	85.0730	<.0001
TRTSEQ_LB	DETEMIR	1	-0.3754	0.1373	7.4798	0.0062
hba1c_change		1	-0.1165	0.1105	1.1102	0.2920

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.472	0.276	0.808
hba1c_change		0.890	0.717	1.105

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.9	Somers' D	0.237
Percent Discordant	37.2	Gamma	0.242
Percent Tied	1.9	Tau-a	0.064
Pairs	23383	c	0.619

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A02event_          Total
   Value     any            Frequency

       1             0            169
       2             1            292

Probability modeled is A02event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	607.863	591.068
SC	611.996	603.468
-2 Log L	605.863	585.068

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20.7949	2	<.0001
Score	20.5550	2	<.0001
Wald	20.0693	2	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	19.3715	<.0001
hba1c_change	1	1.3787	0.2403

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.7889	0.2119	13.8635	0.0002
TRTSEQ_LB	DETEMIR	1	-0.4395	0.0998	19.3715	<.0001
hba1c_change		1	0.1282	0.1092	1.3787	0.2403

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.415	0.281	0.614
hba1c_change		1.137	0.918	1.408

Association of Predicted Probabilities and Observed Responses

Percent Concordant	61.6	Somers' D	0.248
Percent Discordant	36.8	Gamma	0.252
Percent Tied	1.6	Tau-a	0.115
Pairs	49348	c	0.624

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A02event_          Total
   Value     any            Frequency

       1     0                    156
       2     1                    109

Probability modeled is A02event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	360.988	357.630
SC	364.568	368.369
-2 Log L	358.988	351.630

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.3576	2	0.0253
Score	7.2900	2	0.0261
Wald	7.1975	2	0.0274

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	6.9992	0.0082
hba1c_change	1	0.2680	0.6047

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.3327	0.1725	3.7201	0.0538
TRTSEQ_LB	DETEMIR	1	-0.3382	0.1278	6.9992	0.0082
hba1c_change		1	0.0626	0.1209	0.2680	0.6047

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.508	0.308	0.839
hba1c_change		1.065	0.840	1.349

Association of Predicted Probabilities and Observed Responses

Percent Concordant	57.5	Somers' D	0.170
Percent Discordant	40.5	Gamma	0.173
Percent Tied	2.0	Tau-a	0.083
Pairs	17004	c	0.585

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         582
Number of Observations Used         543


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             245
       2           1             298

Probability modeled is A02event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	749.576	752.033
SC	753.874	764.924
-2 Log L	747.576	746.033

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.5438	2	0.4621
Score	1.5431	2	0.4623
Wald	1.5414	2	0.4627

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.4770	0.2242
hba1c_change	1	0.0521	0.8195

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.1688	0.1438	1.3772	0.2406
TRTSEQ_LB	DETEMIR	1	-0.1050	0.0864	1.4770	0.2242
hba1c_change		1	-0.0182	0.0797	0.0521	0.8195

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect			Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR	vs GLARGINE	0.811	0.578	1.137
hba1c_change			0.982	0.840	1.148

Association of Predicted Probabilities and Observed Responses

Percent Concordant	49.0	Somers' D	0.042
Percent Discordant	44.8	Gamma	0.045
Percent Tied	6.1	Tau-a	0.021
Pairs	73010	c	0.521

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             235
       2           1             449

Probability modeled is A02event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	882.135	882.836
SC	886.663	896.420
-2 Log L	880.135	876.836

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	3.2986	2	0.1922
Score	3.2873	2	0.1933
Wald	3.2661	2	0.1953

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.2829	0.5948
hba1c_change	1	2.7236	0.0989

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	0.5364	0.1168	21.0899	<.0001
TRTSEQ_LB	DETEMIR	1	-0.0462	0.0868	0.2829	0.5948
hba1c_change		1	-0.1186	0.0719	2.7236	0.0989

```
Logistic regression for endpoint A02event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

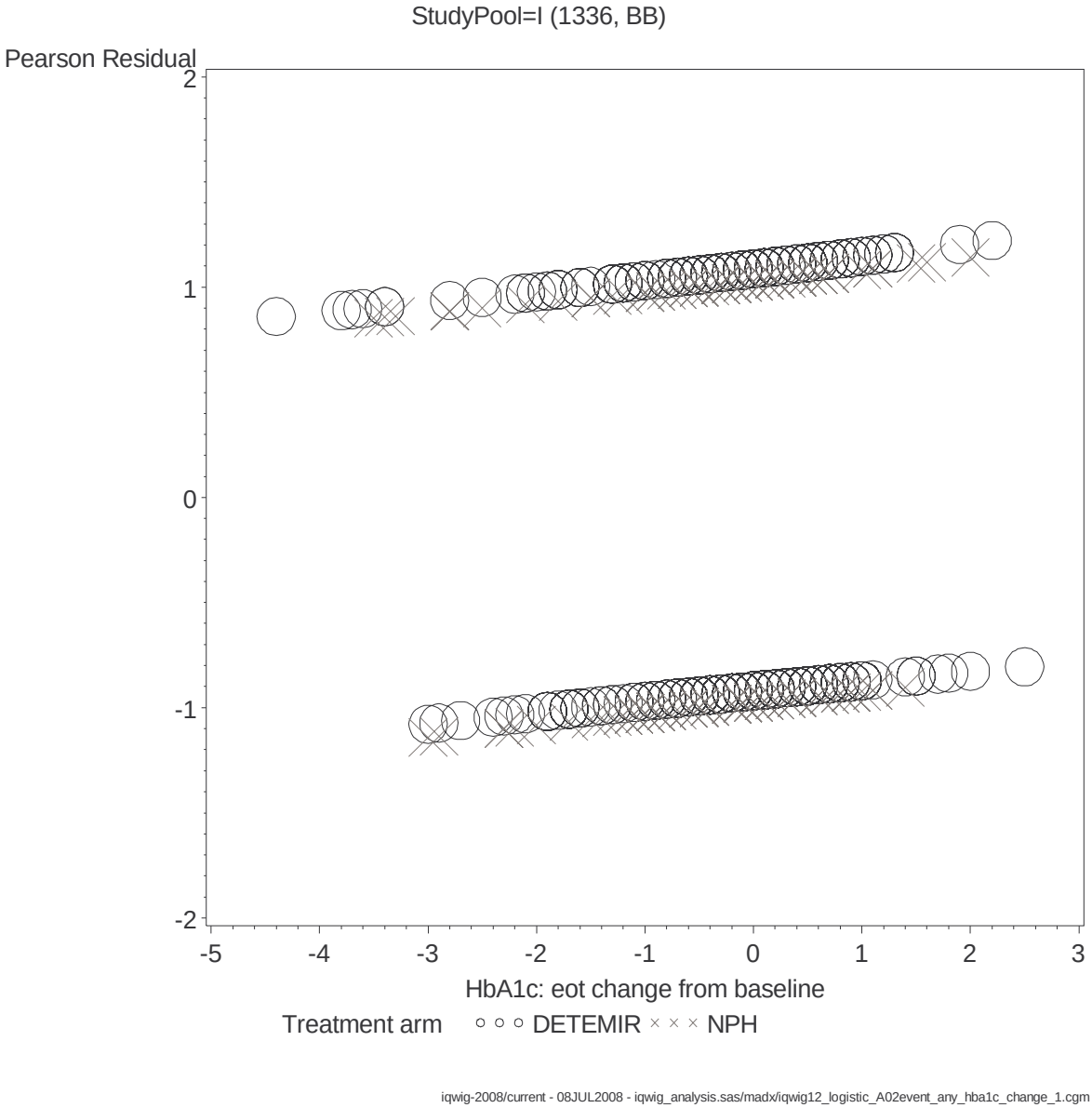
                      Odds Ratio Estimates

                                      Point          95% Wald
Effect                              Estimate      Confidence Limits

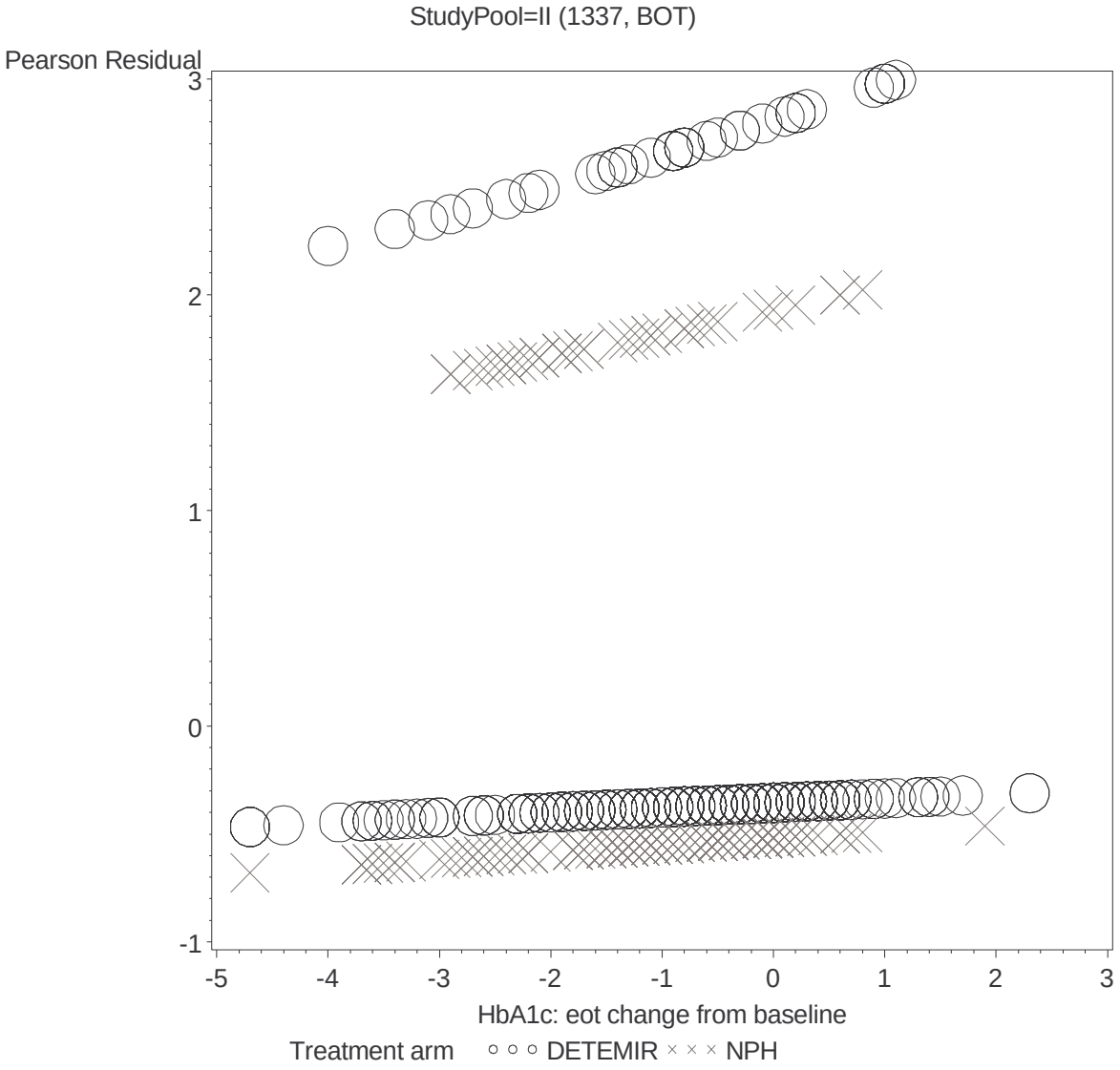
TRTSEQ_LB    DETEMIR  vs GLARGINE       0.912       0.649       1.281
hba1c_change                           0.888       0.771       1.023


Association of Predicted Probabilities and Observed Responses

Percent Concordant      53.5    Somers' D    0.091
Percent Discordant      44.4    Gamma        0.093
Percent Tied             2.0    Tau-a        0.041
Pairs                 105515    c            0.545
```

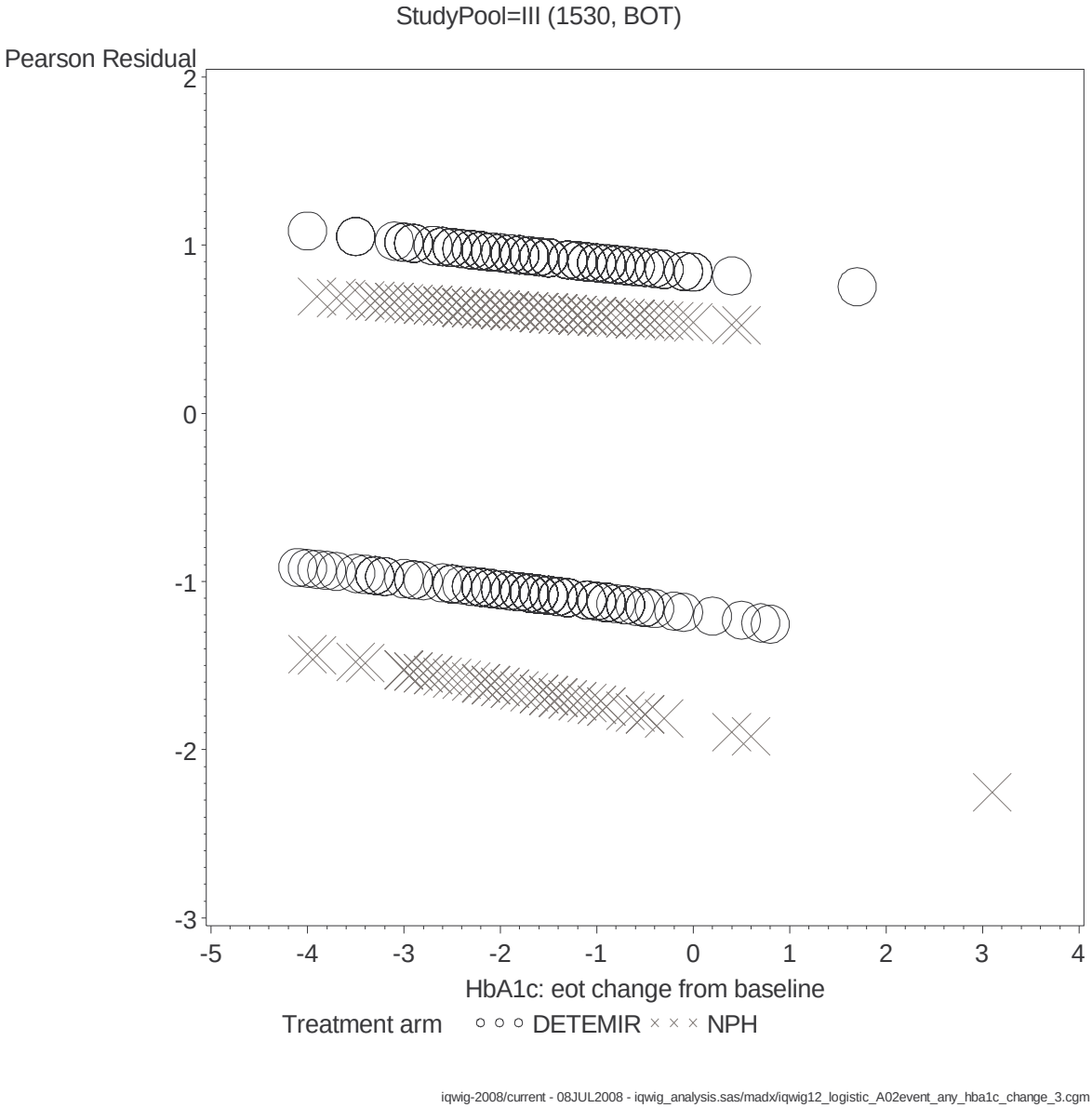


16: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change

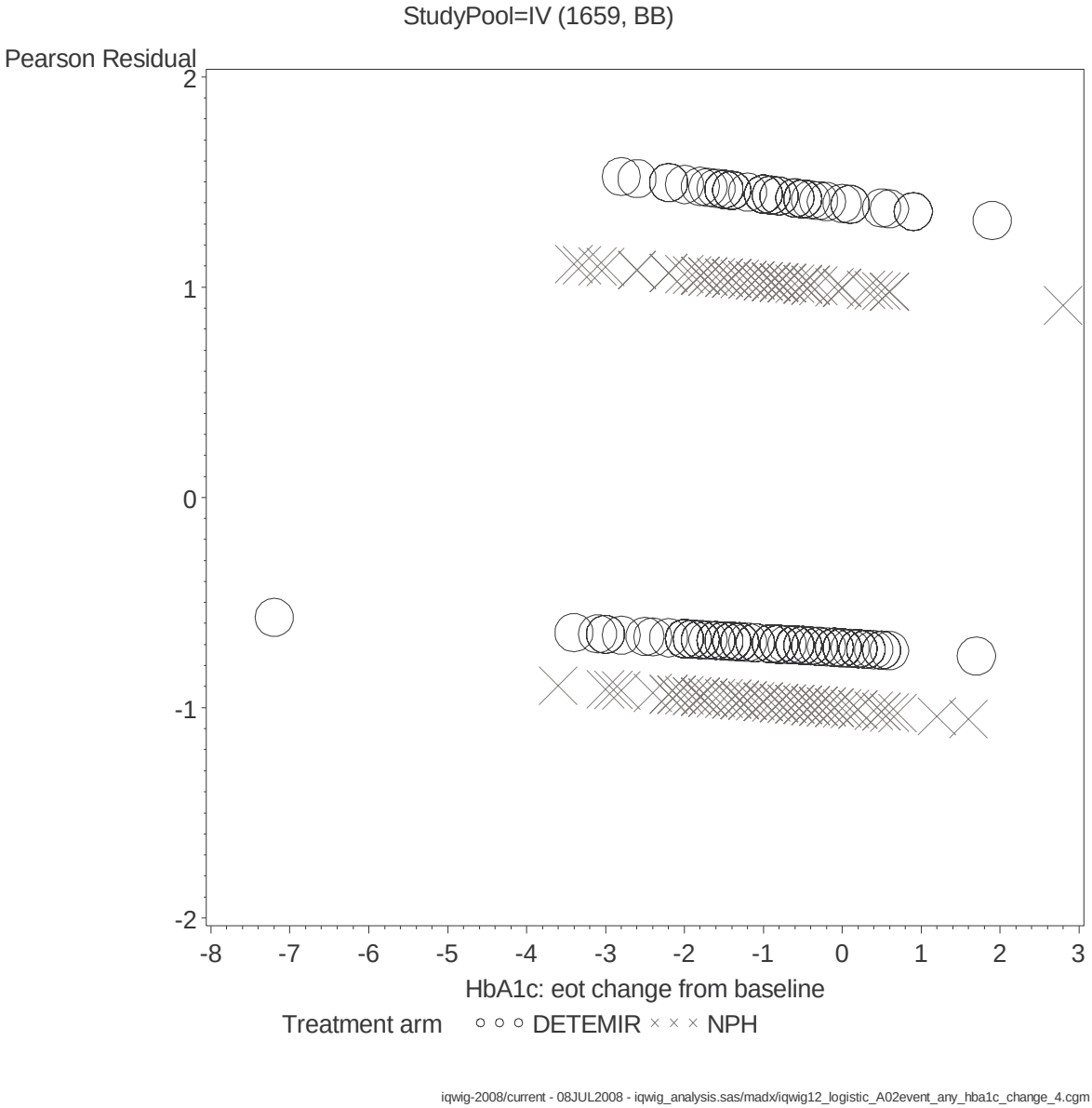


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_change_2.cgm

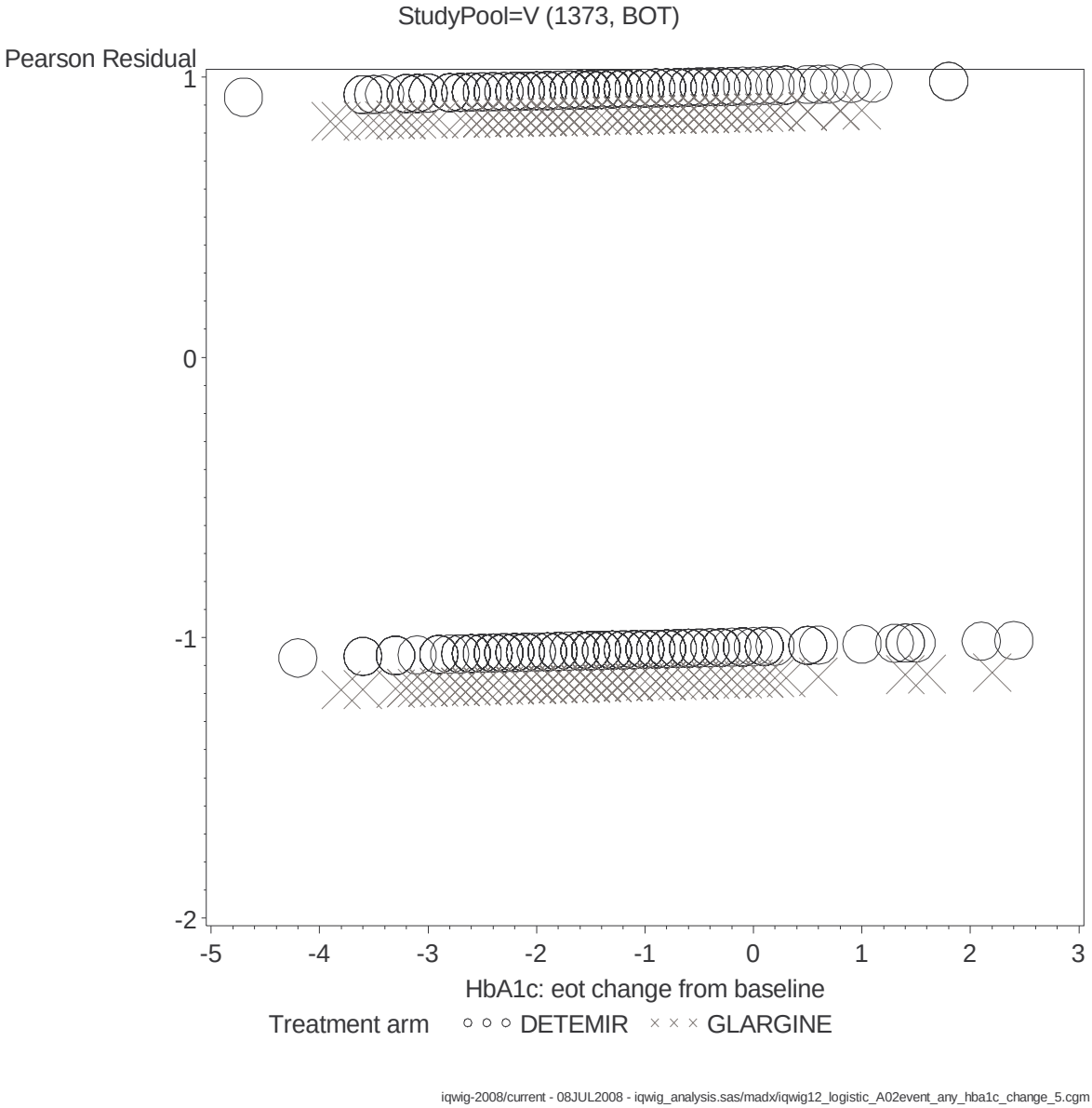
17: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change



18: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change

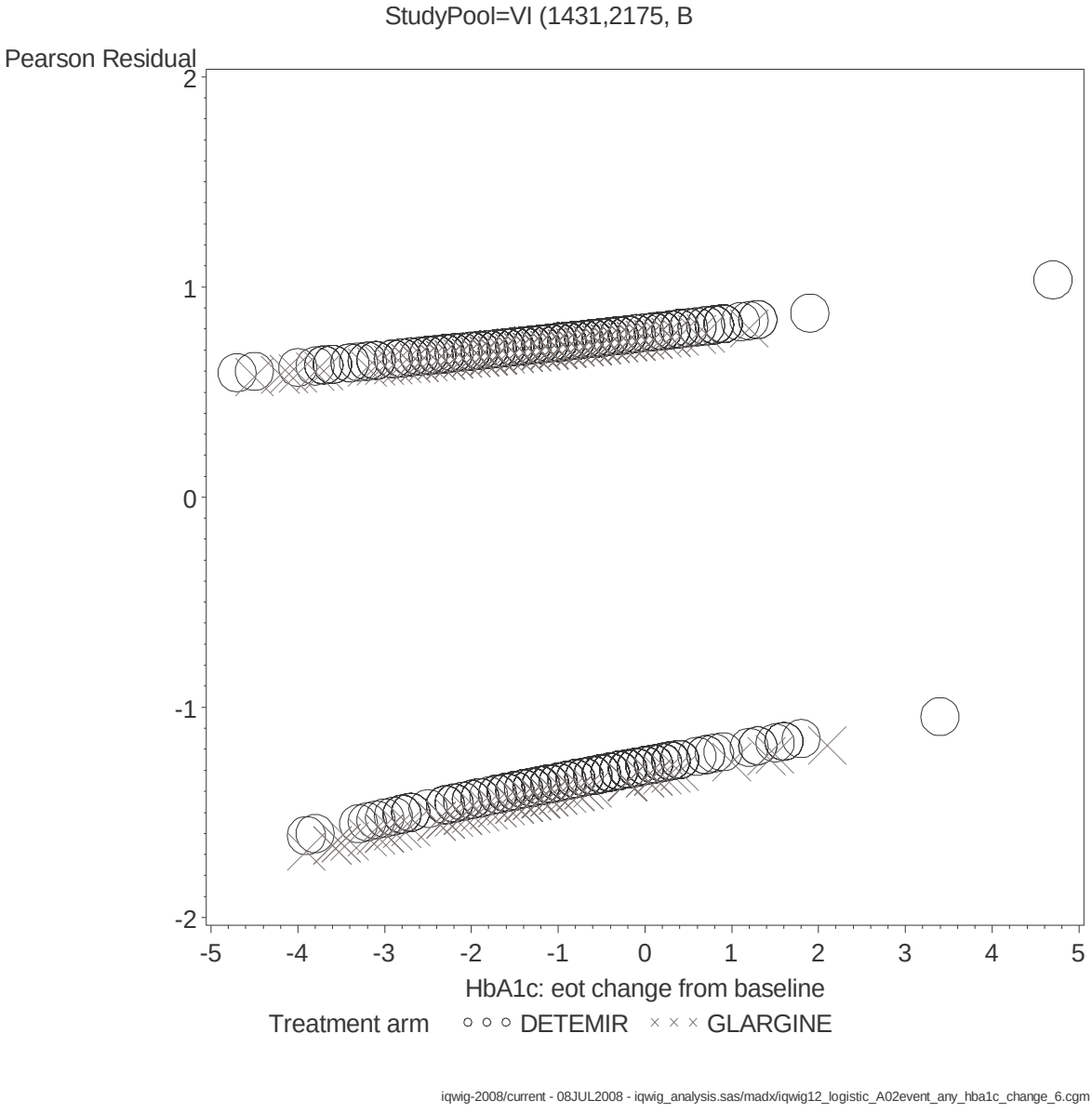


19: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_change_5.cgm

20: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change



21: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_change

22: Logistic regression for endpoint A02event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                                    Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A02event_any          A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A02event_          Total
   Value     any            Frequency

       1     0                    250
       2     1                    227

Probability modeled is A02event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	662.153	659.024
SC	666.320	671.526
-2 Log L	660.153	653.024

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.1295	2	0.0283
Score	7.0478	2	0.0295
Wald	6.9232	2	0.0314

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.1478	0.7006
hba1c_eot	1	6.5297	0.0106

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.6876	0.6959	5.8807	0.0153
TRTSEQ_LB DETEMIR	1	-0.0379	0.0986	0.1478	0.7006
hba1c_eot	1	-0.2357	0.0922	6.5297	0.0106

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.927	0.630	1.364
hba1c_eot	0.790	0.659	0.947

Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.8	Somers' D	0.139
Percent Discordant	41.9	Gamma	0.142
Percent Tied	2.3	Tau-a	0.069
Pairs	56750	c	0.569

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A02event_any	A02: Minor or symptom, daytime (not nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read 467
Number of Observations Used 416

Response Profile

Ordered Value	A02event_any	Total Frequency
1	0	349
2	1	67

Probability modeled is A02event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	369.261	364.019
SC	373.292	376.111
-2 Log L	367.261	358.019

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.2426	2	0.0098
Score	9.5955	2	0.0082
Wald	9.2884	2	0.0096

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	7.3715	0.0066
hba1c_eot	1	0.6240	0.4296

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7799	1.0006	0.6076	0.4357
TRTSEQ_LB DETEMIR	1	-0.3759	0.1385	7.3715	0.0066
hba1c_eot	1	-0.0942	0.1192	0.6240	0.4296

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.471	0.274	0.811
hba1c_eot	0.910	0.720	1.150

Association of Predicted Probabilities and Observed Responses

Percent Concordant	58.5	Somers' D	0.192
Percent Discordant	39.3	Gamma	0.197
Percent Tied	2.2	Tau-a	0.052
Pairs	23383	c	0.596

```
Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                        WORK.ANALYSIS_IQWIG
Response Variable               A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels       2
Model                           binary logit
Optimization Technique          Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             169
       2           1             292

Probability modeled is A02event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              NPH          -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	607.863	591.332
SC	611.996	603.733
-2 Log L	605.863	585.332

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	20.5307	2	<.0001
Score	20.3181	2	<.0001
Wald	19.8591	2	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	17.5769	<.0001
hba1c_eot	1	1.1306	0.2876

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.5758	0.9514	2.7435	0.0977
TRTSEQ_LB DETEMIR	1	-0.4196	0.1001	17.5769	<.0001
hba1c_eot	1	-0.1461	0.1374	1.1306	0.2876

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.432	0.292	0.640
hba1c_eot	0.864	0.660	1.131

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.6	Somers' D	0.231
Percent Discordant	37.5	Gamma	0.235
Percent Tied	1.9	Tau-a	0.107
Pairs	49348	c	0.615

```
Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                                                  Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             156
       2           1             109

Probability modeled is A02event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              NPH          -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	360.988	357.557
SC	364.568	368.297
-2 Log L	358.988	351.557

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.4305	2	0.0243
Score	7.3616	2	0.0252
Wald	7.2658	2	0.0264

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	7.0735	0.0078
hba1c_eot	1	0.3420	0.5587

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.9812	1.0147	0.9351	0.3335
TRTSEQ_LB DETEMIR	1	-0.3404	0.1280	7.0735	0.0078
hba1c_eot	1	0.0747	0.1277	0.3420	0.5587

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.506	0.307	0.836
hba1c_eot	1.078	0.839	1.384

Association of Predicted Probabilities and Observed Responses

Percent Concordant	58.4	Somers' D	0.187
Percent Discordant	39.7	Gamma	0.190
Percent Tied	1.8	Tau-a	0.091
Pairs	17004	c	0.593

```
Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

                                                    Model Information

Data Set                        WORK.ANALYSIS_IQWIG
Response Variable               A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels       2
Model                           binary logit
Optimization Technique          Fisher's scoring


Number of Observations Read         582
Number of Observations Used         543


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             245
       2           1             298

Probability modeled is A02event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	749.576	750.625
SC	753.874	763.516
-2 Log L	747.576	744.625

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.9516	2	0.2286
Score	2.9502	2	0.2288
Wald	2.9343	2	0.2306

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.4151	0.2342
hba1c_eot	1	1.4551	0.2277

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.9623	0.6422	2.2454	0.1340
TRTSEQ_LB DETEMIR	1	-0.1029	0.0865	1.4151	0.2342
hba1c_eot	1	-0.1070	0.0887	1.4551	0.2277

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.814	0.580	1.143
hba1c_eot	0.898	0.755	1.069

Association of Predicted Probabilities and Observed Responses

Percent Concordant	53.4	Somers' D	0.088
Percent Discordant	44.6	Gamma	0.090
Percent Tied	2.0	Tau-a	0.044
Pairs	73010	c	0.544

```
Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                                    Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A02event_any            A02: Minor or symptom, daytime (not
nocturnal), BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A02event_         Total
   Value     any           Frequency

       1           0             235
       2           1             449

Probability modeled is A02event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	882.135	871.118
SC	886.663	884.702
-2 Log L	880.135	865.118

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0170	2	0.0005
Score	15.2271	2	0.0005
Wald	14.7263	2	0.0006

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.1702	0.6800
hba1c_eot	1	14.2104	0.0002

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	2.7574	0.5635	23.9455	<.0001
TRTSEQ_LB DETEMIR	1	-0.0360	0.0872	0.1702	0.6800
hba1c_eot	1	-0.2797	0.0742	14.2104	0.0002

Logistic regression for endpoint A02event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

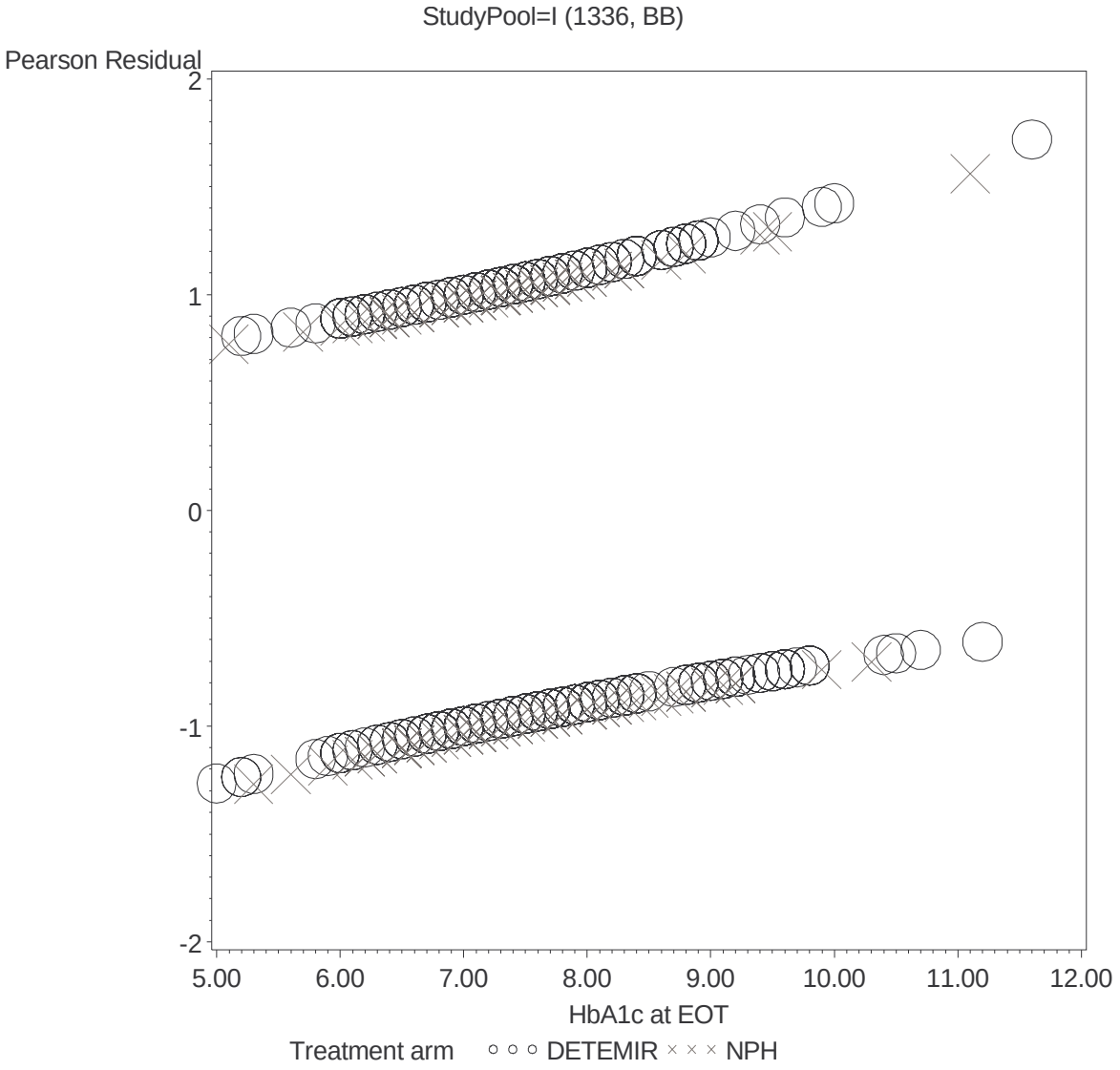
The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.931	0.661	1.310
hba1c_eot	0.756	0.654	0.874

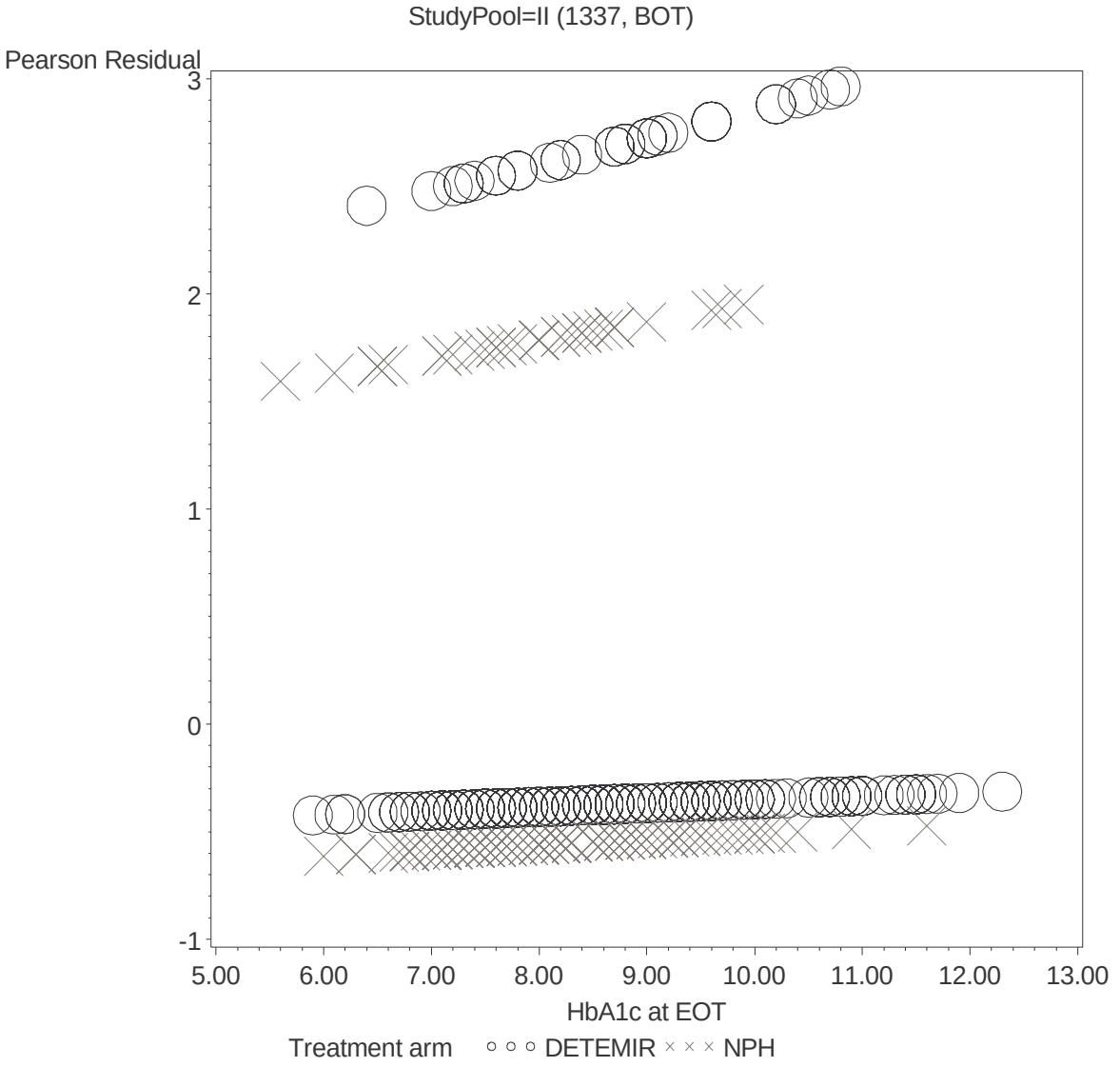
Association of Predicted Probabilities and Observed Responses

Percent Concordant	57.1	Somers' D	0.158
Percent Discordant	41.3	Gamma	0.160
Percent Tied	1.5	Tau-a	0.071
Pairs	105515	c	0.579



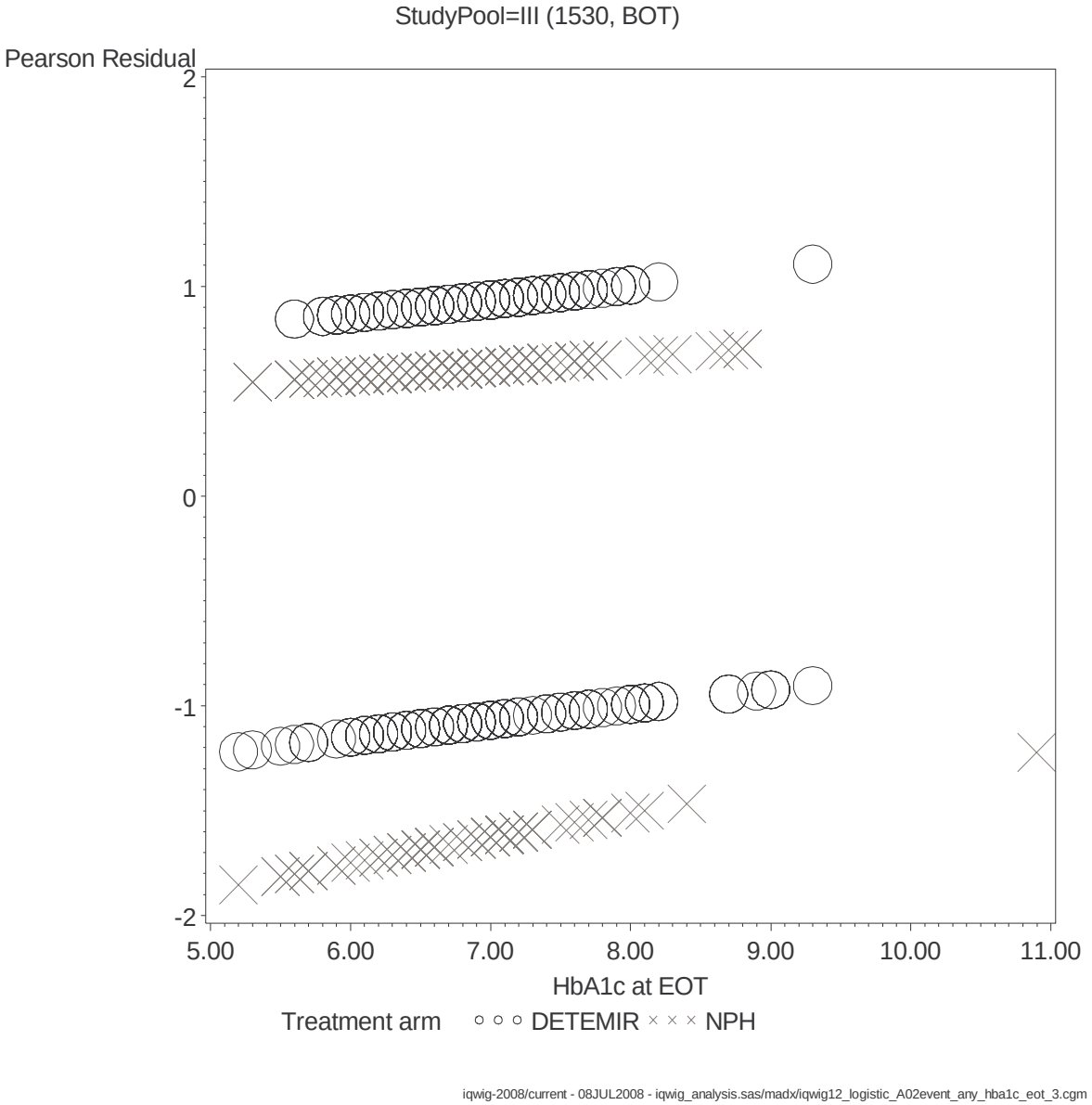
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_1.cgm

23: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot



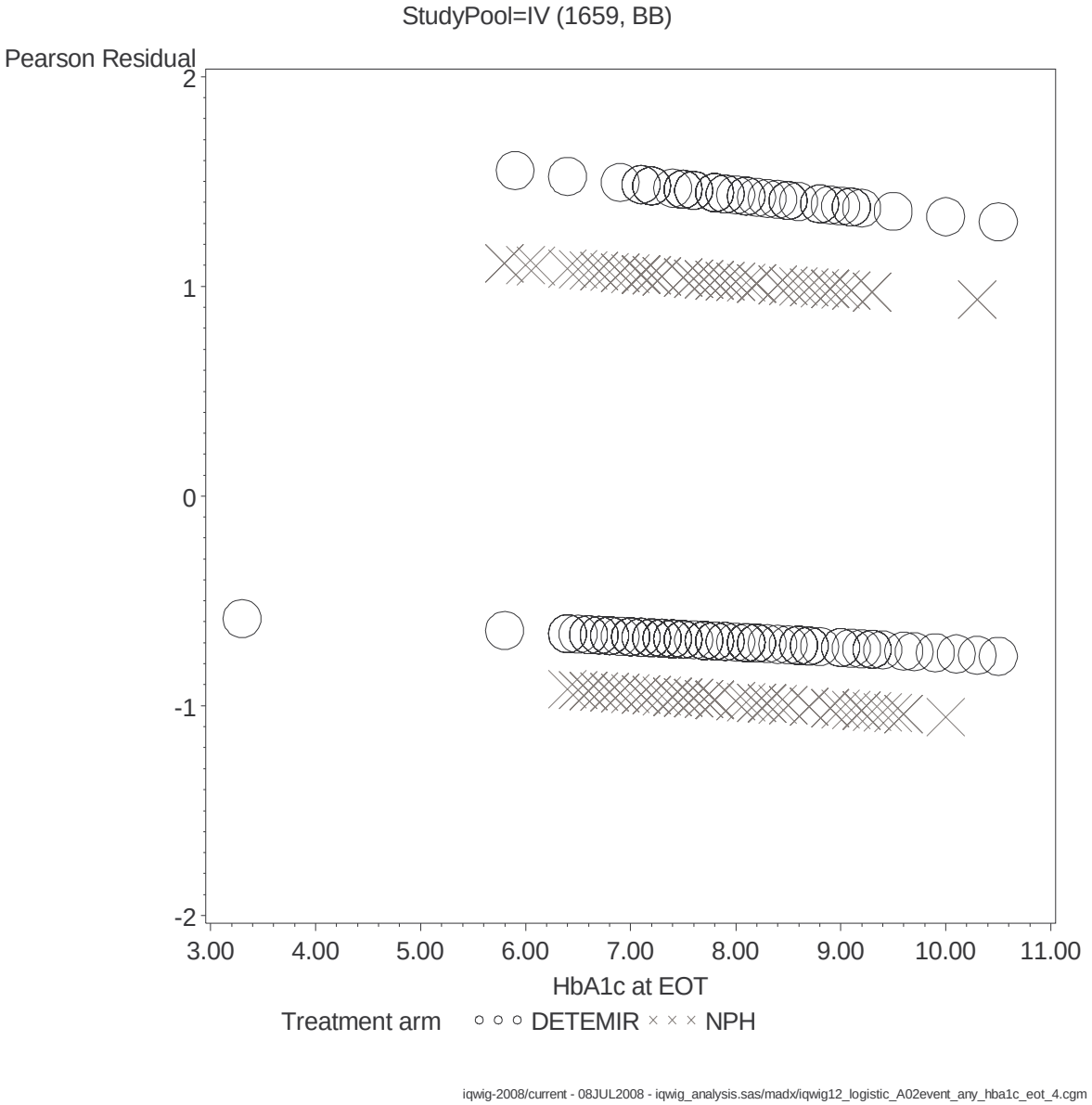
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_2.cgm

24: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot



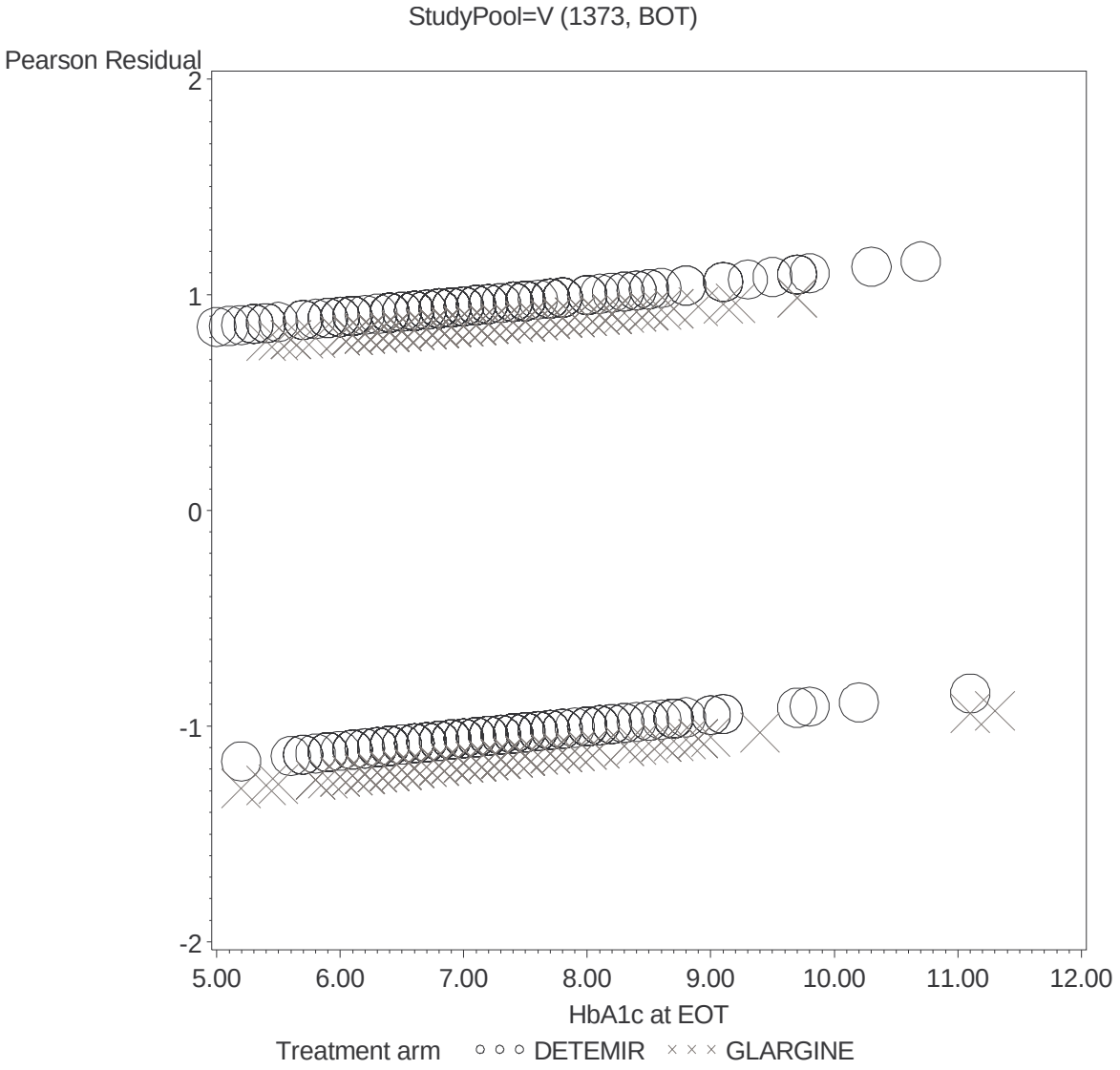
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_3.cgm

25: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot



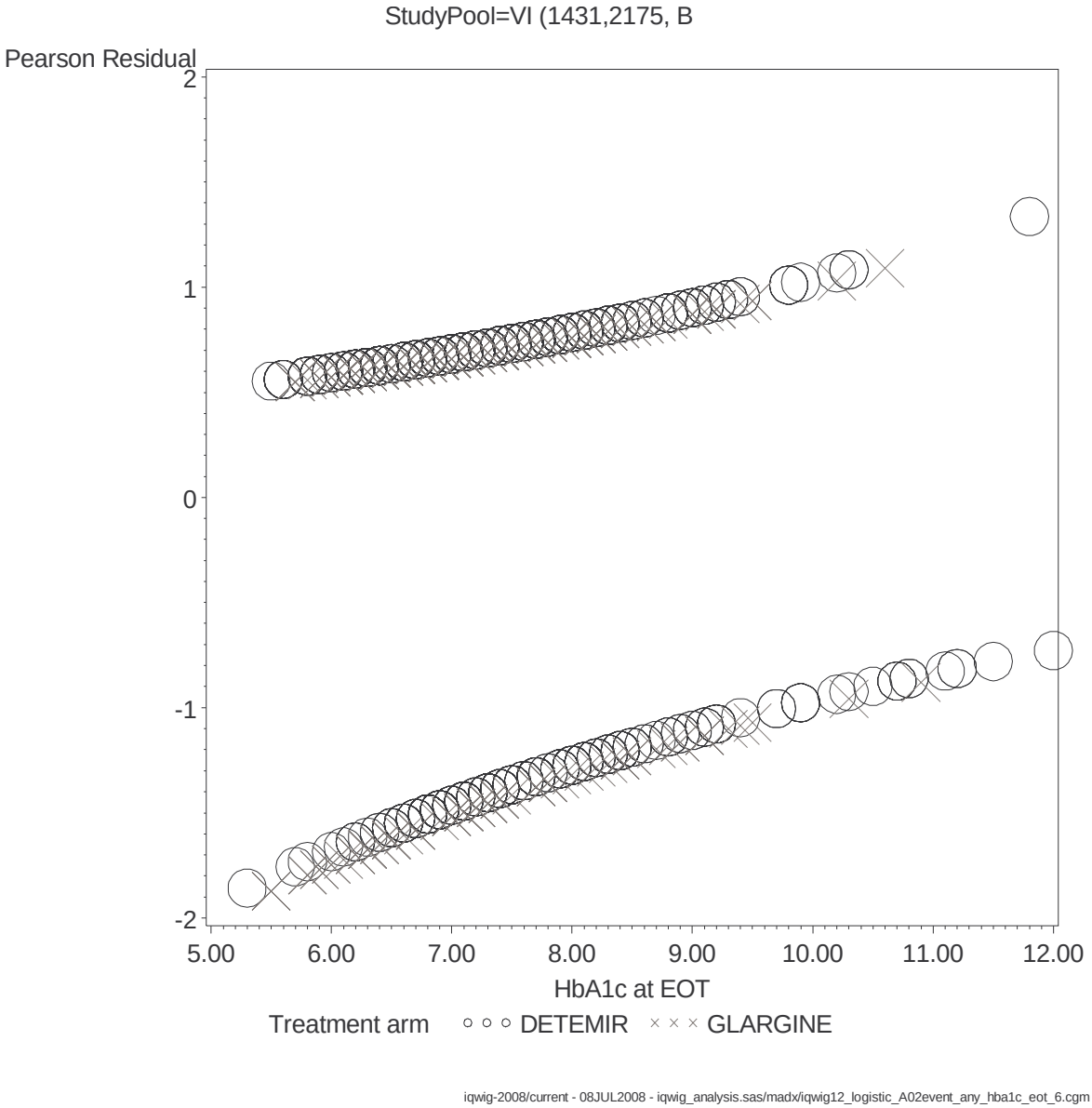
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_4.cgm

26: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_5.cgm

27: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A02event_any_hba1c_eot_6.cgm

28: Logistic regression: residuals for endpoint A02event_any controlling for hba1c_eot

29: Logistic regression for endpoint A03event_any controlling for hba1c_change

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A03event_any           A03: Minor or symptom, nocturnal, BG<2.0
mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A03event_          Total
   Value     any            Frequency

       1     0                    465
       2     1                     12

Probability modeled is A03event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	114.078	109.004
SC	118.246	121.506
-2 Log L	112.078	103.004

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.0746	2	0.0107
Score	9.0976	2	0.0106
Wald	8.1363	2	0.0171

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	5.9834	0.0144
hba1c_change	1	2.8796	0.0897

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.6001	0.3153	130.3899	<.0001
TRTSEQ_LB	DETEMIR	1	-0.7651	0.3128	5.9834	0.0144
hba1c_change		1	0.5883	0.3467	2.8796	0.0897

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.217	0.064	0.738
hba1c_change	1.801	0.913	3.553

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.4	Somers' D	0.477
Percent Discordant	24.7	Gamma	0.492
Percent Tied	3.0	Tau-a	0.023
Pairs	5580	c	0.739

```
Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                          Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A03event_any           A03: Minor or symptom, nocturnal, BG<2.0
mmol/L (36 mg/dl)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A03event_          Total
   Value     any            Frequency

       1             0              413
       2             1                3

Probability modeled is A03event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value         Variables

TRTSEQ_LB     DETEMIR        1
              NPH           -1


                    Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	37.571	34.166
SC	41.601	46.258
-2 Log L	35.571	28.166

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.4052	2	0.0247
Score	6.6890	2	0.0353
Wald	0.7406	2	0.6905

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0055	0.9407
hba1c_change	1	0.7352	0.3912

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-9.4854	82.2795	0.0133	0.9082
TRTSEQ_LB	DETEMIR	1	-6.1234	82.2792	0.0055	0.9407
hba1c_change		1	0.4772	0.5566	0.7352	0.3912

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	<0.001	<0.001	>999.999
hba1c_change		1.612	0.541	4.798

Association of Predicted Probabilities and Observed Responses

Percent Concordant	88.5	Somers' D	0.780
Percent Discordant	10.5	Gamma	0.788
Percent Tied	1.0	Tau-a	0.011
Pairs	1239	c	0.890

```
Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A03event_any            A03: Minor or symptom, nocturnal, BG<2.0
mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A03event_          Total
   Value     any            Frequency

       1             0            454
       2             1              7

Probability modeled is A03event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	74.518	76.996
SC	78.651	89.397
-2 Log L	72.518	70.996

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.5216	2	0.4673
Score	1.4817	2	0.4767
Wald	1.3999	2	0.4966

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.2514	0.2633
hba1c_change	1	0.1926	0.6608

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.9895	0.7476	28.4803	<.0001
TRTSEQ_LB	DETEMIR	1	-0.4715	0.4215	1.2514	0.2633
hba1c_change		1	0.1762	0.4015	0.1926	0.6608

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.389	0.075	2.032
hba1c_change		1.193	0.543	2.620

Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.6	Somers' D	0.263
Percent Discordant	29.2	Gamma	0.311
Percent Tied	15.2	Tau-a	0.008
Pairs	3178	c	0.632

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	251
2	1	14

Probability modeled is A03event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	111.586	113.009
SC	115.166	123.748
-2 Log L	109.586	107.009

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.5769	2	0.2757
Score	2.4622	2	0.2920
Wald	2.3894	2	0.3028

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.8509	0.1737
hba1c_change	1	0.6046	0.4368

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-2.8033	0.3676	58.1467	<.0001
TRTSEQ_LB	DETEMIR	1	-0.4121	0.3029	1.8509	0.1737
hba1c_change		1	0.2106	0.2708	0.6046	0.4368

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.439	0.134	1.438
hba1c_change		1.234	0.726	2.099

Association of Predicted Probabilities and Observed Responses

Percent Concordant	62.6	Somers' D	0.285
Percent Discordant	34.1	Gamma	0.295
Percent Tied	3.3	Tau-a	0.029
Pairs	3514	c	0.643

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read 582
Number of Observations Used 543

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	534
2	1	9

Probability modeled is A03event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	93.648	95.115
SC	97.945	108.006
-2 Log L	91.648	89.115

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.5329	2	0.2818
Score	2.3738	2	0.3052
Wald	2.3956	2	0.3019

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0917	0.7620
hba1c_change	1	2.3297	0.1269

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.9759	0.7499	44.0271	<.0001
TRTSEQ_LB	DETEMIR	1	-0.1027	0.3392	0.0917	0.7620
hba1c_change		1	-0.5152	0.3375	2.3297	0.1269

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.814	0.215	3.077
hba1c_change	0.597	0.308	1.158

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.9	Somers' D	0.283
Percent Discordant	32.6	Gamma	0.303
Percent Tied	6.5	Tau-a	0.009
Pairs	4806	c	0.642

```
Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A03event_any            A03: Minor or symptom, nocturnal, BG<2.0
mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A03event_          Total
   Value     any            Frequency

       1             0            659
       2             1             25

Probability modeled is A03event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	216.529	220.412
SC	221.057	233.996
-2 Log L	214.529	214.412

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.1171	2	0.9432
Score	0.1180	2	0.9427
Wald	0.1179	2	0.9428

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0689	0.7929
hba1c_change	1	0.0350	0.8515

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.2923	0.3011	119.5809	<.0001
TRTSEQ_LB	DETEMIR	1	-0.0562	0.2140	0.0689	0.7929
hba1c_change		1	-0.0338	0.1803	0.0350	0.8515

Logistic regression for endpoint A03event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

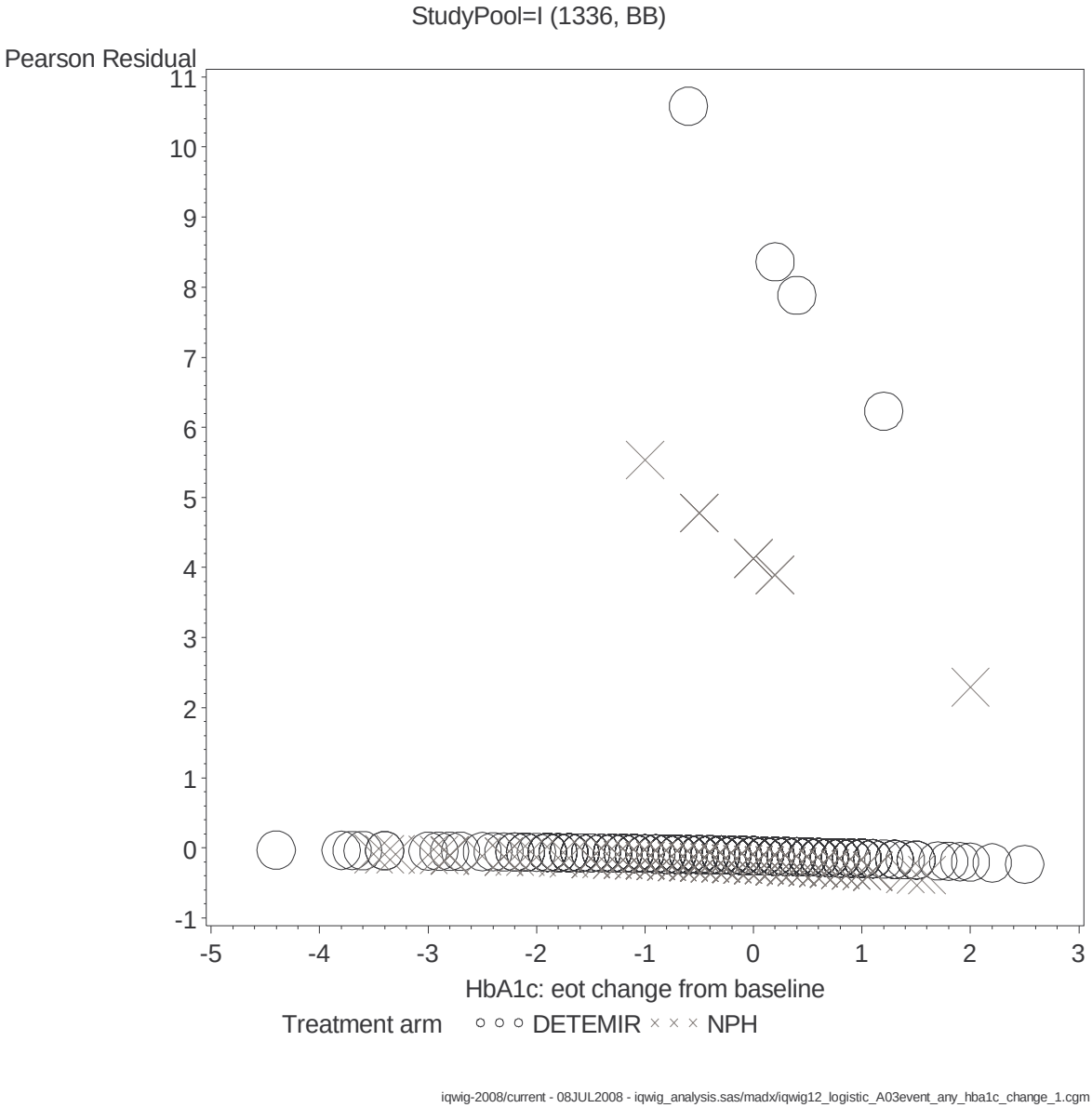
The LOGISTIC Procedure

Odds Ratio Estimates

Effect			Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR	vs GLARGINE	0.894	0.386	2.068
hba1c_change			0.967	0.679	1.377

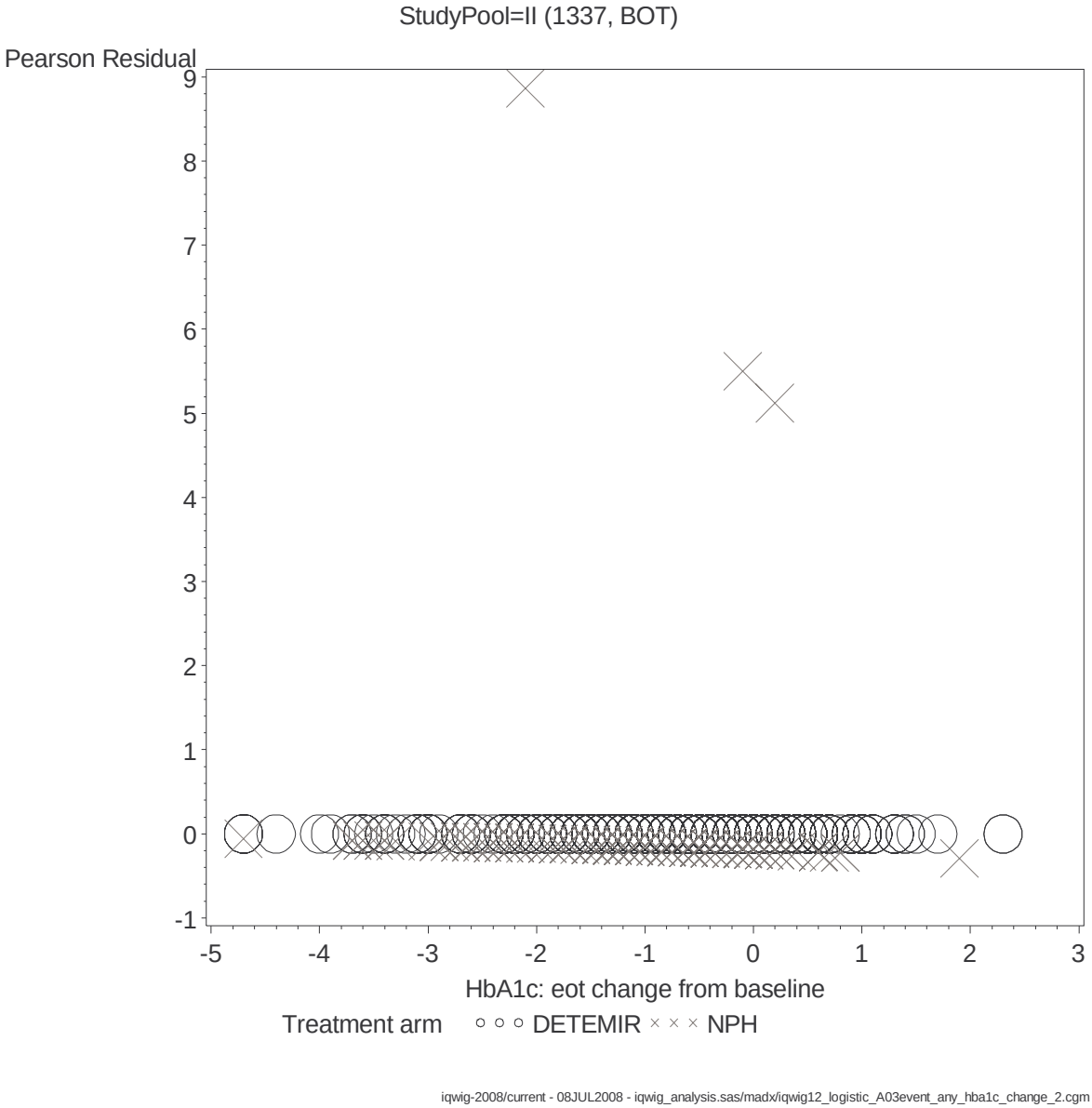
Association of Predicted Probabilities and Observed Responses

Percent Concordant	35.7	Somers' D	-.031
Percent Discordant	38.9	Gamma	-.042
Percent Tied	25.4	Tau-a	-.002
Pairs	16475	c	0.484



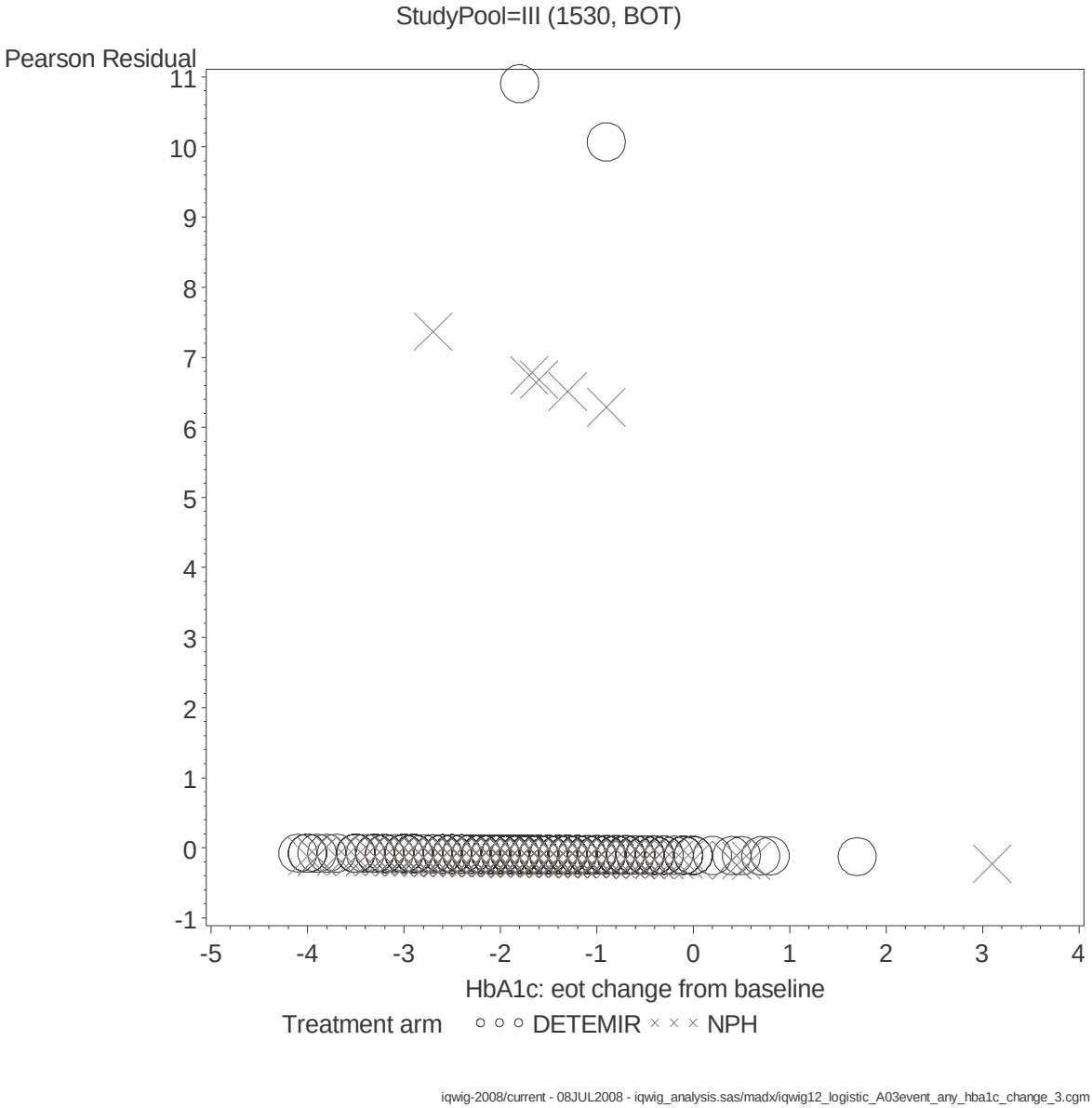
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_change_1.cgm

30: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change

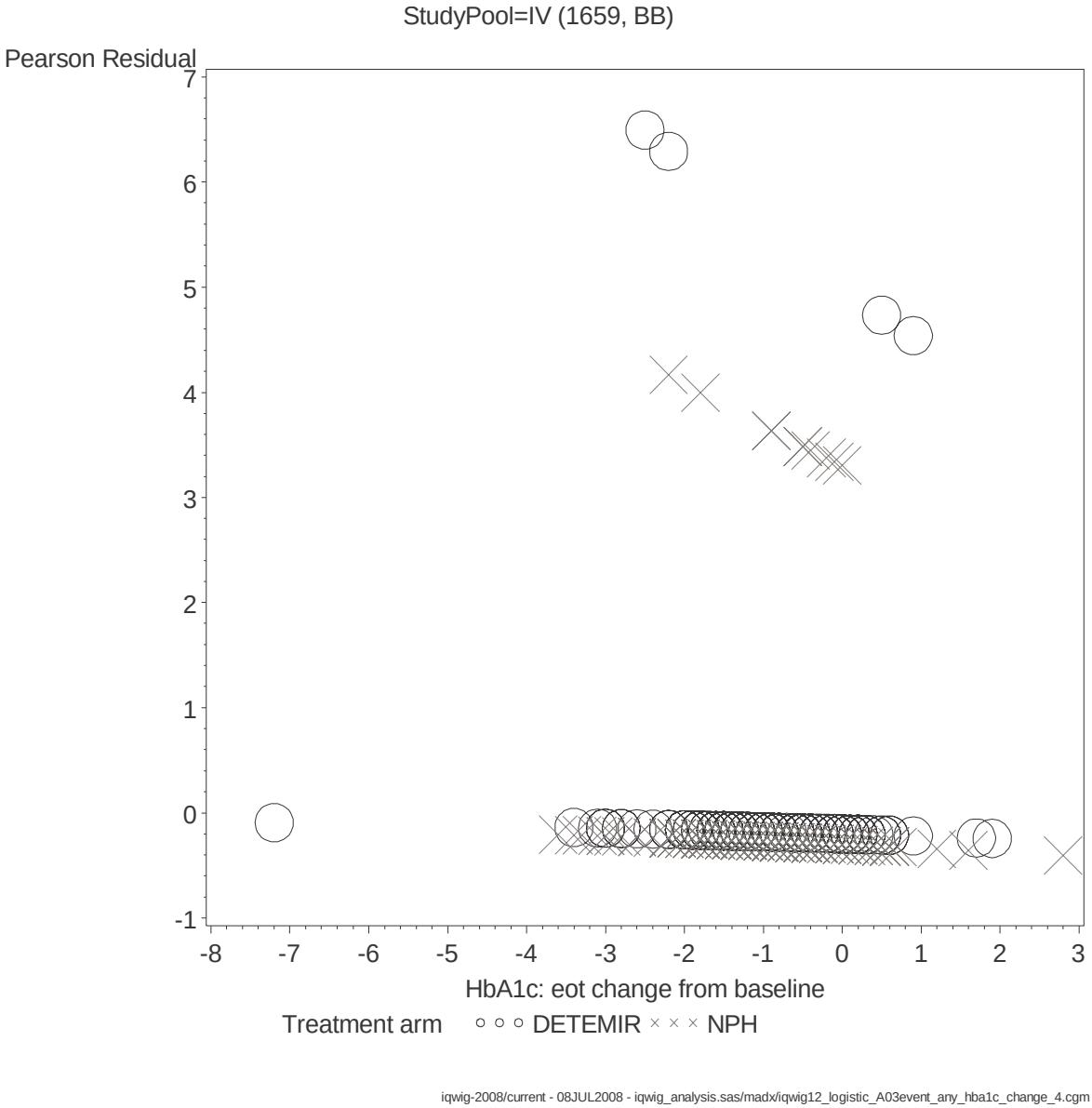


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_change_2.cgm

31: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change

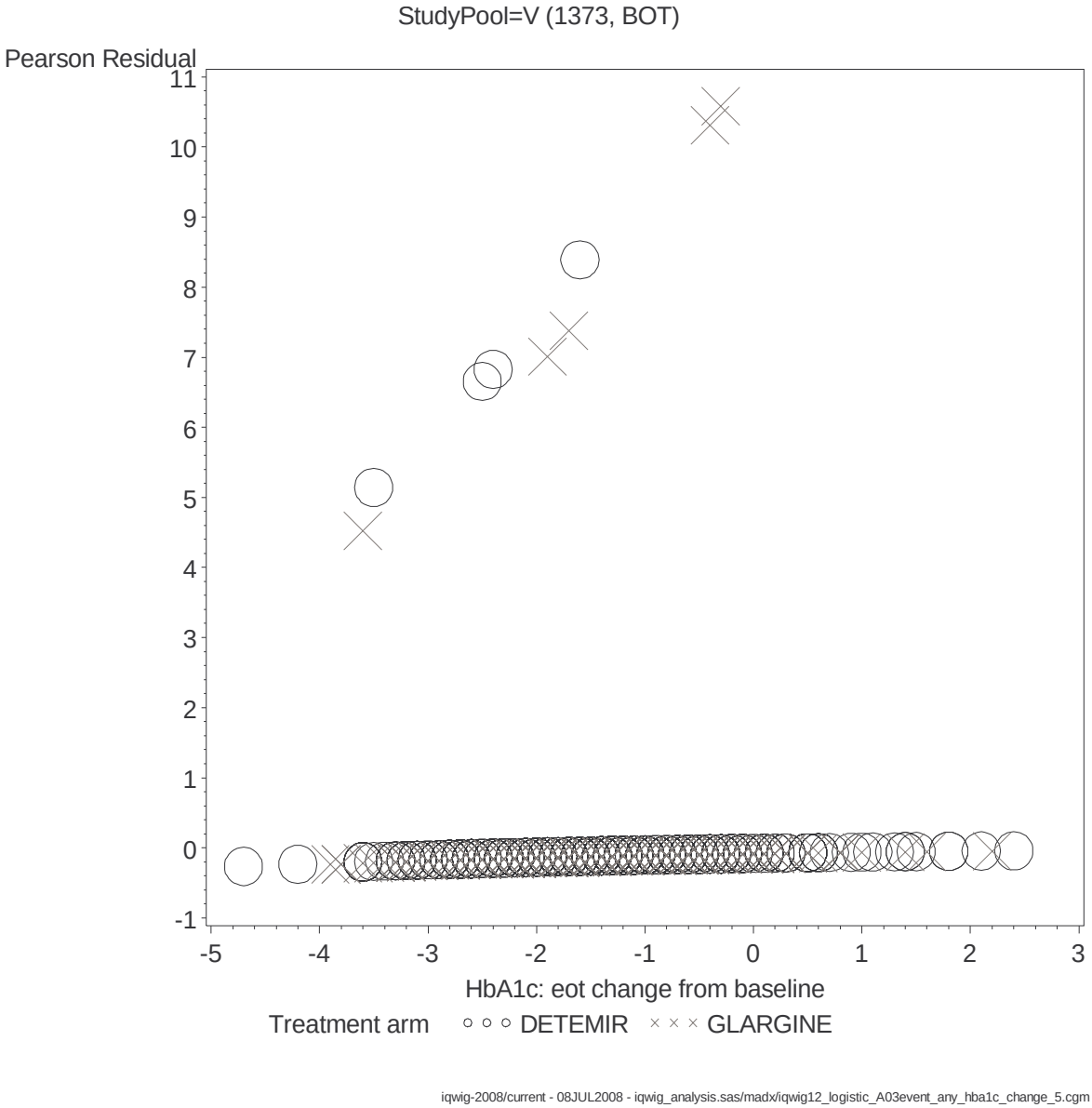


32: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change

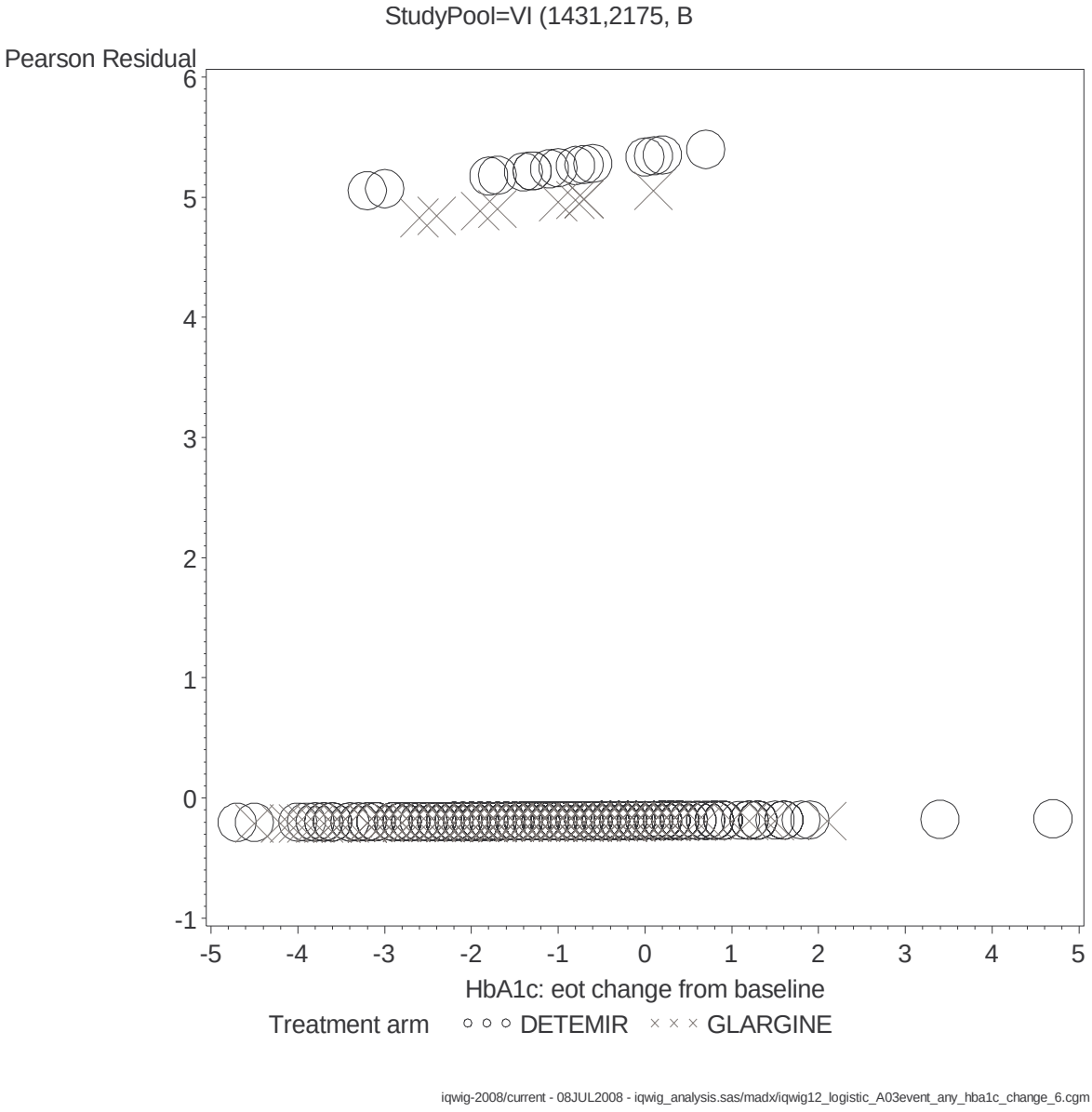


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_change_4.cgm

33: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change



34: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_change_6.cgm

35: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_change

36: Logistic regression for endpoint A03event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A03event_any           A03: Minor or symptom, nocturnal, BG<2.0
mmol/L (36 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A03event_          Total
   Value     any            Frequency

       1     0                    465
       2     1                     12

Probability modeled is A03event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	114.078	112.053
SC	118.246	124.556
-2 Log L	112.078	106.053

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.0248	2	0.0492
Score	6.5054	2	0.0387
Wald	5.5752	2	0.0616

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	5.5576	0.0184
hba1c_eot	1	0.1709	0.6793

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5502	2.2239	4.1864	0.0407
TRTSEQ_LB DETEMIR	1	-0.7367	0.3125	5.5576	0.0184
hba1c_eot	1	0.1197	0.2896	0.1709	0.6793

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.229	0.067	0.780
hba1c_eot	1.127	0.639	1.988

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.8	Somers' D	0.412
Percent Discordant	25.6	Gamma	0.446
Percent Tied	7.7	Tau-a	0.020
Pairs	5580	c	0.706

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read 467
Number of Observations Used 416

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	413
2	1	3

Probability modeled is A03event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration.
Validity of the model fit is questionable.

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	37.571	31.671
SC	41.601	43.763
-2 Log L	35.571	25.671

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	9.8998	2	0.0071
Score	8.8687	2	0.0119
Wald	3.2893	2	0.1931

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0038	0.9507
hba1c_eot	1	3.2860	0.0699

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-18.5404	107.6	0.0297	0.8633
TRTSEQ_LB DETEMIR	1	-6.6457	107.5	0.0038	0.9507
hba1c_eot	1	0.9335	0.5150	3.2860	0.0699

```
Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           2.543       0.927       6.979


Association of Predicted Probabilities and Observed Responses

Percent Concordant     94.5    Somers' D    0.897
Percent Discordant      4.8    Gamma        0.903
Percent Tied            0.6    Tau-a        0.013
Pairs                  1239    c            0.948
```

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read 475
Number of Observations Used 461

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	454
2	1	7

Probability modeled is A03event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	74.518	76.059
SC	78.651	88.459
-2 Log L	72.518	70.059

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.4594	2	0.2924
Score	2.5631	2	0.2776
Wald	2.5349	2	0.2816

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.4369	0.2306
hba1c_eot	1	1.3118	0.2521

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-7.7515	3.1296	6.1346	0.0133
TRTSEQ_LB DETEMIR	1	-0.5071	0.4230	1.4369	0.2306
hba1c_eot	1	0.4968	0.4337	1.3118	0.2521

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.363	0.069	1.904
hba1c_eot	1.643	0.702	3.845

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.4	Somers' D	0.387
Percent Discordant	27.6	Gamma	0.412
Percent Tied	6.0	Tau-a	0.012
Pairs	3178	c	0.694

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	251
2	1	14

Probability modeled is A03event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	111.586	113.597
SC	115.166	124.336
-2 Log L	109.586	107.597

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.9888	2	0.3699
Score	1.9132	2	0.3842
Wald	1.8251	2	0.4015

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.7878	0.1812
hba1c_eot	1	0.0220	0.8821

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.6582	2.2281	1.4234	0.2328
TRTSEQ_LB DETEMIR	1	-0.4047	0.3027	1.7878	0.1812
hba1c_eot	1	-0.0418	0.2819	0.0220	0.8821

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.445	0.136	1.458
hba1c_eot	0.959	0.552	1.666

Association of Predicted Probabilities and Observed Responses

Percent Concordant	52.2	Somers' D	0.182
Percent Discordant	33.9	Gamma	0.212
Percent Tied	13.9	Tau-a	0.018
Pairs	3514	c	0.591

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	582
Number of Observations Used	543

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	534
2	1	9

Probability modeled is A03event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	93.648	91.648
SC	97.945	104.540
-2 Log L	91.648	85.648

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.9995	2	0.0498
Score	4.7487	2	0.0931
Wald	4.9771	2	0.0830

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.1419	0.7064
hba1c_eot	1	4.9487	0.0261

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	3.0502	3.0819	0.9796	0.3223
TRTSEQ_LB DETEMIR	1	-0.1288	0.3419	0.1419	0.7064
hba1c_eot	1	-1.0527	0.4732	4.9487	0.0261

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.773	0.202	2.952
hba1c_eot	0.349	0.138	0.882

Association of Predicted Probabilities and Observed Responses

Percent Concordant	70.6	Somers' D	0.451
Percent Discordant	25.5	Gamma	0.469
Percent Tied	3.9	Tau-a	0.015
Pairs	4806	c	0.726

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A03event_any A03: Minor or symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read 704
Number of Observations Used 684

Response Profile

Ordered Value	A03event_any	Total Frequency
1	0	659
2	1	25

Probability modeled is A03event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	216.529	220.401
SC	221.057	233.985
-2 Log L	214.529	214.401

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.1278	2	0.9381
Score	0.1280	2	0.9380
Wald	0.1279	2	0.9381

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0718	0.7887
hba1c_eot	1	0.0451	0.8318

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.9523	1.4275	4.2771	0.0386
TRTSEQ_LB DETEMIR	1	-0.0571	0.2132	0.0718	0.7887
hba1c_eot	1	-0.0406	0.1913	0.0451	0.8318

Logistic regression for endpoint A03event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

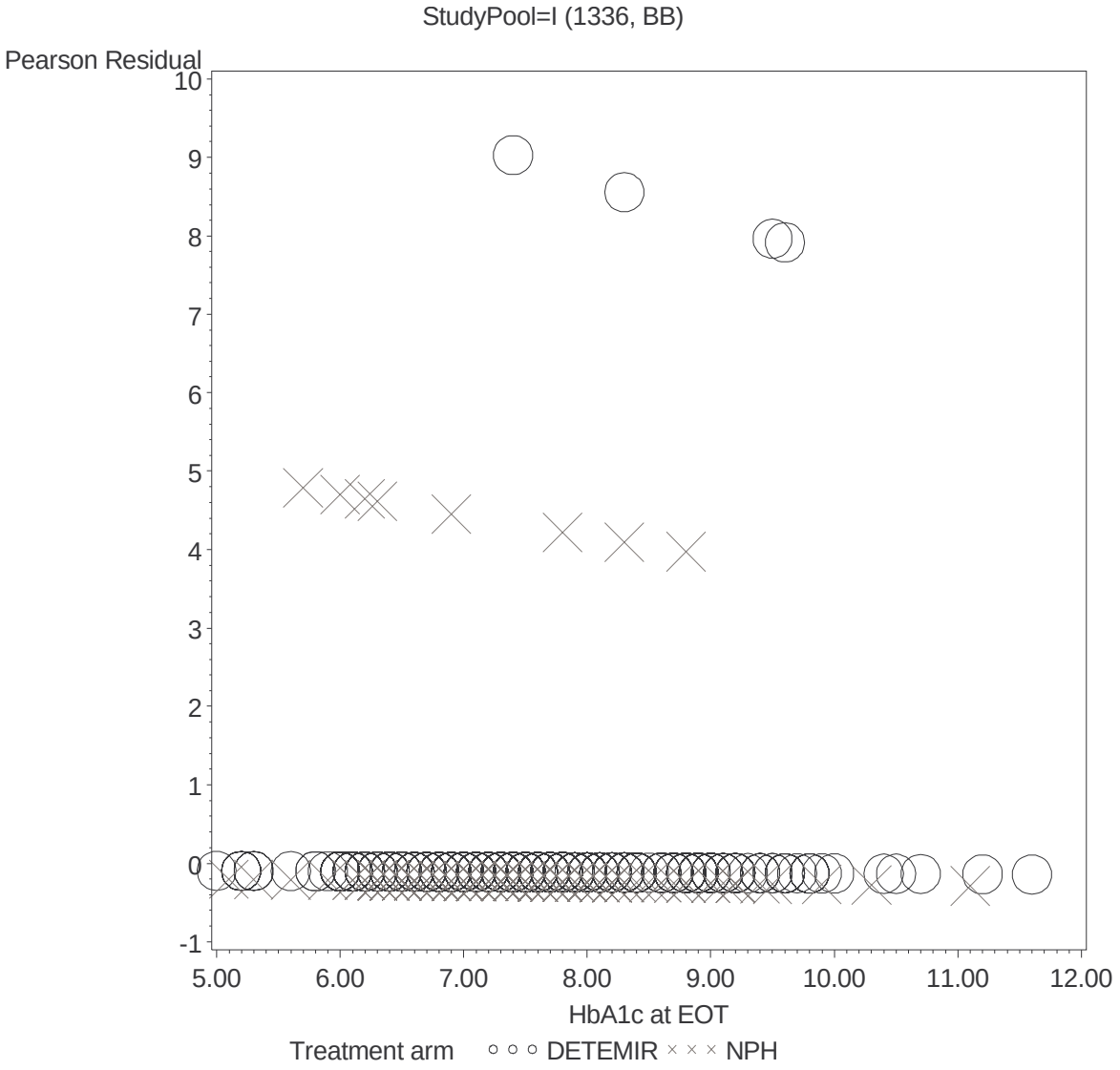
The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.892	0.387	2.058
hba1c_eot	0.960	0.660	1.397

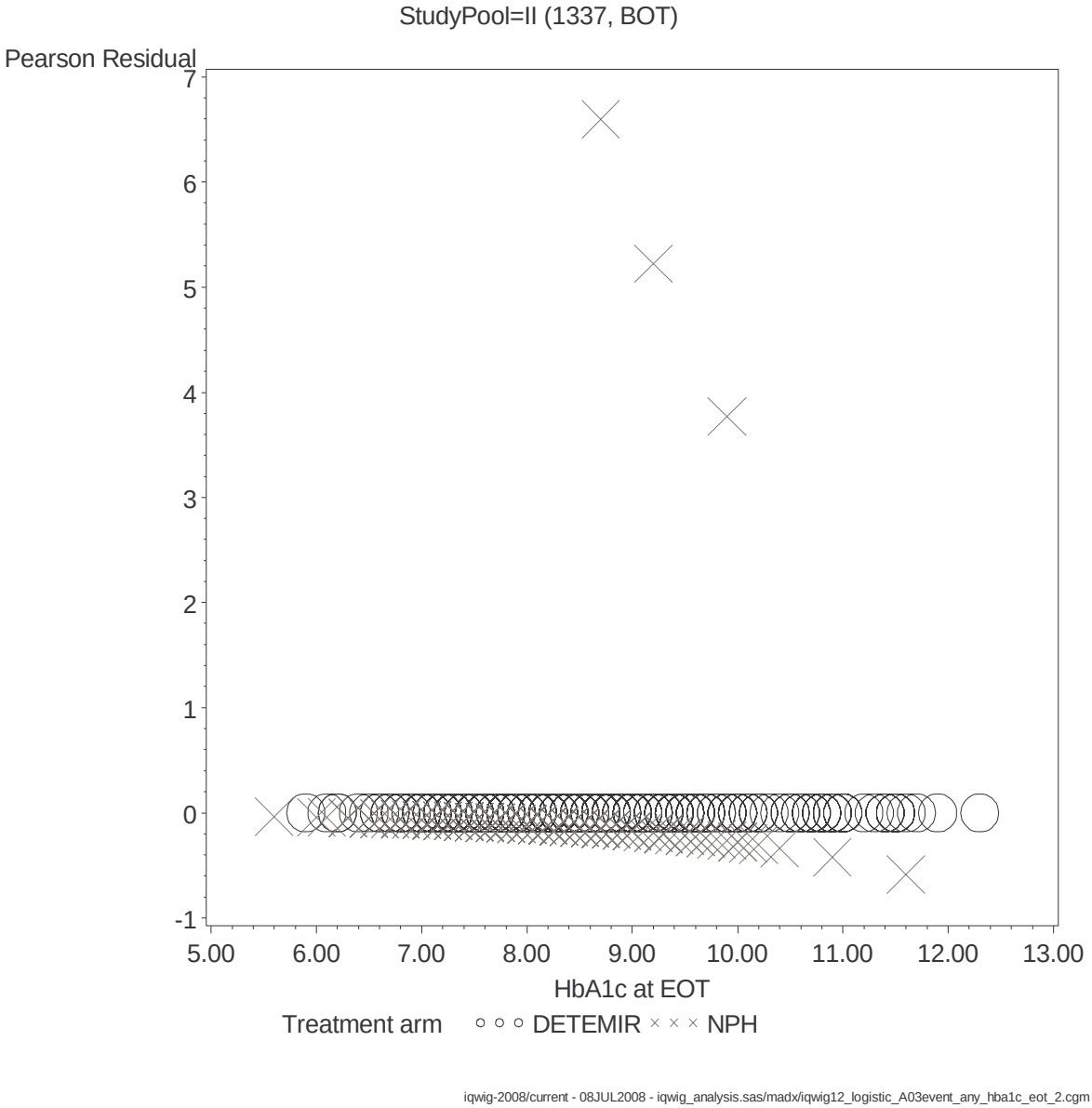
Association of Predicted Probabilities and Observed Responses

Percent Concordant	39.4	Somers' D	0.039
Percent Discordant	35.5	Gamma	0.052
Percent Tied	25.1	Tau-a	0.003
Pairs	16475	c	0.520



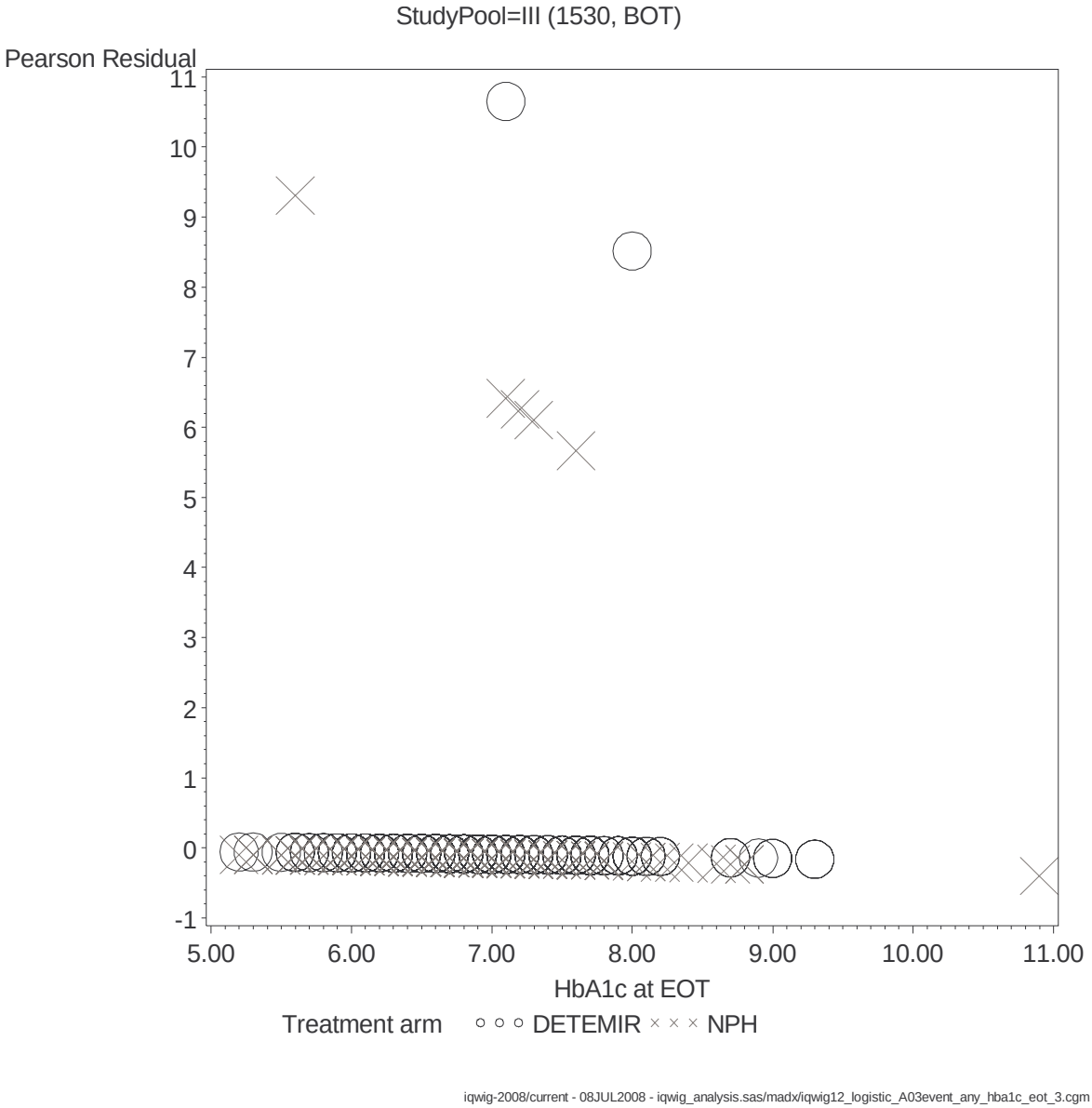
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_eot_1.cgm

37: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot

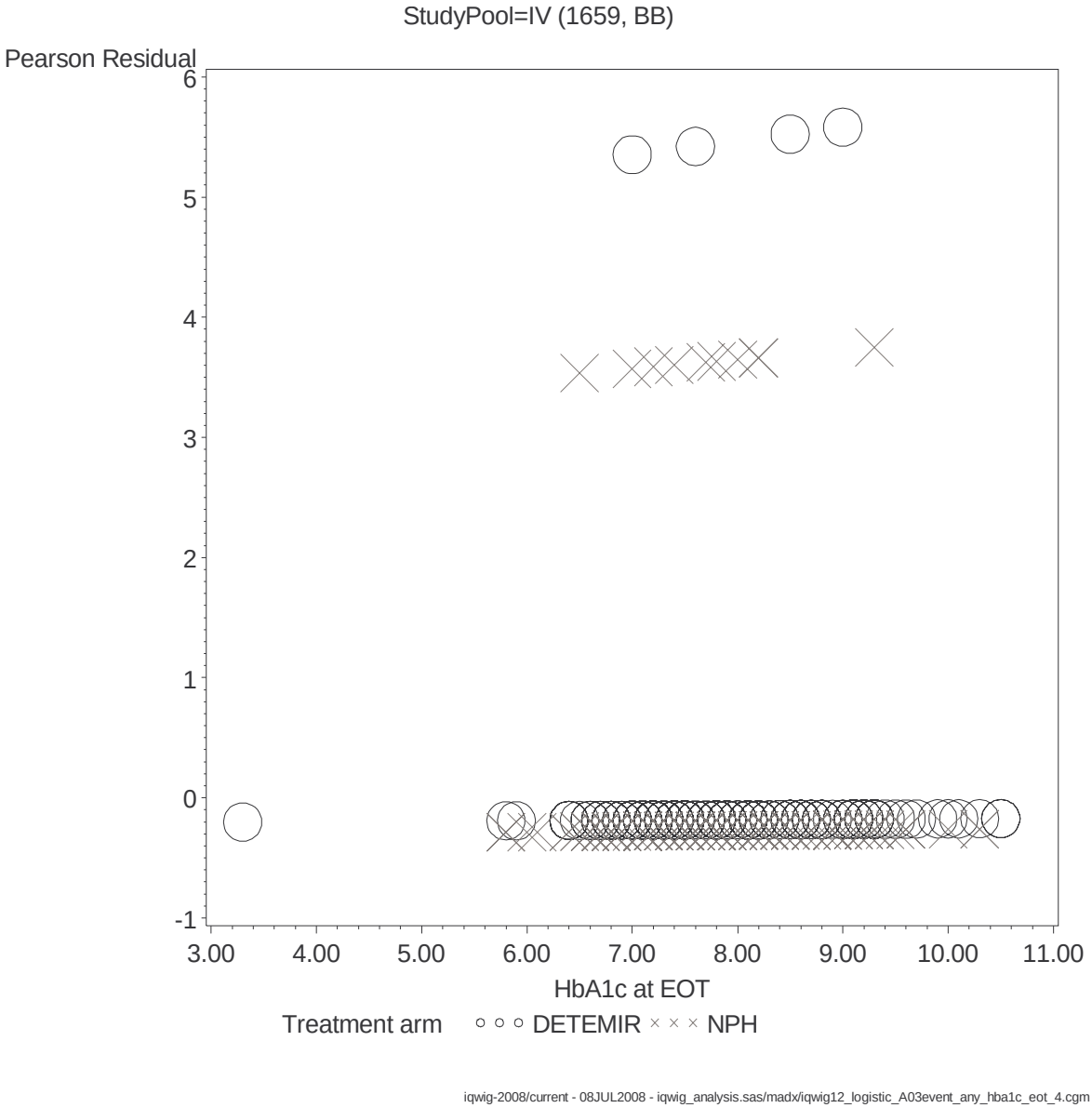


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_eot_2.cgm

38: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot

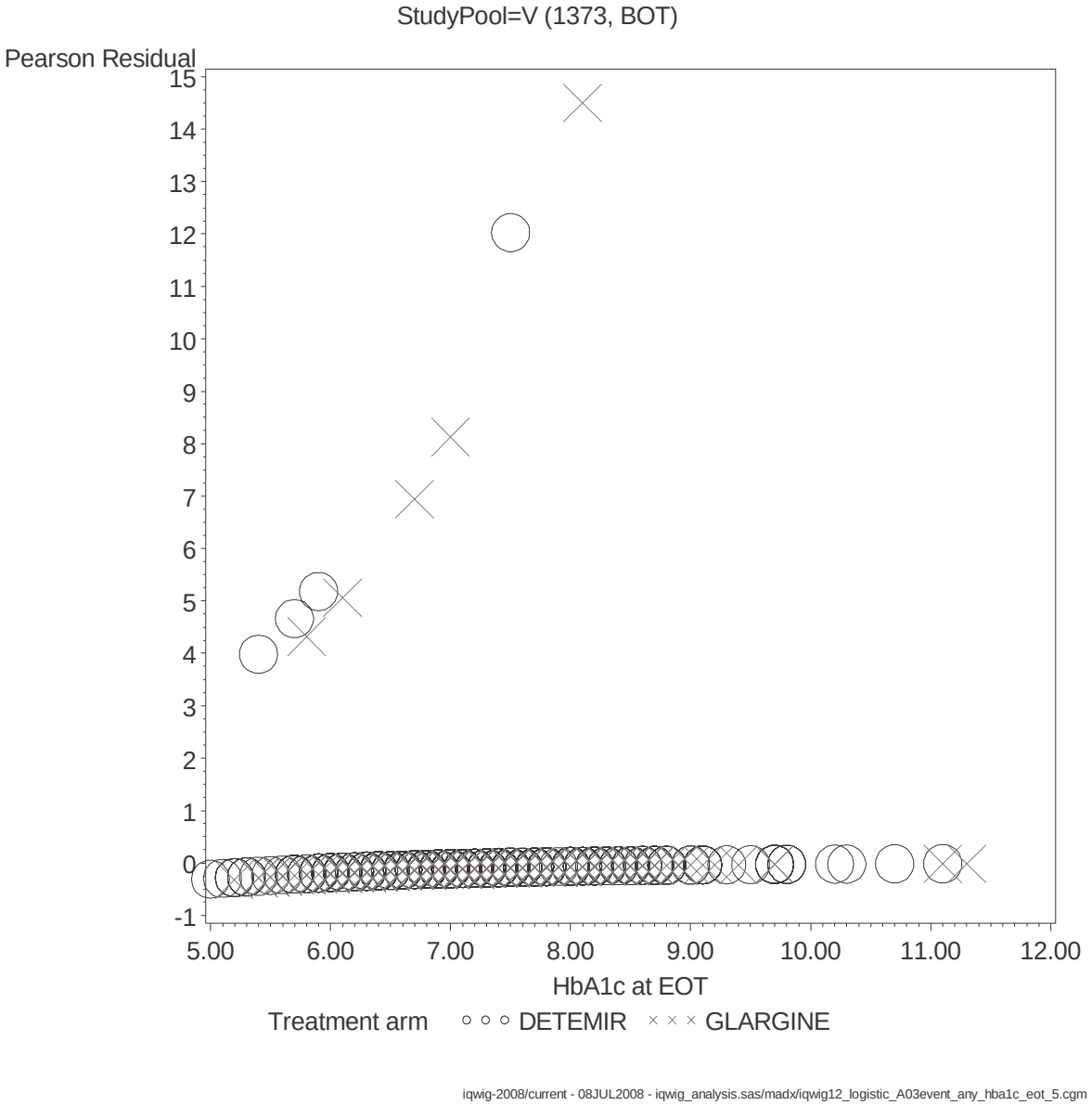


39: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot

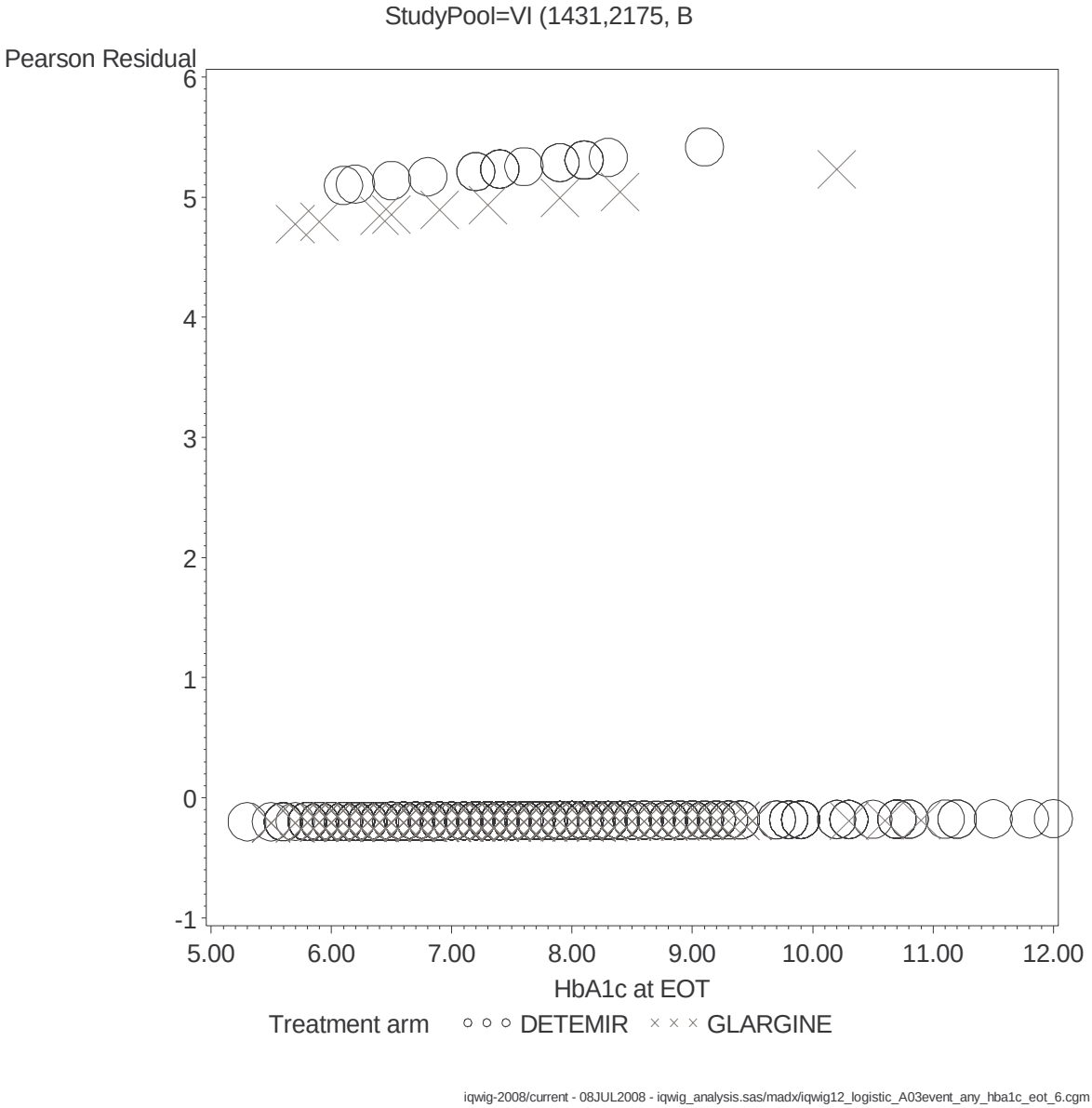


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A03event_any_hba1c_eot_4.cgm

40: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot



41: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot



42: Logistic regression: residuals for endpoint A03event_any controlling for hba1c_eot

43: Logistic regression for endpoint A04event_any controlling for hba1c_change

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any           A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1     0                    382
       2     1                     95

Probability modeled is A04event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	478.273	476.808
SC	482.440	489.311
-2 Log L	476.273	470.808

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	5.4645	2	0.0651
Score	5.6397	2	0.0596
Wald	5.5632	2	0.0619

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	5.5560	0.0184
hba1c_change	1	0.0022	0.9626

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.3155	0.1233	113.8516	<.0001
TRTSEQ_LB	DETEMIR	1	-0.2776	0.1178	5.5560	0.0184
hba1c_change		1	0.00528	0.1124	0.0022	0.9626

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.574	0.362	0.911
hba1c_change	1.005	0.806	1.253

Association of Predicted Probabilities and Observed Responses

Percent Concordant	38.5	Somers' D	0.126
Percent Discordant	25.9	Gamma	0.195
Percent Tied	35.6	Tau-a	0.040
Pairs	36290	c	0.563

```
Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any            A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1             0            371
       2             1             45

Probability modeled is A04event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	287.109	278.324
SC	291.139	290.416
-2 Log L	285.109	272.324

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	12.7848	2	0.0017
Score	13.2392	2	0.0013
Wald	12.4505	2	0.0020

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	11.9305	0.0006
hba1c_change	1	2.2195	0.1363

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.8444	0.1965	88.1458	<.0001
TRTSEQ_LB	DETEMIR	1	-0.5701	0.1650	11.9305	0.0006
hba1c_change		1	0.2125	0.1426	2.2195	0.1363

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.320	0.167	0.611
hba1c_change		1.237	0.935	1.636

Association of Predicted Probabilities and Observed Responses

Percent Concordant	63.1	Somers' D	0.277
Percent Discordant	35.4	Gamma	0.281
Percent Tied	1.5	Tau-a	0.054
Pairs	16695	c	0.638

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A04event_any A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read 475
Number of Observations Used 461

Response Profile

Ordered Value	A04event_any	Total Frequency
1	0	295
2	1	166

Probability modeled is A04event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	604.498	591.889
SC	608.631	604.289
-2 Log L	602.498	585.889

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	16.6088	2	0.0002
Score	16.4683	2	0.0003
Wald	16.1823	2	0.0003

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	15.8476	<.0001
hba1c_change	1	0.1510	0.6975

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.6693	0.2087	10.2847	0.0013
TRTSEQ_LB	DETEMIR	1	-0.3965	0.0996	15.8476	<.0001
hba1c_change		1	-0.0421	0.1082	0.1510	0.6975

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.453	0.306	0.669
hba1c_change		0.959	0.776	1.185

Association of Predicted Probabilities and Observed Responses

Percent Concordant	58.9	Somers' D	0.211
Percent Discordant	37.8	Gamma	0.218
Percent Tied	3.3	Tau-a	0.097
Pairs	48970	c	0.605

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A04event_any A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A04event_any	Total Frequency
1	0	202
2	1	63

Probability modeled is A04event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	292.682	288.106
SC	296.261	298.845
-2 Log L	290.682	282.106

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	8.5754	2	0.0137
Score	8.3722	2	0.0152
Wald	8.1158	2	0.0173

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	6.9408	0.0084
hba1c_change	1	1.2777	0.2583

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.3944	0.2116	43.4341	<.0001
TRTSEQ_LB	DETEMIR	1	-0.4029	0.1529	6.9408	0.0084
hba1c_change		1	-0.1546	0.1368	1.2777	0.2583

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.447	0.245	0.814
hba1c_change		0.857	0.655	1.120

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.4	Somers' D	0.223
Percent Discordant	38.1	Gamma	0.226
Percent Tied	1.5	Tau-a	0.081
Pairs	12726	c	0.612

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A04event_any A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	582
Number of Observations Used	543

Response Profile

Ordered Value	A04event_any	Total Frequency
1	0	375
2	1	168

Probability modeled is A04event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	673.814	672.837
SC	678.111	685.729
-2 Log L	671.814	666.837

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.9771	2	0.0830
Score	4.9086	2	0.0859
Wald	4.8597	2	0.0881

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0583	0.8093
hba1c_change	1	4.8163	0.0282

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-1.0911	0.1642	44.1329	<.0001
TRTSEQ_LB	DETEMIR	1	0.0225	0.0933	0.0583	0.8093
hba1c_change		1	-0.1937	0.0882	4.8163	0.0282

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.046	0.726	1.508
hba1c_change	0.824	0.693	0.980

Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.0	Somers' D	0.116
Percent Discordant	43.4	Gamma	0.118
Percent Tied	1.6	Tau-a	0.050
Pairs	63000	c	0.558

```
Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any            A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1     0                    388
       2     1                    296

Probability modeled is A04event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	937.813	936.960
SC	942.341	950.544
-2 Log L	935.813	930.960

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.8539	2	0.0883
Score	4.8304	2	0.0893
Wald	4.7858	2	0.0914

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0130	0.9091
hba1c_change	1	4.6445	0.0312

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.4327	0.1151	14.1256	0.0002
TRTSEQ_LB	DETEMIR	1	-0.00944	0.0826	0.0130	0.9091
hba1c_change		1	-0.1492	0.0692	4.6445	0.0312

```
Logistic regression for endpoint A04event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

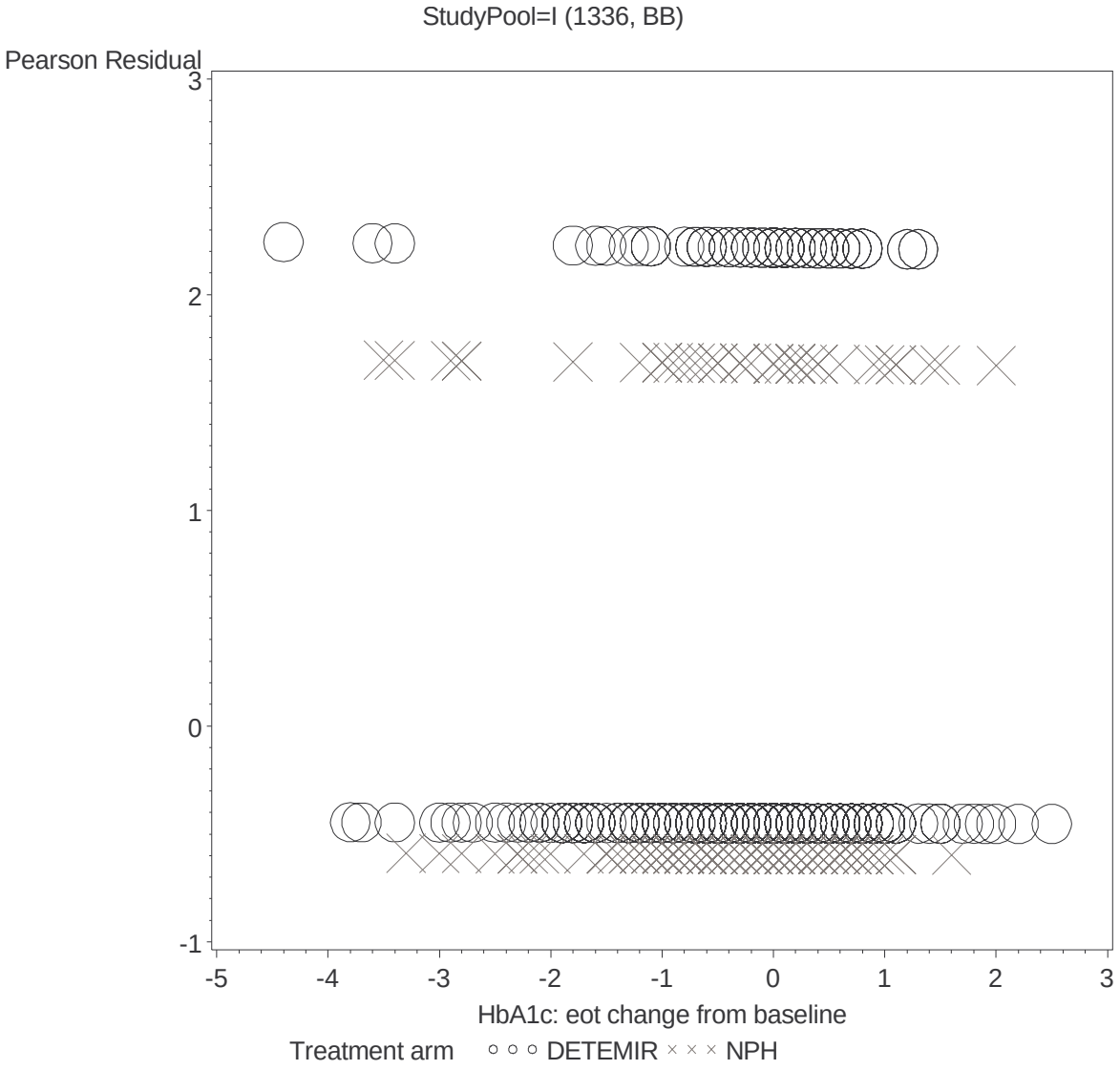
                      Odds Ratio Estimates

                                      Point          95% Wald
Effect                              Estimate      Confidence Limits

TRTSEQ_LB    DETEMIR  vs GLARGINE       0.981       0.710       1.357
hba1c_change                           0.861       0.752       0.987

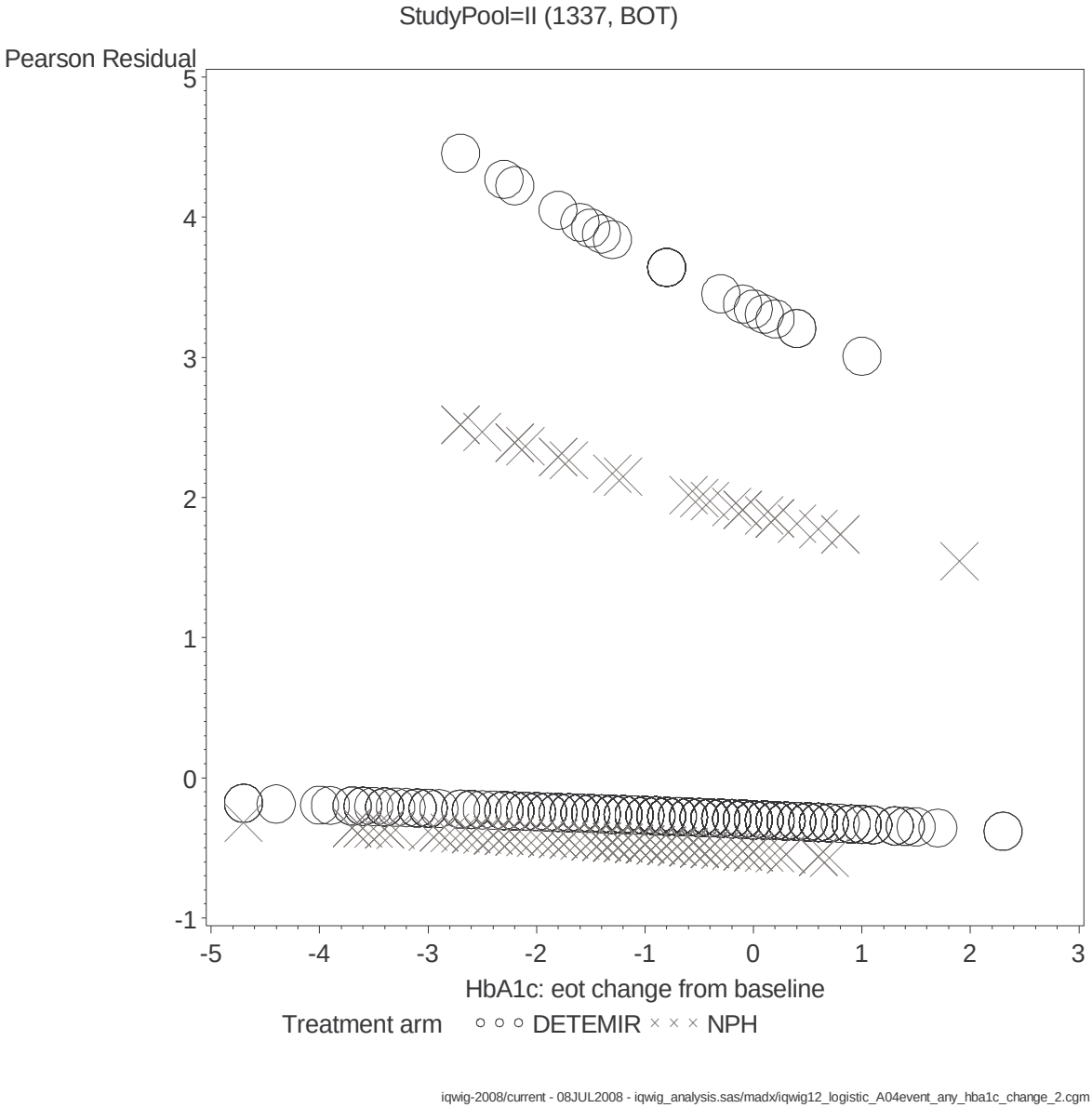

Association of Predicted Probabilities and Observed Responses

Percent Concordant      53.5    Somers' D    0.090
Percent Discordant      44.4    Gamma        0.092
Percent Tied             2.1    Tau-a        0.044
Pairs                 114848    c            0.545
```

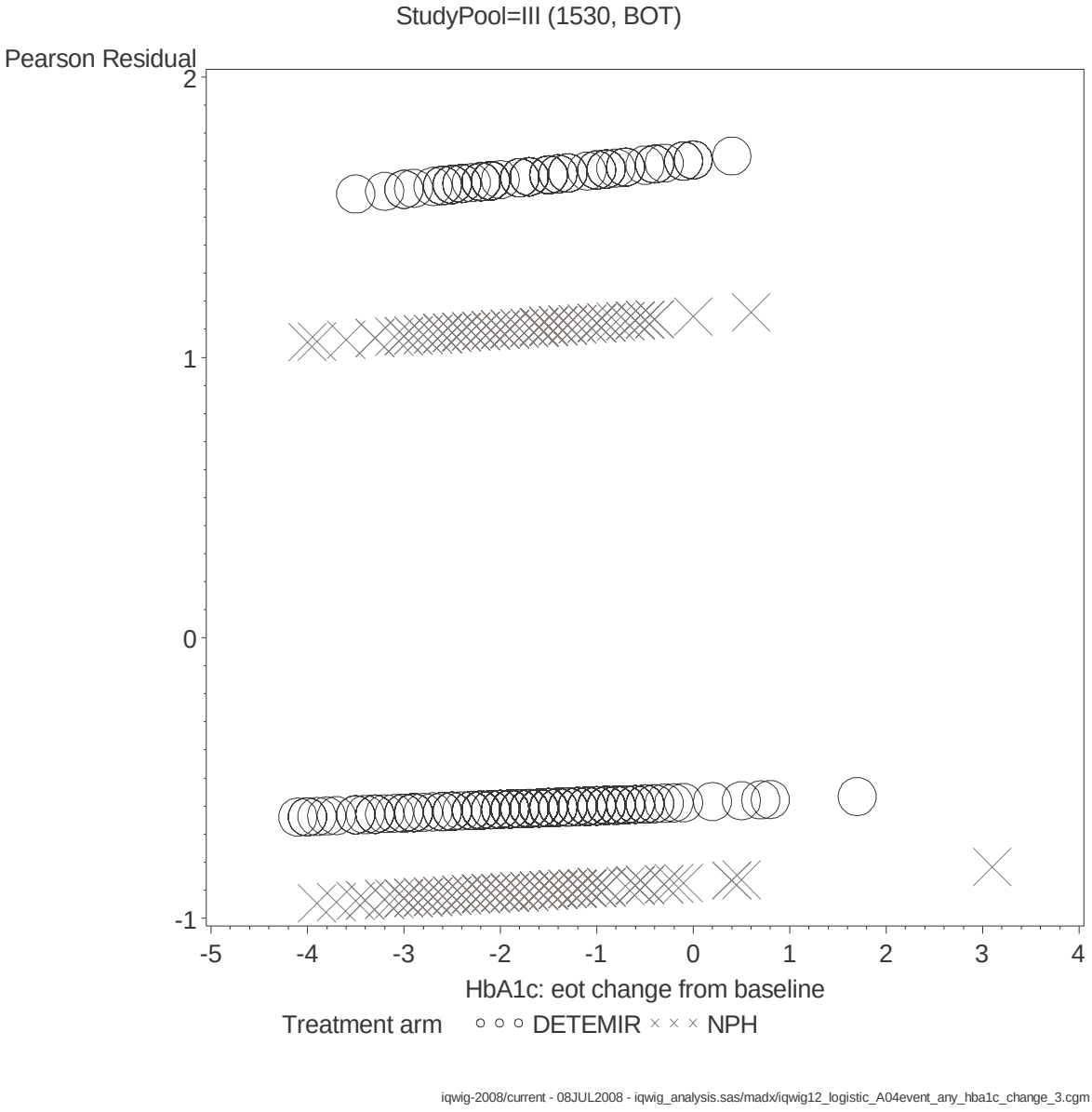


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_change_1.cgm

44: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change

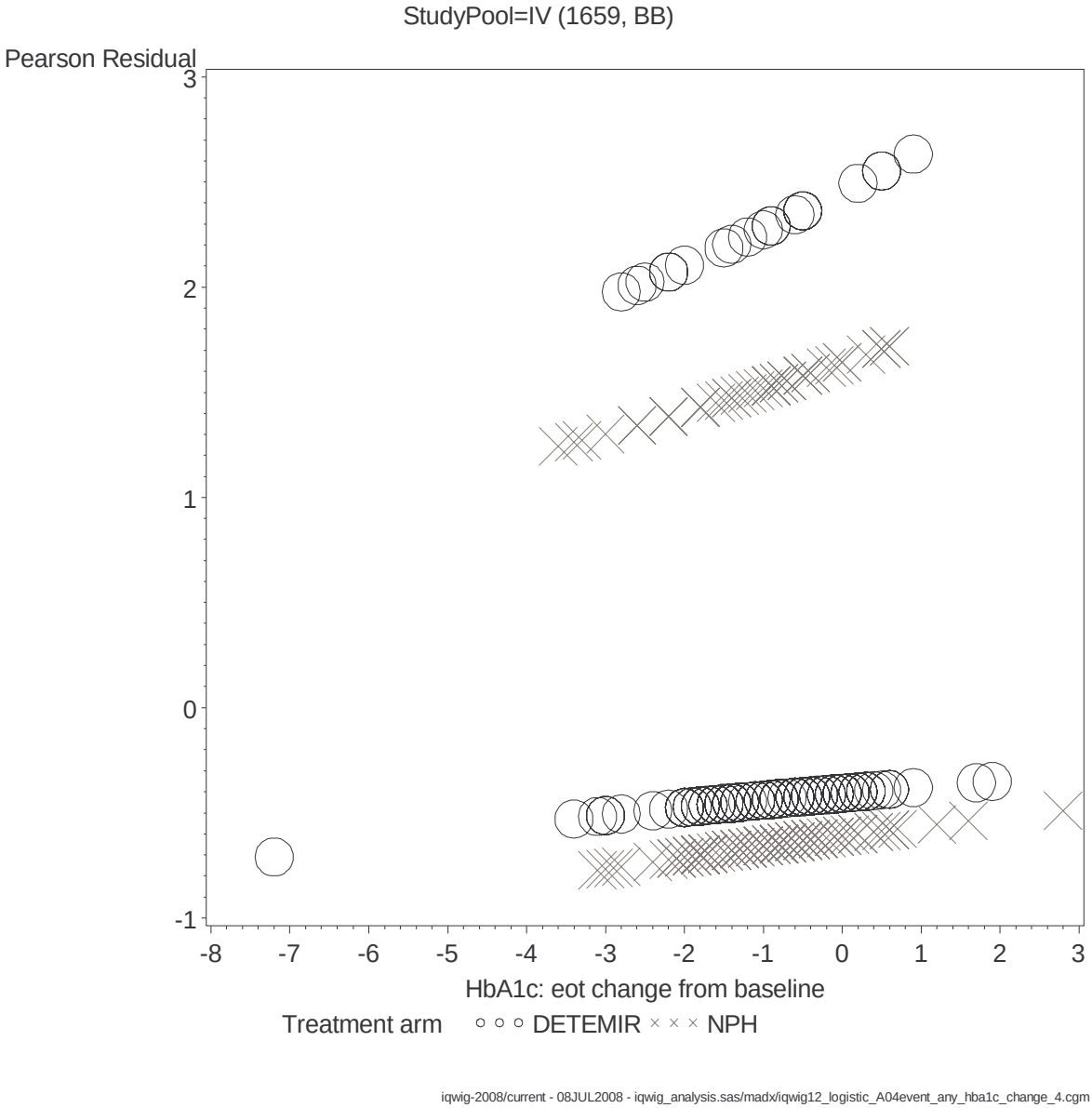


45: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change

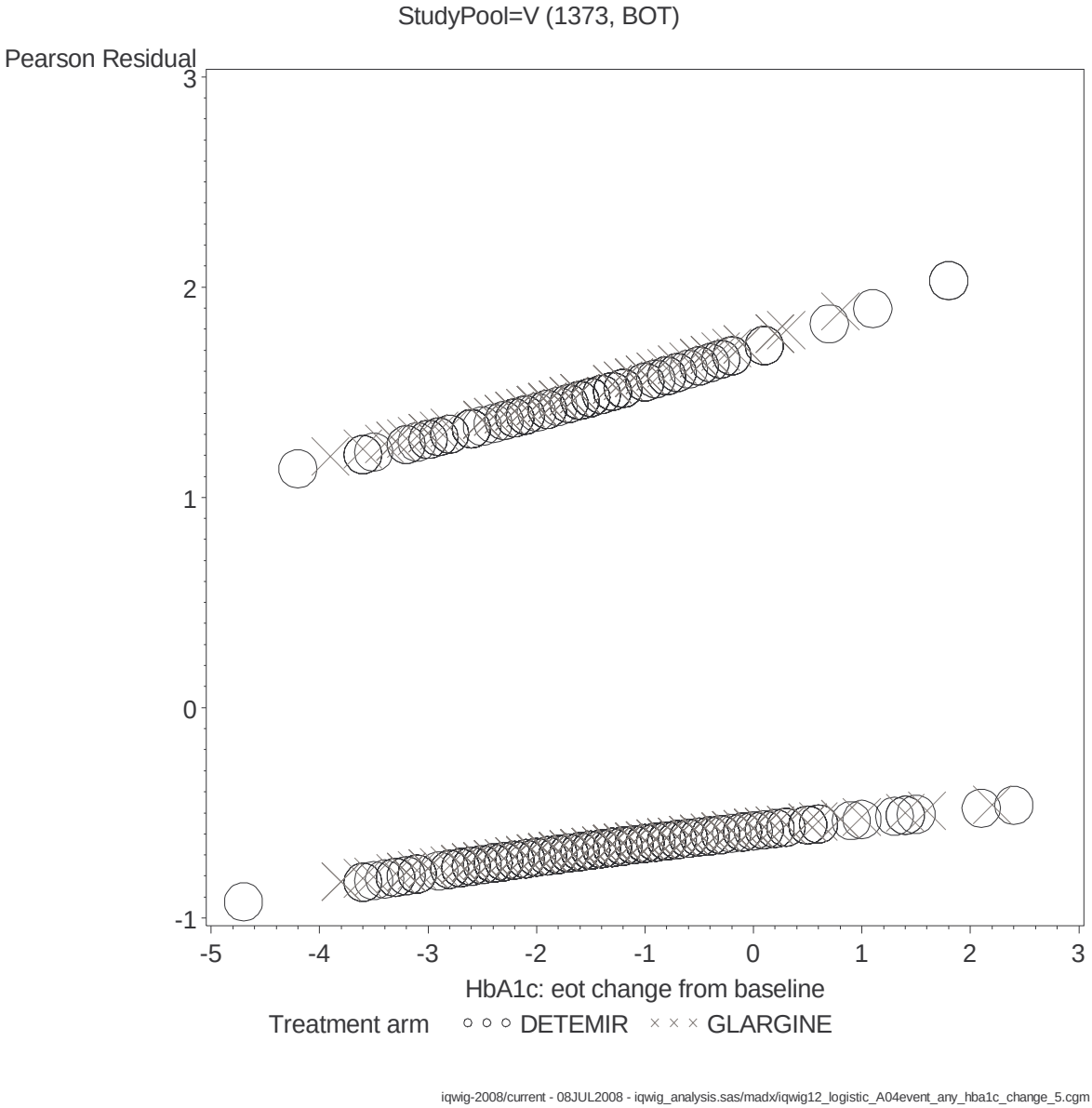


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_change_3.cgm

46: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change

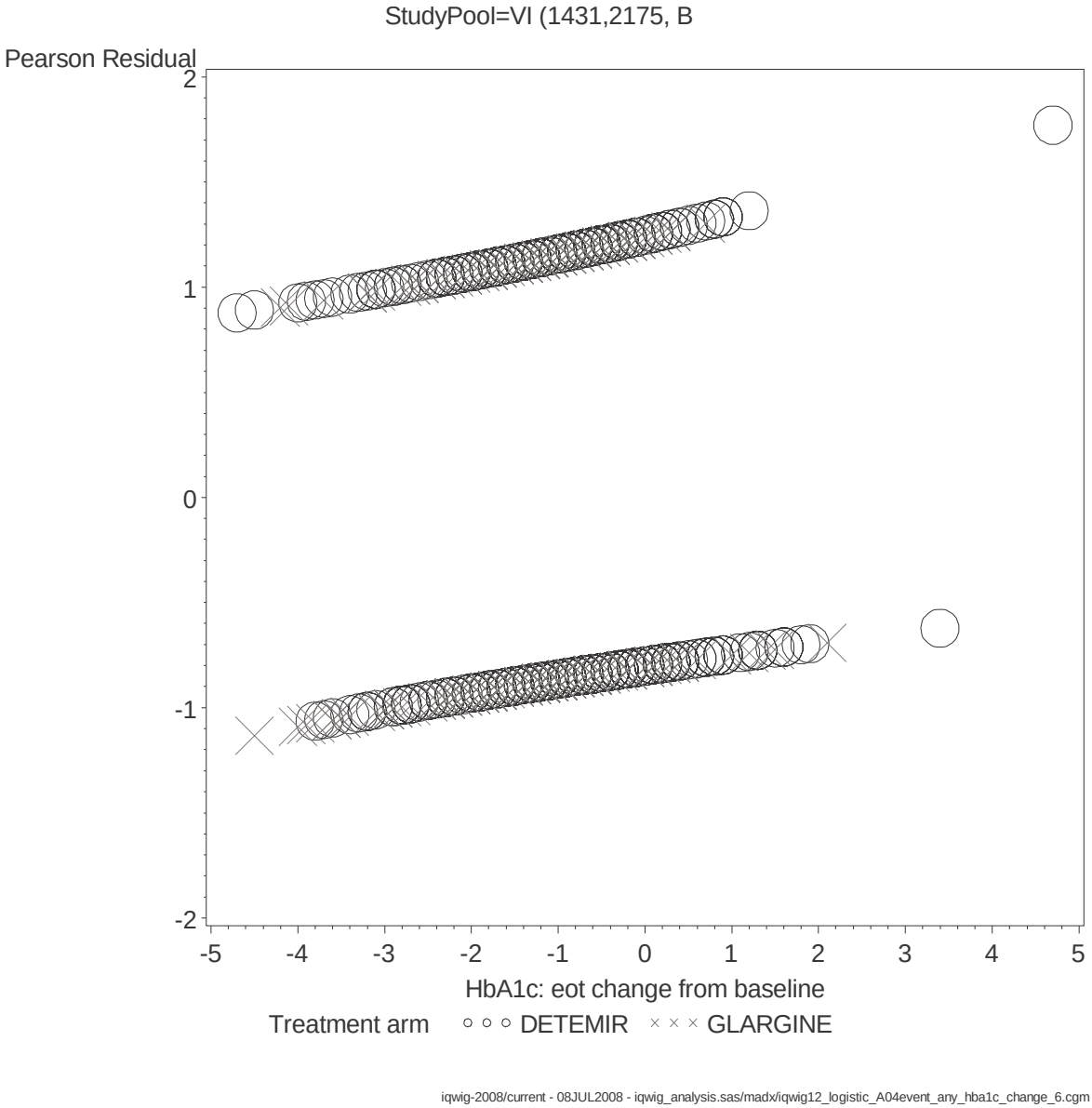


47: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_change_5.cgm

48: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_change_6.cgm

49: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_change

50: Logistic regression for endpoint A04event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any           A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1     0                    382
       2     1                     95

Probability modeled is A04event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	478.273	476.049
SC	482.440	488.551
-2 Log L	476.273	470.049

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	6.2239	2	0.0445
Score	6.3837	2	0.0411
Wald	6.2862	2	0.0431

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	5.9365	0.0148
hba1c_eot	1	0.7681	0.3808

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.0552	0.8533	5.8007	0.0160
TRTSEQ_LB DETEMIR	1	-0.2887	0.1185	5.9365	0.0148
hba1c_eot	1	0.0982	0.1120	0.7681	0.3808

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.561	0.353	0.893
hba1c_eot	1.103	0.886	1.374

Association of Predicted Probabilities and Observed Responses

Percent Concordant	57.3	Somers' D	0.170
Percent Discordant	40.2	Gamma	0.175
Percent Tied	2.5	Tau-a	0.054
Pairs	36290	c	0.585

```
Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any            A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1             0            371
       2             1             45

Probability modeled is A04event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	287.109	280.520
SC	291.139	292.612
-2 Log L	285.109	274.520

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	10.5883	2	0.0050
Score	11.2275	2	0.0036
Wald	10.5565	2	0.0051

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	10.3678	0.0013
hba1c_eot	1	0.1183	0.7308

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-2.4389	1.1875	4.2182	0.0400
TRTSEQ_LB DETEMIR	1	-0.5312	0.1650	10.3678	0.0013
hba1c_eot	1	0.0479	0.1394	0.1183	0.7308

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.346	0.181	0.660
hba1c_eot	1.049	0.798	1.379

Association of Predicted Probabilities and Observed Responses

Percent Concordant	59.3	Somers' D	0.245
Percent Discordant	34.8	Gamma	0.260
Percent Tied	5.8	Tau-a	0.047
Pairs	16695	c	0.622

```
Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any            A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1     0                    295
       2     1                    166

Probability modeled is A04event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	604.498	591.406
SC	608.631	603.806
-2 Log L	602.498	585.406

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.0919	2	0.0002
Score	16.9344	2	0.0002
Wald	16.6093	2	0.0002

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	16.5560	<.0001
hba1c_eot	1	0.6339	0.4259

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.3566	0.9587	2.0021	0.1571
TRTSEQ_LB DETEMIR	1	-0.4090	0.1005	16.5560	<.0001
hba1c_eot	1	0.1103	0.1386	0.6339	0.4259

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.441	0.298	0.654
hba1c_eot	1.117	0.851	1.465

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.6	Somers' D	0.234
Percent Discordant	37.3	Gamma	0.239
Percent Tied	2.1	Tau-a	0.108
Pairs	48970	c	0.617

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A04event_any A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A04event_any	Total Frequency
1	0	202
2	1	63

Probability modeled is A04event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	292.682	288.791
SC	296.261	299.530
-2 Log L	290.682	282.791

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	7.8910	2	0.0193
Score	7.7025	2	0.0213
Wald	7.4861	2	0.0237

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	6.8288	0.0090
hba1c_eot	1	0.5795	0.4465

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.3492	1.1709	0.0889	0.7655
TRTSEQ_LB DETEMIR	1	-0.3991	0.1527	6.8288	0.0090
hba1c_eot	1	-0.1132	0.1487	0.5795	0.4465

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.450	0.247	0.819
hba1c_eot	0.893	0.667	1.195

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.5	Somers' D	0.226
Percent Discordant	37.9	Gamma	0.230
Percent Tied	1.6	Tau-a	0.082
Pairs	12726	c	0.613

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG
Response Variable	A04event_any A04: Minor or symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
Number of Response Levels	2
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	582
Number of Observations Used	543

Response Profile

Ordered Value	A04event_any	Total Frequency
1	0	375
2	1	168

Probability modeled is A04event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	673.814	676.932
SC	678.111	689.823
-2 Log L	671.814	670.932

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.8826	2	0.6432
Score	0.8744	2	0.6458
Wald	0.8725	2	0.6465

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0499	0.8233
hba1c_eot	1	0.8314	0.3619

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.1709	0.6976	0.0600	0.8065
TRTSEQ_LB DETEMIR	1	0.0208	0.0929	0.0499	0.8233
hba1c_eot	1	-0.0884	0.0969	0.8314	0.3619

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.042	0.724	1.501
hba1c_eot	0.915	0.757	1.107

Association of Predicted Probabilities and Observed Responses

Percent Concordant	49.7	Somers' D	0.026
Percent Discordant	47.1	Gamma	0.027
Percent Tied	3.2	Tau-a	0.011
Pairs	63000	c	0.513

```
Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                          Model Information

Data Set                          WORK.ANALYSIS_IQWIG
Response Variable                 A04event_any            A04: Minor or symptom, nocturnal, BG<3.9
mmol/L (70 mg/dl)
Number of Response Levels         2
Model                             binary logit
Optimization Technique            Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A04event_          Total
   Value     any            Frequency

       1     0                    388
       2     1                    296

Probability modeled is A04event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	937.813	928.226
SC	942.341	941.810
-2 Log L	935.813	922.226

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	13.5874	2	0.0011
Score	13.2388	2	0.0013
Wald	12.8981	2	0.0016

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0043	0.9477
hba1c_eot	1	12.7689	0.0004

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	1.7266	0.5599	9.5092	0.0020
TRTSEQ_LB DETEMIR	1	-0.00543	0.0828	0.0043	0.9477
hba1c_eot	1	-0.2687	0.0752	12.7689	0.0004

Logistic regression for endpoint A04event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

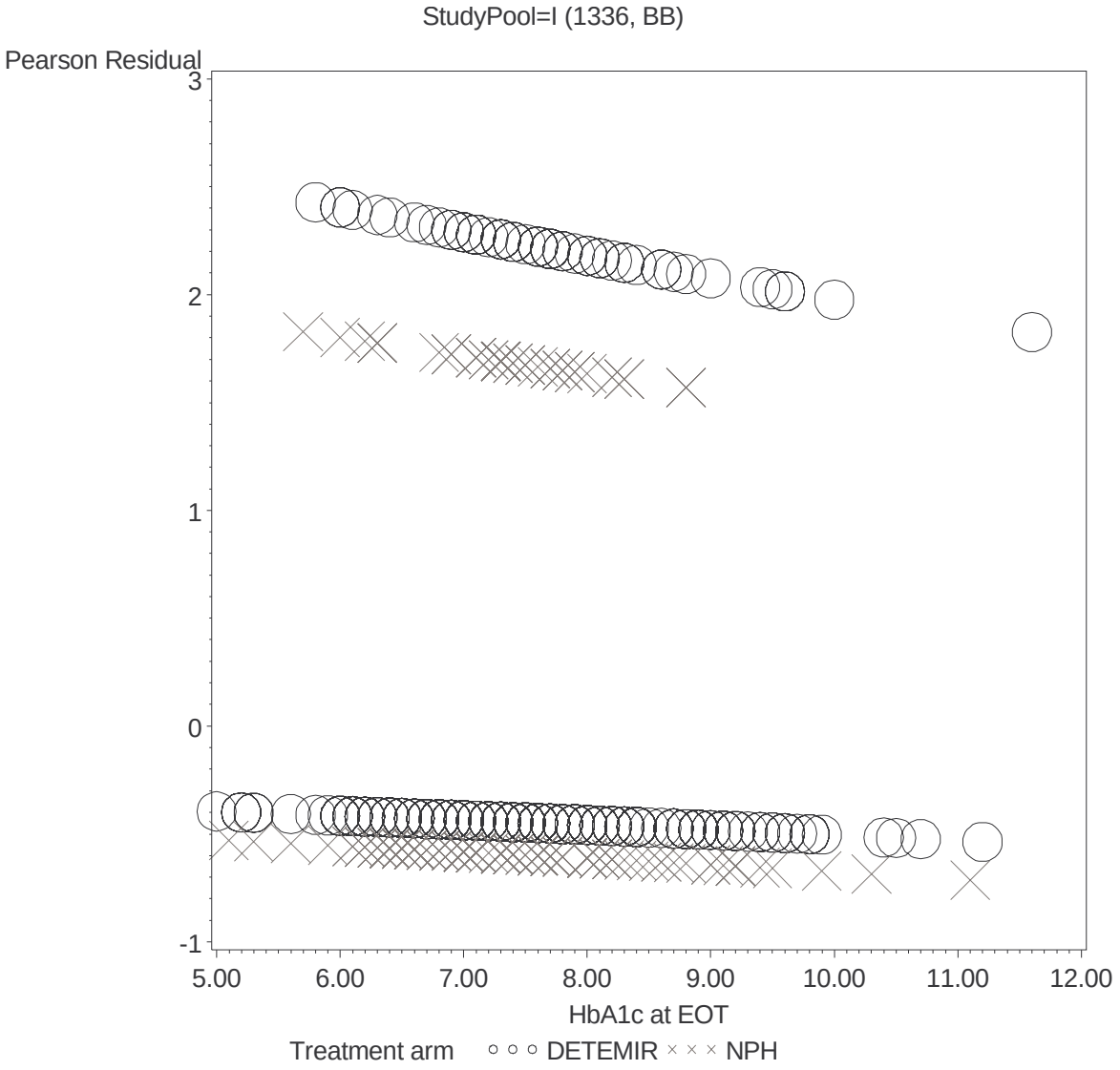
The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.989	0.715	1.368
hba1c_eot	0.764	0.660	0.886

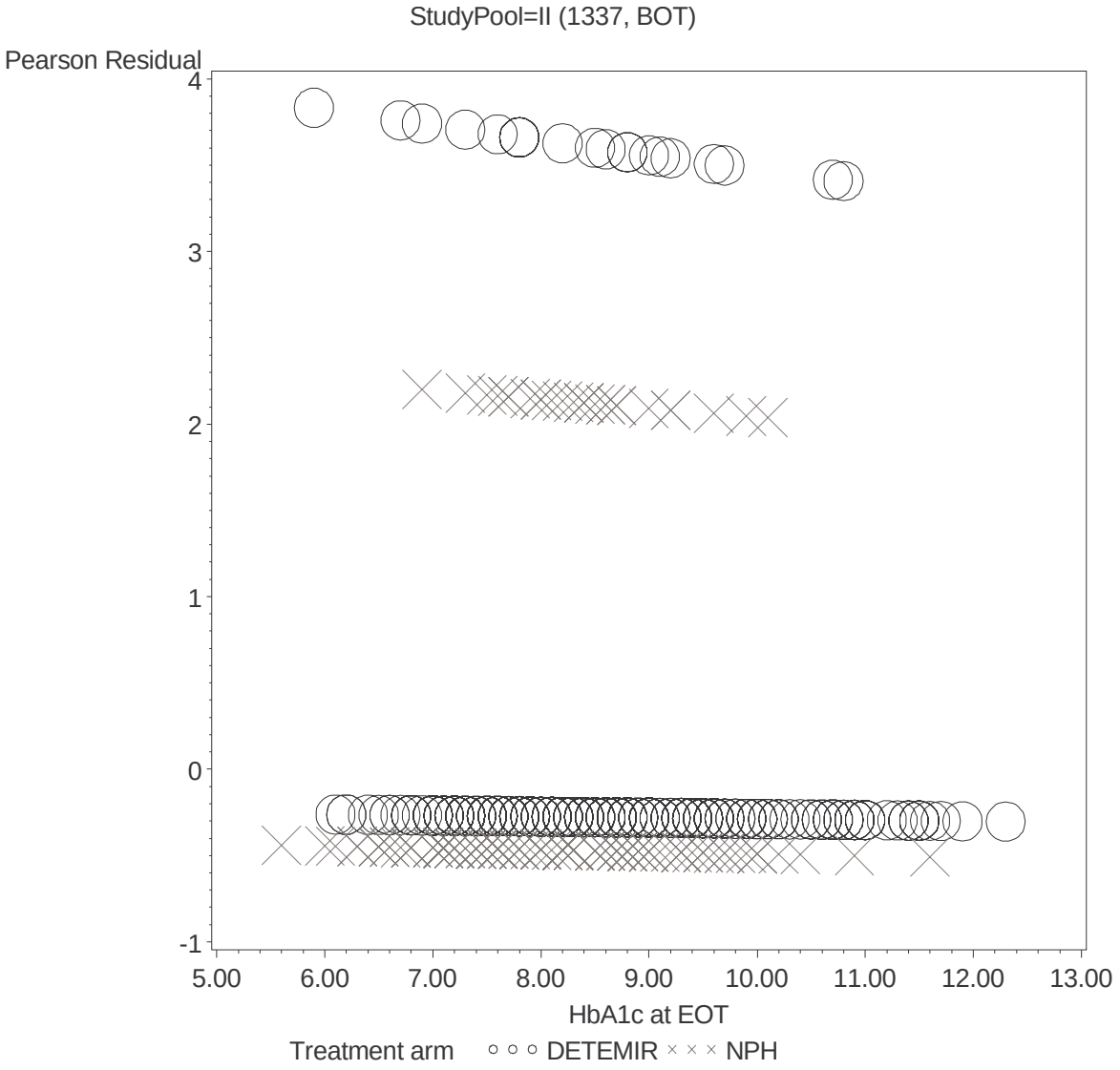
Association of Predicted Probabilities and Observed Responses

Percent Concordant	55.6	Somers' D	0.128
Percent Discordant	42.8	Gamma	0.130
Percent Tied	1.6	Tau-a	0.063
Pairs	114848	c	0.564



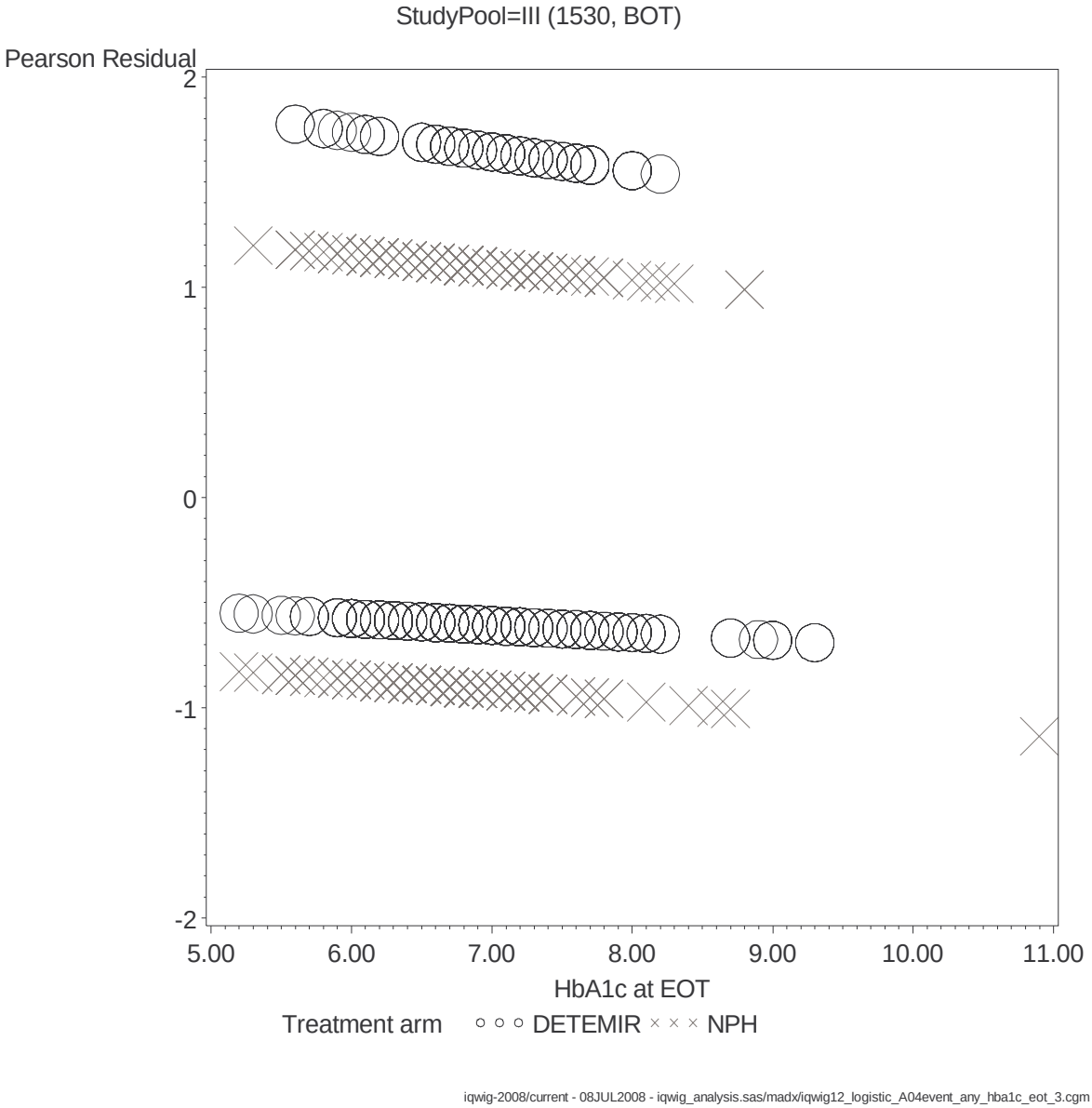
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_eot_1.cgm

51: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot

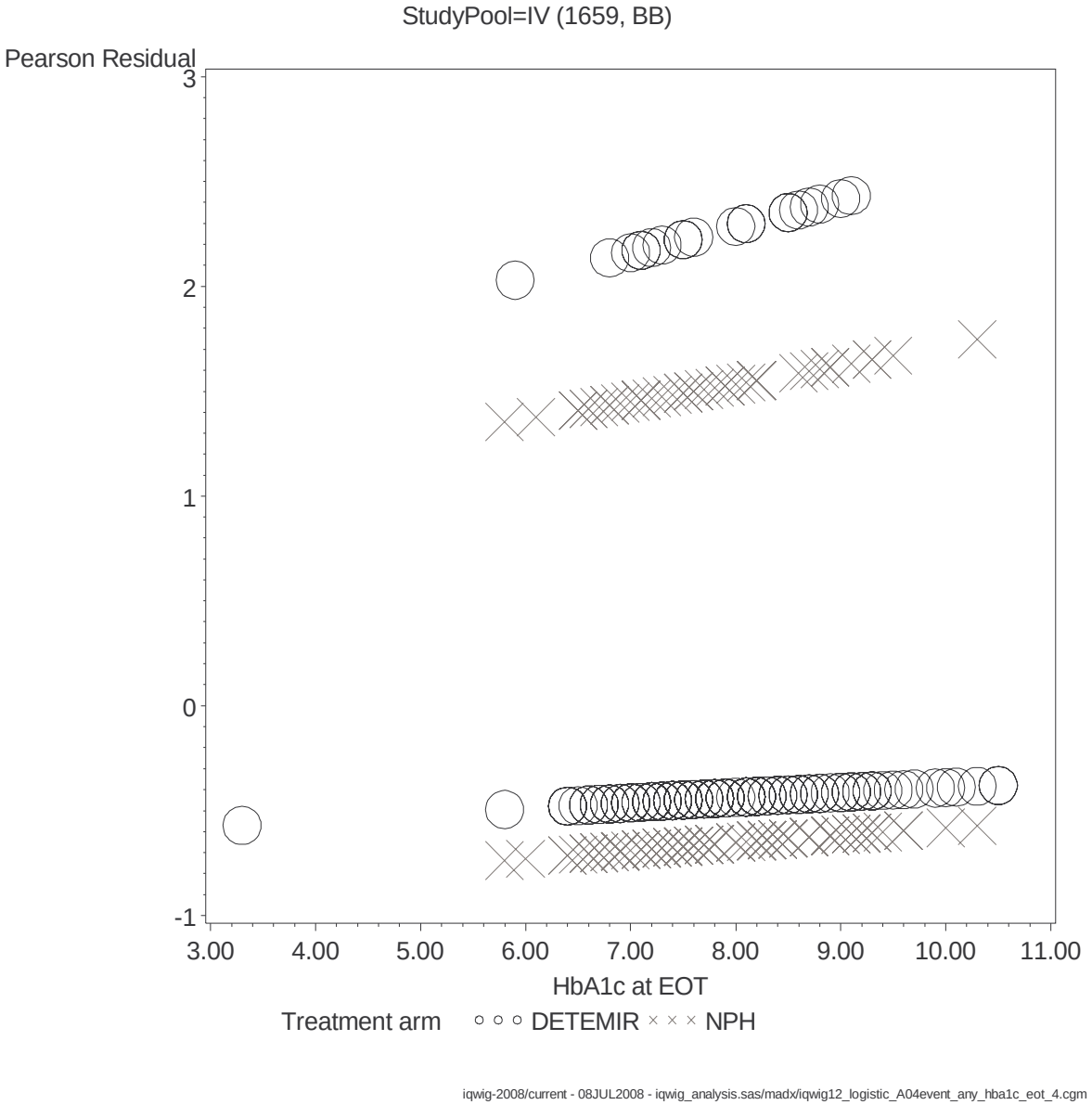


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_eot_2.cgm

52: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot

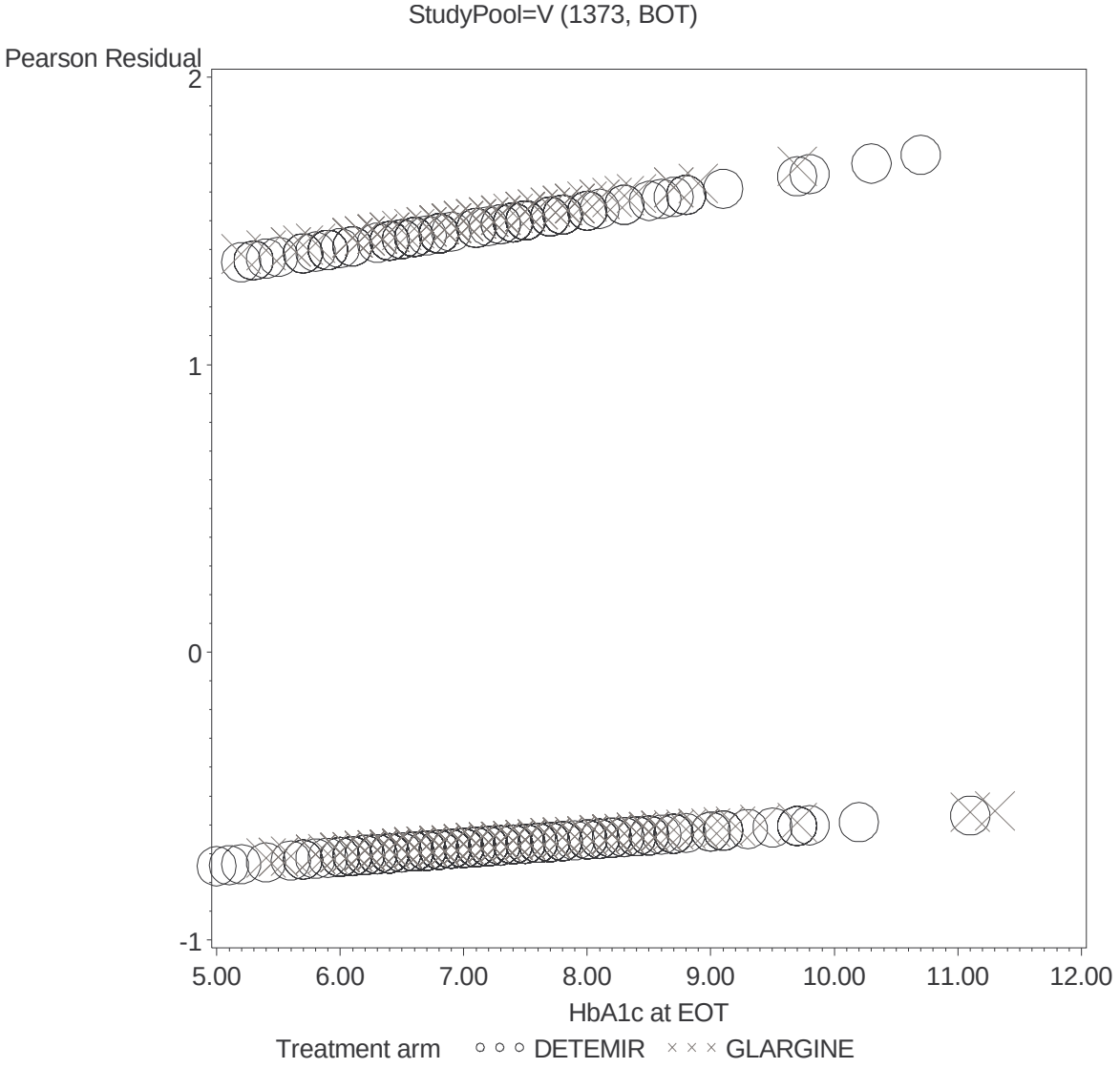


53: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot



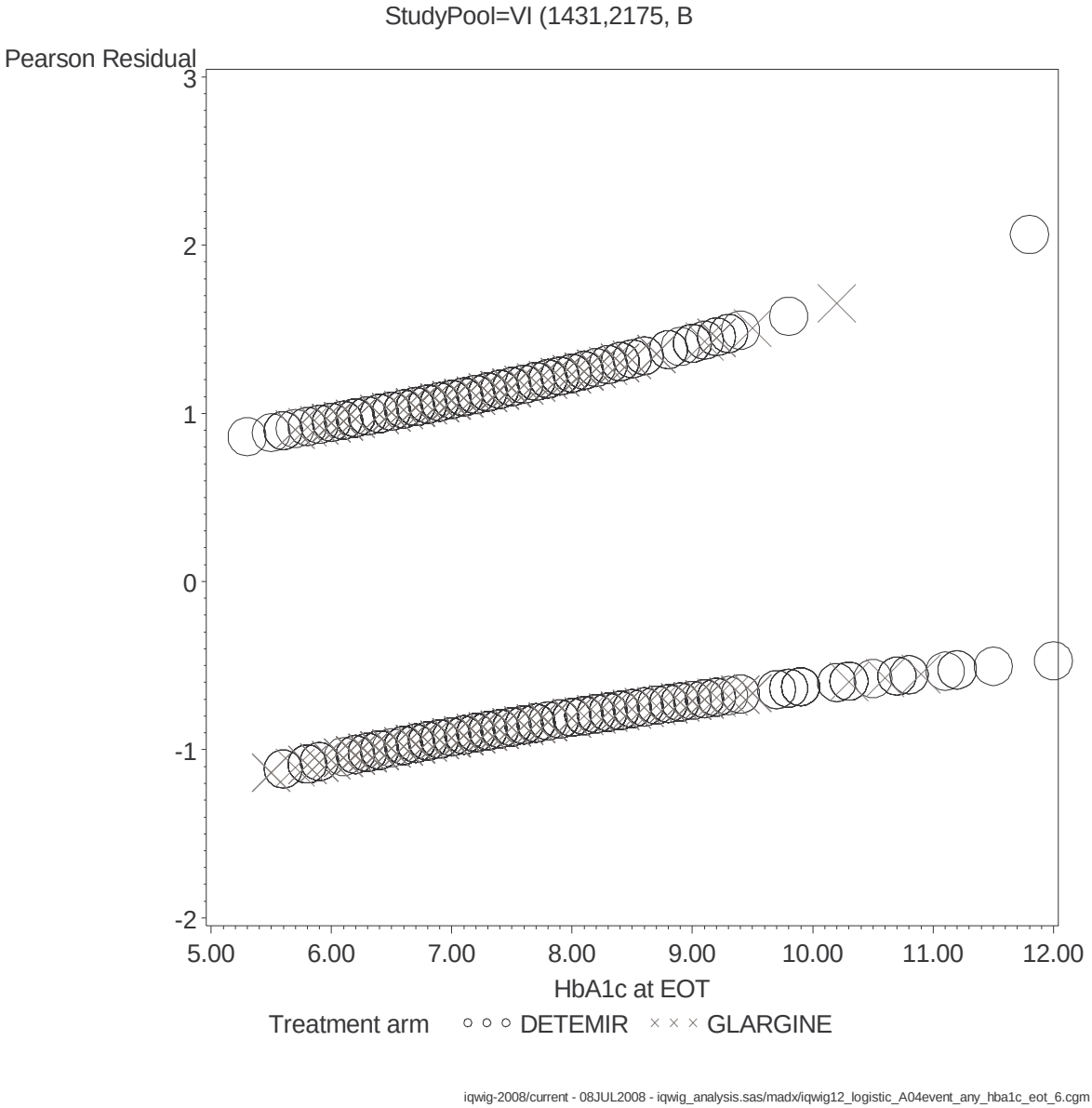
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_eot_4.cgm

54: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_eot_5.cgm

55: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A04event_any_hba1c_eot_6.cgm

56: Logistic regression: residuals for endpoint A04event_any controlling for hba1c_eot

57: Logistic regression for endpoint A05event_any controlling for hba1c_change

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1             0              469
       2             1                8

Probability modeled is A05event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value         Variables

TRTSEQ_LB     DETEMIR        1
              NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	83.274	87.045
SC	87.442	99.548
-2 Log L	81.274	81.045

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.2290	2	0.8918
Score	0.2360	2	0.8887
Wald	0.2360	2	0.8887

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0665	0.7965
hba1c_change	1	0.1591	0.6900

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.0933	0.3999	104.7552	<.0001
TRTSEQ_LB	DETEMIR	1	-0.0951	0.3689	0.0665	0.7965
hba1c_change		1	-0.1310	0.3285	0.1591	0.6900

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	0.827	0.195	3.511
hba1c_change		0.877	0.461	1.670

Association of Predicted Probabilities and Observed Responses

Percent Concordant	47.1	Somers' D	0.171
Percent Discordant	30.0	Gamma	0.222
Percent Tied	22.9	Tau-a	0.006
Pairs	3752	c	0.586

```
Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1     0                    415
       2     1                      1

Probability modeled is A05event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.059	17.861
SC	20.090	29.953
-2 Log L	14.059	11.861

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.1980	2	0.3332
Score	1.9983	2	0.3682
Wald	0.0045	2	0.9977

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0037	0.9513
hba1c_change	1	0.0008	0.9775

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-10.4550	90.9908	0.0132	0.9085
TRTSEQ_LB	DETEMIR	1	-5.5578	90.9870	0.0037	0.9513
hba1c_change		1	0.0247	0.8756	0.0008	0.9775

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	<0.001	<0.001	>999.999
hba1c_change		1.025	0.184	5.702

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.7	Somers' D	0.667
Percent Discordant	0.0	Gamma	1.000
Percent Tied	33.3	Tau-a	0.003
Pairs	415	c	0.834

```
Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1             0            455
       2             1              6

Probability modeled is A05event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	66.021	67.061
SC	70.155	79.461
-2 Log L	64.021	61.061

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.9601	2	0.2276
Score	2.7150	2	0.2573
Wald	2.2086	2	0.3314

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	2.1534	0.1423
hba1c_change	1	0.0312	0.8597

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.7618	0.9738	23.9133	<.0001
TRTSEQ_LB	DETEMIR	1	-0.8073	0.5502	2.1534	0.1423
hba1c_change		1	-0.0811	0.4591	0.0312	0.8597

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.199	0.023	1.719
hba1c_change	0.922	0.375	2.268

Association of Predicted Probabilities and Observed Responses

Percent Concordant	54.8	Somers' D	0.316
Percent Discordant	23.3	Gamma	0.404
Percent Tied	21.9	Tau-a	0.008
Pairs	2730	c	0.658

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A05event_any	A05: Major, daytime (not nocturnal)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A05event_any	Total Frequency
1	0	263
2	1	2

Probability modeled is A05event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration.
Validity of the model fit is questionable.

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	25.531	27.007
SC	29.111	37.746
-2 Log L	23.531	21.007

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.5245	2	0.2830
Score	1.7597	2	0.4149
Wald	0.0207	2	0.9897

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0048	0.9449
hba1c_change	1	0.0159	0.8996

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-9.4856	77.0447	0.0152	0.9020
TRTSEQ_LB	DETEMIR	1	-5.3212	77.0422	0.0048	0.9449
hba1c_change		1	0.0890	0.7053	0.0159	0.8996

```
Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                     Point          95% Wald
Effect                             Estimate      Confidence Limits

TRTSEQ_LB    DETEMIR vs NPH          <0.001      <0.001    >999.999
hba1c_change                          1.093       0.274       4.355


Association of Predicted Probabilities and Observed Responses

Percent Concordant     59.9    Somers' D    0.433
Percent Discordant     16.5    Gamma        0.567
Percent Tied           23.6    Tau-a        0.007
Pairs                   526    c            0.717
```

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A05event_any	A05: Major, daytime (not nocturnal)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	582
Number of Observations Used	543

Response Profile

Ordered Value	A05event_any	Total Frequency
1	0	537
2	1	6

Probability modeled is A05event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	GLARGINE	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	67.998	71.512
SC	72.295	84.404
-2 Log L	65.998	65.512

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.4853	2	0.7845
Score	0.5034	2	0.7775
Wald	0.5044	2	0.7771

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0000	0.9950
hba1c_change	1	0.5025	0.4784

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.1650	0.5791	51.7244	<.0001
TRTSEQ_LB	DETEMIR	1	0.00260	0.4113	0.0000	0.9950
hba1c_change		1	0.2544	0.3589	0.5025	0.4784

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.005	0.200	5.040
hba1c_change	1.290	0.638	2.606

Association of Predicted Probabilities and Observed Responses

Percent Concordant	49.3	Somers' D	0.215
Percent Discordant	27.8	Gamma	0.279
Percent Tied	22.8	Tau-a	0.005
Pairs	3222	c	0.608

```
Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1             0              664
       2             1               20

Probability modeled is A05event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value         Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	182.698	186.614
SC	187.226	200.198
-2 Log L	180.698	180.614

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0845	2	0.9586
Score	0.0847	2	0.9585
Wald	0.0847	2	0.9585

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0167	0.8971
hba1c_change	1	0.0590	0.8080

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.5476	0.3388	109.6464	<.0001
TRTSEQ_LB	DETEMIR	1	-0.0310	0.2398	0.0167	0.8971
hba1c_change		1	-0.0487	0.2006	0.0590	0.8080

```
Logistic regression for endpoint A05event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

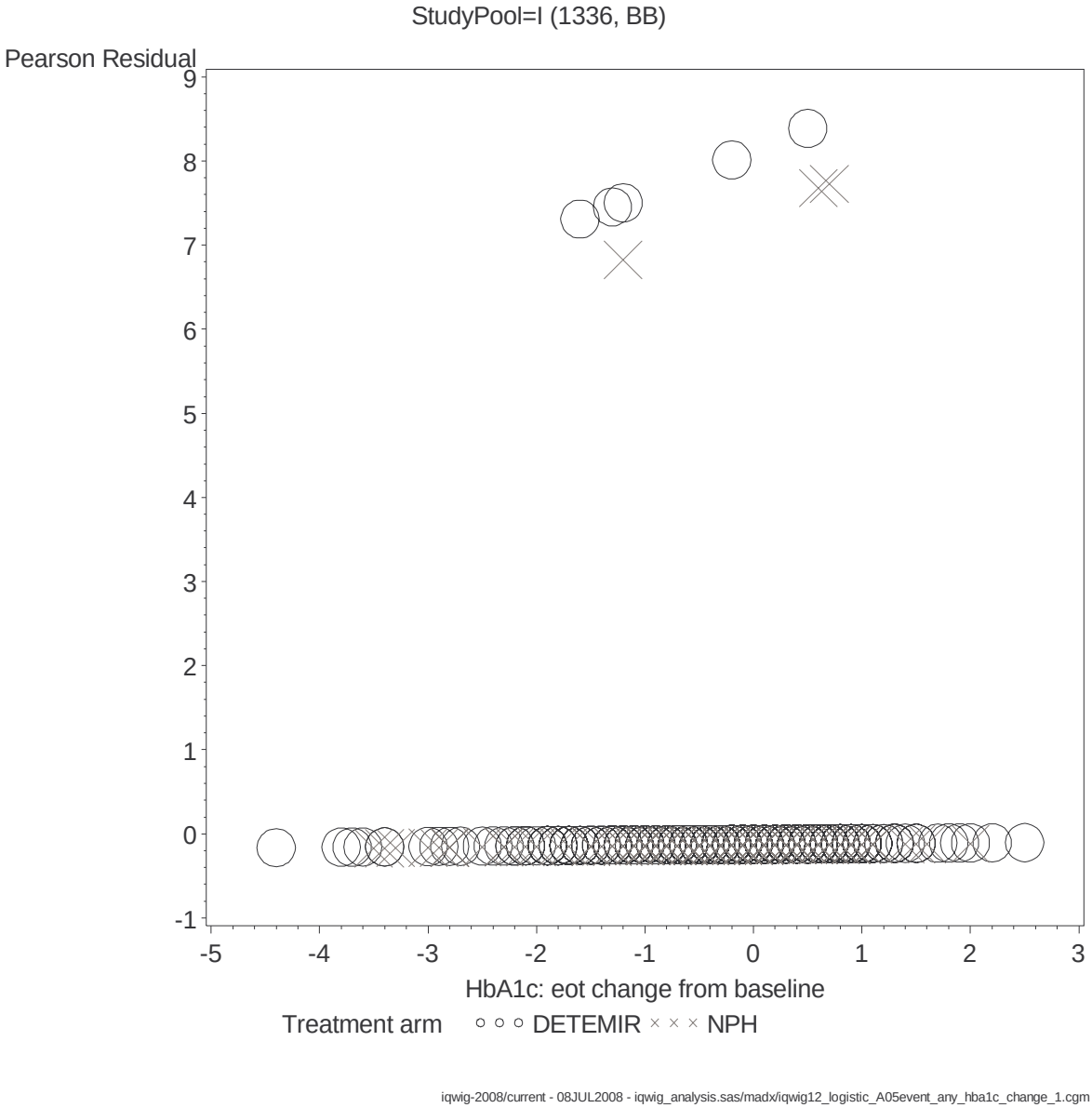
                      Odds Ratio Estimates

                                      Point          95% Wald
Effect                              Estimate      Confidence Limits

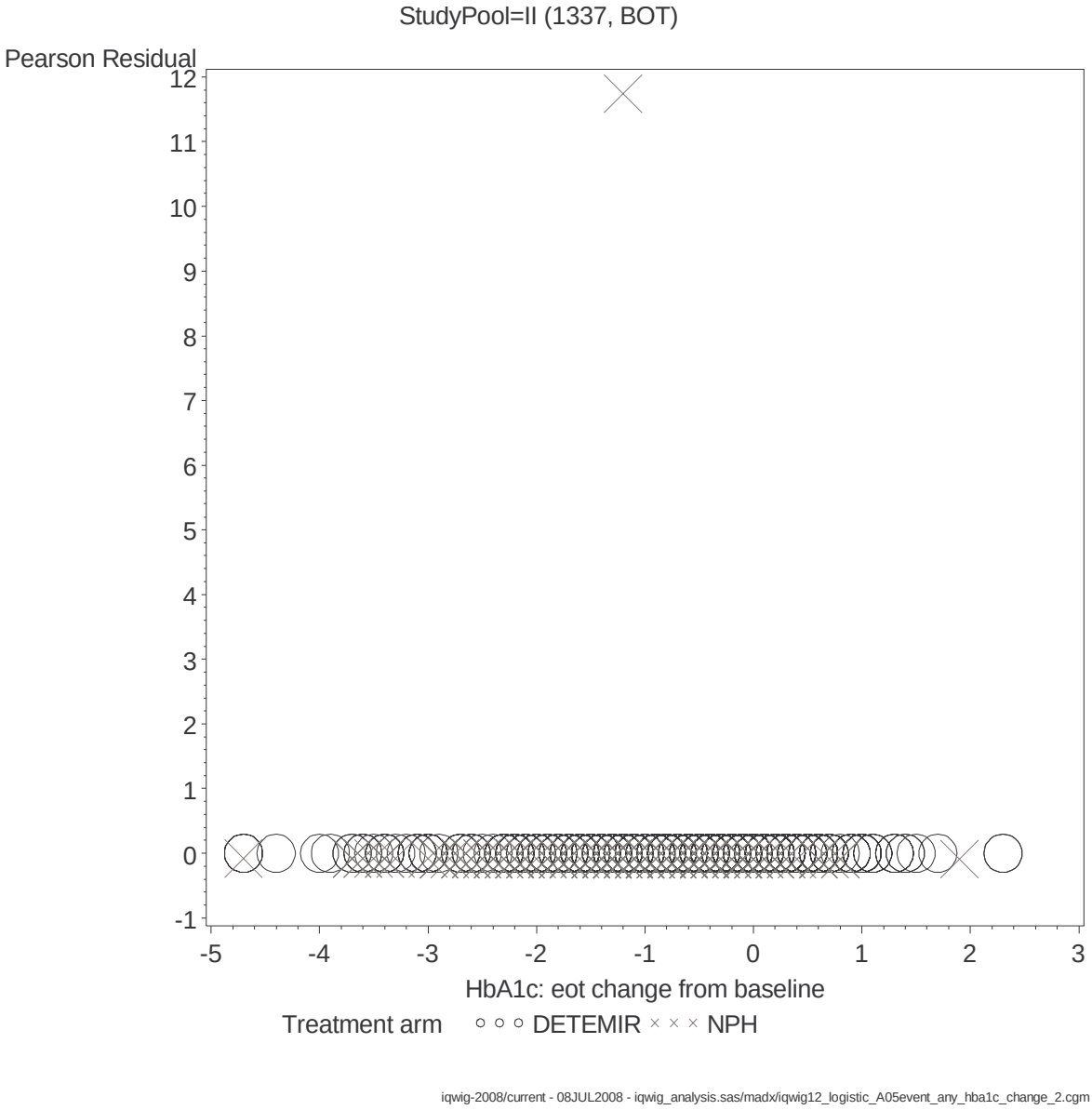
TRTSEQ_LB    DETEMIR  vs GLARGINE       0.940       0.367       2.406
hba1c_change                           0.952       0.643       1.411


Association of Predicted Probabilities and Observed Responses

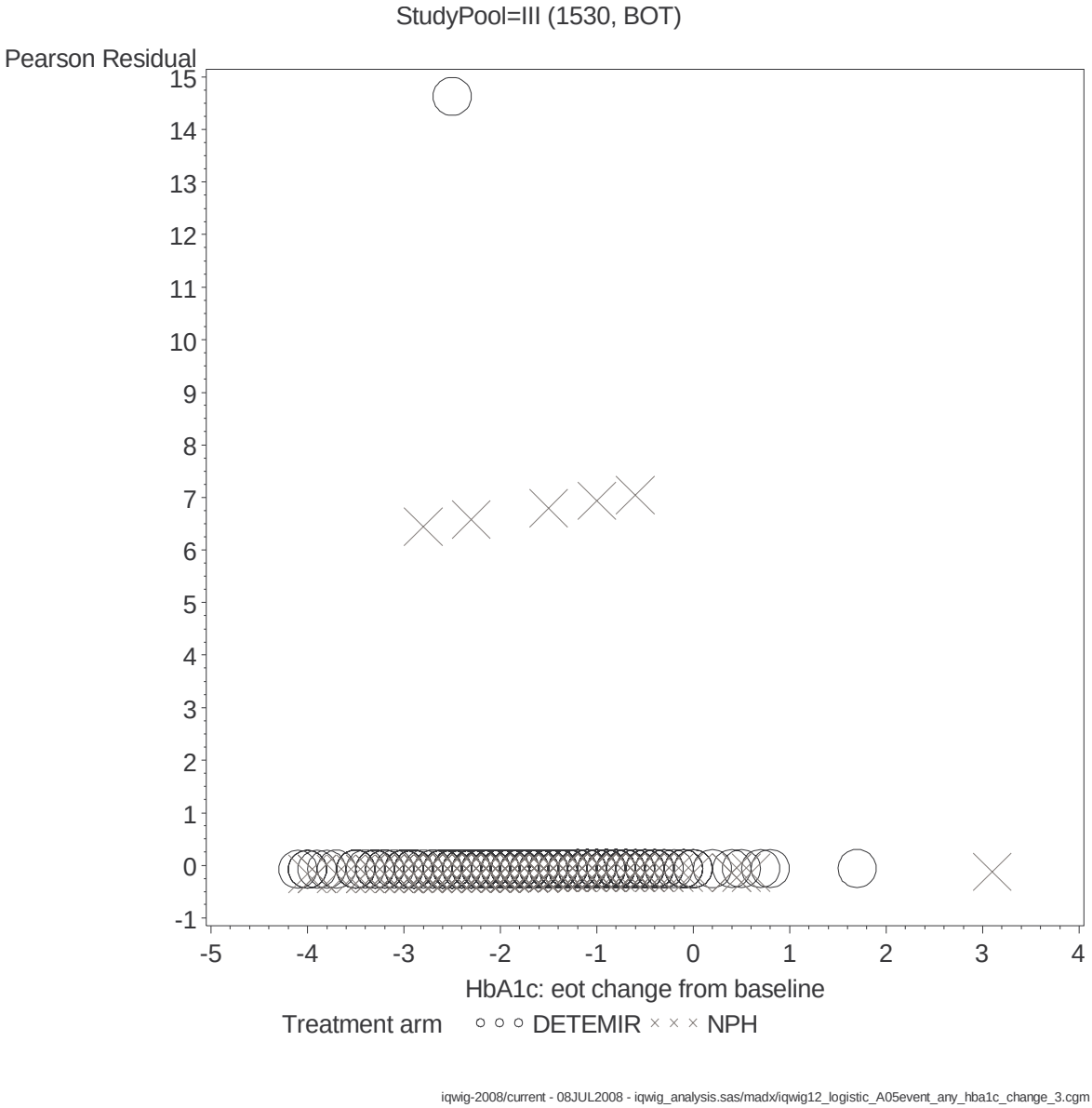
Percent Concordant     35.3    Somers' D    -.020
Percent Discordant     37.3    Gamma        -.027
Percent Tied           27.4    Tau-a        -.001
Pairs                 13280    c             0.490
```



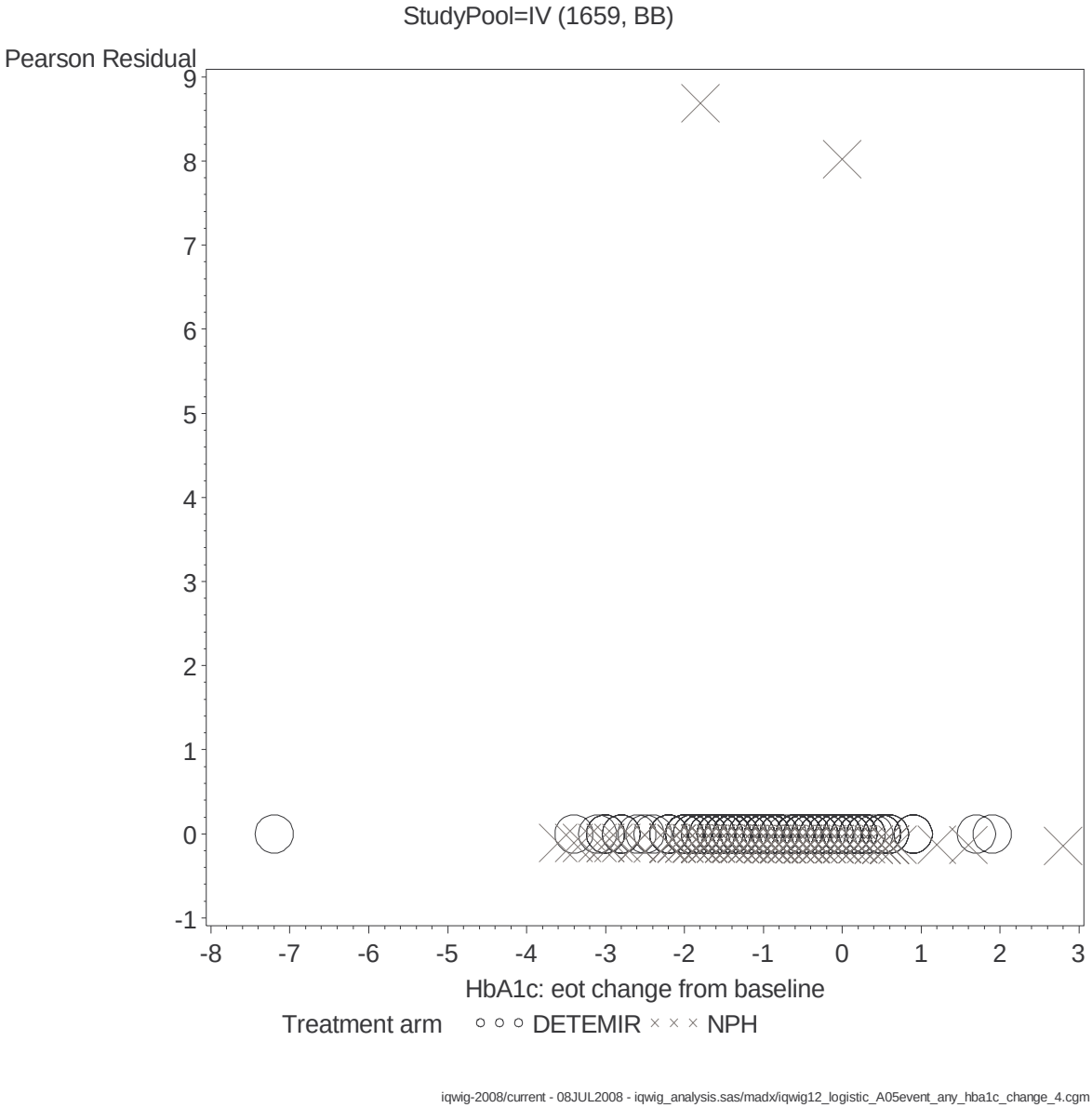
58: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change



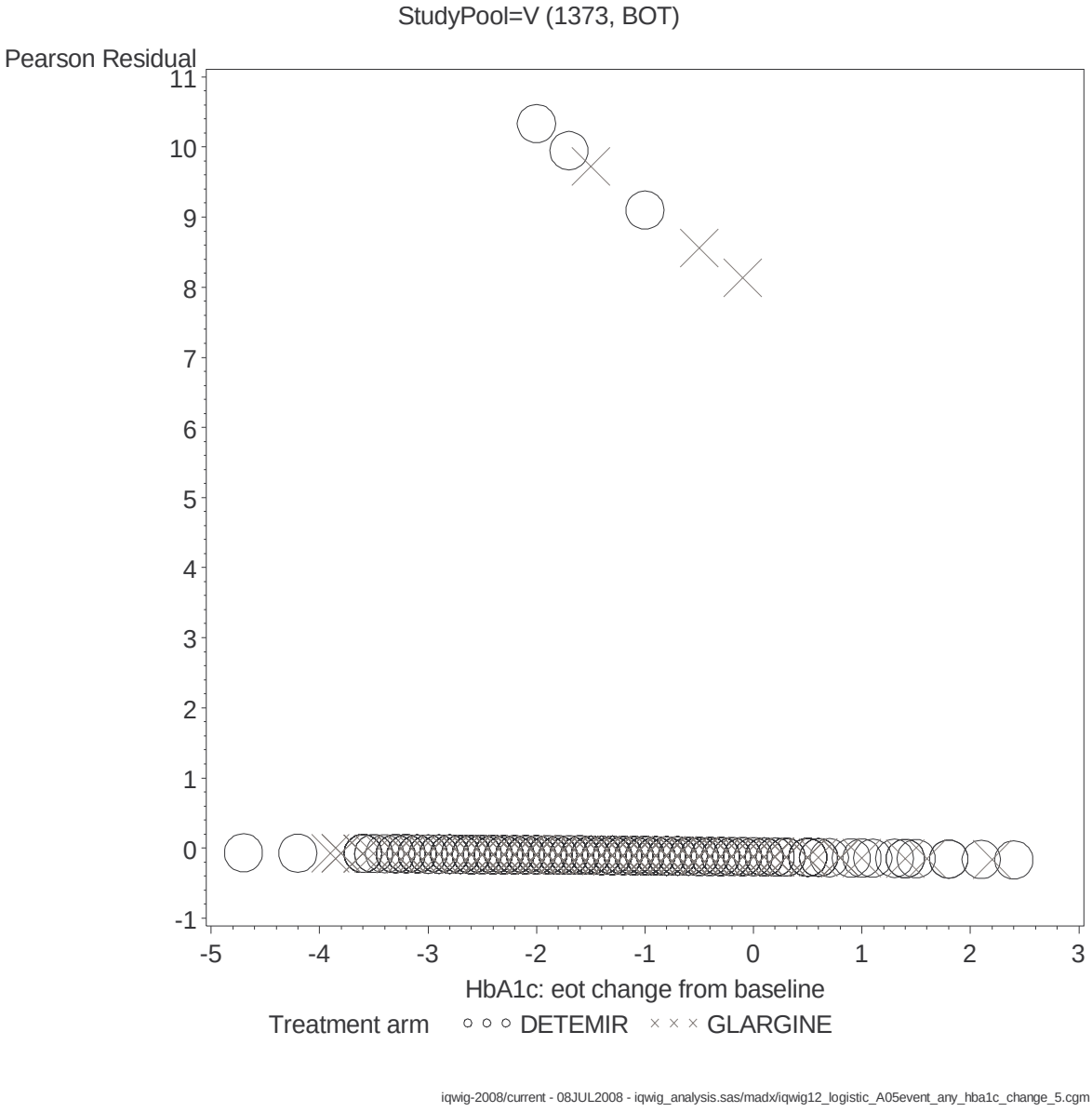
59: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change



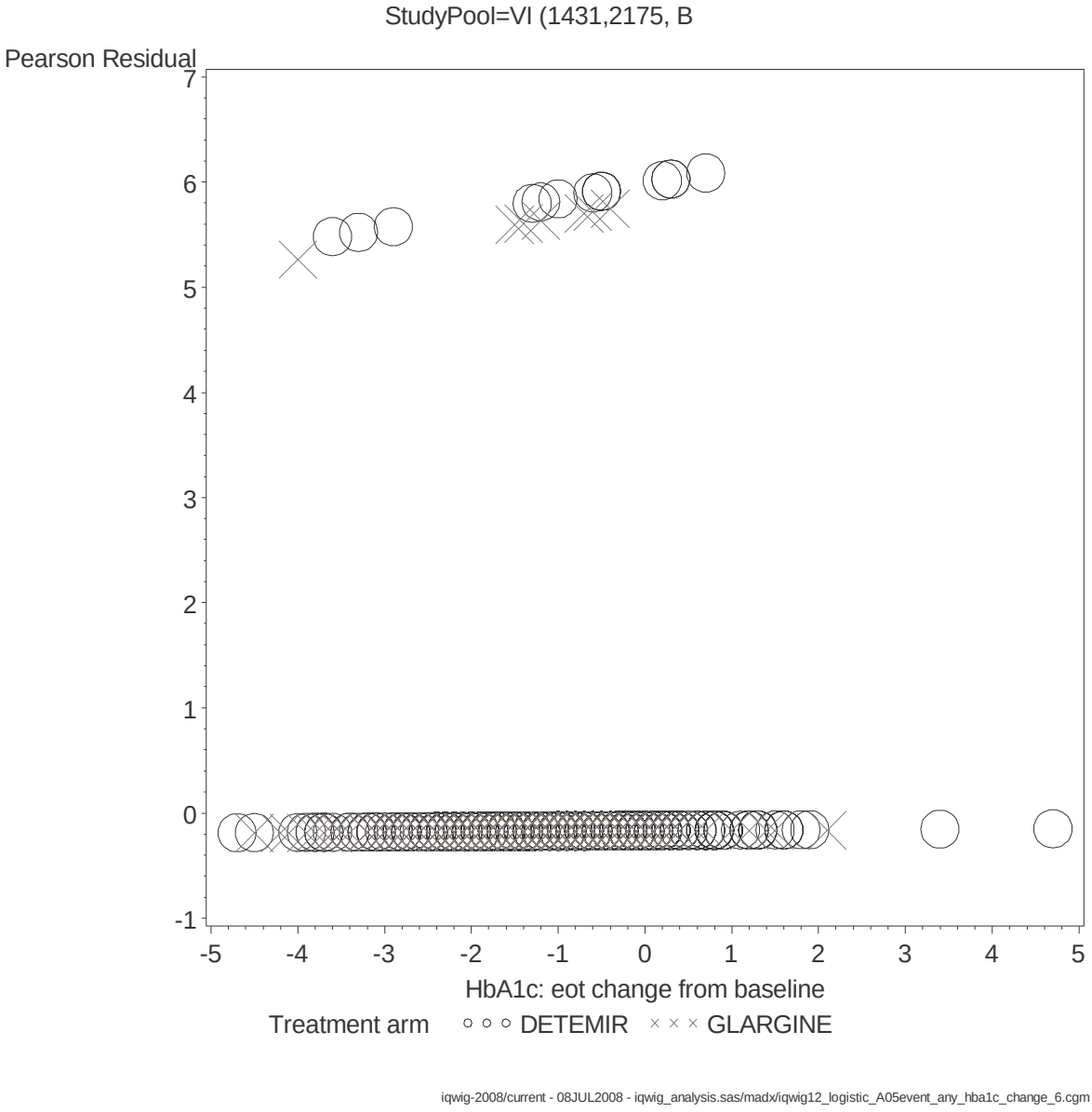
60: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change



61: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change



62: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change



63: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_change

64: Logistic regression for endpoint A05event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1             0              469
       2             1                8

Probability modeled is A05event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value         Variables

TRTSEQ_LB     DETEMIR        1
              NPH           -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	83.274	86.904
SC	87.442	99.407
-2 Log L	81.274	80.904

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.3701	2	0.8311
Score	0.3790	2	0.8274
Wald	0.3791	2	0.8273

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.1125	0.7373
hba1c_eot	1	0.3053	0.5806

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-5.4287	2.5702	4.4613	0.0347
TRTSEQ_LB DETEMIR	1	-0.1245	0.3712	0.1125	0.7373
hba1c_eot	1	0.1831	0.3314	0.3053	0.5806

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.780	0.182	3.340
hba1c_eot	1.201	0.627	2.299

Association of Predicted Probabilities and Observed Responses

Percent Concordant	47.3	Somers' D	0.088
Percent Discordant	38.5	Gamma	0.102
Percent Tied	14.3	Tau-a	0.003
Pairs	3752	c	0.544

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A05event_          Total
   Value     any           Frequency

       1           0             415
       2           1               1

Probability modeled is A05event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              NPH          -1


                   Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.059	17.668
SC	20.090	29.760
-2 Log L	14.059	11.668

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.3908	2	0.3026
Score	2.1599	2	0.3396
Wald	0.2068	2	0.9018

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0042	0.9483
hba1c_eot	1	0.2028	0.6524

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-13.9054	87.1468	0.0255	0.8732
TRTSEQ_LB DETEMIR	1	-5.6309	86.7823	0.0042	0.9483
hba1c_eot	1	0.4006	0.8894	0.2028	0.6524

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           1.493       0.261       8.532


Association of Predicted Probabilities and Observed Responses

Percent Concordant     81.0    Somers' D    0.711
Percent Discordant      9.9    Gamma        0.782
Percent Tied            9.2    Tau-a        0.003
Pairs                   415    c            0.855
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A05event_any	A05: Major, daytime (not nocturnal)
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	475
Number of Observations Used	461

Response Profile

Ordered Value	A05event_any	Total Frequency
1	0	455
2	1	6

Probability modeled is A05event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	66.021	65.801
SC	70.155	78.201
-2 Log L	64.021	59.801

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	4.2206	2	0.1212
Score	3.8040	2	0.1493
Wald	3.3202	2	0.1901

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	1.8662	0.1719
hba1c_eot	1	1.2021	0.2729

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.3260	4.4448	0.0054	0.9415
TRTSEQ_LB DETEMIR	1	-0.7539	0.5519	1.8662	0.1719
hba1c_eot	1	-0.7385	0.6736	1.2021	0.2729

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.221	0.025	1.926
hba1c_eot	0.478	0.128	1.789

Association of Predicted Probabilities and Observed Responses

Percent Concordant	72.0	Somers' D	0.487
Percent Discordant	23.3	Gamma	0.511
Percent Tied	4.7	Tau-a	0.013
Pairs	2730	c	0.743

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A05event_          Total
   Value     any           Frequency

       1           0             263
       2           1               2

Probability modeled is A05event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              NPH          -1


                   Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	25.531	26.794
SC	29.111	37.533
-2 Log L	23.531	20.794

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.7374	2	0.2544
Score	1.9551	2	0.3762
Wald	0.2369	2	0.8883

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0050	0.9435
hba1c_eot	1	0.2319	0.6301

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-12.3718	75.2342	0.0270	0.8694
TRTSEQ_LB DETEMIR	1	-5.3135	74.9939	0.0050	0.9435
hba1c_eot	1	0.3519	0.7307	0.2319	0.6301

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           1.422       0.340       5.954


Association of Predicted Probabilities and Observed Responses

Percent Concordant     76.8    Somers' D    0.605
Percent Discordant     16.3    Gamma        0.649
Percent Tied            6.8    Tau-a        0.009
Pairs                   526    c            0.802
```

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         582
Number of Observations Used         543


          Response Profile

 Ordered     A05event_         Total
   Value     any           Frequency

       1           0             537
       2           1               6

Probability modeled is A05event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                            Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              GLARGINE     -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	67.998	70.417
SC	72.295	83.308
-2 Log L	65.998	64.417

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.5806	2	0.4537
Score	1.8318	2	0.4002
Wald	1.8102	2	0.4045

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0010	0.9749
hba1c_eot	1	1.8058	0.1790

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-7.9018	2.6631	8.8044	0.0030
TRTSEQ_LB DETEMIR	1	-0.0130	0.4124	0.0010	0.9749
hba1c_eot	1	0.4599	0.3423	1.8058	0.1790

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	0.974	0.193	4.907
hba1c_eot	1.584	0.810	3.098

Association of Predicted Probabilities and Observed Responses

Percent Concordant	60.5	Somers' D	0.322
Percent Discordant	28.3	Gamma	0.363
Percent Tied	11.2	Tau-a	0.007
Pairs	3222	c	0.661

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                                  Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A05event_any           A05: Major, daytime (not nocturnal)
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A05event_          Total
   Value     any            Frequency

       1             0            664
       2             1             20

Probability modeled is A05event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                             Design
Class         Value        Variables

TRTSEQ_LB     DETEMIR       1
              GLARGINE     -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	182.698	186.617
SC	187.226	200.201
-2 Log L	180.698	180.617

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0818	2	0.9599
Score	0.0809	2	0.9603
Wald	0.0809	2	0.9604

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0193	0.8895
hba1c_eot	1	0.0553	0.8141

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.1183	1.5943	3.8254	0.0505
TRTSEQ_LB DETEMIR	1	-0.0332	0.2389	0.0193	0.8895
hba1c_eot	1	-0.0503	0.2140	0.0553	0.8141

```
Logistic regression for endpoint A05event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

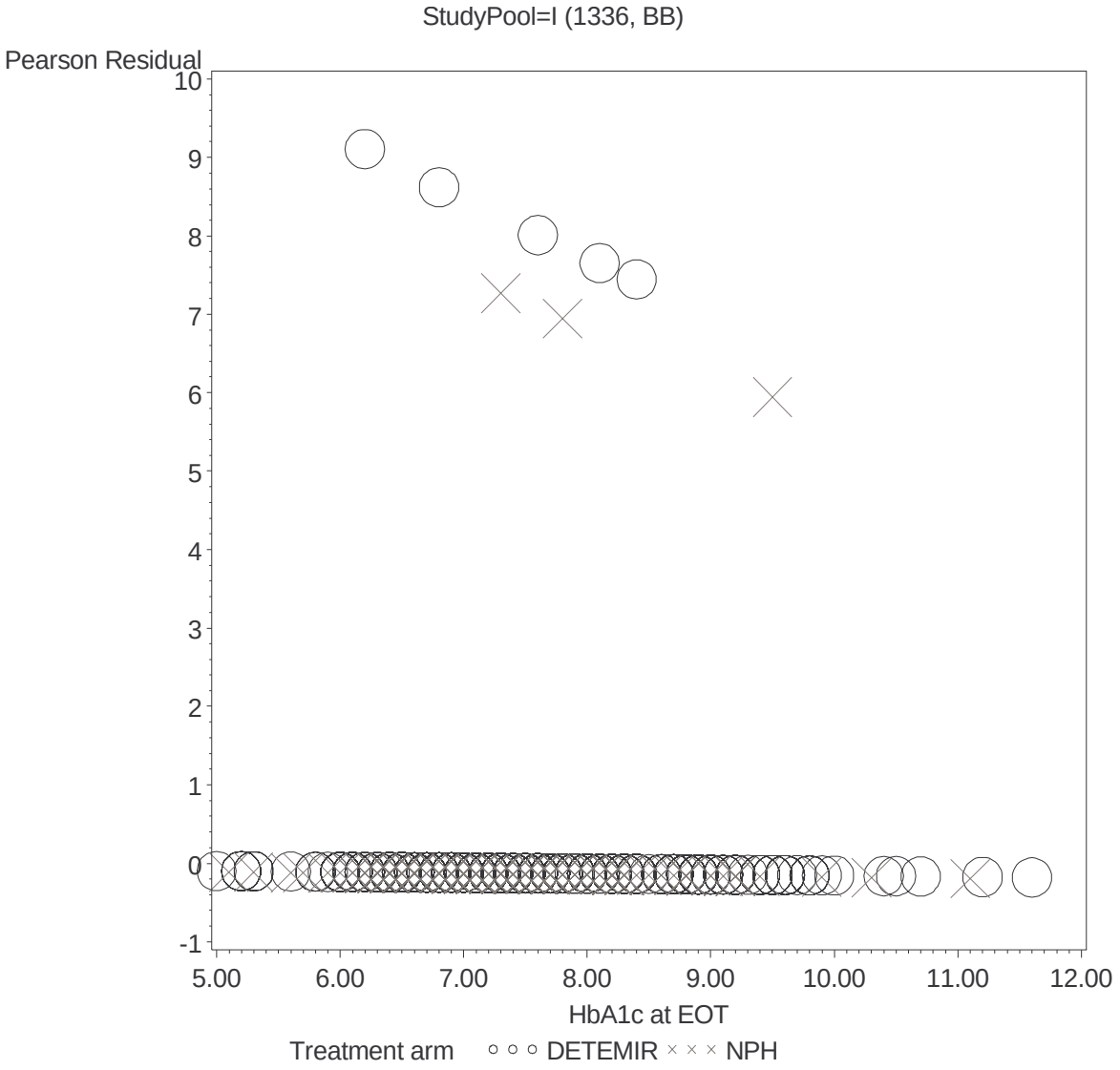
                     Odds Ratio Estimates

                                  Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR  vs GLARGINE       0.936       0.367       2.387
hba1c_eot                            0.951       0.625       1.446

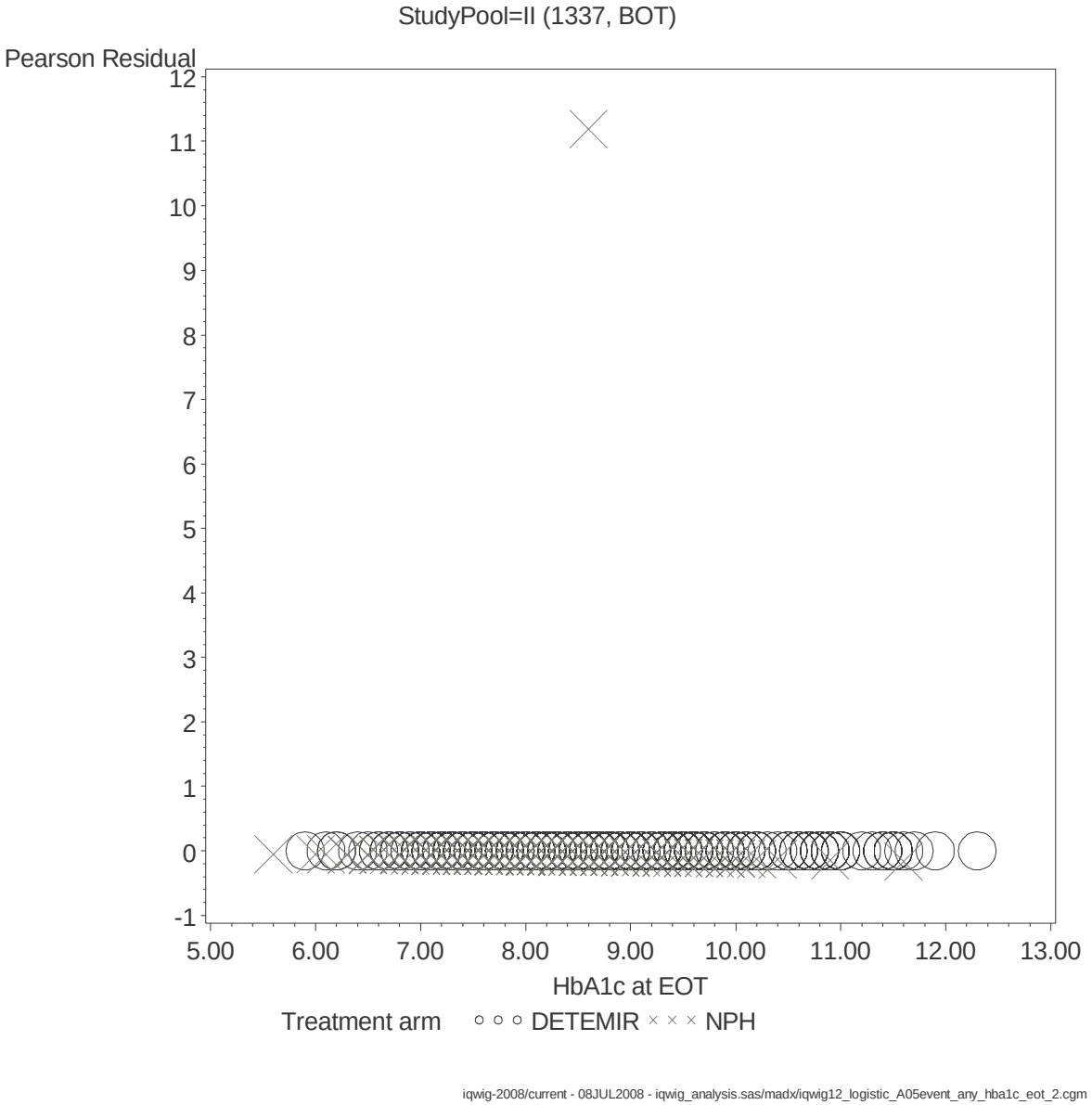

Association of Predicted Probabilities and Observed Responses

Percent Concordant     34.4    Somers' D    -.020
Percent Discordant     36.4    Gamma        -.029
Percent Tied           29.2    Tau-a        -.001
Pairs                 13280    c            0.490
```

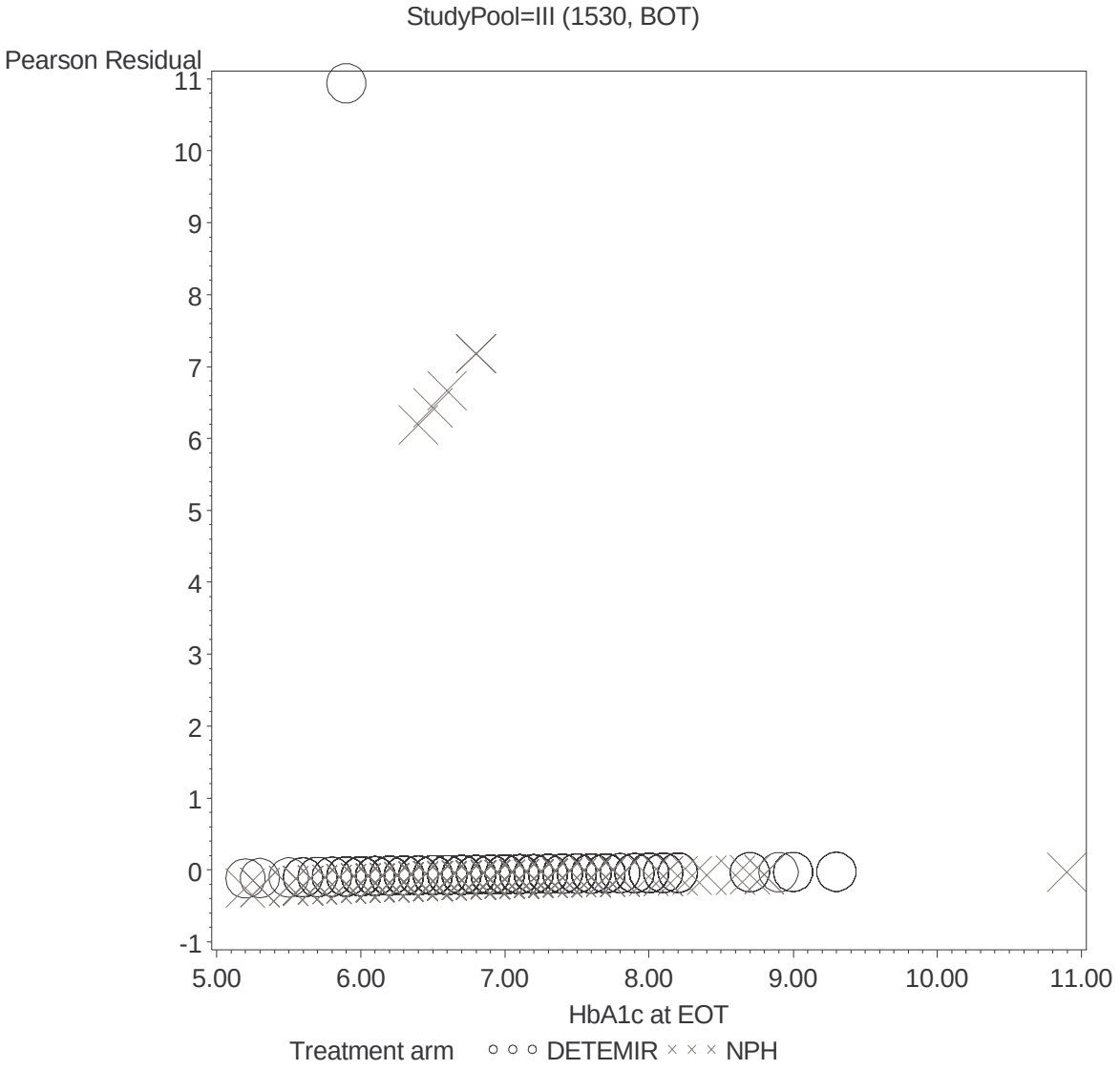


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A05event_any_hba1c_eot_1.cgm

65: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot

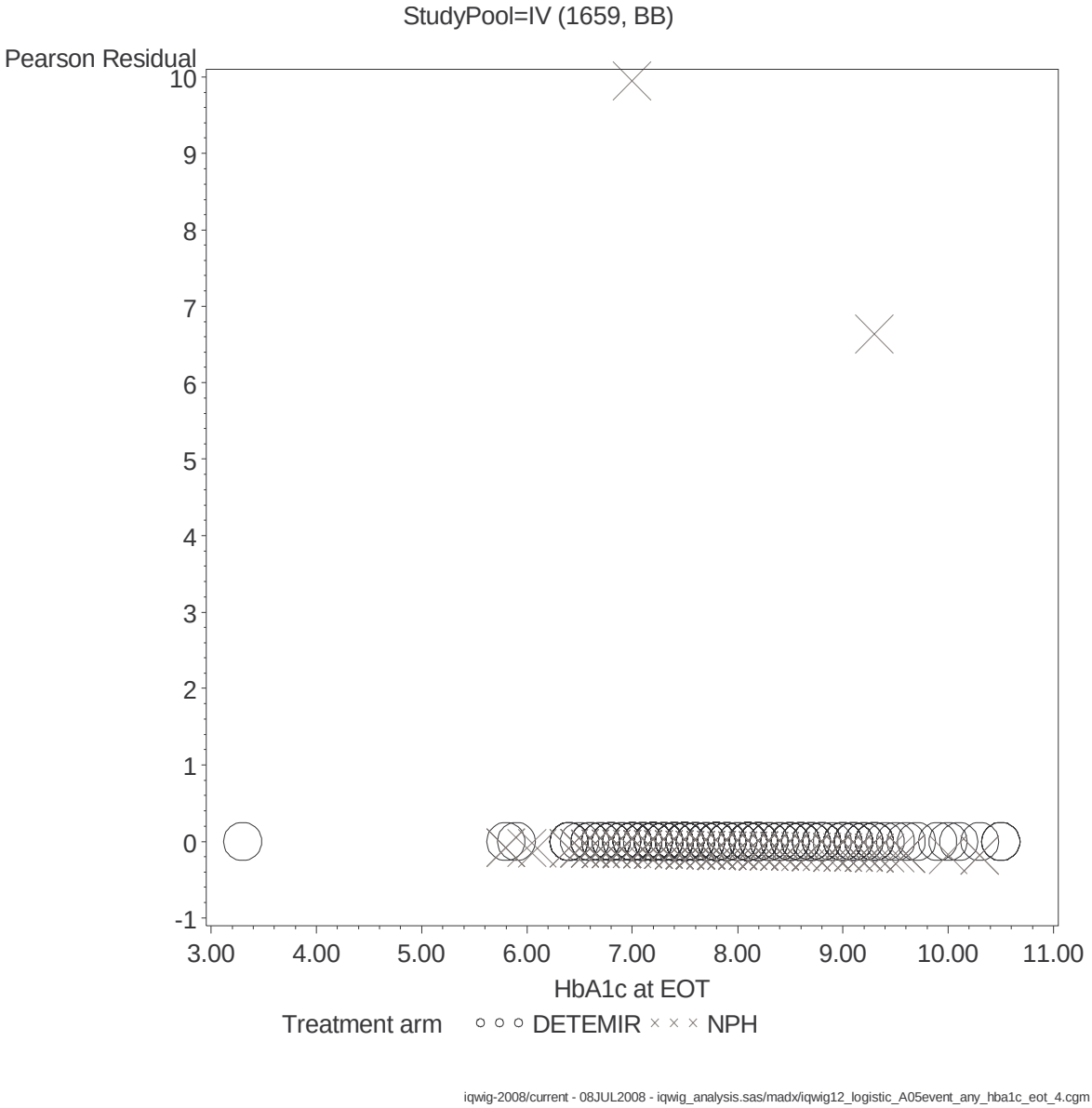


66: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot

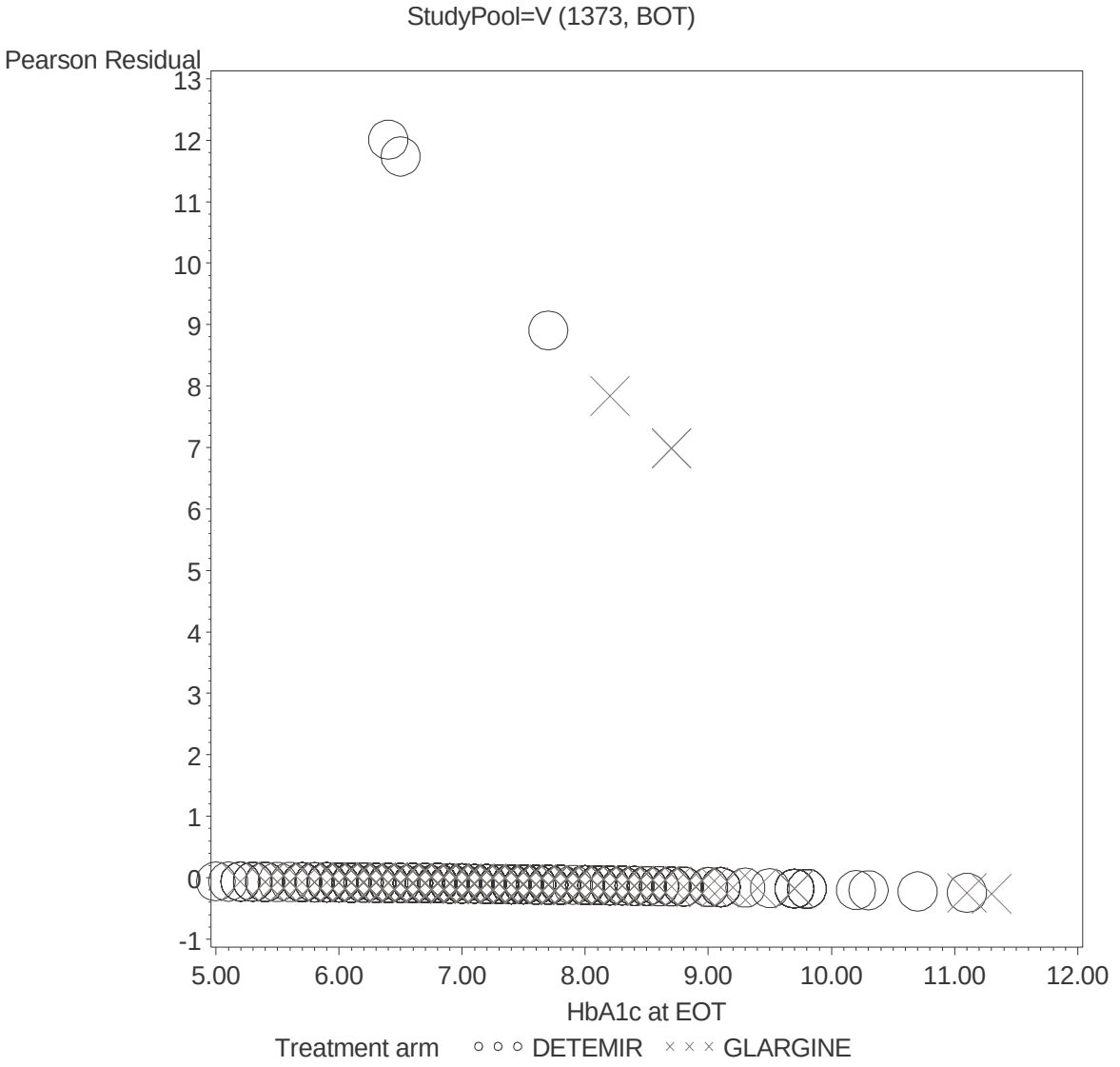


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A05event_any_hba1c_eot_3.cgm

67: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot

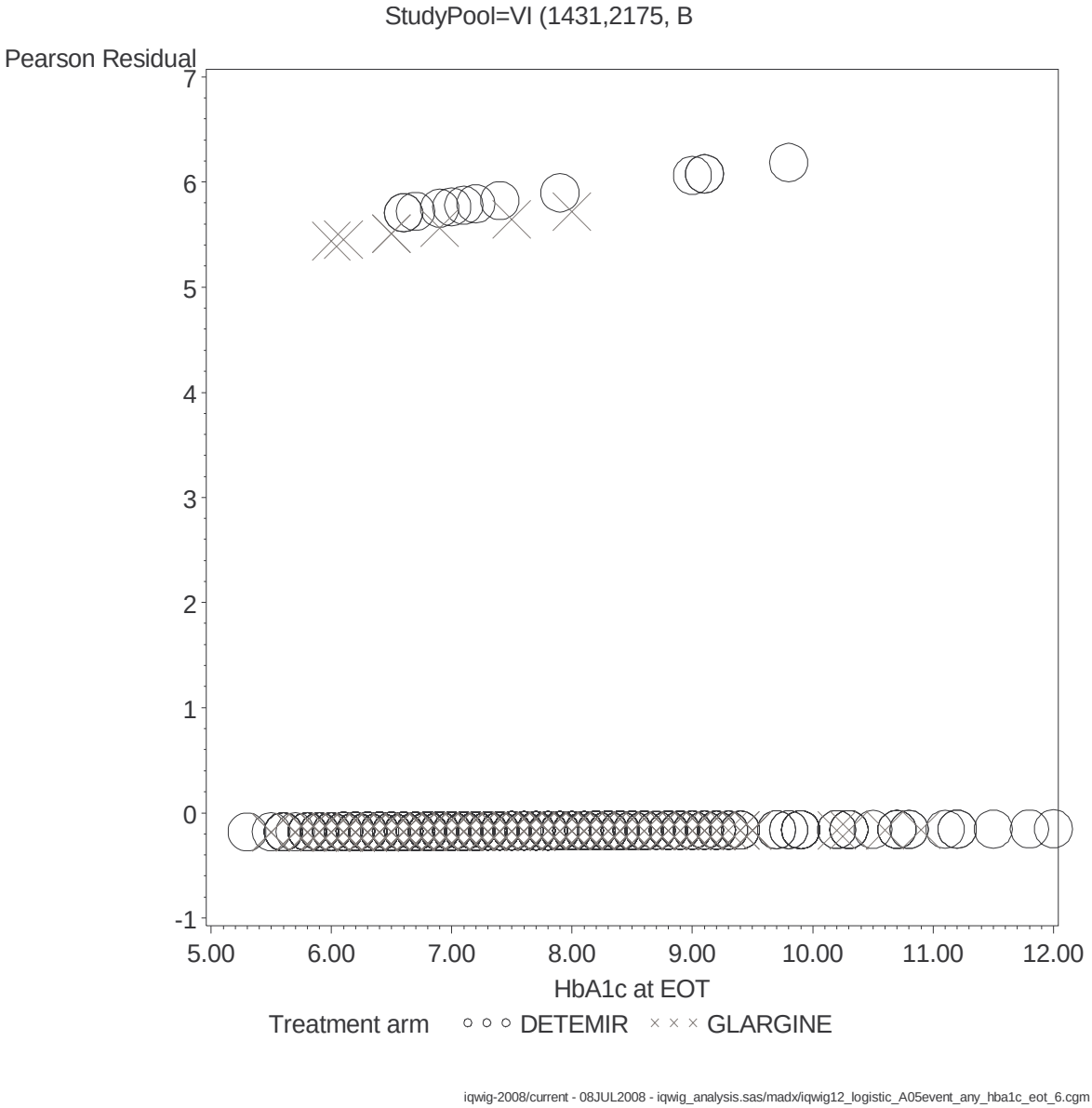


68: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A05event_any_hba1c_eot_5.cgm

69: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A05event_any_hba1c_eot_6.cgm

70: Logistic regression: residuals for endpoint A05event_any controlling for hba1c_eot

71: Logistic regression for endpoint A06event_any controlling for hba1c_change

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              474
       2             1                3

Probability modeled is A06event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	38.395	41.796
SC	42.562	54.299
-2 Log L	36.395	35.796

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.5981	2	0.7415
Score	0.6737	2	0.7140
Wald	0.6682	2	0.7160

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0003	0.9865
hba1c_change	1	0.6682	0.4137

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-5.2760	0.7249	52.9762	<.0001
TRTSEQ_LB	DETEMIR	1	0.0104	0.6151	0.0003	0.9865
hba1c_change		1	-0.3881	0.4748	0.6682	0.4137

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	1.021	0.092	11.381
hba1c_change		0.678	0.267	1.720

Association of Predicted Probabilities and Observed Responses

Percent Concordant	53.9	Somers' D	0.327
Percent Discordant	21.2	Gamma	0.436
Percent Tied	25.0	Tau-a	0.004
Pairs	1422	c	0.664

```
Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         467
Number of Observations Used         416


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1     0                    415
       2     1                      1

Probability modeled is A06event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                   Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.059	17.861
SC	20.090	29.953
-2 Log L	14.059	11.861

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.1980	2	0.3332
Score	1.9983	2	0.3682
Wald	0.0045	2	0.9977

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0037	0.9513
hba1c_change	1	0.0008	0.9775

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-10.4550	90.9908	0.0132	0.9085
TRTSEQ_LB	DETEMIR	1	-5.5578	90.9870	0.0037	0.9513
hba1c_change		1	0.0247	0.8756	0.0008	0.9775

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	<0.001	<0.001	>999.999
hba1c_change		1.025	0.184	5.702

Association of Predicted Probabilities and Observed Responses

Percent Concordant	66.7	Somers' D	0.667
Percent Discordant	0.0	Gamma	1.000
Percent Tied	33.3	Tau-a	0.003
Pairs	415	c	0.834

```
Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              460
       2             1                1

Probability modeled is A06event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.265	18.465
SC	20.398	30.865
-2 Log L	14.265	12.465

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.8000	2	0.4066
Score	1.3803	2	0.5015
Wald	0.4091	2	0.8150

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0030	0.9562
hba1c_change	1	0.4061	0.5239

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-12.1803	95.1950	0.0164	0.8982
TRTSEQ_LB	DETEMIR	1	-5.2264	95.1566	0.0030	0.9562
hba1c_change		1	-0.7499	1.1768	0.4061	0.5239

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=III (1530, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	<0.001	<0.001	>999.999
hba1c_change	0.472	0.047	4.742

Association of Predicted Probabilities and Observed Responses

Percent Concordant	80.0	Somers' D	0.700
Percent Discordant	10.0	Gamma	0.778
Percent Tied	10.0	Tau-a	0.003
Pairs	460	c	0.850

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A06event_any	A06: Major, nocturnal
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	271
Number of Observations Used	265

Response Profile

Ordered Value	A06event_any	Total Frequency
1	0	264
2	1	1

Probability modeled is A06event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration.
Validity of the model fit is questionable.

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	15.156	17.024
SC	18.735	27.763
-2 Log L	13.156	11.024

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.1321	2	0.3444
Score	1.7405	2	0.4189
Wald	1.0496	2	0.5917

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0033	0.9541
hba1c_change	1	1.0463	0.3064

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-9.6240	89.0096	0.0117	0.9139
TRTSEQ_LB	DETEMIR	1	-5.1227	89.0096	0.0033	0.9541
hba1c_change		1	0.8284	0.8099	1.0463	0.3064

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Odds Ratio Estimates

Effect		Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR vs NPH	<0.001	<0.001	>999.999
hba1c_change		2.290	0.468	11.199

Association of Predicted Probabilities and Observed Responses

Percent Concordant	90.9	Somers' D	0.848
Percent Discordant	6.1	Gamma	0.875
Percent Tied	3.0	Tau-a	0.006
Pairs	264	c	0.924

```
Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         582
Number of Observations Used         543


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              537
       2             1                6

Probability modeled is A06event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	67.998	71.919
SC	72.295	84.810
-2 Log L	65.998	65.919

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0785	2	0.9615
Score	0.0776	2	0.9620
Wald	0.0776	2	0.9619

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0013	0.9711
hba1c_change	1	0.0765	0.7821

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-4.6550	0.7309	40.5644	<.0001
TRTSEQ_LB	DETEMIR	1	0.0149	0.4106	0.0013	0.9711
hba1c_change		1	-0.1068	0.3863	0.0765	0.7821

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.030	0.206	5.151
hba1c_change	0.899	0.421	1.916

Association of Predicted Probabilities and Observed Responses

Percent Concordant	29.4	Somers' D	0.116
Percent Discordant	17.8	Gamma	0.247
Percent Tied	52.8	Tau-a	0.003
Pairs	3222	c	0.558

```
Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              670
       2             1               14

Probability modeled is A06event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	138.601	142.109
SC	143.129	155.693
-2 Log L	136.601	136.109

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.4913	2	0.7822
Score	0.4912	2	0.7822
Wald	0.4937	2	0.7813

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0764	0.7822
hba1c_change	1	0.4588	0.4982

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-3.6829	0.3613	103.9284	<.0001
TRTSEQ_LB	DETEMIR	1	-0.0786	0.2846	0.0764	0.7822
hba1c_change		1	0.1609	0.2376	0.4588	0.4982

Logistic regression for endpoint A06event_any controlling for hba1c_change

StudyPool=VI (1431,2175, B

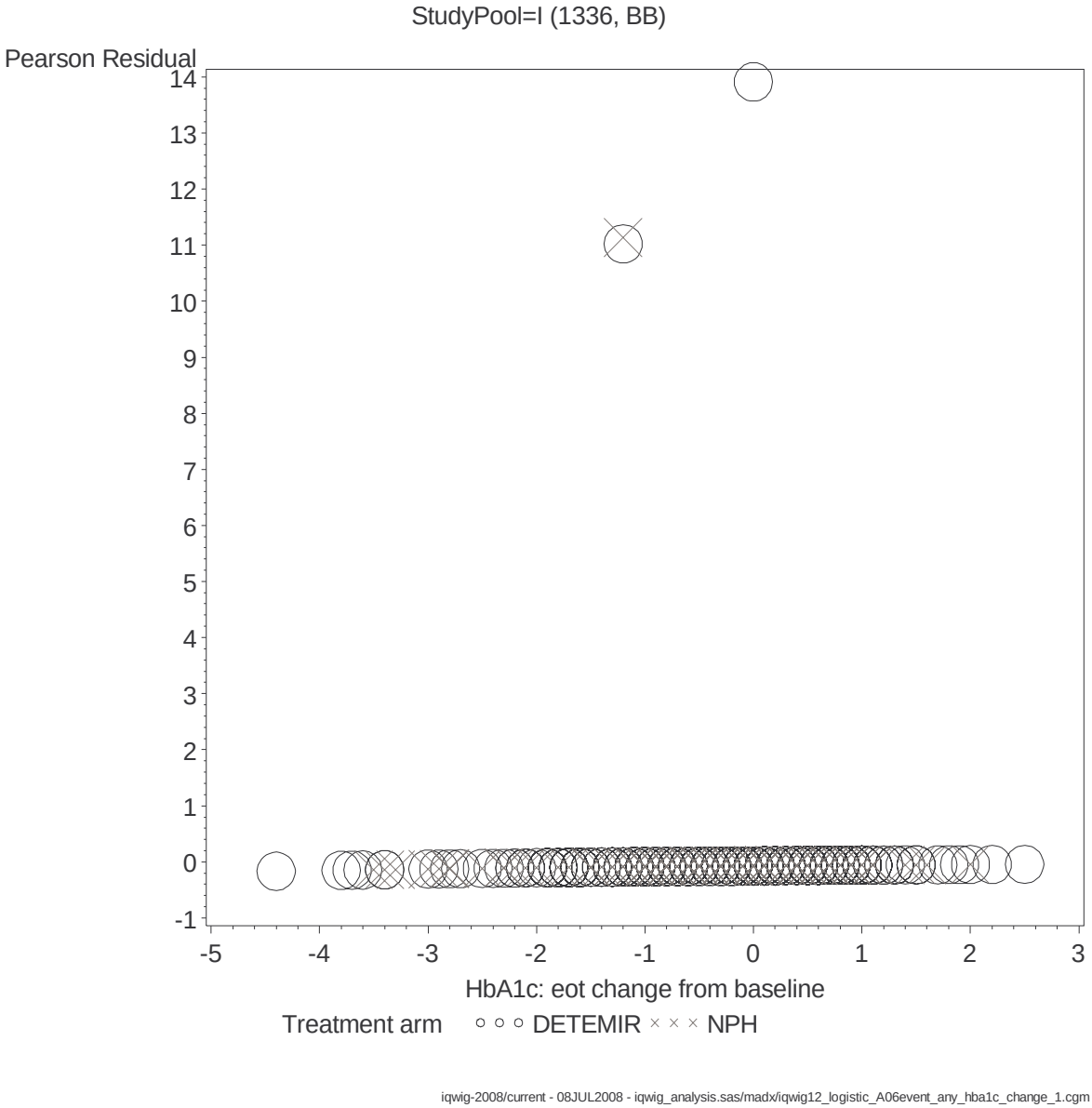
The LOGISTIC Procedure

Odds Ratio Estimates

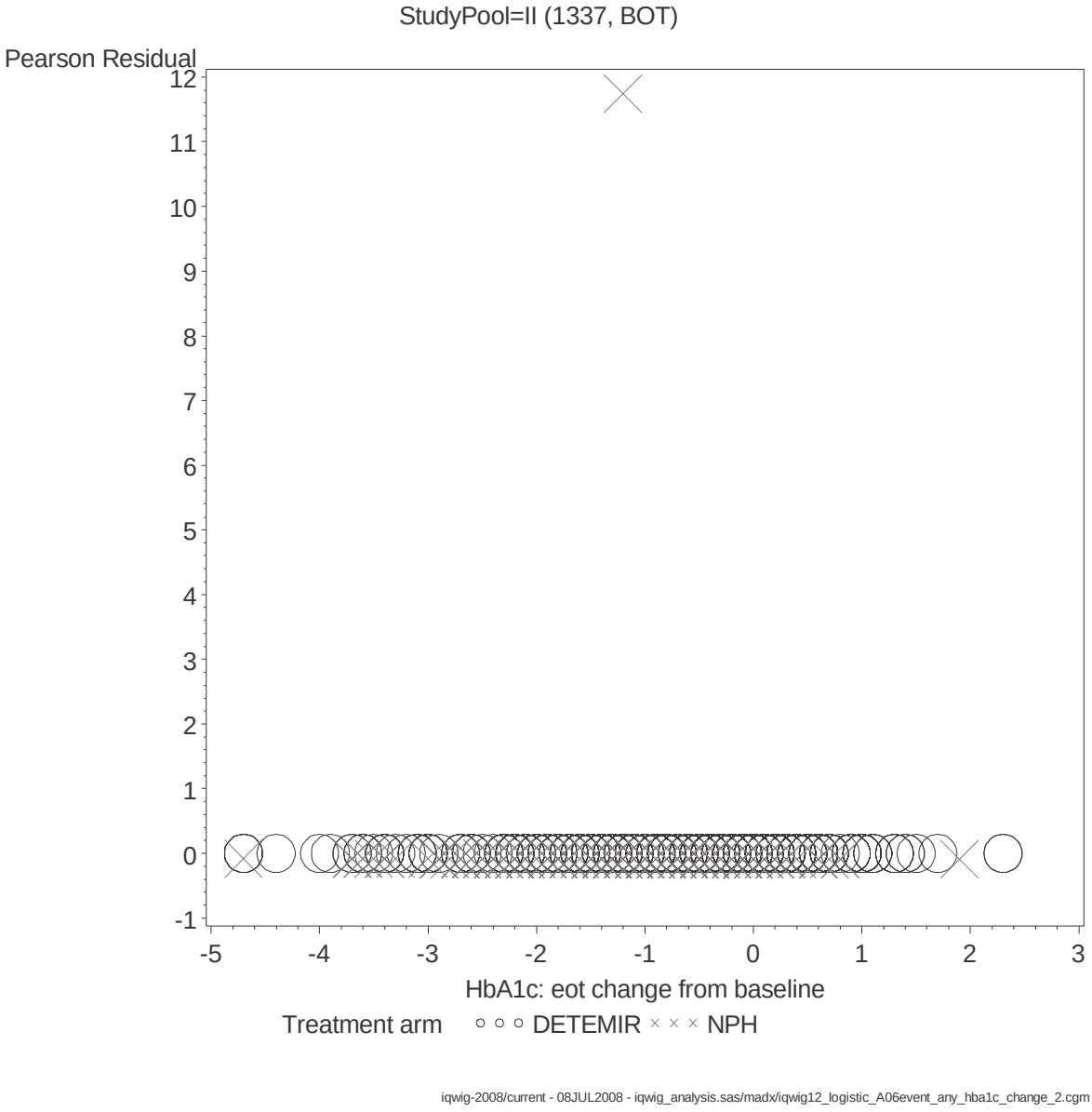
Effect			Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB	DETEMIR	vs GLARGINE	0.854	0.280	2.607
hba1c_change			1.175	0.737	1.871

Association of Predicted Probabilities and Observed Responses

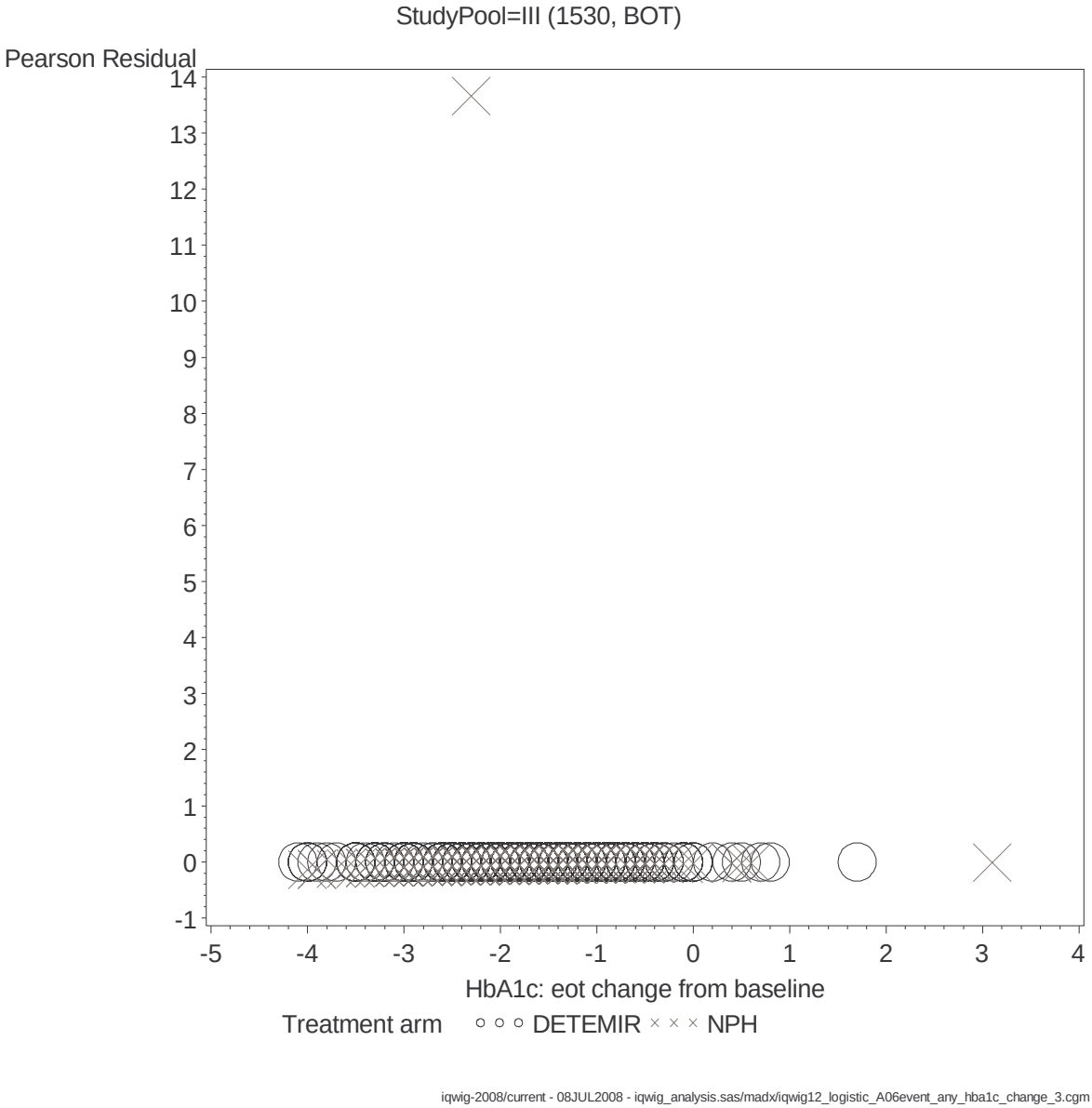
Percent Concordant	48.2	Somers' D	0.155
Percent Discordant	32.7	Gamma	0.192
Percent Tied	19.1	Tau-a	0.006
Pairs	9380	c	0.578



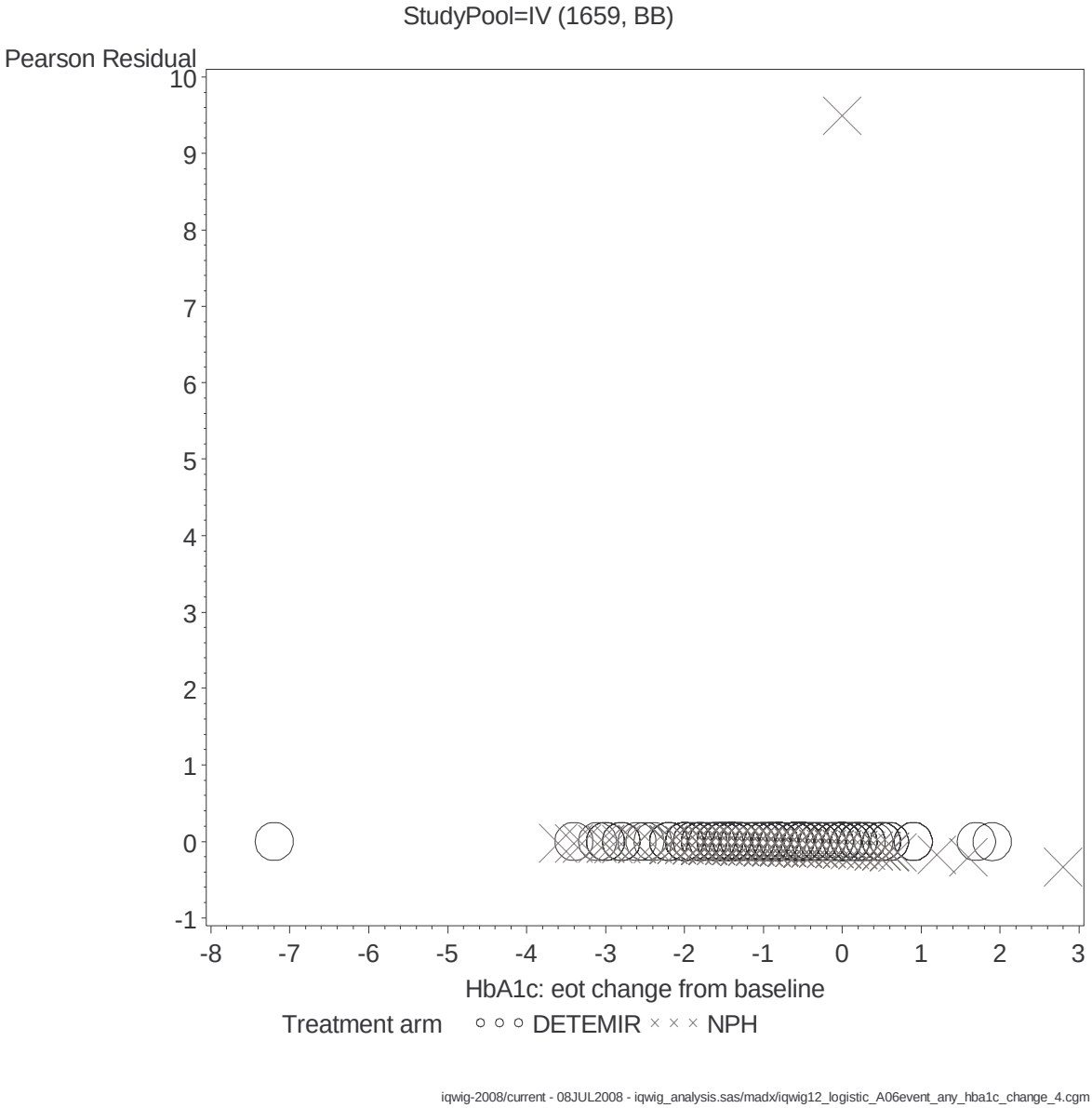
72: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change



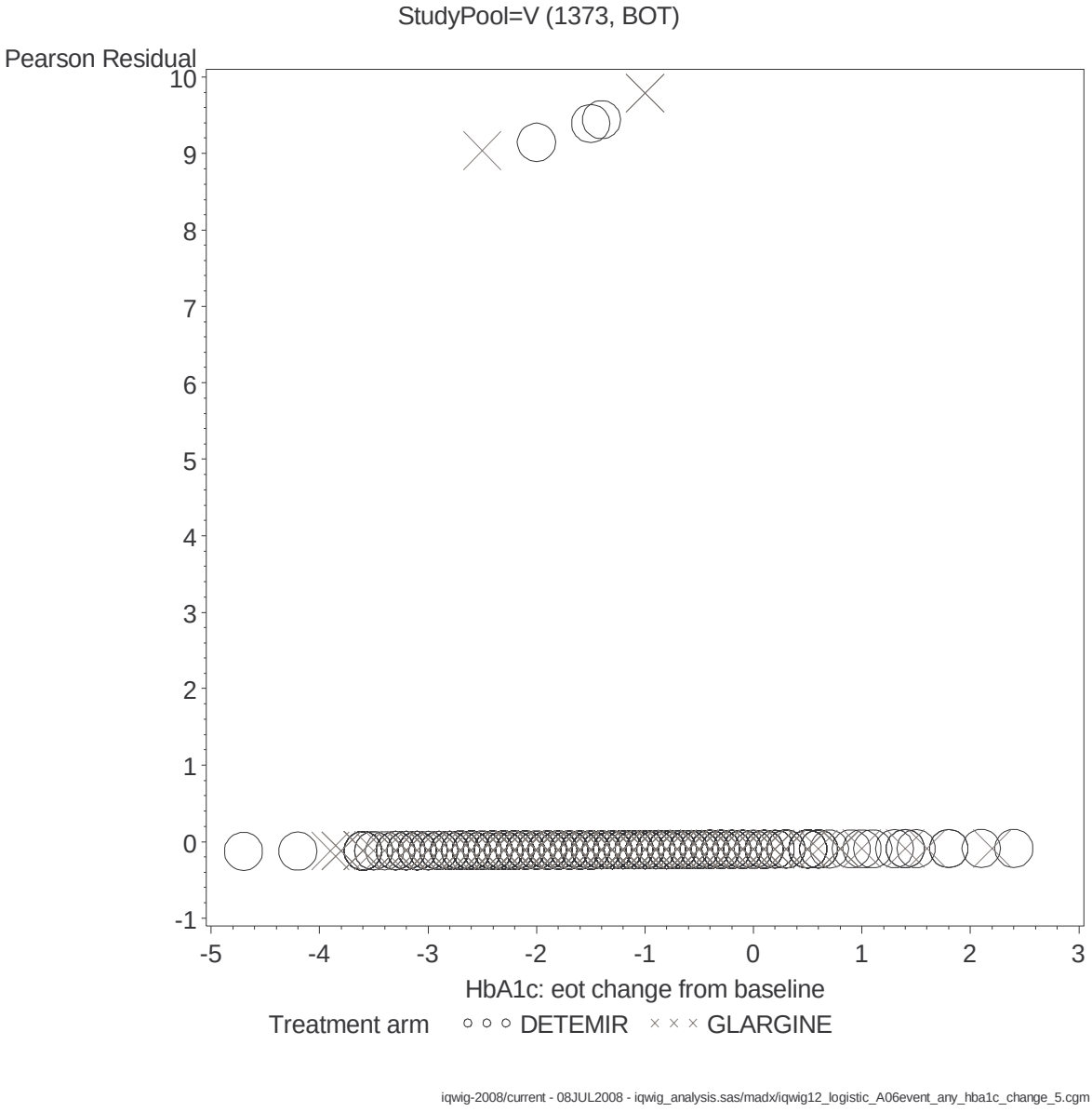
73: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change



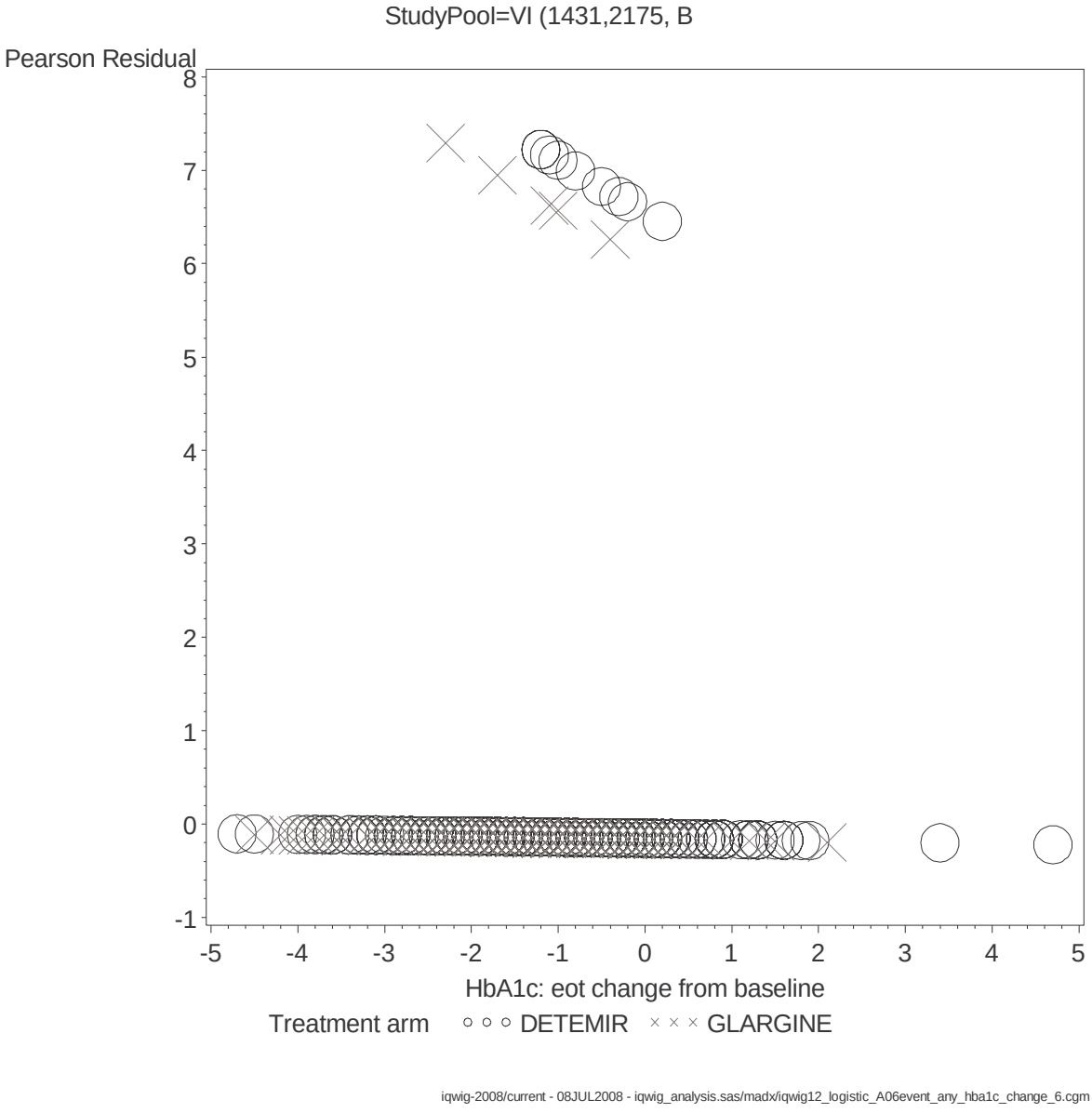
74: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change



75: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change



76: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change



77: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_change

78: Logistic regression for endpoint A06event_any controlling for hba1c_eot

```
StudyPool=I (1336, BB)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         505
Number of Observations Used         477


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              474
       2             1                3

Probability modeled is A06event_any=1.

NOTE: 28 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              NPH            -1


                    Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	38.395	42.134
SC	42.562	54.637
-2 Log L	36.395	36.134

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.2601	2	0.8781
Score	0.2749	2	0.8716
Wald	0.2754	2	0.8714

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0050	0.9434
hba1c_eot	1	0.2753	0.5998

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-7.1386	4.0879	3.0495	0.0808
TRTSEQ_LB DETEMIR	1	-0.0440	0.6190	0.0050	0.9434
hba1c_eot	1	0.2725	0.5193	0.2753	0.5998

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=I (1336, BB)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs NPH	0.916	0.081	10.367
hba1c_eot	1.313	0.475	3.634

Association of Predicted Probabilities and Observed Responses

Percent Concordant	38.1	Somers' D	0.124
Percent Discordant	25.7	Gamma	0.194
Percent Tied	36.1	Tau-a	0.002
Pairs	1422	c	0.562

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure

Model Information

Data Set	WORK.ANALYSIS_IQWIG	
Response Variable	A06event_any	A06: Major, nocturnal
Number of Response Levels	2	
Model	binary logit	
Optimization Technique	Fisher's scoring	

Number of Observations Read	467
Number of Observations Used	416

Response Profile

Ordered Value	A06event_any	Total Frequency
1	0	415
2	1	1

Probability modeled is A06event_any=1.

NOTE: 51 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information

Class	Value	Design Variables
TRTSEQ_LB	DETEMIR	1
	NPH	-1

Model Convergence Status

Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration.
Validity of the model fit is questionable.

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.059	17.668
SC	20.090	29.760
-2 Log L	14.059	11.668

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.3908	2	0.3026
Score	2.1599	2	0.3396
Wald	0.2068	2	0.9018

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0042	0.9483
hba1c_eot	1	0.2028	0.6524

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-13.9054	87.1468	0.0255	0.8732
TRTSEQ_LB DETEMIR	1	-5.6309	86.7823	0.0042	0.9483
hba1c_eot	1	0.4006	0.8894	0.2028	0.6524

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=II (1337, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           1.493       0.261       8.532


Association of Predicted Probabilities and Observed Responses

Percent Concordant     81.0    Somers' D    0.711
Percent Discordant      9.9    Gamma        0.782
Percent Tied            9.2    Tau-a        0.003
Pairs                   415    c            0.855
```

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         475
Number of Observations Used         461


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              460
       2             1                1

Probability modeled is A06event_any=1.

NOTE: 14 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	16.265	18.706
SC	20.398	31.106
-2 Log L	14.265	12.706

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1.5587	2	0.4587
Score	1.1479	2	0.5633
Wald	0.1658	2	0.9204

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0027	0.9586
hba1c_eot	1	0.1630	0.6864

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-6.2916	101.9	0.0038	0.9507
TRTSEQ_LB DETEMIR	1	-5.2574	101.3	0.0027	0.9586
hba1c_eot	1	-0.6640	1.6445	0.1630	0.6864

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=III (1530, BOT)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           0.515       0.021      12.924


Association of Predicted Probabilities and Observed Responses

Percent Concordant     74.6    Somers' D    0.663
Percent Discordant      8.3    Gamma        0.801
Percent Tied           17.2    Tau-a        0.003
Pairs                   460    c            0.832
```

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         271
Number of Observations Used         265


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              264
       2             1                1

Probability modeled is A06event_any=1.

NOTE: 6 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class          Value         Variables

TRTSEQ_LB      DETEMIR        1
               NPH           -1


                    Model Convergence Status

       Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on
the last maximum likelihood iteration.
         Validity of the model fit is questionable.
```

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	15.156	17.024
SC	18.735	27.764
-2 Log L	13.156	11.024

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	2.1314	2	0.3445
Score	1.5605	2	0.4583
Wald	0.7329	2	0.6932

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0041	0.9486
hba1c_eot	1	0.7303	0.3928

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.7218	83.2614	0.0004	0.9835
TRTSEQ_LB DETEMIR	1	-5.3369	82.8612	0.0041	0.9486
hba1c_eot	1	-1.1600	1.3574	0.7303	0.3928

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=IV (1659, BB)

The LOGISTIC Procedure
WARNING: The validity of the model fit is questionable.

                      Odds Ratio Estimates

                                   Point          95% Wald
Effect                          Estimate      Confidence Limits

TRTSEQ_LB DETEMIR vs NPH           <0.001      <0.001    >999.999
hba1c_eot                           0.314       0.022       4.484


Association of Predicted Probabilities and Observed Responses

Percent Concordant     89.0    Somers' D    0.784
Percent Discordant     10.6    Gamma        0.787
Percent Tied            0.4    Tau-a        0.006
Pairs                   264    c            0.892
```

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         582
Number of Observations Used         543


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              537
       2             1                6

Probability modeled is A06event_any=1.

NOTE: 39 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	67.998	71.995
SC	72.295	84.887
-2 Log L	65.998	65.995

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0024	2	0.9988
Score	0.0024	2	0.9988
Wald	0.0024	2	0.9988

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0011	0.9739
hba1c_eot	1	0.0014	0.9704

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.3812	3.0574	2.0534	0.1519
TRTSEQ_LB DETEMIR	1	0.0134	0.4106	0.0011	0.9739
hba1c_eot	1	-0.0157	0.4237	0.0014	0.9704

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=V (1373, BOT)

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
TRTSEQ_LB DETEMIR vs GLARGINE	1.027	0.205	5.137
hba1c_eot	0.984	0.429	2.259

NOTE: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are
indistinguishable when they are classified into intervals of length 0.002.

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

                            Model Information

Data Set                      WORK.ANALYSIS_IQWIG
Response Variable             A06event_any           A06: Major, nocturnal
Number of Response Levels     2
Model                         binary logit
Optimization Technique        Fisher's scoring


Number of Observations Read         704
Number of Observations Used         684


          Response Profile

 Ordered     A06event_          Total
   Value     any            Frequency

       1             0              670
       2             1               14

Probability modeled is A06event_any=1.

NOTE: 20 observations were deleted due to missing values for the response or explanatory variables.


      Class Level Information

                              Design
Class         Value          Variables

TRTSEQ_LB     DETEMIR         1
              GLARGINE       -1


                   Model Convergence Status

         Convergence criterion (GCONV=1E-8) satisfied.
```

Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	138.601	142.542
SC	143.129	156.126
-2 Log L	136.601	136.542

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	0.0582	2	0.9713
Score	0.0588	2	0.9710
Wald	0.0587	2	0.9711

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
TRTSEQ_LB	1	0.0414	0.8388
hba1c_eot	1	0.0224	0.8810

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.1237	1.8445	4.9983	0.0254
TRTSEQ_LB DETEMIR	1	-0.0576	0.2830	0.0414	0.8388
hba1c_eot	1	0.0366	0.2443	0.0224	0.8810

```
Logistic regression for endpoint A06event_any controlling for hba1c_eot

StudyPool=VI (1431,2175, B

The LOGISTIC Procedure

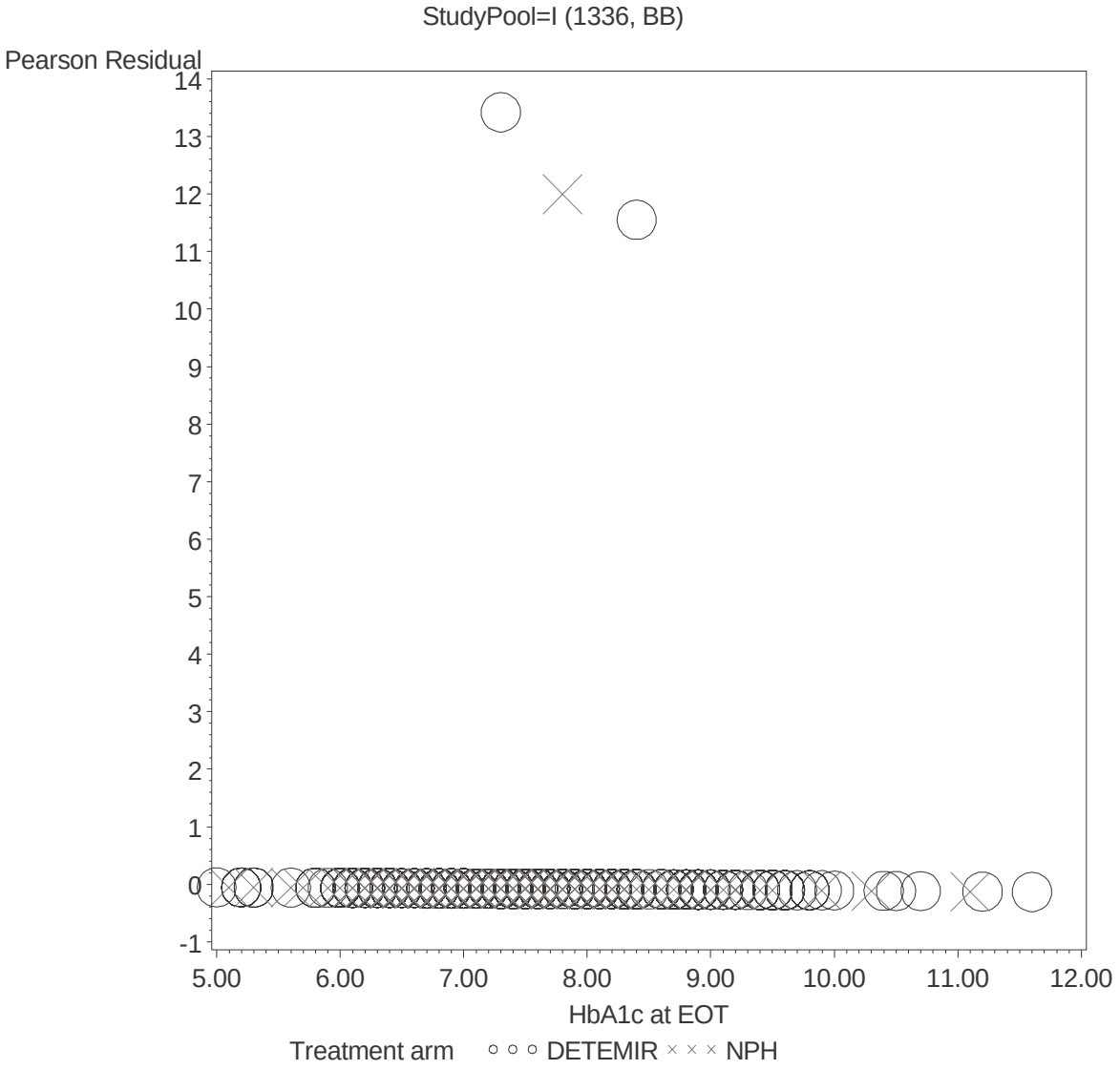
                     Odds Ratio Estimates

                                  Point          95% Wald
Effect                          Estimate      Confidence Limits

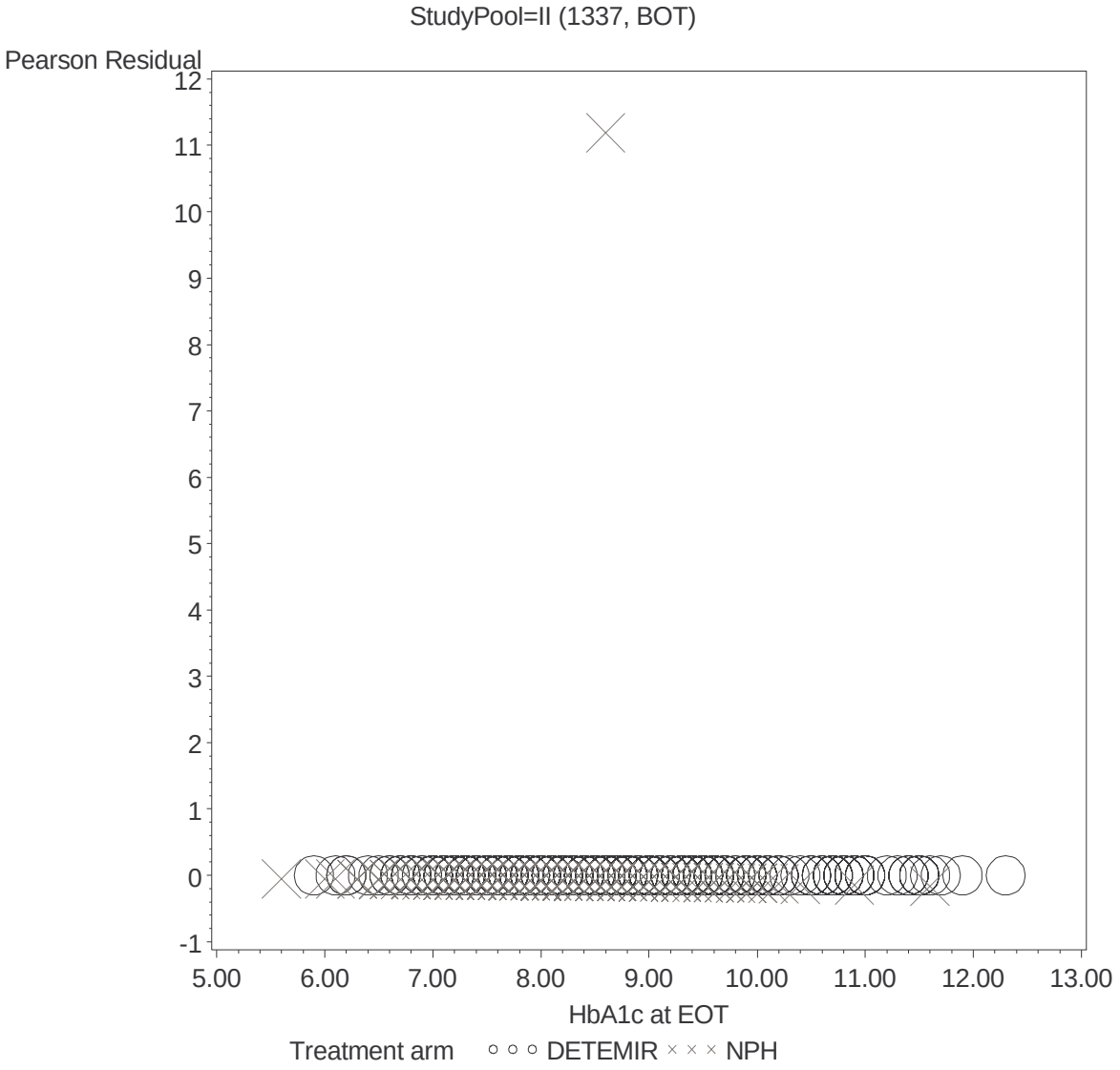
TRTSEQ_LB DETEMIR  vs GLARGINE       0.891       0.294       2.703
hba1c_eot                            1.037       0.643       1.674


Association of Predicted Probabilities and Observed Responses

Percent Concordant     36.4    Somers' D    0.122
Percent Discordant     24.3    Gamma        0.201
Percent Tied           39.3    Tau-a        0.005
Pairs                  9380    c            0.561
```

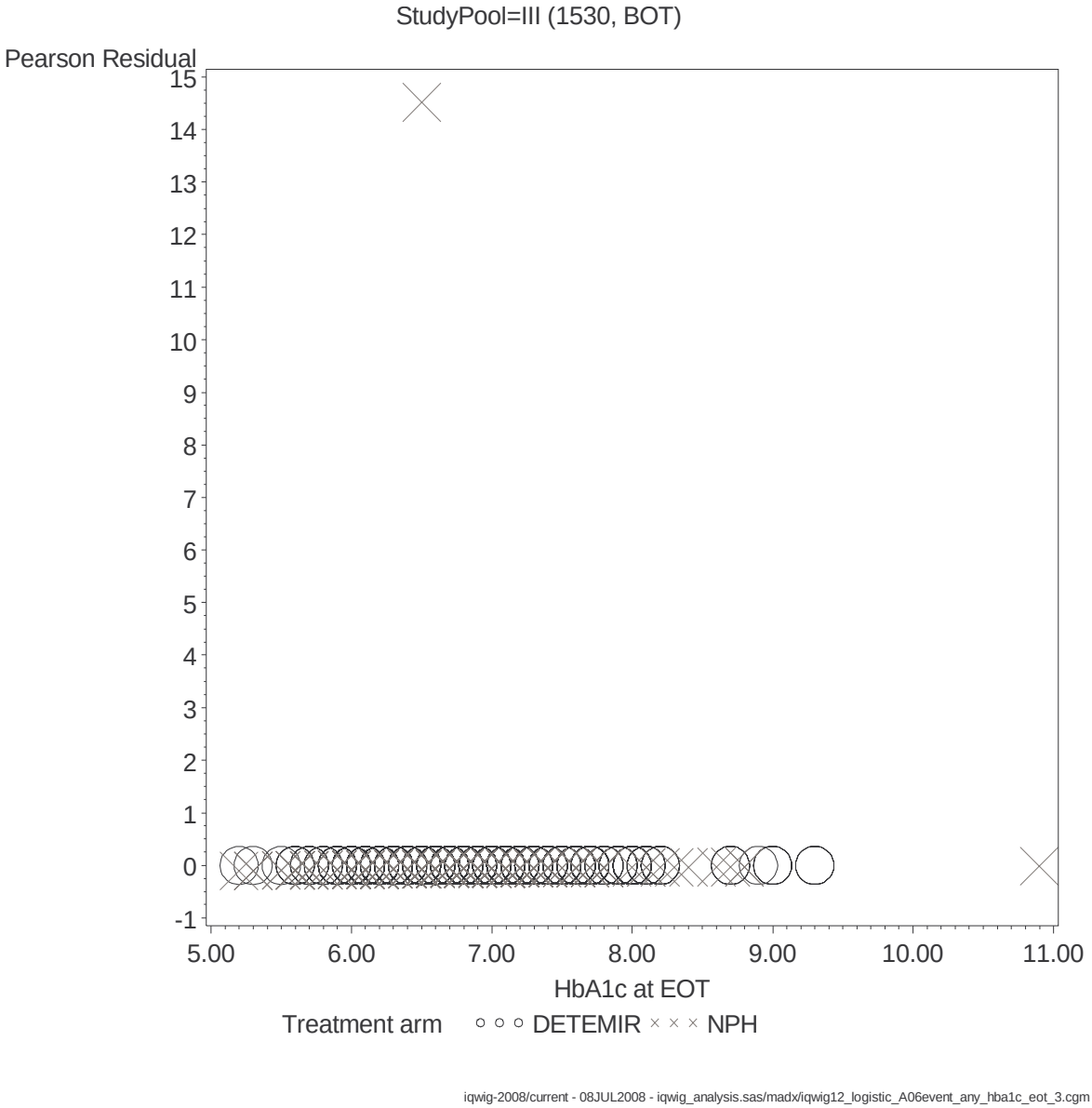


79: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot



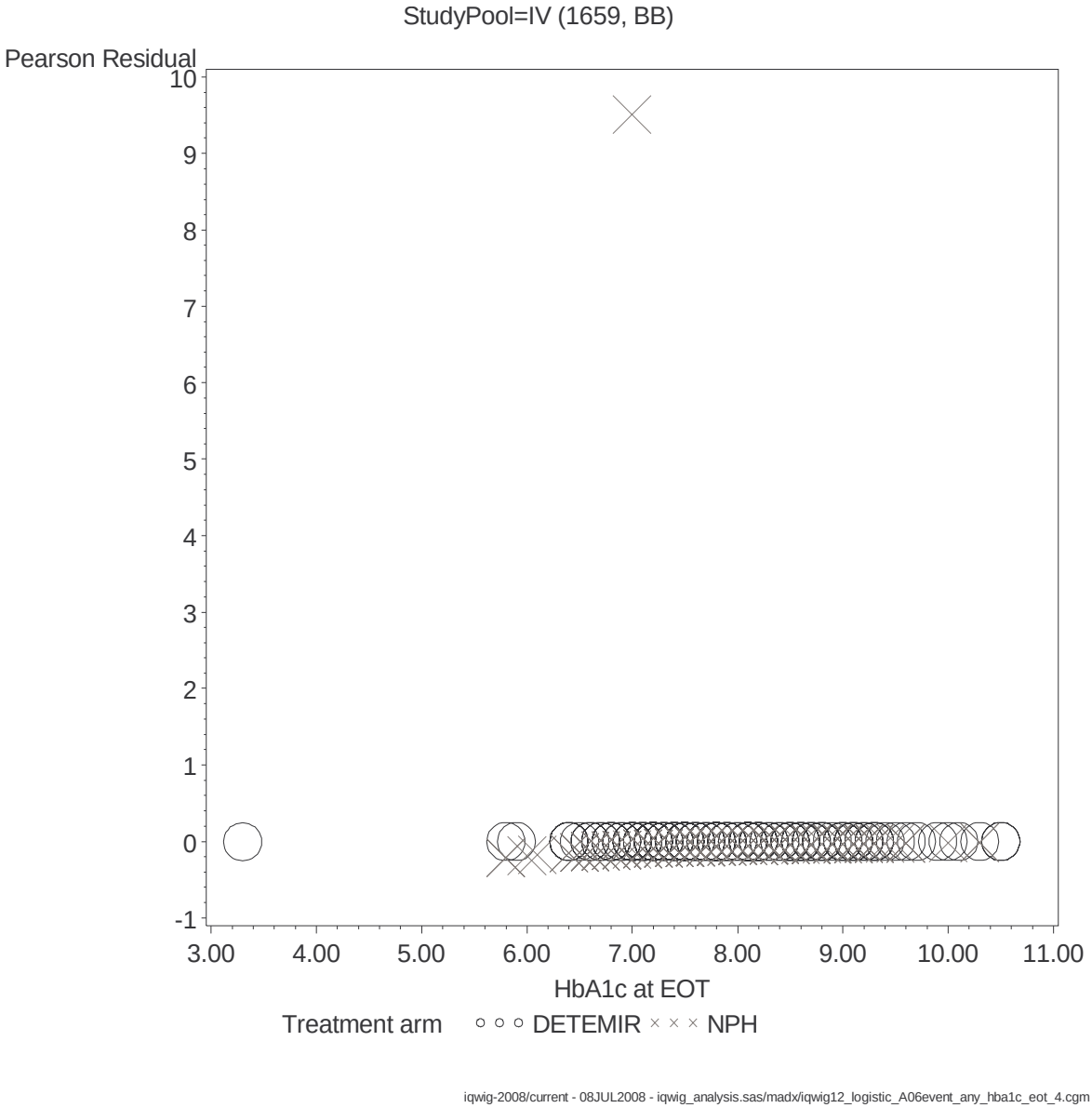
iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A06event_any_hba1c_eot_2.cgm

80: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot

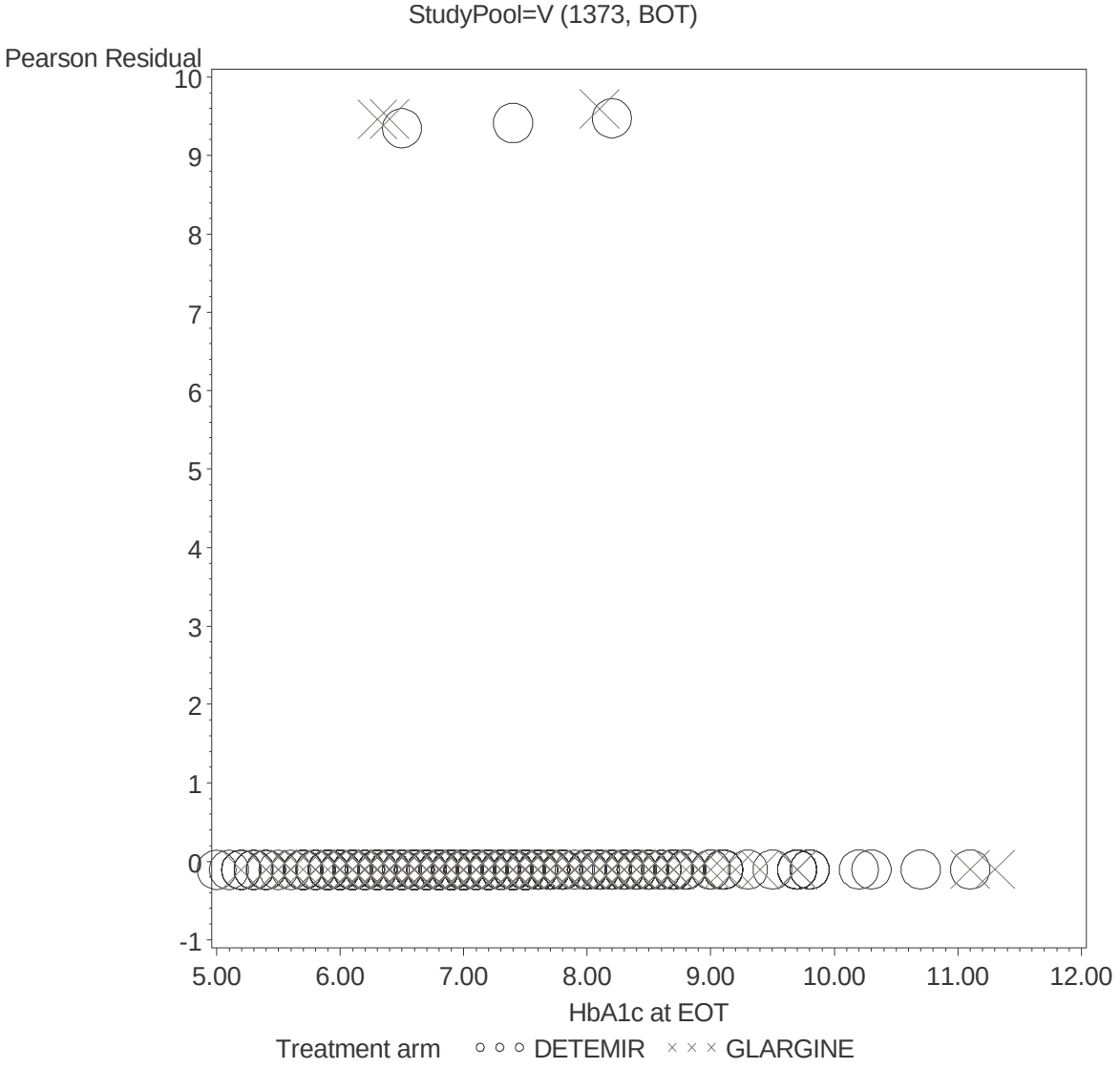


iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A06event_any_hba1c_eot_3.cgm

81: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot

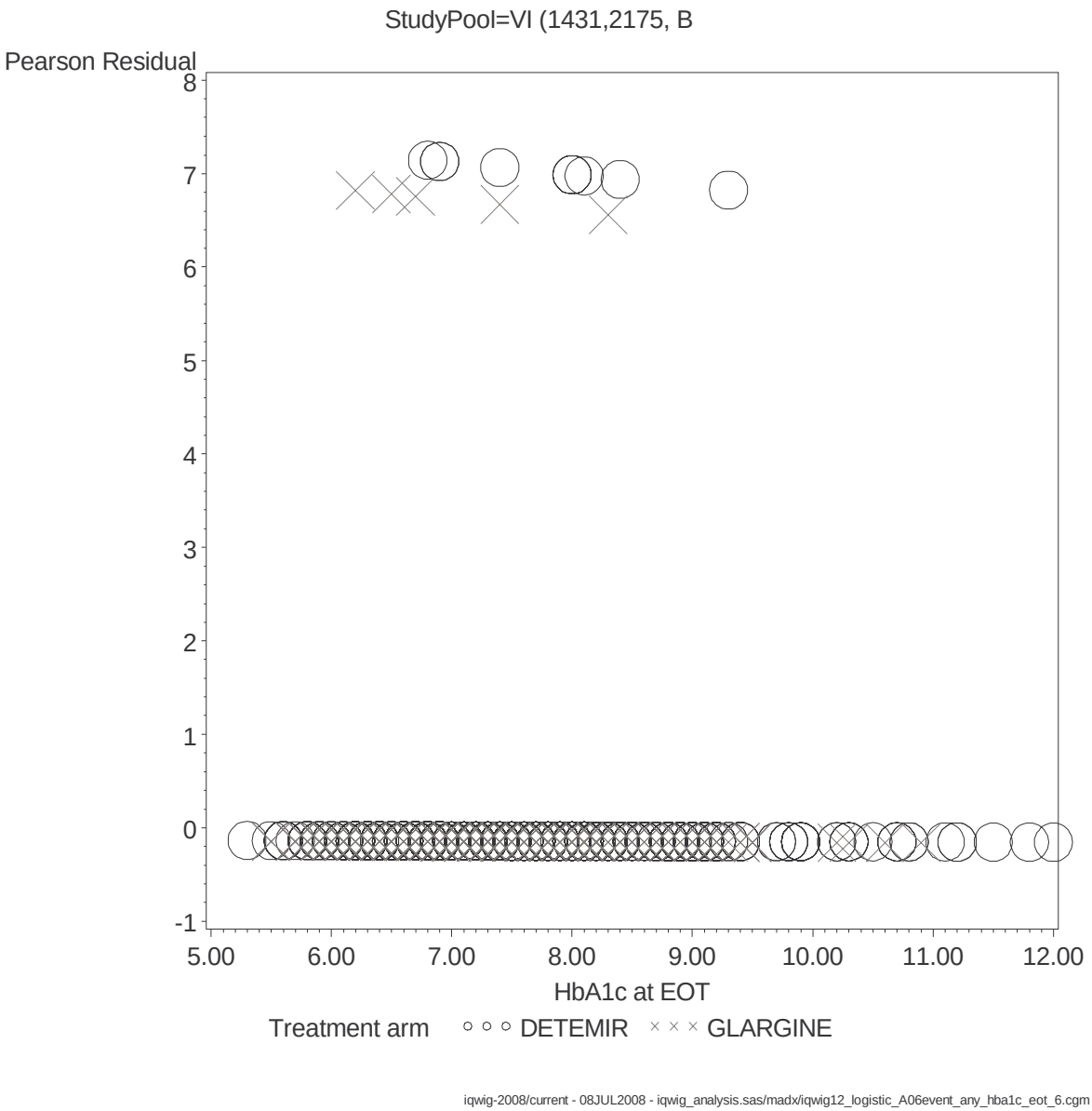


82: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot



iqwig-2008/current - 08JUL2008 - iqwig_analysis.sas/madx/iqwig12_logistic_A06event_any_hba1c_eot_5.cgm

83: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot



84: Logistic regression: residuals for endpoint A06event_any controlling for hba1c_eot

Statistical Documentation

List of Contents

Page

1: Description of analyses 2
2: Overview of endpoints by study(-pools) 3
3: Overview of endpoints by EOT HbA1c and by study(-pools) 8
4: HbA1c summary table: End of Trial change from baseline 18
5: HbA1c summary table: HbA1c at End of Trial 20
6: Summary of odds ratio estimates from logistic regression controlling for change in HbA1c 22
7: Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c 25

1: Description of analyses

Logistic regression adjusting for End of Trial (EOT) HbA1c or change from basline to EOT HbA1c was used to obtain the odds ratios and 95% confidence intervals. Last observation carried forward was used to obtain the HbA1c EOT values. (Note that EOT varies for the different studies, e.g. some studies are 26 weeks durations and others 52 weeks duration.) Only hypoglycaemic episodes where symptoms have been reported are included. Hypoglycaemic episodes with a missing or partial date were treated as treatment emergent. Studies have been pooled where appropriate and the analyses were run by study/study pool and treatment group. Analyses are grouped for each endpoint:

A01: Minor with symptom, daytime (not nocturnal), BG<2.0 mmol/L (36 mg/dl)
A02: Minor with symptom, daytime (not nocturnal), BG<3.9 mmol/L (70 mg/dl)
A03: Minor with symptom, nocturnal, BG<2.0 mmol/L (36 mg/dl)
A04: Minor with symptom, nocturnal, BG<3.9 mmol/L (70 mg/dl)
A05: Major with symptom, daytime (not nocturnal)
A06: Major with symptom, nocturnal
A11: Minor with symptom, BG<2.0 mmol/L (36 mg/dl)
A12: Minor with symptom, BG<3.9 mmol/L (70 mg/dl)
A13: Minor with symptom
A15: Major with symptom

2: Overview of endpoints by study(-pools)

	A01				A02			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	18	5.3	11	6.7	141	41.3	76	46.3
II (1337, BOT)	3	1.0	4	2.5	25	8.1	26	16.5
III (1530, BOT)	2	0.8	12	5.0	112	47.3	164	68.9
IV (1659, BB)	4	3.2	13	8.9	33	26.4	59	40.4
V (1373, BOT)	11	3.8	8	2.7	130	44.7	147	50.5
VI (1431,2175, B	13	2.8	7	3.0	273	58.3	147	62.3
BB (1336,1659)	22	4.7	24	7.7	174	37.3	135	43.5
BOT (1337,1530)	5	0.9	16	4.0	137	25.1	190	48.0

Overview of endpoints by study(-pools)

	A03				A04			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	5	1.5	9	5.5	53	15.5	42	25.6
II (1337, BOT)	.	.	3	1.9	18	5.8	23	14.6
III (1530, BOT)	2	0.8	6	2.5	59	24.9	98	41.2
IV (1659, BB)	4	3.2	10	6.8	18	14.4	43	29.5
V (1373, BOT)	4	1.4	5	1.7	73	25.1	81	27.8
VI (1431,2175, B	15	3.2	9	3.8	192	41.0	99	41.9
BB (1336,1659)	9	1.9	19	6.1	71	15.2	85	27.4
BOT (1337,1530)	2	0.4	9	2.3	77	14.1	121	30.6

Overview of endpoints by study(-pools)

	A05				A06			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	5	1.5	3	1.8	2	0.6	1	0.6
II (1337, BOT)	.	.	.	.	.	.	1	0.6
III (1530, BOT)	1	0.4	6	2.5	.	.	1	0.4
IV (1659, BB)	.	.	2	1.4	.	.	1	0.7
V (1373, BOT)	2	0.7	4	1.4	3	1.0	4	1.4
VI (1431,2175, B	9	1.9	6	2.5	9	1.9	5	2.1
BB (1336,1659)	5	1.1	5	1.6	2	0.4	2	0.6
BOT (1337,1530)	1	0.2	6	1.5	.	.	2	0.5

Overview of endpoints by study(-pools)

	A11				A12			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	21	6.2	18	11.0	152	44.6	83	50.6
II (1337, BOT)	3	1.0	7	4.4	35	11.3	40	25.3
III (1530, BOT)	4	1.7	14	5.9	124	52.3	180	75.6
IV (1659, BB)	6	4.8	18	12.3	38	30.4	66	45.2
V (1373, BOT)	13	4.5	12	4.1	147	50.5	167	57.4
VI (1431,2175, B	24	5.1	14	5.9	306	65.4	161	68.2
BB (1336,1659)	27	5.8	36	11.6	190	40.8	149	48.1
BOT (1337,1530)	7	1.3	21	5.3	159	29.1	220	55.6

Overview of endpoints by study(-pools)

	A13				A15			
	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	101	29.6	63	38.4	6	1.8	3	1.8
II (1337, BOT)	12	3.9	18	11.4	.	.	1	0.6
III (1530, BOT)	84	35.4	143	60.1	1	0.4	6	2.5
IV (1659, BB)	32	25.6	53	36.3	.	.	2	1.4
V (1373, BOT)	113	38.8	134	46.0	4	1.4	8	2.7
VI (1431,2175, B	232	49.6	130	55.1	16	3.4	11	4.7
BB (1336,1659)	133	28.5	116	37.4	6	1.3	5	1.6
BOT (1337,1530)	96	17.6	161	40.7	1	0.2	7	1.8

3: Overview of endpoints by EOT HbA1c and by study(-pools)

Study(-pool)	EOT HbA1c	A01 Detemir Freq	A01 Detemir Perc	A01 Comparator Freq	A01 Comparator Perc	A02 Detemir Freq	A02 Detemir Perc	A02 Comparator Freq	A02 Comparator Perc
I (1336, BB)	<6	2	6.5	2	16.7	7	22.6	5	41.7
	6-<7	7	9.5	2	4.5	36	48.6	24	54.5
	7-<8	7	5.2	5	6.9	62	46.3	38	52.8
	9-<10	.	.	.	.	5	17.9	2	22.2
	10-<11	.	.	.	.	1	25.0	.	.
	>=11	.	.	.	.	1	50.0	1	100.0
II (1337, BOT)	<6	.	.	.	.	1	3.0	1	5.0
	6-<7	.	.	.	.	1	6.7	4	26.7
	7-<8	.	.	1	2.3	3	4.3	7	16.3
	9-<10	.	.	1	4.5	10	15.4	2	9.1
	10-<11	2	7.7	.	.	4	15.4	.	.
III (1530, BOT)	<6	.	.	2	6.7	9	39.1	14	46.7
	6-<7	1	1.0	5	3.9	46	46.0	99	78.0
	7-<8	.	.	4	5.7	52	55.3	44	62.9
	9-<10	.	.	.	.	1	25.0	.	.
IV (1659, BB)	<6	.	.	.	.	1	20.0	2	33.3
	6-<7	.	.	1	3.8	.	.	10	38.5
	7-<8	1	2.0	5	8.9	15	29.4	24	42.9
	9-<10	1	6.3	1	6.3	5	31.3	4	25.0
	10-<11	.	.	1	33.3	1	20.0	2	66.7
V (1373, BOT)	<6	1	2.2	2	6.9	12	26.7	10	34.5
	6-<7	3	2.8	.	.	57	52.8	59	50.0
	7-<8	5	6.0	4	4.1	40	48.2	51	52.6
	9-<10	1	7.1	.	.	7	50.0	4	66.7
	10-<11	.	.	.	.	2	66.7	.	.
VI (1431,2175, B	<6	1	3.1	.	.	16	50.0	8	40.0
	6-<7	3	2.2	3	3.7	89	64.5	57	70.4
	7-<8	4	2.5	2	2.6	101	62.3	47	61.0
	9-<10	1	3.2	.	.	17	54.8	11	68.8
	10-<11	.	.	1	25.0	3	30.0	2	50.0
	>=11	1	16.7	.	.	1	16.7	.	.
BB (1336,1659)	<6	2	5.6	2	11.1	8	22.2	7	38.9
	6-<7	7	8.0	3	4.3	36	40.9	34	48.6
	7-<8	8	4.3	10	7.8	77	41.6	62	48.4

Overview of endpoints by EOT HbA1c and by study(-pools)

		A01				A02			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
BB (1336,1659)	9-<10	1	2.3	1	4.0	10	22.7	6	24.0
	10-<11	.	.	1	25.0	2	22.2	2	50.0
	>=11	.	.	.	.	1	50.0	1	100.0
BOT (1337,1530)	<6	.	.	2	4.0	10	17.9	15	30.0
	6-<7	1	0.9	5	3.5	47	40.9	103	72.5
	7-<8	.	.	5	4.4	55	33.7	51	45.1
	9-<10	.	.	1	4.5	11	15.9	2	9.1
	10-<11	2	7.7	.	.	4	15.4	.	.

Overview of endpoints by EOT HbA1c and by study(-pools)

		A03				A04			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	1	3.2	2	16.7	2	6.5	4	33.3
	6-<7	.	.	4	9.1	10	13.5	8	18.2
	7-<8	1	0.7	1	1.4	24	17.9	23	31.9
	9-<10	2	7.1	.	.	5	17.9	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	1	50.0	.	.
II (1337, BOT)	<6	.	.	.	.	2	6.1	.	.
	6-<7	.	.	.	.	2	13.3	.	.
	7-<8	.	.	.	.	3	4.3	6	14.0
	9-<10	.	.	2	9.1	5	7.7	5	22.7
	10-<11	.	.	.	.	2	7.7	1	16.7
III (1530, BOT)	<6	.	.	2	6.7	5	21.7	10	33.3
	6-<7	.	.	.	.	21	21.0	50	39.4
	7-<8	1	1.1	4	5.7	29	30.9	32	45.7
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	.	.	.	.	1	20.0	2	33.3
	6-<7	.	.	1	3.8	1	7.1	9	34.6
	7-<8	2	3.9	5	8.9	6	11.8	16	28.6
	9-<10	1	6.3	1	6.3	2	12.5	3	18.8
	10-<11	.	.	.	.	.	.	1	33.3
V (1373, BOT)	<6	3	6.7	1	3.4	12	26.7	7	24.1
	6-<7	.	.	2	1.7	25	23.1	33	28.0
	7-<8	1	1.2	1	1.0	24	28.9	32	33.0
	9-<10	.	.	.	.	2	14.3	1	16.7
	10-<11	.	.	.	.	1	33.3	.	.
VI (1431,2175, B	<6	.	.	2	10.0	9	28.1	6	30.0
	6-<7	4	2.9	3	3.7	67	48.6	35	43.2
	7-<8	7	4.3	2	2.6	73	45.1	36	46.8
	9-<10	1	3.2	.	.	11	35.5	4	25.0
	10-<11	.	.	1	25.0	.	.	1	25.0
	>=11	.	.	.	.	1	16.7	.	.
BB (1336,1659)	<6	1	2.8	2	11.1	3	8.3	6	33.3
	6-<7	.	.	5	7.1	11	12.5	17	24.3
	7-<8	3	1.6	6	4.7	30	16.2	39	30.5

Overview of endpoints by EOT HbA1c and by study(-pools)

		A03				A04			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
BB (1336,1659)	9-<10	3	6.8	1	4.0	7	15.9	3	12.0
	10-<11	.	.	.	.	.	.	1	25.0
	>=11	.	.	.	.	1	50.0	.	.
BOT (1337,1530)	<6	.	.	2	4.0	7	12.5	10	20.0
	6-<7	.	.	.	.	23	20.0	50	35.2
	7-<8	1	0.6	4	3.5	32	19.6	38	33.6
	9-<10	.	.	2	9.1	5	7.2	5	22.7
	10-<11	.	.	.	.	2	7.7	1	14.3

Overview of endpoints by EOT HbA1c and by study(-pools)

		A05				A06			
		Detemir		Comparator		Detemir		Comparator	
Study(-pool)	EOT HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	.	.	.	.	.	.	.	.
	6-<7	2	2.7	.	.	.	.	.	.
	7-<8	1	0.7	2	2.8	1	0.7	1	1.4
	9-<10	.	.	1	11.1	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
II (1337, BOT)	<6	.	.	.	.	.	.	.	.
	6-<7	.	.	.	.	.	.	.	.
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
III (1530, BOT)	<6	1	4.3	1	3.3	.	.	.	.
	6-<7	.	.	5	3.9	.	.	1	0.8
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	.	.	.	.	.	.	.	.
	6-<7	.	.	.	.	.	.	.	.
	7-<8	.	.	1	1.8	.	.	1	1.8
	9-<10	.	.	1	6.3	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
V (1373, BOT)	<6	.	.	1	3.4	.	.	1	3.4
	6-<7	1	0.9	.	.	1	0.9	2	1.7
	7-<8	1	1.2	.	.	1	1.2	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
VI (1431,2175, B	<6	.	.	.	.	.	.	.	.
	6-<7	1	0.7	4	4.9	3	2.2	3	3.7
	7-<8	5	3.1	1	1.3	1	0.6	1	1.3
	9-<10	3	9.7	.	.	1	3.2	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
BB (1336,1659)	<6	.	.	.	.	.	.	.	.
	6-<7	2	2.3	.	.	.	.	.	.
	7-<8	1	0.5	3	2.3	1	0.5	2	1.6

Overview of endpoints by EOT HbA1c and by study(-pools)

		A05				A06			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
BB (1336,1659)	9-<10	.	.	2	8.0	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.
	>=11	.	.	.	.	.	.	.	.
BOT (1337,1530)	<6	1	1.8	1	2.0	.	.	.	.
	6-<7	.	.	5	3.5	.	.	1	0.7
	7-<8	.	.	.	.	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
	10-<11	.	.	.	.	.	.	.	.

Overview of endpoints by EOT HbA1c and by study(-pools)

		A11				A12			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	2	6.5	2	16.7	7	22.6	6	50.0
	6-<7	7	9.5	6	13.6	39	52.7	24	54.5
	7-<8	7	5.2	6	8.3	66	49.3	42	58.3
	9-<10	2	7.1	.	.	7	25.0	2	22.2
	10-<11	.	.	.	.	1	25.0	.	.
	>=11	.	.	.	.	1	50.0	1	100.0
II (1337, BOT)	<6	.	.	.	.	2	6.1	1	5.0
	6-<7	.	.	.	.	3	20.0	4	26.7
	7-<8	.	.	1	2.3	5	7.2	12	27.9
	9-<10	.	.	3	13.6	12	18.5	6	27.3
	10-<11	2	7.7	.	.	4	15.4	1	16.7
III (1530, BOT)	<6	.	.	2	6.7	9	39.1	15	50.0
	6-<7	1	1.0	5	3.9	51	51.0	107	84.3
	7-<8	1	1.1	6	8.6	58	61.7	50	71.4
	9-<10	.	.	.	.	1	25.0	.	.
IV (1659, BB)	<6	.	.	.	.	1	20.0	3	50.0
	6-<7	.	.	1	3.8	1	7.1	14	53.8
	7-<8	2	3.9	8	14.3	17	33.3	25	44.6
	9-<10	2	12.5	2	12.5	5	31.3	5	31.3
	10-<11	.	.	1	33.3	1	20.0	2	66.7
V (1373, BOT)	<6	3	6.7	2	6.9	18	40.0	12	41.4
	6-<7	3	2.8	2	1.7	59	54.6	67	56.8
	7-<8	5	6.0	5	5.2	43	51.8	60	61.9
	9-<10	1	7.1	.	.	8	57.1	4	66.7
	10-<11	.	.	.	.	2	66.7	.	.
VI (1431,2175, B	<6	1	3.1	2	10.0	19	59.4	11	55.0
	6-<7	7	5.1	5	6.2	103	74.6	59	72.8
	7-<8	8	4.9	4	5.2	114	70.4	52	67.5
	9-<10	2	6.5	.	.	17	54.8	11	68.8
	10-<11	.	.	1	25.0	3	30.0	2	50.0
	>=11	1	16.7	.	.	1	16.7	.	.
BB (1336,1659)	<6	2	5.6	2	11.1	8	22.2	9	50.0
	6-<7	7	8.0	7	10.0	40	45.5	38	54.3
	7-<8	9	4.9	14	10.9	83	44.9	67	52.3

Overview of endpoints by EOT HbA1c and by study(-pools)

		A11				A12			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
BB (1336,1659)	9-<10	4	9.1	2	8.0	12	27.3	7	28.0
	10-<11	.	.	1	25.0	2	22.2	2	50.0
	>=11	.	.	.	.	1	50.0	1	100.0
BOT (1337,1530)	<6	.	.	2	4.0	11	19.6	16	32.0
	6-<7	1	0.9	5	3.5	54	47.0	111	78.2
	7-<8	1	0.6	7	6.2	63	38.7	62	54.9
	9-<10	.	.	3	13.6	13	18.8	6	27.3
	10-<11	2	7.7	.	.	4	15.4	1	14.3

Overview of endpoints by EOT HbA1c and by study(-pools)

		A13				A15			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
I (1336, BB)	<6	5	16.1	4	33.3	.	.	.	.
	6-<7	23	31.1	17	38.6	2	2.7	.	.
	7-<8	45	33.6	32	44.4	2	1.5	2	2.8
	9-<10	6	21.4	2	22.2	.	.	1	11.1
	10-<11	.	.	.	.	.	.	.	.
	>=11	1	50.0	.	.	.	.	.	.
II (1337, BOT)	<6	1	3.0	.	.	.	.	.	.
	6-<7	.	.	2	13.3	.	.	.	.
	7-<8	1	1.4	2	4.7	.	.	.	.
	9-<10	2	3.1	5	22.7	.	.	.	.
	10-<11	3	11.5	1	16.7	.	.	.	.
III (1530, BOT)	<6	9	39.1	12	40.0	1	4.3	1	3.3
	6-<7	35	35.0	83	65.4	.	.	5	3.9
	7-<8	36	38.3	41	58.6	.	.	.	.
	9-<10	.	.	.	.	.	.	.	.
IV (1659, BB)	<6	1	20.0	2	33.3	.	.	.	.
	6-<7	1	7.1	11	42.3	.	.	.	.
	7-<8	13	25.5	18	32.1	.	.	1	1.8
	9-<10	4	25.0	4	25.0	.	.	1	6.3
	10-<11	1	20.0	2	66.7	.	.	.	.
V (1373, BOT)	<6	14	31.1	9	31.0	.	.	2	6.9
	6-<7	41	38.0	56	47.5	1	0.9	2	1.7
	7-<8	41	49.4	45	46.4	2	2.4	.	.
	9-<10	6	42.9	3	50.0	.	.	.	.
	10-<11	2	66.7	.	.	.	.	.	.
VI (1431,2175, B	<6	13	40.6	8	40.0	.	.	.	.
	6-<7	84	60.9	50	61.7	3	2.2	7	8.6
	7-<8	80	49.4	40	51.9	5	3.1	2	2.6
	9-<10	14	45.2	10	62.5	4	12.9	.	.
	10-<11	2	20.0	2	50.0	.	.	.	.
	>=11	1	16.7	.	.	.	.	.	.
BB (1336,1659)	<6	6	16.7	6	33.3	.	.	.	.
	6-<7	24	27.3	28	40.0	2	2.3	.	.
	7-<8	58	31.4	50	39.1	2	1.1	3	2.3

Overview of endpoints by EOT HbA1c and by study(-pools)

		A13				A15			
	EOT	Detemir		Comparator		Detemir		Comparator	
Study(-pool)	HbA1c	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
BB (1336,1659)	9-<10	10	22.7	6	24.0	.	.	2	8.0
	10-<11	1	11.1	2	50.0	.	.	.	.
	>=11	1	50.0	.	.	.	.	.	.
BOT (1337,1530)	<6	10	17.9	12	24.0	1	1.8	1	2.0
	6-<7	35	30.4	85	59.9	.	.	5	3.5
	7-<8	37	22.7	43	38.1	.	.	.	.
	9-<10	2	2.9	5	22.7	.	.	.	.
	10-<11	3	11.5	1	14.3	.	.	.	.

4: HbA1c summary table: End of Trial change from baseline

StudyPool	Statistic	Detemir	Comparator
I (1336, BB)	N	320	157
	No. missing	21	7
	Mean	-0.28	-0.39
	Standard deviation	1.04	1.02
	Minimum	-4.40	-3.50
	Maximum	2.50	2.00
II (1337, BOT)	N	277	139
	No. missing	32	19
	Mean	-0.78	-1.23
	Standard deviation	1.23	1.16
	Minimum	-4.70	-4.70
	Maximum	2.30	1.90
III (1530, BOT)	N	230	231
	No. missing	7	7
	Mean	-1.63	-1.74
	Standard deviation	0.92	0.91
	Minimum	-4.10	-4.00
	Maximum	1.70	3.10
IV (1659, BB)	N	123	142
	No. missing	2	4
	Mean	-0.96	-0.99
	Standard deviation	1.13	1.01
	Minimum	-7.20	-3.60
	Maximum	1.90	2.80
V (1373, BOT)	N	268	275
	No. missing	23	16
	Mean	-1.42	-1.47
	Standard deviation	1.14	1.03
	Minimum	-4.70	-3.90
	Maximum	2.40	2.20
VI (1431,2175, B	N	456	228
	No. missing	12	8
	Mean	-1.00	-1.29
	Standard deviation	1.15	1.10
	Minimum	-4.70	-4.50
	Maximum	4.70	2.10

HbA1c summary table: End of Trial change from baseline

StudyPool	Statistic	Detemir	Comparator
BB (1336,1659)	N	443	299
	No. missing	23	11
	Mean	-0.47	-0.67
	Standard deviation	1.11	1.05
	Minimum	-7.20	-3.60
	Maximum	2.50	2.80
BOT (1337,1530)	N	507	370
	No. missing	39	26
	Mean	-1.17	-1.55
	Standard deviation	1.18	1.04
	Minimum	-4.70	-4.70
	Maximum	2.30	3.10

5: HbA1c summary table: HbA1c at End of Trial

StudyPool	Statistic	Detemir	Comparator
I (1336, BB)	N	320	157
	No. missing	21	7
	Mean	7.60	7.38
	Standard deviation	1.07	0.93
	Minimum	5.00	5.10
	Maximum	11.60	11.10
II (1337, BOT)	N	277	139
	No. missing	32	19
	Mean	8.66	8.13
	Standard deviation	1.23	1.05
	Minimum	5.90	5.60
	Maximum	12.30	11.60
III (1530, BOT)	N	230	231
	No. missing	7	7
	Mean	6.96	6.78
	Standard deviation	0.72	0.70
	Minimum	5.20	5.20
	Maximum	9.30	10.90
IV (1659, BB)	N	123	142
	No. missing	2	4
	Mean	7.91	7.83
	Standard deviation	1.06	0.95
	Minimum	3.30	5.80
	Maximum	10.50	10.30
V (1373, BOT)	N	268	275
	No. missing	23	16
	Mean	7.19	7.14
	Standard deviation	1.04	0.91
	Minimum	5.00	5.20
	Maximum	11.10	11.30
VI (1431,2175, B	N	456	228
	No. missing	12	8
	Mean	7.52	7.32
	Standard deviation	1.12	1.03
	Minimum	5.30	5.50
	Maximum	12.00	10.90

HbA1c summary table: HbA1c at End of Trial

StudyPool	Statistic	Detemir	Comparator
BB (1336,1659)	N	443	299
	No. missing	23	11
	Mean	7.68	7.59
	Standard deviation	1.07	0.97
	Minimum	3.30	5.10
	Maximum	11.60	11.10
BOT (1337,1530)	N	507	370
	No. missing	39	26
	Mean	7.89	7.28
	Standard deviation	1.34	1.07
	Minimum	5.20	5.20
	Maximum	12.30	11.60

6: Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A01			A02			A03		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.82	0.37	1.85	0.88	0.60	1.30	0.22	0.06	0.74
II (1337, BOT)	0.33	0.07	1.57	0.43	0.23	0.79	0.00	0.00	6E134
III (1530, BOT)	0.17	0.04	0.79	0.40	0.27	0.58	0.39	0.07	2.03
IV (1659, BB)	0.32	0.10	1.01	0.52	0.31	0.87	0.44	0.13	1.44
V (1373, BOT)	1.63	0.62	4.27	0.86	0.61	1.20	0.81	0.22	3.08
VI (1431,2175, B	0.74	0.29	1.93	0.87	0.63	1.22	0.85	0.36	1.98
BB (1336,1659)	0.57	0.31	1.06	0.80	0.59	1.08	0.27	0.12	0.64
BOT (1337,1530)	0.22	0.08	0.63	0.38	0.29	0.51	0.17	0.03	0.81

Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A04			A05			A06		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.59	0.37	0.94	0.83	0.19	3.51	1.02	0.09	11.38
II (1337, BOT)	0.28	0.14	0.56	.	.	.	0.00	0.00	1E150
III (1530, BOT)	0.48	0.32	0.71	0.20	0.02	1.72	0.00	0.00	3E157
IV (1659, BB)	0.41	0.22	0.76	0.00	0.00	3E126	0.00	0.00	1E147
V (1373, BOT)	0.87	0.59	1.26	0.67	0.11	4.04	1.03	0.21	5.15
VI (1431,2175, B	1.00	0.72	1.38	0.75	0.26	2.15	0.85	0.28	2.61
BB (1336,1659)	0.51	0.36	0.73	0.67	0.19	2.34	0.68	0.09	4.87
BOT (1337,1530)	0.38	0.27	0.53	0.16	0.02	1.37	0.00	0.00	8E186

Summary of odds ratio estimates from logistic regression controlling for change in HbA1c

	A11			A12			A13			A15		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.54	0.27	1.06	0.86	0.59	1.27	0.70	0.47	1.04	0.99	0.24	4.02
II (1337, BOT)	0.18	0.04	0.73	0.34	0.20	0.58	0.26	0.12	0.58	0.00	0.00	1E150
III (1530, BOT)	0.29	0.09	0.92	0.34	0.23	0.51	0.36	0.25	0.53	0.20	0.02	1.72
IV (1659, BB)	0.34	0.13	0.89	0.53	0.32	0.88	0.61	0.36	1.03	0.00	0.00	3E126
V (1373, BOT)	1.22	0.54	2.78	0.81	0.58	1.14	0.79	0.56	1.11	0.67	0.19	2.42
VI (1431,2175, B	0.80	0.40	1.59	0.94	0.67	1.33	0.83	0.60	1.15	0.70	0.32	1.55
BB (1336,1659)	0.44	0.26	0.76	0.78	0.58	1.05	0.68	0.50	0.93	0.80	0.24	2.66
BOT (1337,1530)	0.23	0.10	0.56	0.34	0.25	0.45	0.33	0.24	0.45	0.13	0.02	1.08

7: Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A01			A02			A03		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.89	0.40	2.01	0.92	0.62	1.36	0.23	0.07	0.78
II (1337, BOT)	0.29	0.06	1.42	0.43	0.23	0.79	0.00	0.00	2E177
III (1530, BOT)	0.16	0.04	0.75	0.40	0.27	0.59	0.36	0.07	1.90
IV (1659, BB)	0.32	0.10	1.02	0.51	0.30	0.86	0.45	0.14	1.46
V (1373, BOT)	1.62	0.62	4.25	0.86	0.61	1.20	0.77	0.20	2.95
VI (1431,2175, B	0.88	0.35	2.26	0.89	0.64	1.23	0.84	0.36	1.95
BB (1336,1659)	0.60	0.33	1.11	0.81	0.60	1.09	0.28	0.12	0.66
BOT (1337,1530)	0.21	0.07	0.61	0.45	0.33	0.61	0.15	0.03	0.74

Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A04			A05			A06		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.58	0.36	0.93	0.78	0.18	3.34	0.92	0.08	10.37
II (1337, BOT)	0.31	0.15	0.61	.	.	.	0.00	0.00	7E142
III (1530, BOT)	0.47	0.31	0.69	0.22	0.03	1.93	0.00	0.00	7E167
IV (1659, BB)	0.41	0.22	0.76	0.00	0.00	1E123	0.00	0.00	3E136
V (1373, BOT)	0.87	0.59	1.26	0.62	0.10	3.81	1.03	0.21	5.14
VI (1431,2175, B	1.00	0.73	1.39	0.74	0.26	2.13	0.89	0.29	2.70
BB (1336,1659)	0.50	0.35	0.72	0.66	0.19	2.31	0.67	0.09	4.82
BOT (1337,1530)	0.43	0.31	0.60	0.21	0.02	1.87	0.00	0.00	6E185

Summary of odds ratio estimates from logistic regression controlling for EOT HbA1c

	A11			A12			A13			A15		
Study(-pool)	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI	Estimate	Lower CI	Upper CI
I (1336, BB)	0.58	0.29	1.14	0.90	0.61	1.33	0.72	0.48	1.08	0.95	0.23	3.89
II (1337, BOT)	0.14	0.03	0.61	0.36	0.21	0.61	0.23	0.10	0.52	0.00	0.00	7E142
III (1530, BOT)	0.27	0.09	0.85	0.34	0.23	0.52	0.37	0.25	0.54	0.22	0.03	1.93
IV (1659, BB)	0.34	0.13	0.89	0.53	0.32	0.88	0.60	0.36	1.02	0.00	0.00	1E123
V (1373, BOT)	1.23	0.54	2.79	0.81	0.58	1.14	0.79	0.56	1.11	0.65	0.18	2.34
VI (1431,2175, B	0.85	0.43	1.68	0.95	0.67	1.35	0.84	0.61	1.16	0.71	0.32	1.56
BB (1336,1659)	0.47	0.28	0.80	0.80	0.59	1.07	0.69	0.51	0.95	0.80	0.24	2.65
BOT (1337,1530)	0.21	0.08	0.51	0.39	0.29	0.52	0.38	0.28	0.52	0.16	0.02	1.33

A 1.7 Sanofi-Aventis Deutschland GmbH

Autoren:
Frick, Markus, Dr. med.
Hessel, Franz, Dr. med.
Knollmeyer, Johannes, Dr. med.

Adresse:
Dr. med. Johannes Knollmeyer
Sanofi-Aventis Deutschland GmbH
Industriepark Höchst
65926 Frankfurt

1. Einleitung

Am 28.03.2008 hat das IQWiG zu Stellungnahmen zum Vorbericht des Bewertungsauftrages A05-03 aufgefordert. Einleitend und in Übereinstimmung mit unserer Stellungnahme zum geänderten Berichtsplan stellen wir fest, dass die gesetzlichen Vorgaben des GKV-WSG weiterhin nicht erfüllt werden. Die gesetzlich vorgeschriebene Übereinstimmung mit internationalen Standards, die notwendige Transparenz und Partizipation sind bei den bisherigen Anhörungsverfahren in folgenden Punkten nicht erfüllt worden:

- Die Änderung des Berichtsplans und damit der Methodik während der laufenden Nutzenbewertung unmittelbar nach einer Veröffentlichung zum gleichen Thema durch Mitarbeiter des IQWiG ist methodisch ein gravierender Verstoß gegen fundamentale biometrische Regeln (1)
- Ausgrenzung Sachverständiger von der mündlichen Anhörung zum Berichtsplan
- Begrenzung des Protokolls der mündlichen Anhörung auf deutschsprachige Sachvorträge
- Fehlender internationaler wissenschaftlicher Diskurs mit allen Sachverständigen
- Für Fachkreise extrem kurze Frist von 4 Wochen, um über 300 Seiten Vorbericht zu analysieren und eigene Stellungnahmen abzufassen.

Weiterhin werden die internationalen Standards der Nutzenbewertung verletzt:

- Die Komparatoren sind bei der Festlegung der Einschlusskriterien in 4.1.2 so stark eingegrenzt worden, dass die externe Validität der Untersuchung in das Ergebnis verzerrender Weise beeinträchtigt wird. Die im Versorgungsalltag häufigsten Therapieregime sind nicht berücksichtigt. Die Präzisierung des Auftrages wurde nicht anhand der PICO-Kriterien (Population, Intervention, Comparatoren, Outcome) vorgenommen.
- Zahlreiche Studien wurden von der Nutzenbewertung ausgeschlossen. Folgen sind fehlende Übertragbarkeit auf den Versorgungsalltag und ein verzerrtes Ergebnis (Selection Bias), Verletzung der prädefinierten Kriterien zum Studieneinschluss und Studienausschluss.
- Die Mindest-Studiendauer für Studien im Cross-over-design ist mit 2x24 Wochen unangemessen hoch angelegt worden.
- Die Anerkennung nur der schweren Unterzuckerungen als Zielkriterium und die hierfür gegebene Begründung ist klinisch und wissenschaftlich nicht akzeptabel und steht im Widerspruch zu internationalen HTAs. Diese berücksichtigen die Unterschiede in der Häufigkeit symptomatischer und auch nächtlicher Unterzuckerungen und erkennen diesen Vorteil von Insulin Glargin im Vergleich zu NPH-Insulin auch an.
- Vorteile von Insulin Glargin im Vergleich zu Insulin Detemir werden insbesondere bei der Analyse der HBA_{1c} Veränderungen unzureichend gewürdigt.
- Anwendungsbezogene Unterschiede zwischen den unterschiedlichen Insulinen wurden vom Vorbericht ignoriert, teilweise durch selektive Einschlusskriterien, teilweise durch Verstoß gegen die eigenen Kriterien.
- Analyse und Interpretation der PRO-Daten erfolgt anhand falscher Maßstäbe. Fehlerhafte Interpretation und Nichtberücksichtigung der Daten zu Lebensqualität bzw. Patientenzufriedenheit führen zu einer nachweisbaren Falschbewertung des Patientennutzens. Die Qualifizierung statistisch signifikant verbesserter Patientenzufriedenheit unter Lantus als „irrelevant“ ist keine wissenschaftliche Bewertung, sondern eine subjektive Meinung des IQWiG, die nicht Teil einer rein wissenschaftlichen Bewertung sein kann und für die das IQWiG nicht mandatiert ist.

Die genannten methodischen Mängel und die nicht gesetzeskonforme Durchführung der Nutzenbewertung haben dazu geführt, dass der Vorbericht bezüglich wichtiger Endpunkte zu Ergebnissen kommt, die in Widerspruch zu internationalen Veröffentlichungen stehen. Besonders bemerkenswert ist, dass einige Endpunkte sogar in Widerspruch zu der bereits zitierten Publikation unter Ko-Autorenschaft der IQWiG-Mitarbeiter stehen(2).
Der im Berichtsplan und Vorbericht angesprochene Mangel an Langzeitdaten und Unklarheiten von potenziellen Risiken (Retinopathie) trifft für Insulin Glargin nicht mehr zu. In einer großen

Langzeitstudie (4016) wurde Insulin Glargin im Vergleich mit NPH-Insulin untersucht. Diese Studie zeigt im Nichtunterlegenheitsdesign keinen statistisch signifikanten Unterschied im primären Endpunkt zwischen NPH Insulin und Insulin Glargin und ist vom IQWiG zu berücksichtigen.

2. Schlussfolgerungen des Vorberichtes stehen nicht im Einklang mit den Bewertungsergebnissen internationaler HTA's; es liegt ein wertendes Ergebnis vor, das nicht internationalen Standards entspricht

Die vom IQWiG identifizierten 33 Evidenzquellen der Kategorie 1a wurden lediglich auf relevante zusätzliche Primärstudien geprüft. Das ist unzureichend, um ein dem internationalen Standard der evidenzbasierten Medizin entsprechendes Bewertungsergebnis zu erarbeiten. Die Auflistung im Anhang C zeigt, dass sich das IQWiG inhaltlich mit den darin erarbeiteten Ergebnissen nicht ausreichend auseinandergesetzt hat. Dieses Verzeichnis ist zudem unvollständig. Es sind die internationalen Bewertungen nur teilweise enthalten:

a. CADTH

Der im Oktober 2007 publizierte kanadische HTA bezieht Behandlungskonzepte auf der Basis von kurzwirksamen Insulinanaloga, biphasischen Insulinen, Exenatide und auch Metaanalysen in das Assessment ein. Die Einteilung und Berücksichtigung des Endpunktes Hypoglykämie erfolgt differenzierter als diejenige des IQWiG. Es werden 3 Kategorien unterschieden und bewertet (overall, severe, nocturnal). Bei den 5 Vergleichsstudien, bei denen zusätzlich OAD verabreicht wurden, fand sich unter Insulin Glargin ein relatives Risiko von 0,87 (95% CI: 0,80-0,95) im Vergleich zu NPH Insulin. Bei nächtlichen Hypoglykämien bestand ebenfalls ein deutlicher und statistisch hochsignifikanter Unterschied. Das relative Risiko (RR) beträgt 0,52 (95% CI: 0,43-0,64). Auf dieser Basis werden Number Needed to Treat errechnet, eine anerkannte Maßzahl der evidenzbasierten Medizin (3). Auf das medizinische Risiko nächtlicher Unterzuckerungen wird explizit eingegangen. Ein Bias des offenen Studiendesigns auf die Häufigkeitsangaben der Unterzuckerungen wird zwar diskutiert, im Endergebnis wird aber im Unterschied zum IQWiG ein Vorteil hinsichtlich des Unterzuckerungsrisikos für Insulin Glargin eindeutig konstatiert(4).

b. Cochrane Collaboration

Die Veröffentlichung der Cochrane Collaboration im April 2007 erfolgte während des laufenden Nutzenbewertungsverfahrens unter der Ko-Autorenschaft von IQWiG Mitarbeitern (S. Ebrahim, T. Kaiser). In der Publikation wird angegeben, dass IQWiG habe auch als „External source of support“ fungiert.
Angesichts der offensichtlich parallel zur Nutzenbewertung erfolgten Publikation überrascht es nicht, dass die Ein- und Ausschlusskriterien weitgehend übereinstimmen. Die unterschiedliche Ergebnisdarstellung in der Cochrane Publikation im Vergleich zum Vorbericht muss allerdings erstaunen. In der Cochrane Publikation wird der signifikante Vorteil von Insulin Glargin im Vergleich zu NPH- Insulin sowohl bei symptomatischen Hypoglykämien (RR=0,84; 95% CI: 0,75-0,95) als auch bei nächtlichen Hypoglykämien (RR=0,66; 95% CI: 0,55-0,80) klar beschrieben. Diese Endpunkte werden nach den Modifikationen des Berichtsplans im Vorbericht nicht mehr explizit erwähnt (2).

c. NICE

NICE hat im Jahre 2002 eine Bewertung der Kosten und des Nutzens von Insulin Glargin vorgenommen. Das NICE spricht für folgende Patientengruppen mit Typ 2 Diabetes eine Behandlungsempfehlung mit Insulin Glargin aus:

- Patienten, die bei der Injektion Fremdhilfe benötigen
- Patienten, deren Lebensführung durch symptomatische Hypoglykämien eingeschränkt wird
- Patienten, die bei Einsatz anderer Insuline zwei Insulininjektionen pro Tag in Kombination mit oralen Antidiabetika benötigen würden

Das NICE würdigt somit die Vorteile der Möglichkeit einer Einmalapplikation und der niedrigeren Rate symptomatischer Hypoglykämien (5).

d. ASFSSAPS / HAS

Die französische Transparenzkommission hat ebenfalls Vorteile von Insulin Glargin im Vergleich zu NPH-Insulin beschrieben (6). Als Vorteil langwirksamer Insulinanaloga gegenüber NPH-Insulin wird die

Reduktion nächtlicher Unterzuckerungen bewertet. Für Insulin Glargin werden die 24 Stunden-Wirkung und die einmal tägliche Gabe genannt. Als weiterer Vorteil von Insulin Glargin wird die flexible Wahl des Injektionszeitpunktes im Tagesverlauf eingestuft. Ein Vorteil der langwirksamen Insulinanaloga wird auch in der fehlenden Notwendigkeit der Resuspension gesehen. Diese bei NPH-Insulinen erforderliche Resuspension führt durch Konzentrationsschwankungen häufig zu Dosierfehlern (7;8).

e. PBAC

In Australien wurde der Vorteil der relativen Risikoreduktion symptomatischer und schwerer Unterzuckerungen anerkannt. Sie gab dort den Ausschlag für die Aufnahme von Insulin Glargin auf die Arzneimittelliste im Jahr 2006. Für die Entscheidung wurden auch Evidenz aus Beobachtungsstudien herangezogen. Die durch die längere Wirkdauer bedingte einmal tägliche Gabe ist ebenfalls gewürdigt worden (9).

f. IDF

Nach der Leitlinie zur postprandialen Glukoseeinstellung (10) und der Global Guideline for Type 2 Diabetes (11) sind die Zielgrößen der vorliegenden Nutzenbewertung im Hinblick auf die Glykämie unzureichend ausgestaltet. Mit einer evidenzbasierten Leitlinie der International Diabetes Federation wurde aufgezeigt, dass erhöhte nüchtern- und postprandiale Blutglukosespiegel potenziell schädlich für den Typ 2 Diabetiker sind und im Rahmen des klinischen Diabetesmanagements eindeutig definierte Nüchtern und 2 Stunden postprandiale Glukose-Zielwerte einzuhalten sind. Eine geeignete Kenngröße für die Nutzenbewertung ist die Quote der Diabetiker, die die Zielvorgaben unter dem jeweiligen Behandlungskonzept erreichen konnten.

Die Global Guideline spricht sich für die Gabe von langwirksamen Insulinanaloga im Vergleich zu NPH Insulin aus, weil das Unterzuckerungsrisiko mit NPH-Insulin erhöht ist.

g. EMEA

Der EPAR erkennt das „peakless profile“ und die Möglichkeit der einmal täglichen Gabe ausdrücklich als Charakteristika langwirksamer Insulinanaloga an. Als Langzeitstudien werden Untersuchungen mit einer Dauer von 16-52 Wochen bezeichnet. Es wird festgestellt, dass Glargin behandelte Typ 2 Diabetiker weniger nächtliche Hypoglykämien berichtet hatten (12).

3. Die Methodenbeschreibung im Vorbericht weicht von dem gültigen Berichtsplan ab; dies verändert das Ergebnis des Vorberichts

Im Laufe der Nutzenbewertung sind Änderungen der Methodik vorgenommen worden, auch wenn dies vom IQWiG explizit bestritten wird (s. S 39 „Im Laufe der Bearbeitung des Projekts ergaben sich keine Änderungen und Ergänzungen des Vorgehens bei der Nutzenbewertung im Vergleich zu der im Berichtsplan Version 2.0 dargestellten Methodik.“). Es wurden Veränderungen an den Abschnitten „Behandlung des Diabetes mellitus Typ 2“, Abschnitt 4.1.2, 4.2, 4.4.3, 4.5.1, vorgenommen. Diese Änderungen wurden weder benannt noch begründet. Die fehlende Begründung der Methodik und der fehlende Hinweis darauf sind inakzeptabel. Die Änderung der Methodik nach Kenntnis der Ergebnisse (Mitarbeiter des IQWiG waren sogar Autoren des vorab veröffentlichten Cochrane-Reports) verstößt gegen die Grundregeln der Biometrie. Dies wiegt umso schwerer, da dies im vorliegenden Fall erheblichen Einfluss auf das Ergebnis ausübt. Es widerspricht ferner der IQWiG-eigenen Methodik, dass die Meta-Analysen ausschließlich auf der Basis von Modellen mit zufälligen Effekten durchgeführt werden; das aktuelle Methodenpapier (Version 2) des IQWiG nennt für die Durchführung von Metaanalysen die Verwendung von Modellen mit festen oder mit zufälligen Effekten ohne Präferenz für eine der beiden Methoden und fordert die Berücksichtigung beider: „Es kommen zum einen Modelle mit festen Effekten zum Einsatz, die (zum Beispiel durch die Inverse der Varianz) gewichtete Mittelwerte der Effektstärken liefern. Es werden zum anderen aber auch häufig Modelle mit zufälligen Effekten gewählt, bei denen eine Schätzung der Varianz zwischen den einzelnen Studien (Heterogenität) berücksichtigt wird. Die Frage, in welchen Situationen welches Modell eingesetzt werden soll, wird seit langem kontrovers diskutiert und muss für jede Analyse einzeln beantwortet werden. Wenn es keine eindeutige wissenschaftliche Begründung für die Wahl zwischen zufälligen und festen Effekten gibt, wird das Institut immer beide Methoden anwenden und im Fall von divergenten Ergebnissen diese als solche beschreiben.“ (13)

Diese Festlegungen hätten bereits im Berichtsplan zur Diskussion gestellt werden müssen. Die vom IQWiG vorgenommene ex – post Festlegung ist potenziell datengetrieben und damit unzulässig.

4. Begrenzung des Anhörungsverfahrens auf die deutsche Sprache entspricht nicht den internationalen Standards der evidenzbasierten Medizin und steht nicht im Einklang mit §§ 35b und 139a SGB V

Auch in diesem Stellungnahmeverfahren werden erneut nur Stellungnahmen in deutscher Sprache zugelassen. Wir haben bereits während der mündlichen Anhörung dagegen protestiert, dass einem international ausgewiesenen englischen Experten eine Übersetzung verweigert worden ist und seine Beiträge daher trotz persönlicher Teilnahme an der Anhörung in Köln keine Aufnahme ins Protokoll gefunden haben. Diese Vorgehensweise des IQWiG steht sowohl in klarem Widerspruch zum internationalen wissenschaftlichen Standard der Nutzenbewertung als auch zu dem vom Kölner Institut proklamierten internationalen Anspruch.

5. Inkonsistenzen innerhalb des Vorberichtes belegen ein unzureichendes Verständnis der Datenlage und wirken sich inhaltlich auf das Ergebnis des Vorberichtes aus

Aus dem von Sanofi-Aventis zur Verfügung gestellten Studienbericht der Studie 3006 hätte das IQWiG die vermeintlich strittige Detailfrage der Gabe oraler Antidiabetika als zusätzliche blutglukosesenkende Therapie klären können. Die Gabe von oralen Antidiabetika führt zum Ausschluss von der Studienteilnahme. Unsachgerecht interpretiert ist die auf Seite 36 des Vorberichts zitierte Aussage :“Während weder die Publikation Rosenstock noch der Studienbericht explizite Angaben zu einer OAD-Behandlung macht (eine Vorbehandlung mit OAD jedoch als Ausschlusskriterium aus der Studie angegeben ist) wurde laut der Scientific Discussion der EMEA ein additive OAD-Behandlung durchgeführt (12)“ ist unsachgerecht interpretiert.

Die EMEA-Aussage bezieht sich auf 3 Patienten aus dem NPH-Arm, die fälschlicherweise oral behandelt wurden. Sie erfüllten damit jedoch ein Ausschlusskriterium und wurden entsprechend als Protokollverletzer behandelt. Die Studie 3006 ist folglich eindeutig als intensivierte Insulintherapie bei Typ 2 Diabetes zu kategorisieren. Alle korrespondierenden Aussagen im Vorbericht sind entsprechend zu korrigieren.

Wenn das IQWiG das Fazit zieht, dass „vor allem die vorliegenden Studien zu Insulin Glargin und die Studien zum Direktvergleich die Praxis der Insulintherapie kaum oder in unklarem Maße abbilden“, befremdet diese Bewertung besonders vor dem Hintergrund, dass vorhandene Studien falsch beurteilt bzw. im Vornherein durch mangelhafte Festlegung der Einschlusskriterien teilweise (4001 Morgenarm) oder vollständig (Yokoyama(14), Janka(15), Bretzel(16) Studien im Vergleich zu biphasischen Insulinen, Glitazonen und Exenatide) ausgeschlossen worden sind.

Auch die Feststellungen zum Studiendesign hinsichtlich der Applikationsfrequenz von NPH-Insulin sind nicht zutreffend. Die Studie 3006 hat die zweimal tägliche Gabe von NPH-Insulin zugelassen. Da die einmal tägliche Gabe von NPH-Insulin bei der kombinierten Behandlung mit OAD gebräuchlich ist, ist dieses Design für die Studien in Kombination mit oralen Antidiabetika sehr realitätsnah. Die HAS beschreibt die Insulinneueinstellung sowohl unter NPH-Insulin als auch unter Insulin Glargin zunächst als einmal tägliche Insulingabe (6). Alle vom IQWiG eingeschlossenen Studien sind hinsichtlich der Applikationsfrequenz und auch bezüglich der Titrationsvorgaben realitätsnah gestaltet worden.

Daten aus der Routineversorgung aus Deutschland sind in erheblichem Umfang erhoben und veröffentlicht worden. Das IQWiG hat unserer Anregung und Forderung bisher fortgesetzt widersprochen, diese Daten ,wie z.B. vom NICE (5)und PBAC (9) durchgeführt, in die Nutzenbewertung einfließen zu lassen.

6. Die Einschränkung der Vergleichstherapien berücksichtigt nicht die vom IQWiG anerkannten und in der Praxis relevanten Therapiekonzepte; es liegt sowohl eine Inkonsistenz innerhalb des Vorberichtes als auch eine unsachgemäße Einschränkung des Auftrags vor

In dem derzeit vom IQWiG zur Diskussion gestellten Methodenpapier fordert das Institut: „Die Auswahl der zu bewertenden Gesundheitstechnologien muss dergestalt erfolgen, dass sie eher den gegebenen Markt in Deutschland zum Zeitpunkt der Evaluation wiedergibt, als auf Interventionen eingegrenzt zu werden, für die direkte vergleichende Studien vorliegen.“

Das IQWiG stellt im Vorbericht die klinisch gebräuchlichen Behandlungskonzepte des Diabetes mellitus Typ II zusammenfassend zutreffend dar

a) Die Gabe eines kurzwirksamen Insulins zu den Mahlzeiten (prandiale Insulintherapie)
b) Die einmal oder mehrmals tägliche Gabe eines Verzögerungsinsulins als „Basalunterstützung“
c) Die konventionelle Insulintherapie (CIT)
d) Die intensivierte Insulintherapie (IIT)

Der ersten (eigenen) Forderung folgend muss das IQWiG nunmehr prüfen, ob diese Behandlungskonzepte „den gegebenen Markt in Deutschland zum Zeitpunkt der Evaluation“ widerspiegeln oder ob es möglicherweise Überlappungen zwischen den Behandlungskonzepten gibt. Es muss darüber hinaus geprüft werden, ob es Studien gibt bei denen ein Vergleich verschiedener Behandlungskonzepte erfolgt ist.

Nichteinhaltung der Vorgaben des Berichtsplans bei der Studienauswahl verursacht einen Selektionsbias.

Es stellt einen schwersten Verstoß gegen fundamentale biometrische Regeln dar, von den prädefinierten Ein- und Ausschlusskriterien bei der Studienauswahl abzuweichen. Diesen Fehler hat das IQWiG im Falle der Studie 4001 gemacht. Der Morgenarm der Studie 4001, der wie die gesamte Studie unter Berücksichtigung aller Ein- und Ausschlusskriterien vollständig in die Analyse hätte einfließen müssen, wurde nur partiell in die meta-analytischen Betrachtungen eingeschlossen. Als Begründung für den Ausschluss der Ergebnisse der Studie 4001 werden sogenannte „Tageseffekte“ angegeben. In dem Berichtsplan der Nutzenbewertung A05-03 ist dies aber a priori nicht als Ausschlussgrund festgelegt worden. Die publizierte und dem IQWiG zugängliche Studie 4009 (17) zeigt eindeutig, dass die Wahl des Applikationszeitpunkts keinen signifikanten Einfluss auf den HbA_{1c} sowie die Frequenz der Hypoglykämien ausübt. Damit ist der Ausschluss der für Insulin Glargin dokumentierten günstigen Studienergebnisse methodisch fehlerhaft, unzutreffend begründet und nicht gerechtfertigt. Die Abweichung von den IQWiG-eigenen Vorgaben birgt zudem die große Gefahr einer datengetriebenen Analyse und damit der Ergebnisverzerrung. Die entsprechenden Auswertungen sind um die Ergebnisse des Morgenarms der Studie 4001 zu erweitern.

Auch der Ausschluss der Studie von Yokoyama (26) ist ungerechtfertigt. Im Cochrane Review (2) wird die Studie diskutiert – sie kann nach Aussage der Autoren aber aufgrund des fehlenden Verstoßes gegen die Ein- und Ausschlusskriterien nicht von der Bewertung ausgeschlossen werden. Sie wird in den Meta-Analysen allerdings nicht berücksichtigt. Der Vorbericht des IQWiG geht hier einen Schritt weiter und schließt die Studie über das Kriterium „sonstige Blutzucker senkende Behandlung unterschiedlich“ (siehe Seite 279 Vorbericht) von der Bewertung aus. Wie es zu diesem Ausschluss kommt ist nicht einfach nachvollziehbar:
Tabelle 1 der Arbeit von Yokoyama (14) S. 37 zeigt, dass 12 bzw. 14 von jeweils 31 Patienten mit Glimepirid und 20 bzw. 19 von 31 Patienten mit Metformin eine „sonstige blutzuckersenkende Behandlung“ erhielten. Darüber hinaus erhielten 26 bzw. 27 von 31 Patienten Insulin Aspart, 5 bzw. 4 von 31 Patienten Insulin Lispro. Dass die Dosis des kurzwirksamen Insulins bei gleichzeitiger Steigerung der Insulin Glargin Dosis reduziert wurde, während das Verhältnis beim NPH Insulin gleich blieb (S. 38, Tabelle 2 der Publikation) ist ein unzureichendes Argument, wird jedoch auf Seite 22 des Vorberichts als Begründung bemüht. Die Anfrage des IQWiG an den Autor vom 14.09.2006 ergab

(siehe Seite 289 Vorbericht), dass die Titration der NPH-Insulindosis durchgeführt wurde, aber aufgrund der Befürchtung nächtlicher Hypoglykämien nicht fortgesetzt wurde.
Diese Erkenntnis wird in den Ergebnissen und in der Diskussion nicht ausreichend gewürdigt. Die Möglichkeit der morgendlichen Gabe in Kombination mit der Titration des Analoginsulins im Hinblick auf die Erreichung der HbA_{1c} Werte bei fehlendem Anstieg der Hypoglykämierate ist hier der entscheidende Vorteil von Insulin Glargin gegenüber NPH-Insulin."

Ausschluss von wichtigen Vergleichsstudien durch fehlerhaft definierte Einschlusskriterien.

APOLLO Studie (16):

Die Ergebnisse der APOLLO-Studie sind überzeugend im Hinblick auf die therapeutische Gleichwertigkeit einer supplementären prandialen Insulintherapie mit einer basalunterstützten oralen Therapie mit Insulin Glargin. Die Rate der Unterzuckerungen unterscheidet sich erheblich. Unter der Kombination von Glargin mit OAD werden 5 symptomatische Unterzuckerungen pro Patientenjahr beschrieben, die korrespondierende Häufigkeit bei der prandialen Therapie beträgt 24 Unterzuckerungen pro Patientenjahr bei vergleichbarer Blutzuckereinstellung. Die Patientenzufriedenheit unter der Therapie mit Insulin Glargin ist der multiplen Injektionstherapie mit kurzwirksamen Insulin signifikant und relevant überlegen. Im Hinblick auf potenziellen Schaden durch die unterschiedliche Häufigkeit der Insulinapplikation und Häufigkeit der erforderlichen Blutzuckerselbstkontrolle sind mit Insulin Glargin erhebliche Vorteile für den Patienten bei gleicher Blutzuckereinstellung zu erreichen.

Laptop Studie (15):

Mit der Laptop Studie wurde der Nachweis erbracht, dass mit OAD vorbehandelte Diabetiker mit einer einmal täglichen Gabe von 28 IU Insulin Glargin unter Fortführung der OAD's Vorteile gegenüber der zweimal täglichen Gabe eines 30/70 Mischinsulins bietet. Dies gilt im Hinblick auf den HbA_{1c}, die Häufigkeit von Unterzuckerungen (gesamt, symptomatisch, nächtlich) sowie den Ressourcenverbrauch. Die Insulindosis der Mischinsulinbehandlung betrug 64 IU pro Tag.

7. Der Vorbericht verstößt durch fehlende Berücksichtigung von verfügbaren Studiendaten gegen die internationalen Standards der evidenzbasierten Medizin

Die fehlenden Angaben in der Ergebnisdarstellung klinischer Studien hätten unschwer durch Nachfrage bei den pharmazeutischen Unternehmen lückenlos eruiert werden können. Das IQWiG hat bei den Herstellern diese verfügbaren Ergebnisdaten nicht abgefragt, obwohl Sanofi-Aventis dies ausdrücklich angeboten hatte. Eine aktive Einreichung durch die Hersteller lassen die Einreichungsregeln des IQWiG leider nicht zu [http://www.iqwig.de/download/IQWiG-VFA-Mustervertrag.pdf; Zugriff 24.04.2008 12:43].

Die metaanalytischen Auswertungen sollten um nachstehende Ergebnisse ergänzt werden:

Hospitalisierungen	Angegeben sind jeweils die Anzahl der Ereignisse			
	3102	GA 4		NPH 5
	4001	GA 18	GM 15	NPH 13
	4012	G A 5		NPH 12
	4013	G A 8		NPH 4
	6001	G A 3		NPH 6

GM = Glargin morgens
GA = Glargin abends
NPH = NPH Insulin

Nächtliche schwere Hypoglykämien	Angegeben sind jeweils Anzahl Patienten/Anzahl Episoden		
	3006	Glargin 1/2	NPH 2/2
	3102	Glargin 0	NPH 0
	4001 (M+A)	Glargin 2/3	NPH 4/4
	4012	Glargin 0	NPH 0
	4013	Glargin 4/5	NPH 6/11
	6001	Glargin 0	NPH 0
Schwere Hypoglykämien	3102	Glargin 1	NPH 0
Tabelle 46	Studie 4012	0 [G] 0 [NPH]	Hypo’s / Patientenjahr
Tabelle 49	Studie 4001	5,3 [G], 5,5 [NPH]	Hypo’s / Patientenjahr
	Studie 4012	4,9 [G] 8,6 [NPH]	Hypo’s / Patientenjahr
	Studie 6001	5,5 [G] 7,9 [NPH]	Hypo’s / Patientenjahr

Auch dieser Vorbericht entspricht nicht den Standards der evidenzbasierten Medizin (EbM), da die Suchstrategie ausschließlich auf randomisierte Studien ausgerichtet ist. Die evidenzbasierte Medizin fordert die Verwendung der jeweils besten verfügbaren Evidenz. Dies bedeutet, dass beim Fehlen von Resultaten aus randomisierten Studien für eine Fragestellung die Evidenz aus Untersuchungen höheren Evidenzgrades herangezogen werden muss. Dies ist ein allgemeines Prinzip der EbM (18). Wir möchten hierzu auf unsere Stellungnahme zum Berichtsplan und auf unseren Briefwechsel zu diesem Thema verweisen, wie auch auf das von Ihnen zu diesem Punkt angeführte BMJ-Editorial von Sackett et al., das unmissverständlich konstatiert, dass die beste verfügbare Evidenz („best available evidence“) verwendet werden muss und dass bei fehlender Evidenz einer höheren Evidenzstufe, die Berücksichtigung von Ergebnisse niedriger Evidenzklassen obligat ist (19). Die entsprechenden Schlüsselsätze des Editorials lauten:

„Evidence based medicine is not restricted to randomised trials and meta-analyses. It involves tracking down the best external evidence with which to answer our clinical questions.”

und „However, some questions about therapy do not require randomised trials (successful interventions for otherwise fatal conditions) or cannot wait for trials to be conducted. And if no randomised trial has been carried out for our patients´ predicament, we must follow the trail to the next best external evidence and work from there“.

Um diesem Standardvorgehen der EbM gerecht zu werden, darf die Suchstrategie nicht auf RCTs beschränkt bleiben. Sie muss, sofern eine Fragestellung nicht durch Daten aus randomisierten Studien beantwortet werden kann, Daten aus Studienformen anderen Evidenzgrades berücksichtigen. Eine solche Suchstrategie hätte vorab definiert werden müssen. Die Vorgabe des Berichtsplanes, die auch vom vorliegenden Entwurf des Vorberichtes („4.1.4 Studientypen“, Seite 10) wiederholt wird: „Für alle unter 4.1.3 genannten Therapieziele und alle unter 4.1.2 genannten Interventionen ist eine Evaluation im Rahmen von randomisierten kontrollierten Studien möglich und praktisch durchführbar. Für den zu erstellenden Bericht werden daher ausschließlich RCTs als relevante wissenschaftliche Literatur in die Nutzenbewertung einfließen.“ ignoriert, dass die Standards der EbM aus guten Gründen die Berücksichtigung der besten verfügbaren Evidenz fordert und nicht der IQWiG-eigenen Lesart der ausschließlichen Akzeptanz der besten theoretisch denkbaren Evidenz folgt.

8. Die faktisch ausschließliche Berücksichtigung von schweren bzw. schwerwiegenden Hypoglykämien ist medizinisch nicht zu rechtfertigen und lässt wesentliche Aspekte des patientenrelevanten Nutzens außer Acht; damit berücksichtigt der Vorbericht nicht den allgemein anerkannten Stand der medizinischen Erkenntnisse und entspricht des weiteren nicht den Vorgaben des § 35b SGB V

Alle Unterzuckerungen sind bei der Bewertung einer blutzuckersenkenden Therapie als Zielgröße zu berücksichtigen. Es ist nicht zu rechtfertigen, dass bei der vorliegenden Nutzenbewertung nur schwere Unterzuckerzuckerungen als Zielkriterium definiert wurden und damit andere Schweregrade der Unterzuckerung nicht berücksichtigt wurden. Das IQWiG behandelt den gleichen Aspekt bei Typ 2 Diabetes anders als bei Typ 1 Diabetes. Es besteht ein offensichtlicher Unterschied zwischen Berichtsplan A05-03 und A05-01 zur Bewertung des Nutzens langwirksamer Insulinanaloga bei Typ 1 Diabetes. Da sich die Frequenzen der Unterzuckerungen zwischen Typ 1 und Typ 2 Diabetikern nach 5 Jahren Insulinbehandlung annähern, ist dieser Unterschied nicht zu rechtfertigen.

Die Argumentation des IQWiG, dass aufgrund offener Studiendesigns „nicht schwerwiegende Hypoglykämien nicht sicher erfasst werden“, ist nicht zutreffend. Sie unterscheidet auch nicht zwischen den beiden Diabetesformen. Symptomatische Unterzuckerungen sind auch bei offenen Studien eine valide an verschiedenen Kriterien festzumachende, unerwünschte Wirkung einer Insulinbehandlung. Die Raten symptomatischer Unterzuckerungen mit zeitnah erhobenen pathologischen Blutzuckerwerten sind in den Studien 3002, 3006, 4001 und 4002 genau beschrieben worden. Auf Basis der Protokollkriterien werden diese durch klinische Monitore anhand der Dokumentation in der Patientenakte (Klinik und Blutzuckerbefund) geprüft. Der vom IQWiG erhobene Vorwurf einer subjektiven Ergebnisverzerrung rechtfertigt nicht den Ausschluss der Daten. Die Metaanalysen auf der Basis patientenindividueller Daten von Rosenstock et al (20) und Mullins (21) zeigen auf dem höchsten Evidenzlevel deutliche Vorteile von Insulin Glargin gegenüber NPH-Insulin für die gesamten, symptomatischen und die nächtlichen Unterzuckerungen.

9. Fehlerhafte Interpretation und Nichtberücksichtigung der Daten zu Lebensqualität bzw. Patientenzufriedenheit führen zu einer nachweisbaren Falschbewertung des Patientennutzens. Die Qualifizierung statistisch signifikant verbesserter Patientenzufriedenheit unter Insulin Glargin als „irrelevant“ ist keine wissenschaftliche Bewertung, sondern eine subjektive Meinung des IQWiG, die nicht Teil einer rein wissenschaftlichen Bewertung sein kann und für die das IQWiG nicht mandatiert ist.

Ein wesentlicher Vorteil von Insulin Glargin liegt in der im Vergleich zu anderen Insulinen niedrigeren Applikationsfrequenz. Die gesundheitsbezogene Lebensqualität (HRQL) und die Behandlungszufriedenheit der Patienten (PTS) werden grundsätzlich auch vom IQWiG als patientenrelevante Parameter anerkannt.

Ausschluss relevanter Studien:

Der Ausschluss relevanter Studien aufgrund der in Abschnitt 5 aufgeführten Vorgehensweise des IQWiG betrifft auch die Beurteilung der Auswirkungen auf HRQL und Behandlungszufriedenheit. Für die Versorgung von Diabetikern relevante Therapieregime und Vergleiche von Therapie-Regimes finden somit auch bezüglich dieser Parameter keine Berücksichtigung.

Nichtberücksichtigung von nicht randomisierten Erhebungen zur Abbildung der Routineversorgung:

Auch durch Zulassungsstudien vom Typ RCT wird immer nur ein artifizielles, nicht der Routineversorgung entsprechendes Setting abgebildet werden. Für Kostenerstattungsentscheidungen aber spielt die Abbildung der Routineversorgung eine entscheidende Rolle. Nicht randomisierte kontrollierte Studiendesigns haben hierfür ihren Stellenwert als ergänzende Informationsquelle. Die strikte Beschränkung auf RCT’s wird die Routineversorgung von Diabetikern in Deutschland auch bezüglich HRQL und PTS nicht ausreichend abgebildet. So konnte allein aufgrund der strikten Einschlusskriterien des IQWiG eine Beobachtungsstudie mit über 5000 Typ 2 Diabetikern in Deutschland, die eine klare Verbesserung der Behandlungszufriedenheit unter Insulin Glargin aufzeigte, nicht Berücksichtigung finden (22). Weitere die Zufriedenheit mit der jeweiligen Insulintherapie abbildende Versorgungsstudien sind abgeschlossen und werden kurzfristig publiziert werden.

Methodische Aspekte:

Die vom IQWiG durchgeführte metaanalytische Zusammenfassung der Auswertung der beiden Versionen des DTSQ, d.h. den $DTSQ_s$ und $DTSQ_c$ Ergebnisse aus verschiedenen Studien auf der Ebene der Effektstärke nach Cohen wird den Merkmalen der Erhebungsinstrumente nicht gerecht. Bei den beiden Varianten des DTSQ handelt es sich um Instrumente mit verschiedenen psychometrischen Eigenschaften, so dass die Anwendung bei ein und denselben Patienten zu Effekten verschiedener Stärke führt (23).Der $DTSQ_c$ bildet die Veränderung ab und wurde entwickelt um die möglichen Deckeneffekte, die bei alleiniger Verwendung des den Status messenden $DTSQ_s$ zu verringern. Eine Zusammenfassung der beiden Instrumente erscheint nur nach Adjustierung der Ergebnisse bzw. einer Ableitung der Veränderungsrate aus zwei Status-Messungen mit DTSQs sinnvoll.

Diese für eine metaanalytische Zusammenfassung zentralen Aspekte wurden vom IQWiG nicht beachtet. Die mit dem Vorbericht aufgeführte Abschätzung der Behandlungszufriedenheit ist somit fehlerhaft.

Interpretation der Ergebnisse der Metaanalyse:

Das Vorgehen des IQWiG bei den Metaanalysen der DTSQ-Ergebnisse selbst weicht von seinem sonst vertretenen, streng an biometrischen Kriterien ausgerichteten Ansatz ab. Die Interpretation der Ergebnisse der DTSQ - Metaanalysen steht im Widerspruch zu seiner an anderen Stellen angewendeten Interpretation der Ergebnisse. Das statistisch signifikante Ergebnis wird als „irrelevant“ klassifiziert. Wörtlich wird zusammengefasst, dass es „bez. der Behandlungszufriedenheit keinen Unterschied zwischen Insulin Glargin und NPH-Insulin“ gebe. Als Begründung für diese eigenmächtige Klassifizierung und Urteilssetzung werden zwei Aspekte aufgeführt: (1) Der Effekt sei zu klein, und (2) es bestünden methodische Schwächen.

Zu (1): Die Effektstärke nach Cohen (24) ist ein verbreiteter Parameter um das Ausmaß der Wirkung abzubilden. Cohen selbst führt bezüglich seiner Klassifikation große, kleine und mittlere Effekte an. Die Interpretation der Effektstärke ist immer von der Gesamtsituation abhängig (25). Die Klassifikation eines signifikanten Effektes durch das IQWiG als „irrelevant“ stellt eine subjektive Interpretation und damit die Meinung des IQWiG dar. Die subjektive Meinung des IQWiG wird nicht durch die Literatur gestützt und kann somit nicht als allgemein akzeptiert gelten.

Zu (2): Die aufgeführten methodischen Aspekte sind zum Teil dem Charakter von patient-reported Outcomes geschuldet. Sie sind allgemein für alle Parameter gültig, sie scheinen jedoch teilweise gezielt nur bei HRQL und PTS aufgeführt zu werden. So werden allgemein für sekundäre Endpunkte keine eigenen Fallzahlschätzungen vorgenommen und auch das Problem des multiplen Testens stellt sich ganz allgemein. Es ist bei allen biometrischen Auswertungen des IQWiG und anderer Experten bei allen Parametern in gleicher Weise zu beachten.

Für klinische Parameter liegen in den wenigsten Fällen im Vorfeld der Untersuchung allgemein festgeschriebene Mindestgrenzen für die klinische Relevanz vor. Die Entwickler des DTSQ halten selbst eine starre Minimal Important Difference für den DTSQ für nicht angebracht (26). Die Stärke des Effektes nachgewiesener Unterschiede – und bei einem signifikanten Ergebnis handelt es sich um einen Nachweis – muss immer im Gesamtkontext gesehen werden. Der Einschätzung eines signifikanten Vorteils von Insulin Glargin hinsichtlich Behandlungszufriedenheit durch das IQWiG als „kein Unterschied“ muss widersprochen werden.

Die vom IQWiG geäußerte Kritik am Umgang mit fehlenden Werten ist zurückzuweisen. In allen Studien wurde der Umgang mit fehlenden Werten transparent berichtet und anerkannte Vorgehensweisen gewählt. Der Verweis auf die Draft-Guidance der FDA, die die Empfehlung abgibt, mehrere Verfahren parallel anzuwenden verkennt, dass es sich hierbei lediglich um einen zur Diskussion gestellten Entwurf handelt, der erstmalig im Jahr 2006 veröffentlicht wurde (27). Studien, die in den Jahren 1998 bis 2005 geplant oder durchgeführt worden sind, können diese damals gar nicht vorliegende Draft-Guideline auch nicht erfüllen. Daraus eine Nichtberücksichtigung abzuleiten ist ungerechtfertigt.

Im Zusammenhang mit fehlenden Werten ist auch anzumerken, dass in der vom IQWiG angeführten Nichberücksichtigungsrate eine Reihe von Subgruppen zusammengefasst ist, die das Ergebnis in unterschiedlicher Art und Weise beeinflussen können. So birgt eine Nichterfassung der Behandlungszufriedenheit in einem teilnehmenden Land aufgrund der fehlenden Verfügbarkeit einer übersetzten und validierten Version des Fragebogens eine geringere Gefahr der Verzerrung als fehlende Angaben von Studienabbrechern. Diese unterschiedlichen Gründe der Nichtberücksichtigung sind in der aggregierten Darstellung des IQWiG weder ausreichend dargelegt noch ausreichend und differenziert begründet.

Fehlende Mandatierung des IQWiG zur Festlegung eigener Schwellenwerte:

Das Mandat des IQWiG erstreckt sich auf das Assessment der Folgen des Einsatzes einer Medizintechnologie, d.h. auf das systematische Detektieren und Aufbereiten der verfügbaren Daten. Die hier erfolgte willkürliche Fällung eines Werturteils, basierend auf einer eigenmächtig gesetzten Mindestgrenze, bedarf eines gesellschaftlichen und demokratischen Mandates, welches das IQWiG nicht besitzt. Eine umfassende, transparente Beschreibung der Ergebnisse differenziert nach den einzelnen Instrumenten und eine sachliche Beschreibung der vorhandenen, signifikanten Effekte wären sachgerecht und erforderlich.

10. Fehlende Verblindung von Patient Reported Outcomes (PRO) und durch Patienten festgestellte Hypoglykämien darf nicht zu deren Nichtberücksichtigung führen.

Der Vorbericht erkennt an, dass beim Vergleich langwirksamer Insulinanaloga mit NPH-Insulin eine Verblindung von Patienten und Behandlern nicht möglich ist. Er nimmt deshalb konsequenterweise wegen fehlender Verblindung auch keine Herabstufung der formalen Studienqualität vor (S. 74): „Als Grund für diese Nichtverblindung wird das unterschiedliche Aussehen von NPH-Insulin im Vergleich zu beiden Insulinanaloga angegeben. Während beide Insulinanloga in klarer Lösung vorliegen, ist die Zubereitung von NPH-Insulin trübe. Dadurch wird es unmöglich, eine für den Patienten und Behandelnden verblindete Studie durchführen zu können, ohne eine Double-Dummy-Technik einzusetzen. Aufgrund dessen führte die fehlende Verblindung der Studien zum Vergleich Insulinanalogon vs. NPH-Insulin, bezogen auf die Gesamtstudien, nicht zur Abstufung der Qualitätsbewertung“.

10.1 Nichtberücksichtigung der unter Insulin Glargin verbesserten Patientenzufriedenheit:

Bei den PRO-Daten der randomisierten Vergleiche von Insulin Glargin mit NPH-Insulin zeigt die vom IQWiG durchgeführte Metaanalyse einen Vorteil zugunsten von Insulin Glargin. Dieser Vorteil wird vom IQWiG nicht anerkannt. Das Institut nennt hierfür als Begründung die „nichtverblindete Erhebung der PRO-Daten“ [S. 187-8]: „Ein weiteres wichtiges Kriterium für die Bewertung der Ergebnisse zur Behandlungszufriedenheit ist die Verblindung der Studienteilnehmer“, sowie [S. 196]: „Zusammenfassend liegt zwar eine statistische Signifikanz zu Gunsten von Insulin Glargin gegenüber NPH-Insulin vor, die klinische Relevanz dieses Unterschieds ist jedoch fraglich. Zieht man in Betracht, dass die Patienten in keiner Studie verblindet waren, relativiert sich dieses Ergebnis zusätzlich. So kann man keinen Hinweis auf eine Über- oder Unterlegenheit einer der beiden Therapieoptionen erkennen“.

Hier macht der Bericht die Einhaltung einer Anforderung zur Voraussetzung der Berücksichtigung eines Patientennutzens, die selbst theoretisch nicht erfüllbar ist. Ein wesentliches Merkmal der Erhebung von PRO’s wie der gesundheitsbezogenen Lebensqualität und der Behandlungszufriedenheit ist die Selbstauskunft der Patienten. Der Bericht selbst konstatiert, dass eine Verblindung der Patienten in diesem Fall nicht möglich ist und eine verblindete Endpunkterhebung bei PRO hier ebenfalls schon rein theoretisch ausgeschlossen ist. Das IQWiG ist in seiner eigenen Argumentation in diesem Zusammenhang nicht widerspruchsfrei.

Der Vollständigkeit halber ist anzumerken, dass der Einsatz einer Double-Dummy-Technik zur Verblindung der Patienten bei der Erhebung von PROs ebenfalls ausscheiden muss, weil eine Double-Dummy-Technik hier fehlerhafte Ergebnisse herbeiführen würde. Teil der verbesserten Patientenzufriedenheit oder HRQoL ist die seltenere Spritzhäufigkeit, welche dann aus methodischen Gründen nicht erfassbar ist.

Wir weisen darauf hin, dass durch sprachlich feine Verschiebung der in einer Metaanalyse erbrachte Nachweis der Überlegenheit vom IQWiG zum Fehlen eines Hinweises umgedeutet wird.

10.2 Nichtberücksichtigung des unter Insulin Glargin gesenkten Hypoglykämierisikos:

Einen analogen Bewertungsfehler macht der Entwurf bei der Nichtberücksichtigung von durch die Patienten selbst beobachtete Hypoglykämiesymptome bzw. selbst gemessene Blutzuckerwerte, die lapidar als „ohne nennenswerte Aussagekraft“ disqualifiziert werden [„5.3.4.6 Hypoglykämierate insgesamt“ S. 156]: „Für keine Studie fanden sich Hinweise darauf, dass eine unabhängige und gegenüber der Behandlungsgruppe verblindete Endpunkterhebung erfolgt war. Aufgrund der daraus resultierenden sehr hohen Ergebnisunsicherheit ... sind die Ergebnisse, sowohl die der Gesamtraten als auch die der nächtlichen Hypoglykämien – ohne nennenswerte Aussagekraft und lediglich der Vollständigkeit halber dargestellt.“ Auch hier wird das Fehlen einer Verblindung, als Grund des Ausschlusses von Studienergebnissen angeführt. Dies geschieht, obwohl das IQWiG selbst konstatiert, dass eine Verblindung der Therapie bei dem von IQWiG bzw. G-BA gewählten Vergleich NPH vs. Analoga faktisch unmöglich ist und dass eine verblindete Endpunkterhebung bei von Patienten erfassten Parametern nicht möglich ist, da man einen Menschen nicht sich selbst gegenüber verblinden kann.

Es wird bei den Hypoglykämien die Nichterreichung eines in der dargestellten Form nicht überprüfbaren Zieles zum Anlass genommen, patientenrelevante Studienergebnisse von der

Bewertung auszuschließen [„5.3.4.6 Hypoglykämierate insgesamt“ S. 156]: „Auch statistisch signifikante Unterschiede zwischen den Behandlungsgruppen … können aufgrund der hohen Ergebnisunsicherheit nicht als Nachweis eines Vorteils bzw. eines Nachteils einer Behandlungsoption gewertet werden.“. Dies verstößt gegen den EbM-Standard der Berücksichtigung der besten verfügbaren Evidenz, vor allem gegen die IQWiG-eigene Prämisse der ausschließlichen Berücksichtigung der besten theoretisch erzielbaren Evidenz. Das Ausmaß dieses Fehlers ist schwerwiegend, weil es hier um Hypoglykämien bzw. sogar nächtlichen Hypoglykämien geht und somit nicht um einen Nutzen, sondern um ein Schadenspotential. Laut IQWiG-Definition (13) S. 115: „Das Schadenspotential beschreibt das Risiko für das Auftreten unerwünschter Wirkungen“. Die Methodik des IQWiG sieht für diese Fälle unmissverständlich vor, der Patientensicherheit Vorrang zu geben und deshalb bei bestehender Unsicherheit das Schadenspotenzial bis zum Nachweis des Gegenteils als gegeben anzunehmen (13) S. 115: „Im Interesse der Patientensicherheit und des ärztlichen Axioms des „primum nihil nocere“ wird das Institut [das IQWiG, Anmerkung des Verfassers] bei der Aufarbeitung und Darstellung unerwünschter Effekte folgende Grundsätze beachten:“ Liegen aus den oben genannten. Datenquellen Hinweise auf das Vorliegen einer unerwünschten Wirkung vor, so wird diese so lange als möglich angesehen, bis diese durch die Ergebnisse zielgerichteter Forschung mit hinreichender Sicherheit ausgeschlossen werden kann“ In diesem Kontext legt das IQWiG sich fest, in dieser Weise sogar Ergebnisse aus Studienformen hoher Ergebnisunsicherheit von Kohortenstudien bis hin zu Tierexperimenten zu berücksichtigen. Es ist gedanklich inkonsistent, einen Verstoß gegen die IQWiG-eigene Methodik und gegen den EbM-Standard der Berücksichtigung der besten verfügbaren Evidenz, dass das IQWiG Daten zum Risikopotenzial bzw. unerwünschten Wirkungen aus akzeptierten randomisierten Studien ignorieren will, nur weil eine Verblindung in diesem Fall unmöglich ist.

Selbst die Verblindung in den Studien 4001 (28) und 4013 (29) durch ein „nachträgliches verblindetes Review“ bleibt unberücksichtigt, weil [„5.3.4.6 Hypoglykämierate insgesamt“ S. 156]: „Die Quelle für das hohe Verzerrungspotential dieses Endpunktes [ist] bereits auf Ebene der Patienten zu sehen, die selbst bewerten, ob sie eine Hypoglykämie erlitten oder nicht, und das in Kenntnis der Therapie, die sie erhalten“. Hier bestätigt der Bericht nochmals, dass eine wirksame Verblindung in der Realität nicht durchgeführt werden kann.

Das hier vorgestellte Vorgehen führt im vorliegenden Falle dazu, dass der Bericht ignoriert, dass Insulin Glargin in den vom IQWiG akzeptierte RCTs statistisch signifikant weniger sowohl Gesamthypoglykämien, als auch nächtliche Hypoglykämien verursacht als NPH-Insulin.

11. Fehlerhafte Beurteilung der Applikationsfrequenzen führt zu einer unvollständigen Bewertung des Patientennutzens

Eine auf mehrere Tagesdosen verteilte Applikationsfrequenz stellt für Patienten und Personen die Hilfestellungen geben eine zusätzliche Belastung dar und stellt zudem eine mögliche Gefahrenquelle für Fehldosierungen dar. Die verlängerte Wirkdauer und die damit reduzierte Applikationsfrequenz von Insulin Glargin sind eo ipso ein patientenrelevanter Vorteil gegenüber anderen Insulinen und müssen auch vom IQWiG entsprechend anerkannt werden. Dies hätte durch einen eigenen Endpunkt erfasst und in die Bewertung einbezogen werden müssen. Dies gewinnt zusätzliche Relevanz durch die Setzungen des IQWiG, gesundheitsökonomische Bewertungen von Arzneimitteln nur in den Fällen vorzunehmen, in denen eine isolierte Nutzenbewertung einen Zusatznutzen konstatiert hat und dann ausschließlich die dort erfassten Nutzenaspekte in die ökonomische Bewertung einfließen zu lassen Nach Angaben des Arzneiverordnungsreports 2007 werden in Deutschland mehrfach täglich zu applizierende verzögerungs- und Mischinsuline sehr häufig angewendet (30):

NPH –Insulin	117 Mio DDDs
Detemir	17 Mio DDDs
Mischinsuline	175 Mio DDDs

Die vom IQWiG verwendeten Datenquellen geben keine Auskunft über die Applikationsfrequenzen beim Diabetes mellitus Typ 2. Die vom IQWiG angegebene Datenquelle DPV Wissensdatenbank verfügt lediglich über Daten von Typ 1 Diabetikern. Das IQWiG hat das Ergebnis der Recherche aus dieser Datenquelle darüber hinaus nicht genannt. Da sich die Behandlungskonzepte von Typ 1 und

Typ 2 Diabetikern unterscheiden, ist diese Informationsquelle zur Beurteilung der in der Versorgungspraxis verwendeten Applikationsfrequenzen ungeeignet.
Auf der Basis von IMS-Daten sowie aus Versorgungsforschungsstudien lassen sich für die Behandlung von Typ 2 Diabetikern die folgenden durchschnittlichen Applikationsfrequenzen ermitteln:

Insulin Glargin	1,09 Injektionen / Behandlungstag
NPH Insulin	1,47 Injektionen / Behandlungstag
Insulin Detemir	1,44 Injektionen / Behandlungstag

Nach Erhebungen von IMS wird NPH Insulin in einer Frequenz von 1,47 Injektionen / Behandlungstag angewendet. Das bedeutet, dass auf diese Behandlungen 309 Mio DDD's x 1,47 Injektionen = 445 Millionen Injektionen entfallen. Die Substitution durch Insulin Glargin würde bedeuten, dass 309 Mio DDDs x 1,09 Injektionen = 336 Mio Injektionen ausreichen würden.
Die Verwendung von Insulin Glargin statt des verbreiteten Einsatzes von NPH-Insulin und Insulin Detemir vermag in Deutschland jährlich etwa 109 Millionen Insulininjektionen einzusparen. Ein klarer medizinischer Zusatznutzen und im Rahmen der Ermittlung des Schadenspotenzials anzuerkennender Vorteil, besonders auch, weil mit jeder Insulinapplikation ein Infektionsrisiko, die Möglichkeit einer lokalen Reaktion an der Einstichstelle und für den Diabetiker ein Schmerzerleben verknüpft ist.

Für die supplementäre Insulintherapie konnte gezeigt werden, dass die einmal tägliche Gabe von Insulin Glargin der dreimal täglichen Gabe von Insulin Lispro therapeutisch gleichwertig ist (16). Durch die teilweise Substitution der prandialen Therapie mit der basalunterstützten oralen Therapie mit Insulin Glargin ergibt sich eine noch größere Reduktion des Schadenspotenzials durch vermeidbare Insulininjektionen und Blutzuckerselbstkontrollen. Die Nutzenbewertung des IQWiG blendet diesen für die Diabetiker äußerst relevanten Aspekt vollständig aus.

Insulin Glargin und Insulin Detemir liegen als klare Lösungen vor. Die Notwendigkeit der Resuspension stellt eine mögliche Fehlerquelle der NPH-Insuline dar (7;8). Aus diesem Grunde ist die Erfordernis der Resuspension als Schadenspotential zu beschreiben. Bei Fremdapplikation von langwirksamen Insulinanaloga ist der Aufwand für das betreuende Personal aus dreierlei Gründen reduziert:

1.) seltenere Applikation
2.) Resuspension nicht erforderlich
3.) Integration des Applikationszeitpunktes in Arbeitsabläufe und Tourenplanungen der Pflegedienste

Ebenfalls wurde vom IQWiG übersehen, dass bei Insulinbehandlung mit Insulin Glargin der Injektionszeitpunkt im Tagesverlauf frei gewählt werden kann, dann aber immer der gleiche Tageszeitpunkt für die Injektion verwendet werden sollte. Mit den Studien 4001 (29) und 4009 (31) ist die Datengrundlage für eine entsprechende Zulassung geschaffen worden. Durch die fehlerhaft formulierten Einschlusskriterien hat das IQWiG die entsprechenden Studienergebnisse systematisch ausgeblendet.

12. Fazit

Das IQWiG hat die Studienauswahl so getroffen, dass realitätsnahe Studiendesigns vollständig ausgeblendet worden sind. Die selektive Nichtberücksichtigung von allen und den nächtlichen Unterzuckerungen entspricht nicht dem internationalen Standard und blendet patientenrelevanten Nutzen aus. Da Mitarbeiter des IQWiG bereits an einer Veröffentlichung mit einer ähnliche Fragestellung beteiligt waren, liegt der Verdacht einer datengetriebenen post-hoc Analyse nahe.

Die signifikanten Vorteile der Therapiezufriedenheit unter Insulin Glargin wurden vom IQWiG unsachgemäß und ohne eine wissenschaftliche Begründung als irrelevant abqualifiziert.

Das IQWiG bemängelt einerseits Erkenntnislücken im Vergleich zu verbreiteten Behandlungskonzepten, andererseits trägt es selbst durch die falschen Ein- und Ausschlusskriterien dazu bei, dass Studien wichtiger Behandlungskonzepte systematisch ausgeschlossen sind. Dies wird durch den Ausschluss weiterer Studien oder Teilstudienverstärkt, die ausgeschlossen worden sind, obwohl sie alle Kriterien

erfüllen. Diese Vorgehensweise ist unbegründet und führt zu einem verzerrten und fehlerhaften Bewertungsergebnis.

Das IQWiG stützt seine eigenen Aussagen auf metaanalytische Auswertungen, die auf einer eigenwilligen und nicht konsistent aus Ein- und Ausschlusskriterien ableitbaren Studienselektion basieren. Damit kann das IQWiG keine validen Aussagen über den aktuell gesicherten therapeutischen Nutzen der beiden zur Bewertung anstehenden langwirksamen Insulinanaloga treffen. Wesentliche anwendungsnahe Nutzenkriterien sind bisher vom IQWiG durch methodische Unzulänglichkeiten nicht erfasst worden. Wesentliche Feststellungen des IQWiG zur Verfügbarkeit von Langzeitstudien, zur Langzeitsicherheit und auch über den nicht nachgewiesenen zusätzlichen Nutzen müssen vom IQWiG revidiert werden, bevor der Abschlussbericht an den G-BA übergeben wird.
Der Vorbericht in der vorliegenden Form ist nicht geeignet, den gesamten Nutzen von Insulin Glargin abzubilden, da wesentliche patientenrelevante Nutzenaspekte und versorgungsrelevante Vorteile gar nicht beschrieben wurden, als irrelevant eingestuft oder als nicht hinreichend ergebnissicher kategorisiert wurden. Der medizinische Mehrnutzen von Insulin Glargin für die Diabetikerversorgung ist international zutreffend beschrieben worden, der Vorbericht des IQWiG weicht davon erheblich ab.

Das vorläufige Ergebnis der IQWiG Nutzenbewertung entspricht nicht den sozialrechtlichen Anforderungen.

Reference List

(1) ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 1999 August 15;18(15):1905-42.
Ref ID: 9

(2) Horvath K, Jeitler K, Berghold A, Ebrahim SH, Gratzer TW, Plank J et al. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Database Syst Rev 2007;(2):CD005613.
Ref ID: 12

(3) Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. BMJ 1995 February 18;310(6977):452-4.
Ref ID: 28

(4) Tran K, Banerjee S, Huimini L, Danemann D, Simpson S, Campbell K. Long-Acting Insulin Analogues for Diabetes Mellitus: Meta-analysis of Clinical Outcomes and Assessment of Cost-Effectiveness. Internet 2007 October 1;Available from: URL: http://cadth.ca/index.php/en/publication/757
Ref ID: 11

(5) National Institute for Clinical Excellence. Guidance on the use of long-acting insulin analogues for the treatment of diabetes-insulin glargine. NICE Technology appraisal Guidance No 53. Internet 2008;Available from: URL: http://www.nice.org.uk/usingguidance/
Ref ID: 18

(6) asfssaps, has. TRAITEMENT MEDICAMENTEUX DU DIABETE DE TYPE 2 (ACTUALISATION) : ARGUMENTAIRE. Internet 2008;Available from: URL: http://www.has-sante.fr/portail/display.jsp?id=j_5
Ref ID: 17

(7) Brown A, Steel JM, Duncan C, Duncan A, McBain AM. An assessment of the adequacy of suspension of insulin in pen injectors. Diabet Med 2004 June;21(6):604-8.
Ref ID: 26

(8) Jehle PM, Micheler C, Jehle DR, Breitig D, Boehm BO. Inadequate suspension of neutral protamine Hagendorn (NPH) insulin in pens. Lancet 1999 November 6;354(9190):1604-7.
Ref ID: 27

(9) PBAC. **Insulin glargine, injection, 100 units per mL, 10 mL vials and 3 mL x 5 cartridges, Lantus®, March 2006**. Internet 2006 June 30;Available from: URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/pbac-psd-insulinglargine-mar06
Ref ID: 19

(10) Ceriello A, Colagiuri S, Gerich J, Tuomilehto J. Leitlinie für die postprandiale Glukoseeinstellung. Internet 2007;Available from: URL: http://www.idf.org/home/index.cfm?unode=185108C7-1E27-4A03-9B73-01D54087E32E
Ref ID: 15

(11) Home P, Colagiuri S. IDF Clinical Guidelines Task Force. Global guideline for Type 2 diabetes. Brussels: International Diabetes Federation, 2005. Internet 2008;Available from: URL: http://www.idf.org/home/index.cfm?node=1457
Ref ID: 16

(12) European Medicines Agency. Lantus: European public assessment report; scientific discussion. Internet 2005;Available from: URL: http://www.emea.europa.eu/
Ref ID: 20

(13) IQWiG. Methoden Version 2.0. Internet 2006 December 19;(Version 2.0):28-29.
Ref ID: 22

(14) Yokoyama H, Tada J, Kamikawa F, Kanno S, Yokota Y, Kuramitsu M. Efficacy of conversion from bedtime NPH insulin to morning insulin glargine in type 2 diabetic patients on basal-prandial insulin therapy. Diabetes Research and Clinical Practice 2006 July;73(1):35-40.
Ref ID: 6

(15) Janka HU, Plewe G, Riddle MC, Kliebe-Frisch C, Schweitzer MA, Yki-Jarvinen H. Comparison of basal insulin added to oral agents versus twice-daily premixed insulin as initial insulin therapy for type 2 diabetes. Diabetes Care 2005 February;28(2):254-9.
Ref ID: 13

(16) Bretzel RG, Nuber U, Landgraf W, Owens DR, Bradley C, Linn T. Once-daily basal insulin glargine versus thrice-daily prandial insulin lispro in people with type 2 diabetes on oral hypoglycaemic agents (APOLLO): an open randomised controlled trial. The Lancet371(9618):1073-84.
Ref ID: 3

(17) Standl E, Maxeiner S, Raptis S. Once-daily Insulin Glargine Administration in the Morning Compared to Bedtime in Combination with Morning Glimepiride in Patients with Type 2 Diabetes: An Assessment of Treatment Flexibility. Hormone and Metabolic Research 2006;(03):172-7.
Ref ID: 79

(18) CEBM, Centre for Evidence-Based Medicine O. Levels of evidence (May 2001). Internet 2001 May;Available from: URL: [http://www.cebm.net/index.aspx?o=1025
Ref ID: 23

(19) Sackett DL RWGJHRRW. Evidence-Based medicine: What it is and what it isn`t. BMJ 2008;1996(312):71.-72.
Ref ID: 24

(20) Rosenstock J, Dailey G, Massi-Benedetti M, Fritsche A, Lin Z, Salzman A. Reduced hypoglycemia risk with insulin glargine: a meta-analysis comparing insulin glargine with human NPH insulin in type 2 diabetes. Diabetes Care 2005 April;28(4):950-5.
Ref ID: 14

(21) Mullins P, Sharplin P, Yki-Jarvinen H, Riddle MC, Haring HU. Negative binomial meta-regression analysis of combined glycosylated hemoglobin and hypoglycemia outcomes across eleven Phase III and IV studies of insulin glargine compared with neutral protamine Hagedorn insulin in type 1 and type 2 diabetes mellitus. Clin Ther 2007 August;29(8):1607-19.
Ref ID: 29

(22) Daikeler R, Högy B, Donaubauer B. Auswirkungen einer Basis-Bolus-Therapie (ICT) mit Insulin glargin und Insulinglulisin auf patientenbezogene Endpunkte (PRO) bei Diabetes mellitus. Diabetologie und Stoffwechsel 2007;2:308-14.
Ref ID: 30

(23) Bradley CF, Plowright RF, Stewart JF, Valentine JF, Witthaus E. The Diabetes Treatment Satisfaction Questionnaire change version (DTSQc) evaluated in insulin glargine trials shows greater responsiveness to improvements than the original DTSQ.(1477-7525 (Electronic)).
Ref ID: 31

(24) Cohen J. A Power Primer. Psychological Bulletin 1992 May 2;112(1):155-9.
Ref ID: 32

(25) Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2 ed. Lawrence Erlbaum Associates; 1988.
Ref ID: 36

(26) C.Bradley. **Diabetes Treatment Satisfaction Questionnaire (DTSQ)**. Internet 2006;Available from: URL: www.musc.edu/dfm/RCMAR/DTSQ.html
Ref ID: 5

(27) Bradley C. Feedback on the FDA's February 2006 draft guidance on Patient Reported Outcome (PRO) measures from a developer of PRO measures.(1477-7525 (Electronic)).
Ref ID: 34

(28) Fritsche A, Schweitzer MA, Haring HU. Glimepiride combined with morning insulin glargine, bedtime neutral protamine hagedorn insulin, or bedtime insulin glargine in patients with type 2 diabetes. A randomized, controlled trial. Ann Intern Med 2003 June 17;138(12):952-9.
Ref ID: 8

(29) Eliaschewitz FG, Calvo C, Valbuena H, Ruiz M, Aschner P, Villena J et al. Therapy in type 2 diabetes: insulin glargine vs. NPH insulin both in combination with glimepiride. Arch Med Res 2006 May;37(4):495-501.
Ref ID: 10

(30) Mengel K. Antidiabetika. In: Schwabe U, Paffrath D, editors. Arzneiverordnungsreport 2007.Heidelberg: Springer; 2007. p. 327-46.
Ref ID: 21

(31) Standl E, Maxeiner S, Raptis S, Karimi-Anderesi Z, Schweitzer MA, the HOE. Good Glycemic Control With Flexibility in Timing of Basal Insulin Supply: A 24-week comparison of insulin glargine given once daily in the morning or at bedtime in combination with morning glimepiride. Diabetes Care 2005 February 1;28(2):419-20.
Ref ID: 1

A 1.7.1 Analysen auf Basis von individuellen Patientendaten (IPD)

Nachfolgend werden die von der Firma Sanofi-Aventis Deutschland GmbH im Rahmen des Stellungnahmeverfahrens zur Verfügung gestellten IPD-Meta-Analysen zur gemeinsamen Betrachtung von Hypoglykämien und der langfristigen Blutzuckersenkung dargestellt.



Wissenschaftliche Erörterung
Vorbericht A05-03

Zusätzliche Informationen, Meta-Analysen individueller Patientendaten

27.08.2008

IQWiG – Auswertungsanfrage vom 24.06.2008

1. Auswertung individueller Patientendaten zur gemeinsamen Betrachtung von Hypoglykämien und dem HbA1c-Wert

a. Anfrage folgender Auswertungen:

Auswertung Nr.	Endpunkt	Zeitpunkt der Hypoglykämie (gemäß Definition in den Studien)	Hypoglykämie bestätigt durch Blutzucker-messung	Adjustiert für
1	Anteil Patienten mit mindestens einer Hypoglykämie (Kombination aus Symptomen und Blutzuckermessung)	tagsüber (disjunkt zu nächtlichen)	< 36 mg/dl	Änderung des HbA1c von Studienbeginn bis -ende
2			< 70 mg/dl	Änderung des HbA1c von Studienbeginn bis -ende
3		nächtlich	< 36 mg/dl	Änderung des HbA1c von Studienbeginn bis -ende
4			< 70 mg/dl	Änderung des HbA1c von Studienbeginn bis –ende
5	Anteil Patienten mit mindestens einer schweren Hypoglykämie (Gemäß Definition in den Studien)	tagsüber (disjunkt zu nächtlichen)	-	Änderung des HbA1c von Studienbeginn bis –ende
6		nächtlich	-	Änderung des HbA1c von Studienbeginn bis –ende

b. Für folgende Studien(-pools):

I	II	III	IV	V
3002 4001 (Vergleich Glargin abends vs. NPH-Insulin) 4002 4012 4013 6001	4001 (Vergleich Glargin morgens vs. NPH-Insulin)	3102	4016	3006

c. Benötigte Informationen

- Auswertung der Analysen 1 bis 6 mittels logistischer Regression auf Basis von individuellen Patientendaten für die IIT-Population der Studien(-pools) I bis V.
- Benötigte Informationen:
  - Anzahl der in die Auswertung eingehenden Patienten pro Studie
  - Umgang mit fehlenden Werten (z.B. Ersetzung fehlender Werte für HbA1c mittels „last observation carried forward")
  - Distanzmaß Odds ratio inkl. 95%-Konfidenzintervall
  - Angabe von Kriterien für die Modellgüte, insbesondere der Residuenplots

2. Unklarheiten in der Definition für symptomatische Hypoglykämien in den Studien 4001 und 4013

Der Text des Studienberichts lautet: Hypoglycemia was either symptomatic (…) or asymptomatic (…) and mild (between 50 mg/dl and 75 mg/dl) or moderate (below 50 mg/dl).

Aus diesem Satz geht unseres Erachtens nicht eindeutig hervor, ob für jede symptomatische Hypoglykämie die Bestätigung durch eine Blutzuckermessung mindestens unter 75 mg/dl obligat war. Wir interpretieren diese Formulierung folgendermaßen:

- Die Hypoglykämien werden unterteilt in symptomatische und asymptomatische, die jeweils entweder mild oder moderat sein können, je nach Wert der Blutzuckerbestätigung. Es ergäben sich somit die folgenden 4 Kategorien:
  - Symptomatische milde H.
  - Symptomatische moderate H.
  - Asymptomatische milde H.
  - Asymptomatische moderate H.

Ebenso wäre jedoch folgende Interpretation denkbar:

- Die Hypoglykämien werden unterteilt in symptomatische und asymptomatische und, unabhängig davon milde und moderate. Es ergäben sich somit die folgenden 4 Kategorien:
  - Symptomatische H. (mit oder ohne Blutzuckerbestätigung)
  - Asymptomatische H.
  - Milde H.
  - Moderate H.

Wir bitten um Klarstellung dieses Punktes.

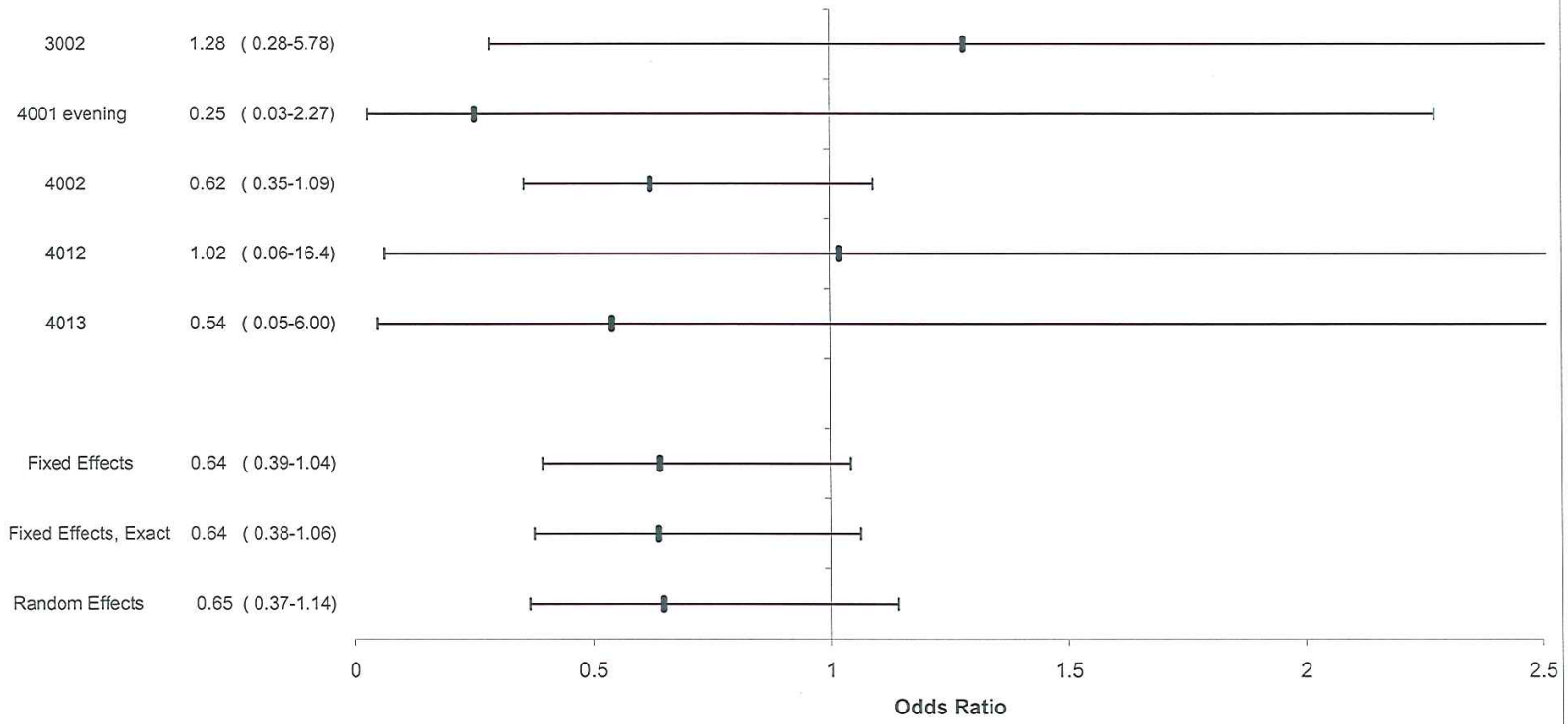
Meta Analyses Studypool I

Percentage of subjects having at least one hypoglycemic event: Daytime Symptomatic SMBG < 36mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)		P Value
3002	4/280 (1.4%)	3/266 (1.1%)	1.28	(0.28-5.78)	0.7476
4001 evening	1/224 (0.4%)	4/227 (1.8%)	0.25	(0.03-2.27)	0.2192
4002	21/355 (5.9%)	35/380 (9.2%)	0.62	(0.35-1.09)	0.0975
4012	1/218 (0.5%)	1/219 (0.5%)	1.02	(0.06-16.4)	0.9902
4013	1/226 (0.4%)	2/246 (0.8%)	0.54	(0.05-6.00)	0.6163
Total	28/1303 (2.1%)	45/1338 (3.4%)			
Results from subject level (fixed effects) model					
Slope: HbA1c Change			Non-significant change in hypo odds		0.5912
Odds Ratio Homogeneity					0.7773
Odds Ratio Summaries					
Fixed Effects*			0.64	(0.39-1.04)	0.0726
Fixed Effects, Exact Inference**			0.64	(0.38-1.06)	0.0874
Random Effects***			0.65	(0.37-1.14)	0.1006

70 subjects missing their HbA1c change from baseline were excluded, and none of these subjects had a hypo event of this type.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

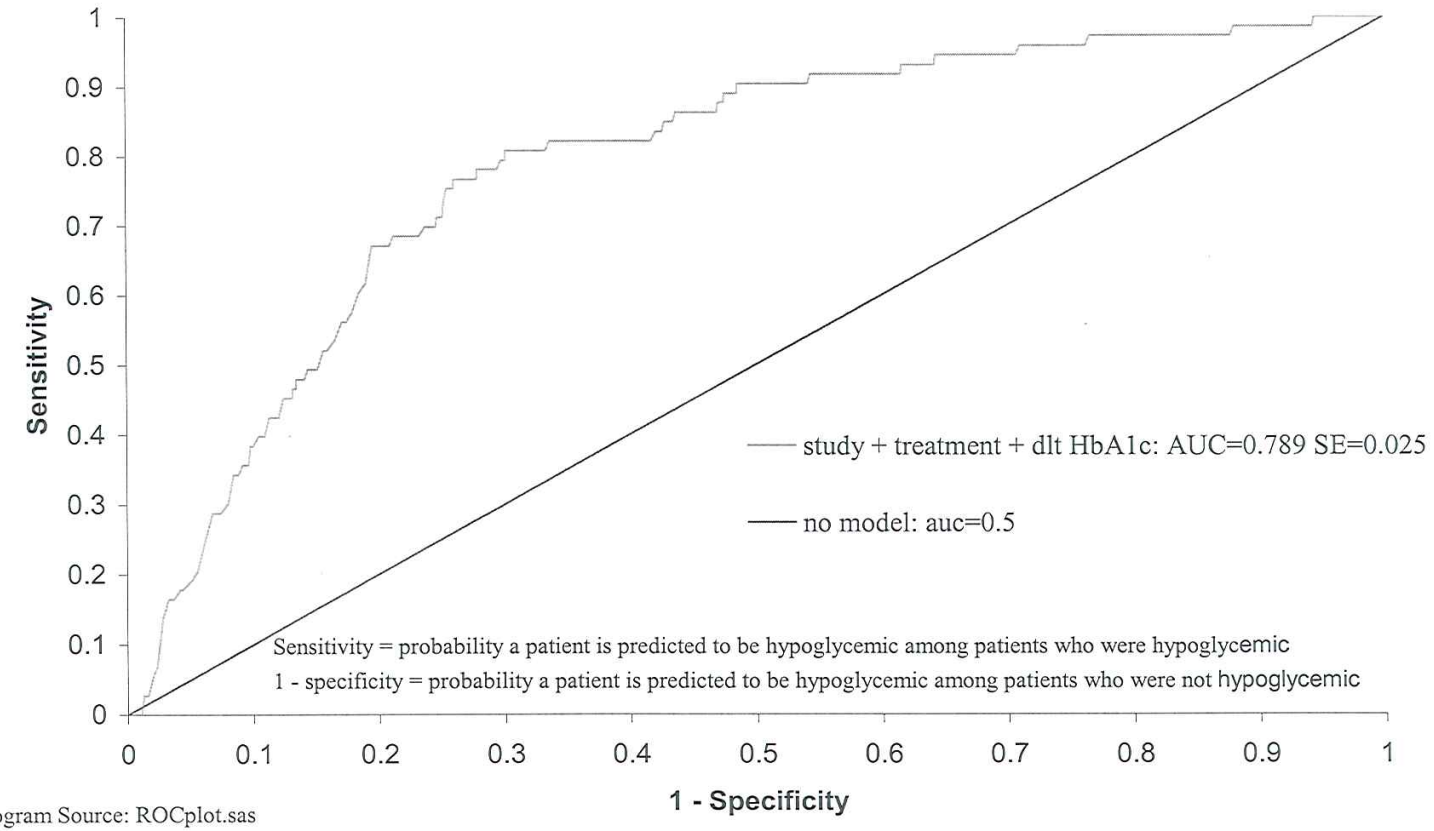
Percentage of subjects having at least one hypoglycemic event:
Daytime Symptomatic SMBG < 36mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



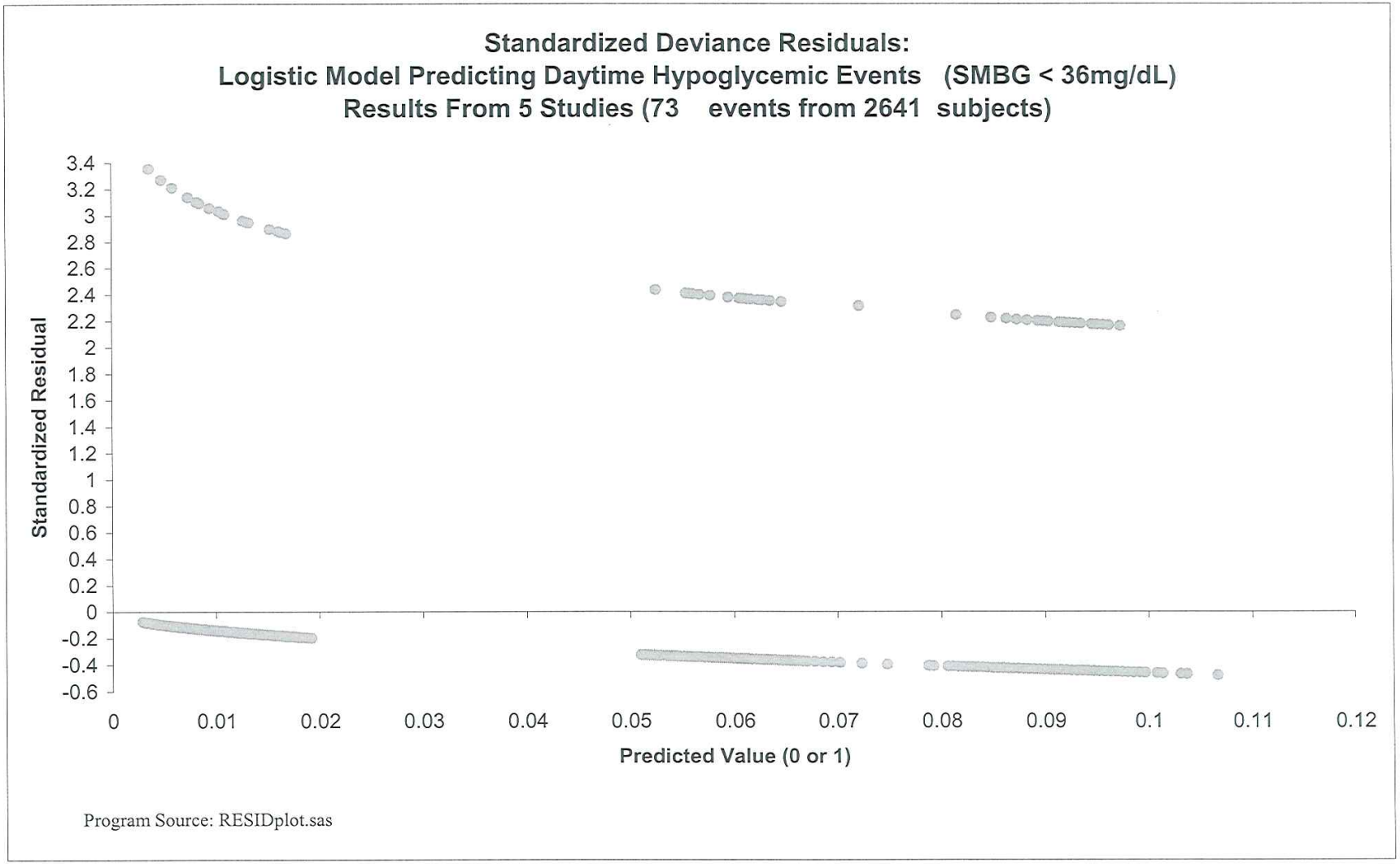
70 subjects missing their HbA1c change from baseline were excluded: none of these subjects had a hypo event of this type

Program Source: forest70d.sas

ROC Curves: Logistic Models Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
Results From 5 Studies (73 events from 2641 subjects)



Program Source: ROCplot.sas

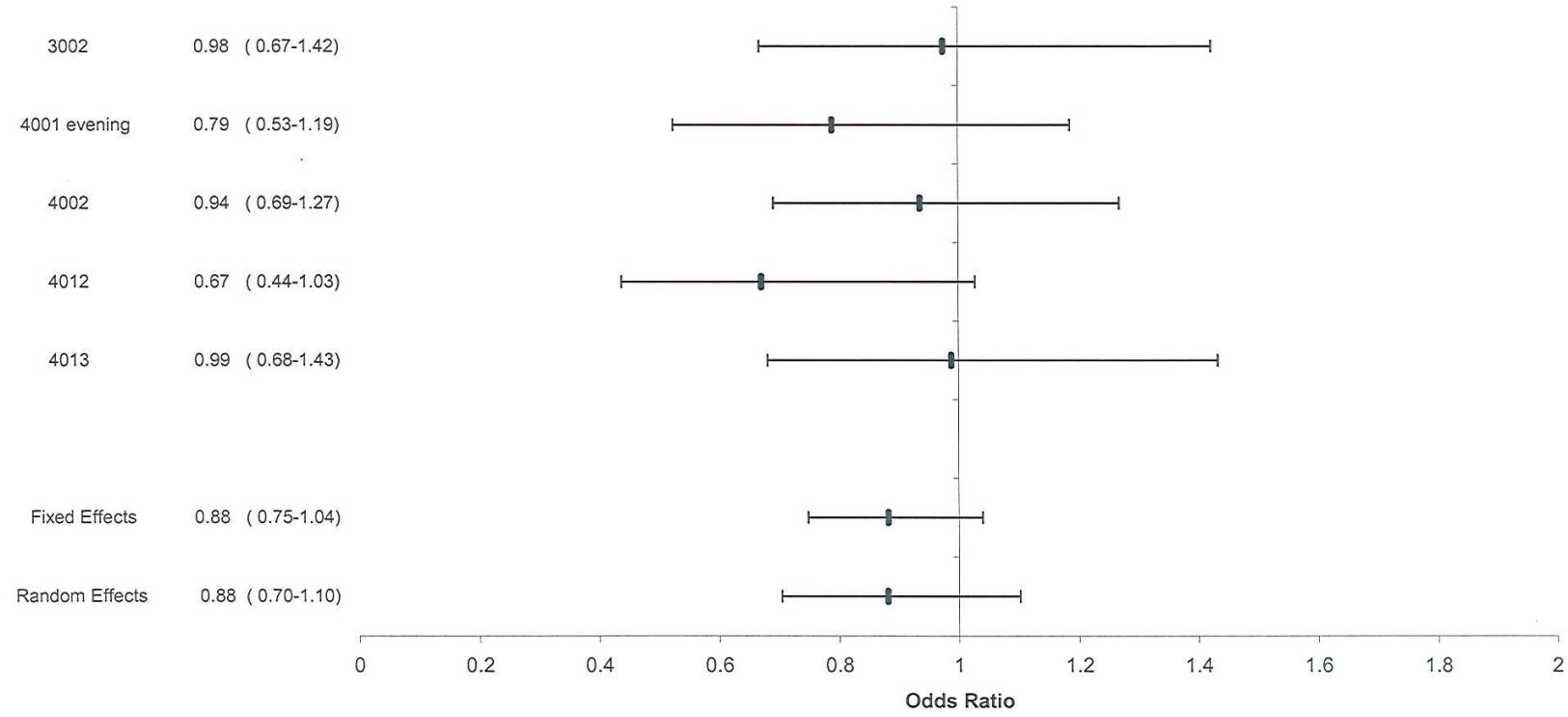

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
Results From 5 Studies (73 events from 2641 subjects)
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
0.11
0.12
Predicted Value (0 or 1)
Program Source: RESIDplot.sas

Percentage of subjects having at least one hypoglycemic event: Daytime Symptomatic SMBG < 70mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)	P Value
3002	77/280 (27.5%)	73/266 (27.4%)	0.98 (0.67-1.42)	0.8972
4001 evening	61/224 (27.2%)	72/227 (31.7%)	0.79 (0.53-1.19)	0.2556
4002	228/355 (64.2%)	249/380 (65.5%)	0.94 (0.69-1.27)	0.6692
4012	51/218 (23.4%)	67/219 (30.6%)	0.67 (0.44-1.03)	0.0669
4013	88/226 (38.9%)	97/246 (39.4%)	0.99 (0.68-1.43)	0.9501
Total	505/1303 (38.8%)	558/1338 (41.7%)		
Results from subject level (fixed effects) model				
Slope: HbA1c Change			17% increase in hypo odds per 1% decrease in A1c	< 0.0001
Odds Ratio Homogeneity				0.6178
Odds Ratio Summaries				
Fixed Effects*			0.88 (0.75-1.04)	0.1358
Random Effects**			0.88 (0.70-1.10)	0.1924

70 subjects missing their HbA1c change from baseline were excluded, including 11 subjects with a hypo. event: 3 glargine and 8 NPH subjects.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment.
** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

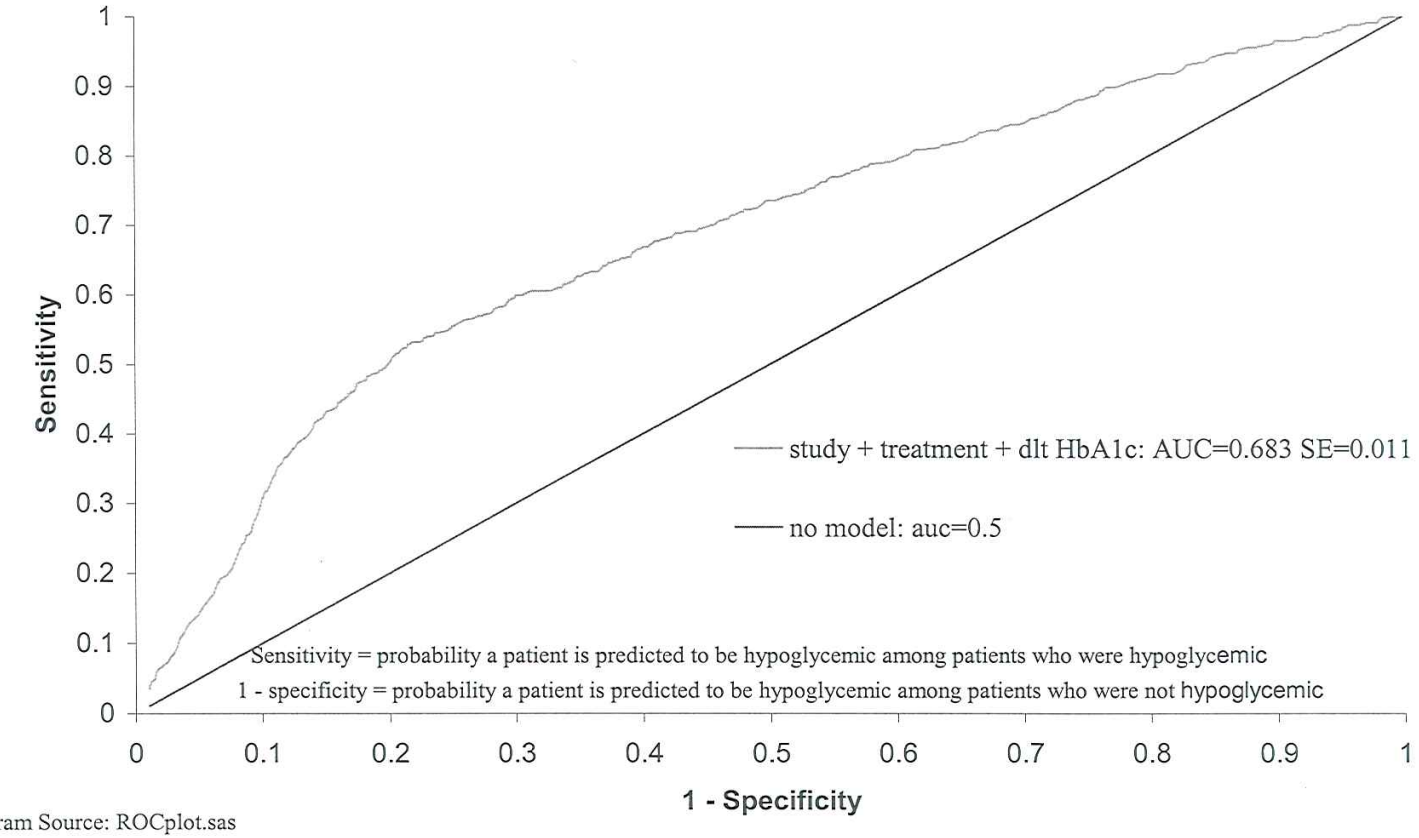
Percentage of subjects having at least one hypoglycemic event: Daytime Symptomatic SMBG < 70mg/dL Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



70 subjects missing their HbA1c change from baseline were excluded: including 11 subjects with a hypo. event: 3 glargine and 8 NPH

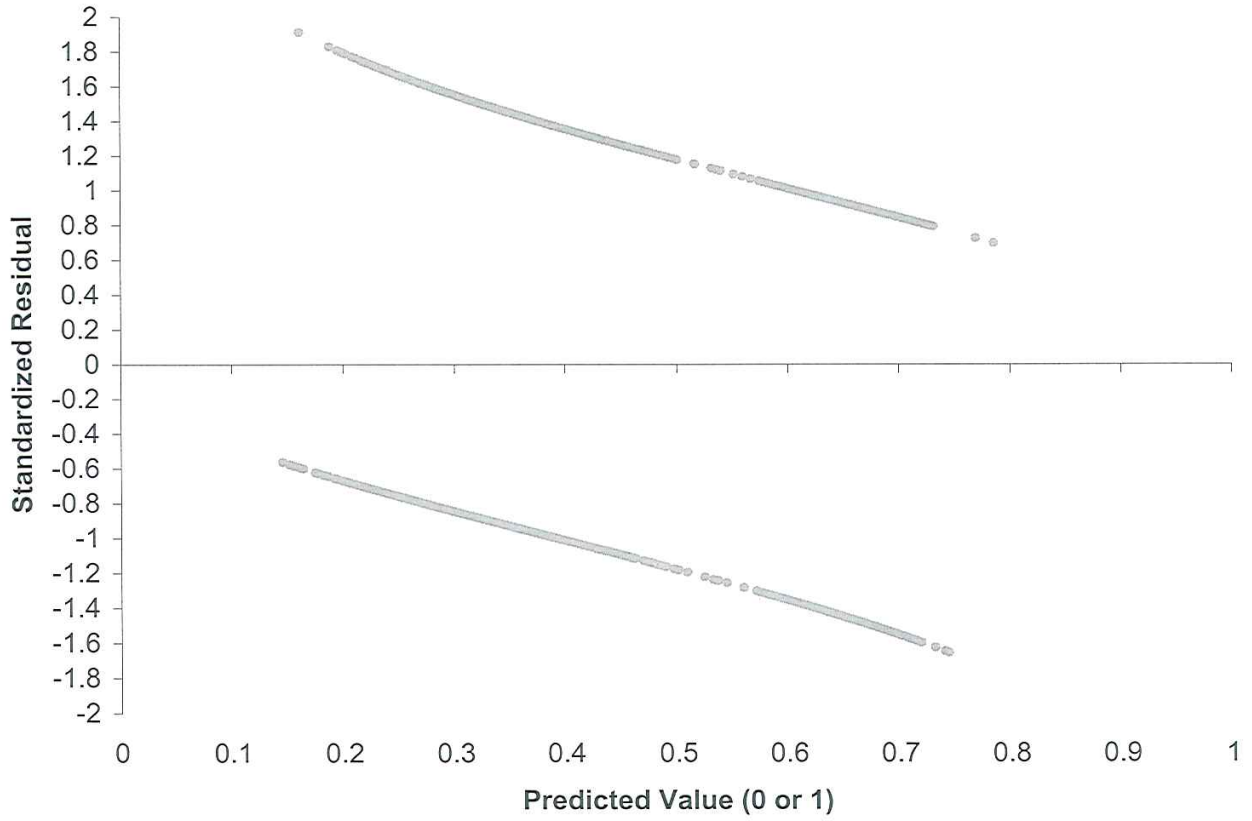
Program Source: forest70D.sas

ROC Curves: Logistic Models Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
Results From 5 Studies (1063 events from 2641 subjects)



Program Source: ROCplot.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
Results From 5 Studies (1063 events from 2641 subjects)



Program Source: RESIDplot.sas

Percentage of subjects having at least one hypoglycemic event: Nighttime Symptomatic SMBG < 36mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)		P Value
3002	2/280 (0.7%)	2/266 (0.8%)	0.94	(0.13-6.71)	0.9495
4001 evening	1/224 (0.4%)	2/227 (0.9%)	0.50	(0.05-5.56)	0.5731
4002	14/355 (3.9%)	31/380 (8.2%)	0.46	(0.24-0.88)	0.0191
4012	0/218 (0.0%)	2/219 (0.9%)			
4013	2/226 (0.9%)	7/246 (2.8%)	0.31	(0.06-1.49)	0.1424
Total	19/1303 (1.5%)	44/1338 (3.3%)			
Results from subject level (fixed effects) model					
Slope: HbA1c Change			Non-significant change in hypo odds		0.5766
Odds Ratio Homogeneity					0.6848
Odds Ratio Summaries					
Fixed Effects*			0.44	(0.25-0.76)	0.0034
Fixed Effects, Exact Inference**			0.44	(0.24-0.78)	0.0038
Random Effects***			0.46	(0.22-0.98)	0.0461

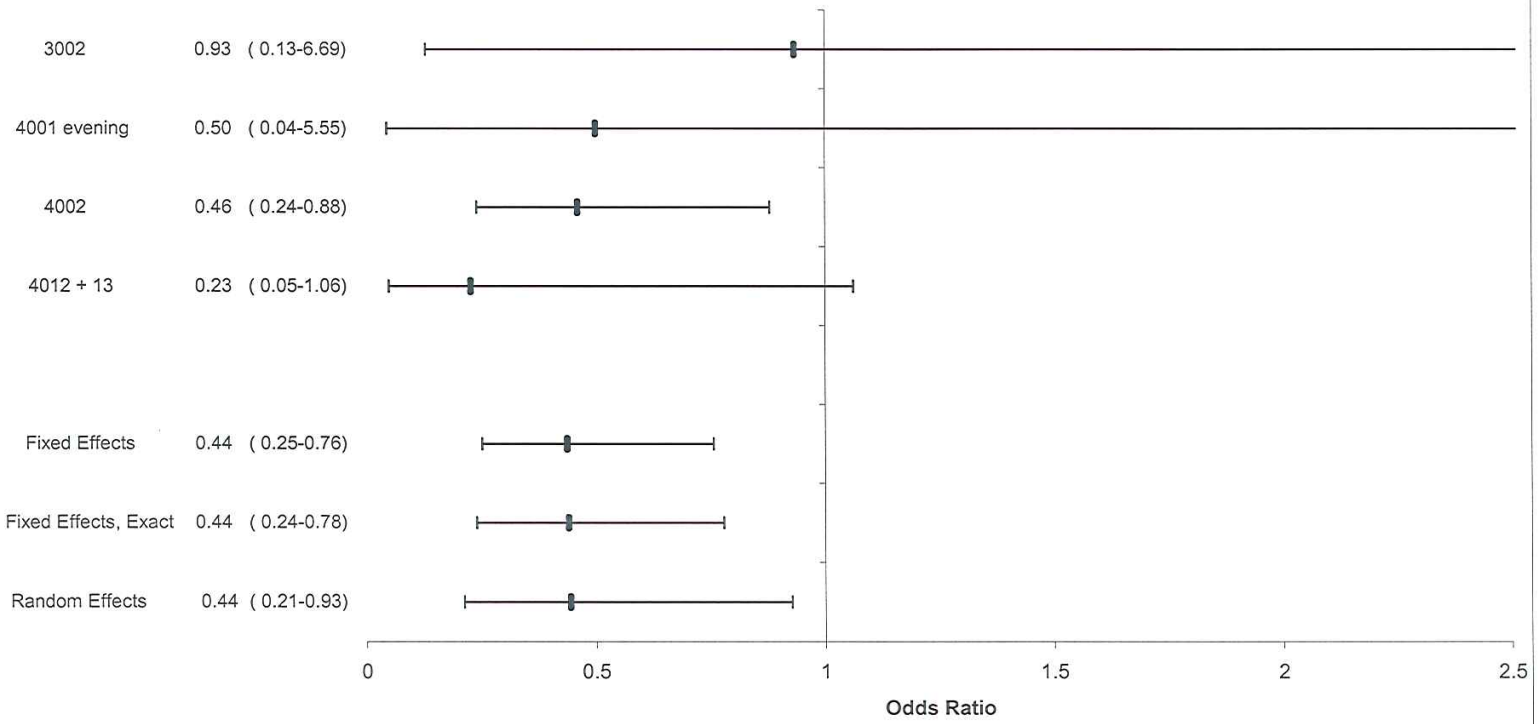
70 subjects missing their HbA1c change from baseline were excluded, including 2 subjects with a hypo. event: 0 glargine and 2 NPH subjects.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

Percentage of subjects having at least one hypoglycemic event: Nighttime Symptomatic SMBG < 36mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study (combining sudies 4012 and 4013)

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)		P Value
3002	2/280 (0.7%)	2/266 (0.8%)	0.93	(0.13-6.69)	0.9459
4001 evening	1/224 (0.4%)	2/227 (0.9%)	0.50	(0.04-5.55)	0.5717
4002	14/355 (3.9%)	31/380 (8.2%)	0.46	(0.24-0.88)	0.0188
4012+13	2/444 (0.5%)	9/465 (1.9%)	0.23	(0.05-1.06)	0.0597
Total	19/1303 (1.5%)	44/1338 (3.3%)			
Results from subject level (fixed effects) model					
Slope: HbA1c Change			Non-significant change in hypo odds		0.4623
Odds Ratio Homogeneity					0.7117
Odds Ratio Summaries					
Fixed Effects*			0.44	(0.25-0.76)	0.0031
Fixed Effects, Exact Inference**			0.44	(0.24-0.78)	0.0036
Random Effects***			0.44	(0.21-0.93)	0.0247

70 subjects missing their HbA1c change from baseline were excluded, including 2 subjects with a hypo. event: 0 glargine and 2 NPH subjects.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

Percentage of subjects having at least one hypoglycemic event:
Nighttime Symptomatic SMBG < 36mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

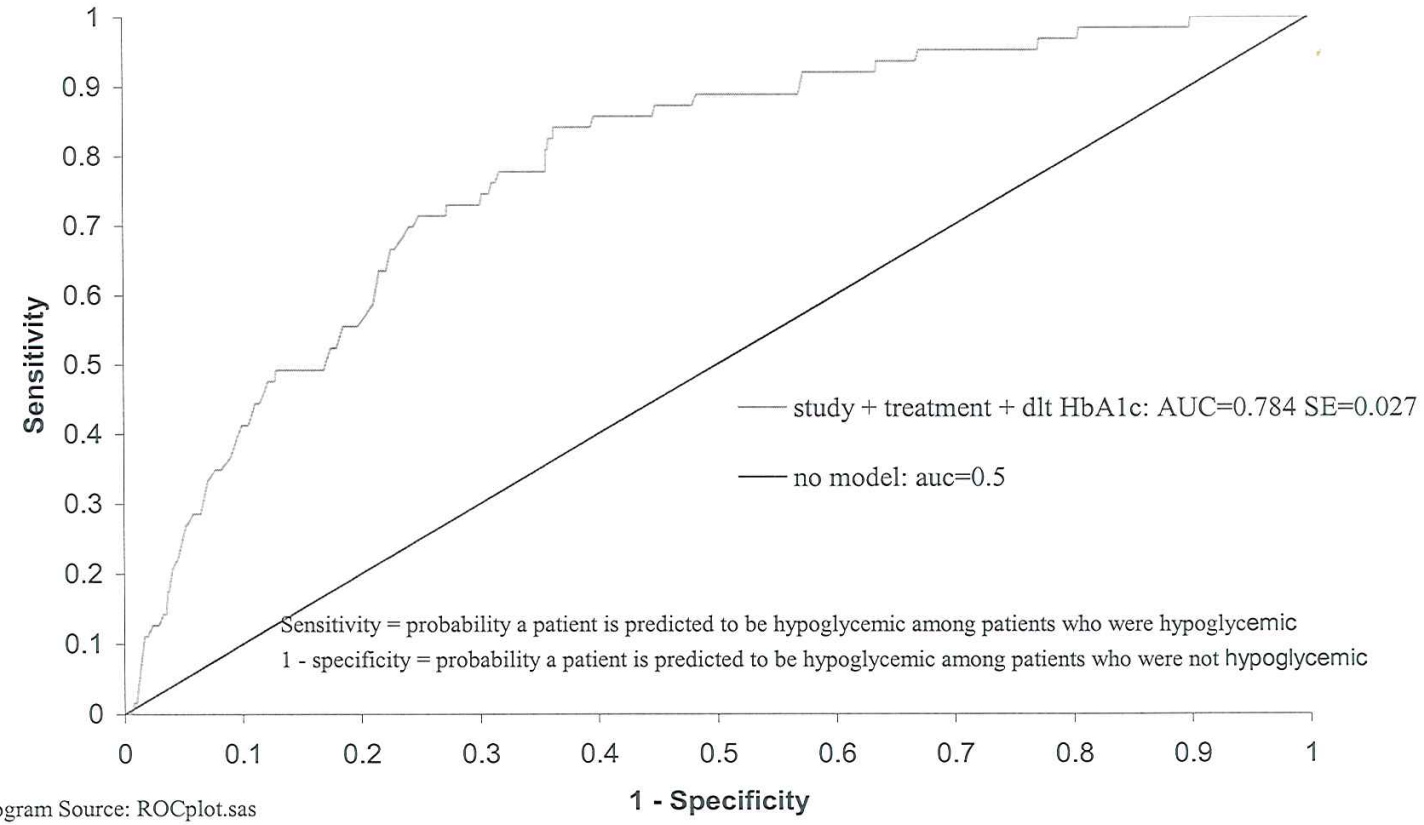


Events from 4012 were combined with 4013 since the glargine arm had zero events, 4013 used a similar study protocol, and all studies showed a glargine numerical advantage.

70 subjects missing their HbA1c change from baseline were excluded: including 2 subjects with a hypo. event: 0 glargine and 2 NPH

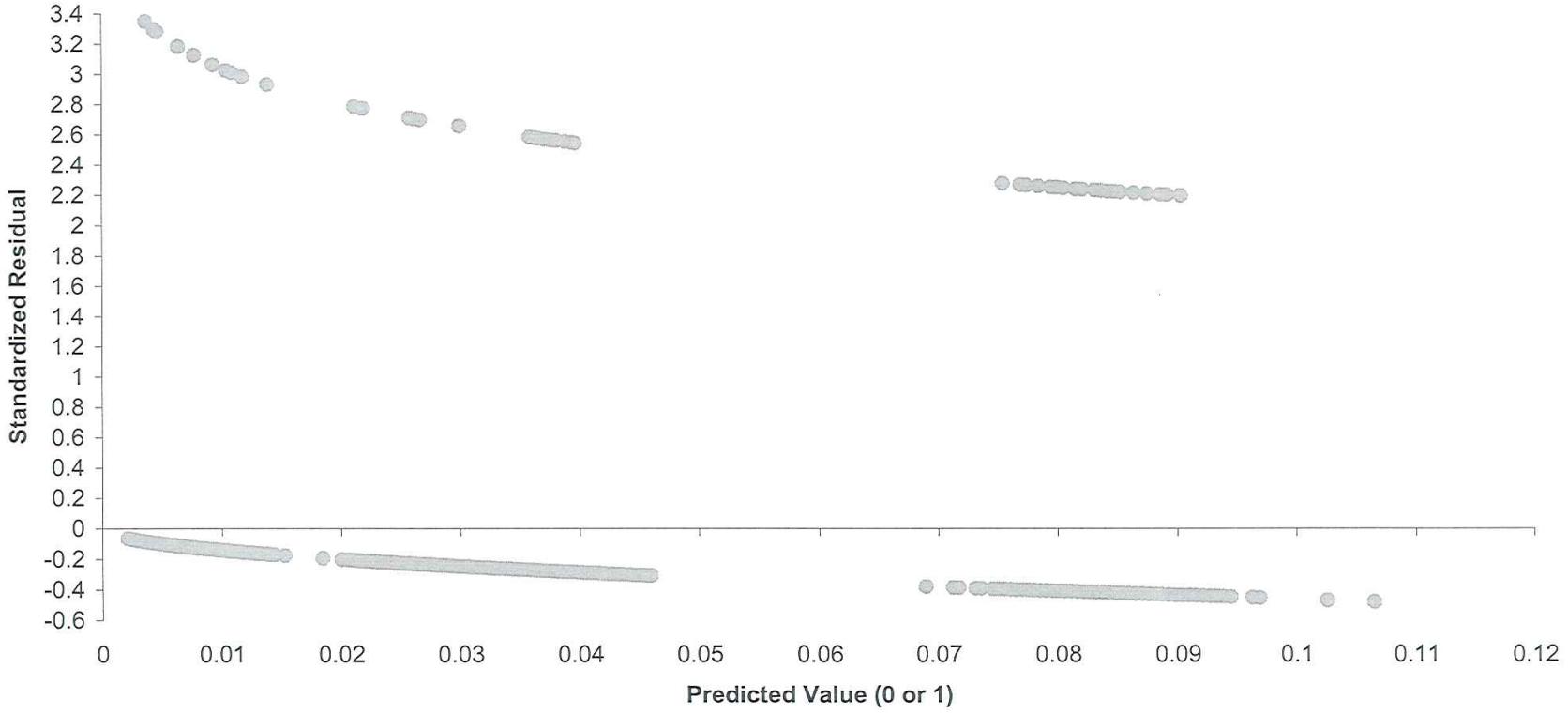
Program Source: forest36n.sas

ROC Curves: Logistic Models Predicting Evening Hypoglycemic Events (SMBG < 36mg/dL)
Results From 5 Studies (63 events from 2641 subjects)



Program Source: ROCplot.sas

Standardized Deviance Residuals:
Logistic Model Predicting Evening Hypoglycemic Events (SMBG < 36mg/dL)
Results From 5 Studies (63 events from 2641 subjects)



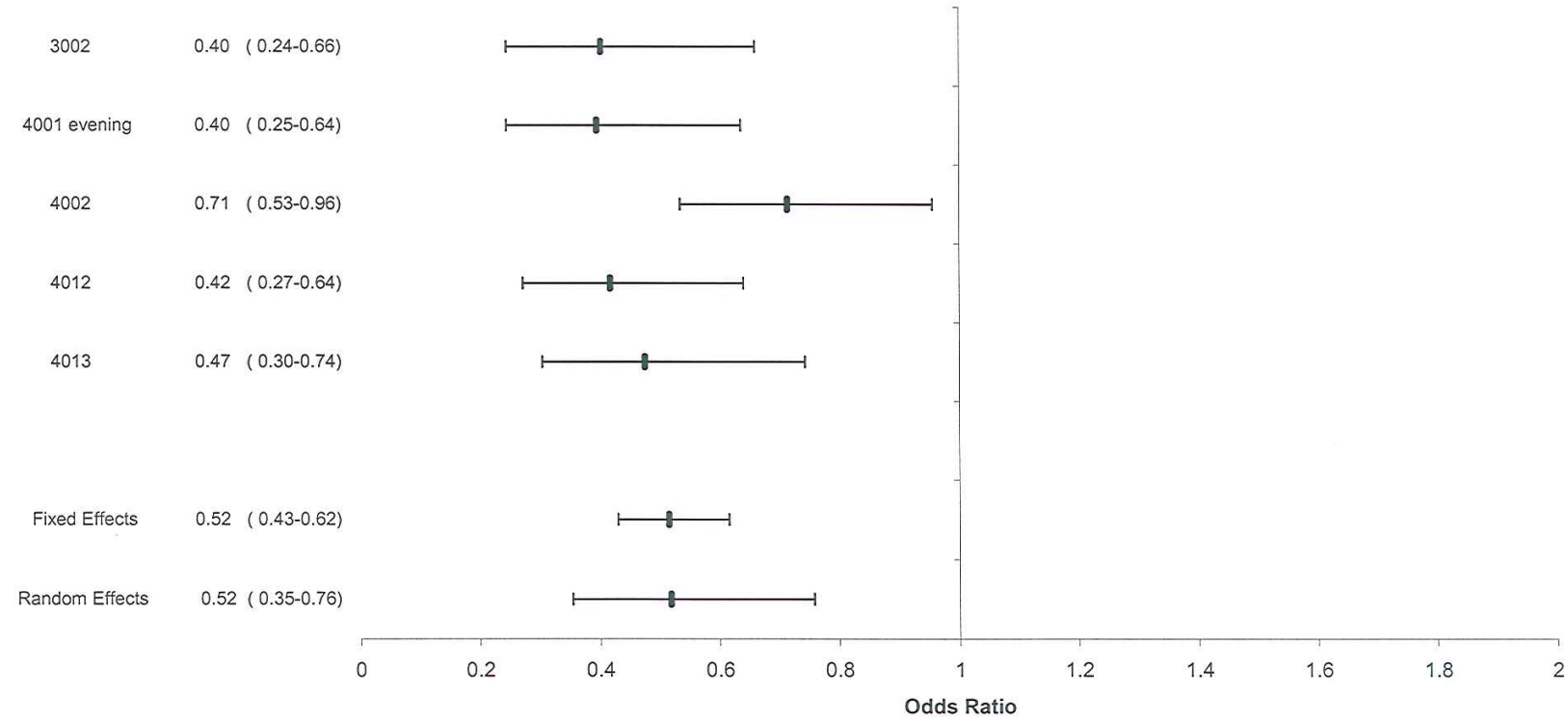
Program Source: RESIDplot.sas

Percentage of subjects having at least one hypoglycemic event: Nighttime Symptomatic SMBG < 70mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)	P Value
3002	27/280 (9.6%)	55/266 (20.7%)	0.40 (0.24-0.66)	0.0003
4001 evening	31/224 (13.8%)	65/227 (28.6%)	0.40 (0.25-0.64)	0.0001
4002	153/355 (43.1%)	195/380 (51.3%)	0.71 (0.53-0.96)	0.0234
4012	45/218 (20.6%)	83/219 (37.9%)	0.42 (0.27-0.64)	< 0.0001
4013	37/226 (16.4%)	72/246 (29.3%)	0.47 (0.30-0.74)	0.0011
Total	293/1303 (22.5%)	470/1338 (35.1%)		
Results from subject level (fixed effects) model				
Slope: HbA1c Change			10% increase in hypo odds per 1% decrease in A1c	0.0148
Odds Ratio Homogeneity				0.0891
Odds Ratio Summaries				
Fixed Effects*			0.52 (0.43-0.62)	< 0.0001
Random Effects**			0.52 (0.35-0.76)	0.0086

70 subjects missing their HbA1c change from baseline were excluded, including 12 subjects with a hypo. event: 2 glargine and 10 NPH subjects.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

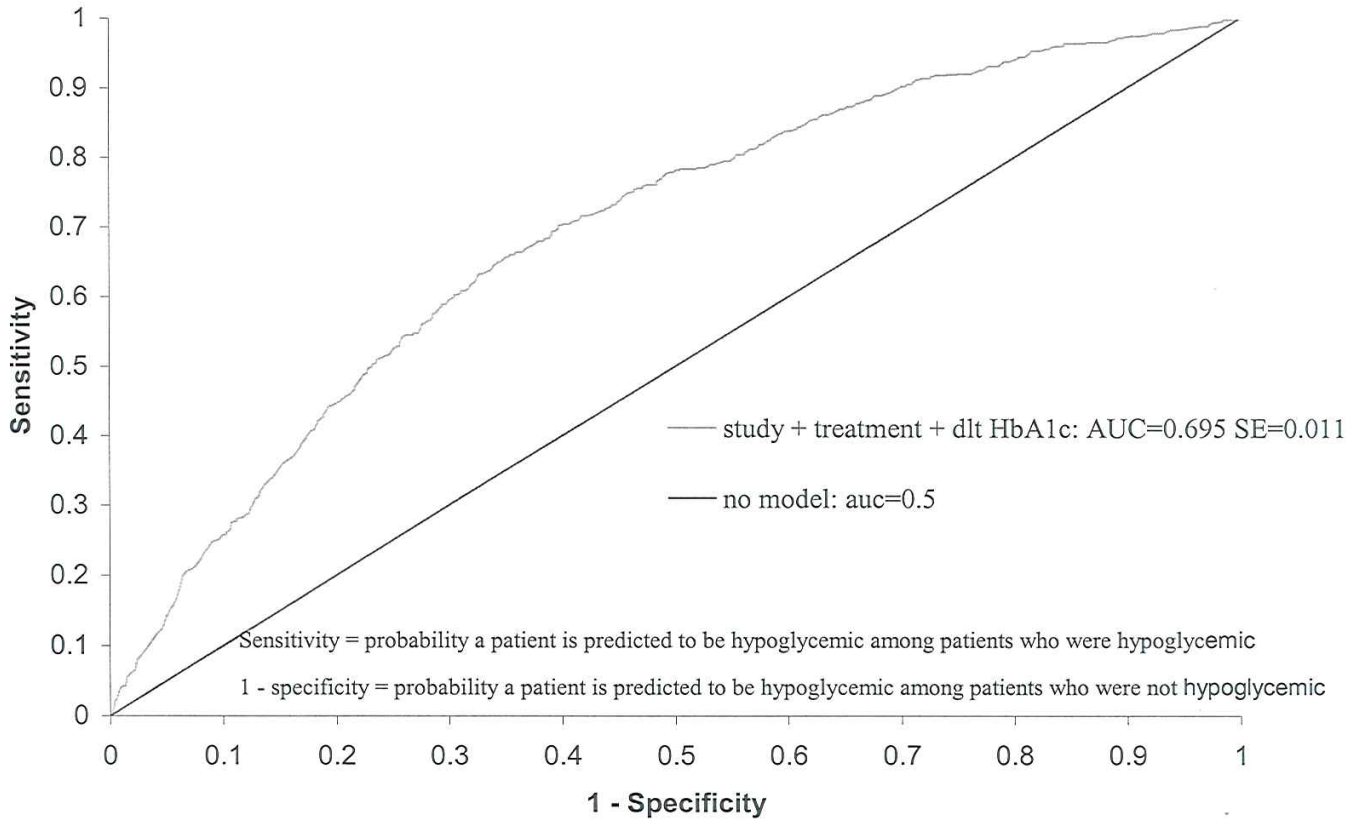
Percentage of subjects having at least one hypoglycemic event:
Nighttime Symptomatic SMBG < 70mg/dL
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



70 subjects missing their HbA1c change from baseline were excluded: including 12 subjects with a hypo. event: 2 glargine and 10 NPH

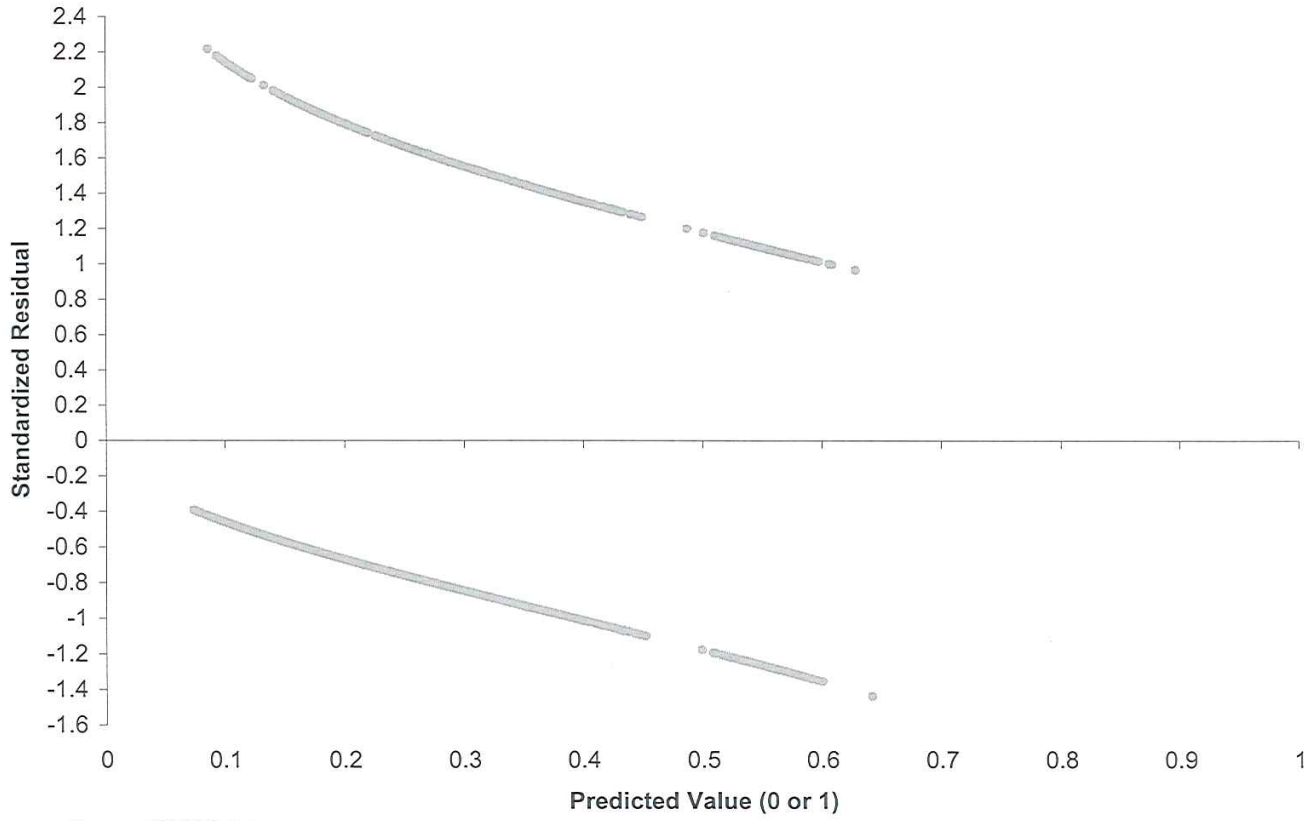
Program Source: forest70n.sas

ROC Curves: Logistic Models Predicting Evening Hypoglycemic Events (SMBG < 70mg/dL)
Results From 5 Studies (763 events from 2641 subjects)



Program Source: ROCplot.sas

Standardized Deviance Residuals:
Logistic Model Predicting Evening Hypoglycemic Events (SMBG < 70mg/dL)
Results From 5 Studies (763 events from 2641 subjects)



Program Source: RESIDplot.sas

Percentage of subjects having at least one severe hypoglycemic event during the daytime hours: Severe defined by study
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)		P Value
3002	3/280 (1.1%)	0/266 (0.0%)			
4001 evening	2/224 (0.9%)	3/227 (1.3%)	0.65	(0.11-3.95)	0.6416
4002	6/355 (1.7%)	3/380 (0.8%)	2.12	(0.53-8.56)	0.2899
4012	2/218 (0.9%)	4/219 (1.8%)	0.47	(0.09-2.62)	0.3933
4013	2/226 (0.9%)	5/246 (2.0%)	0.44	(0.08-2.28)	0.3251
Total	15/1303 (1.2%)	15/1338 (1.1%)			
Results from subject level (fixed effects) model					
Slope: HbA1c Change			Non-significant change in hypo odds		0.0712
Odds Ratio Homogeneity					0.1319
Odds Ratio Summaries					
Fixed Effects*			1.02	(0.49-2.09)	0.9672
Fixed Effects, Exact Inference**			1.04	(0.47-2.29)	> 0.99
Random Effects***			0.80	(0.30-2.14)	0.5576

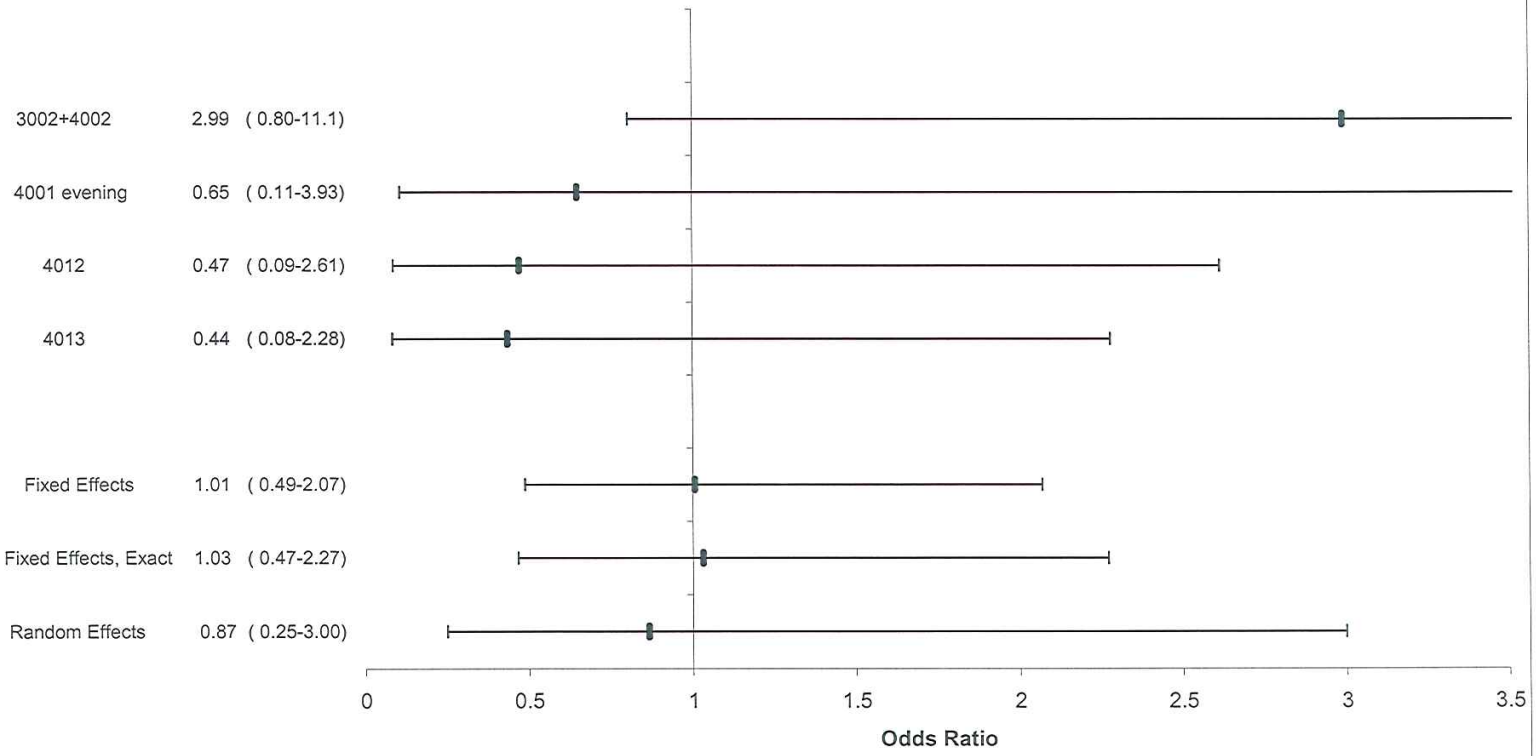
70 subjects missing their HbA1c change from baseline were excluded, and none of these subjects had a severe daytime hypo event.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

Percentage of subjects having at least one severe hypoglycemic event during the daytime hours: Severe defined by study
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study (pooling studies 3002, 4002)

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)	P Value
3002+4002	9/635 (1.4%)	3/646 (0.5%)	2.99 (0.80-11.1)	0.1019
4001 evening	2/224 (0.9%)	3/227 (1.3%)	0.65 (0.11-3.93)	0.6391
4012	2/218 (0.9%)	4/219 (1.8%)	0.47 (0.09-2.61)	0.3907
4013	2/226 (0.9%)	5/246 (2.0%)	0.44 (0.08-2.28)	0.3259
Total	15/1303 (1.2%)	15/1338 (1.1%)		
Results from subject level (fixed effects) model				
Slope: HbA1c Change			36% increase in hypo odds per 1% decrease in A1c	0.0382
Odds Ratio Homogeneity				0.1609
Odds Ratio Summaries				
Fixed Effects*			1.01 (0.49-2.07)	0.9878
Fixed Effects, Exact Inference**			1.03 (0.47-2.27)	> 0.99
Random Effects***			0.87 (0.25-3.00)	0.7302

70 subjects missing their HbA1c change from baseline were excluded, and none of these subjects had a severe daytime hypo event.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

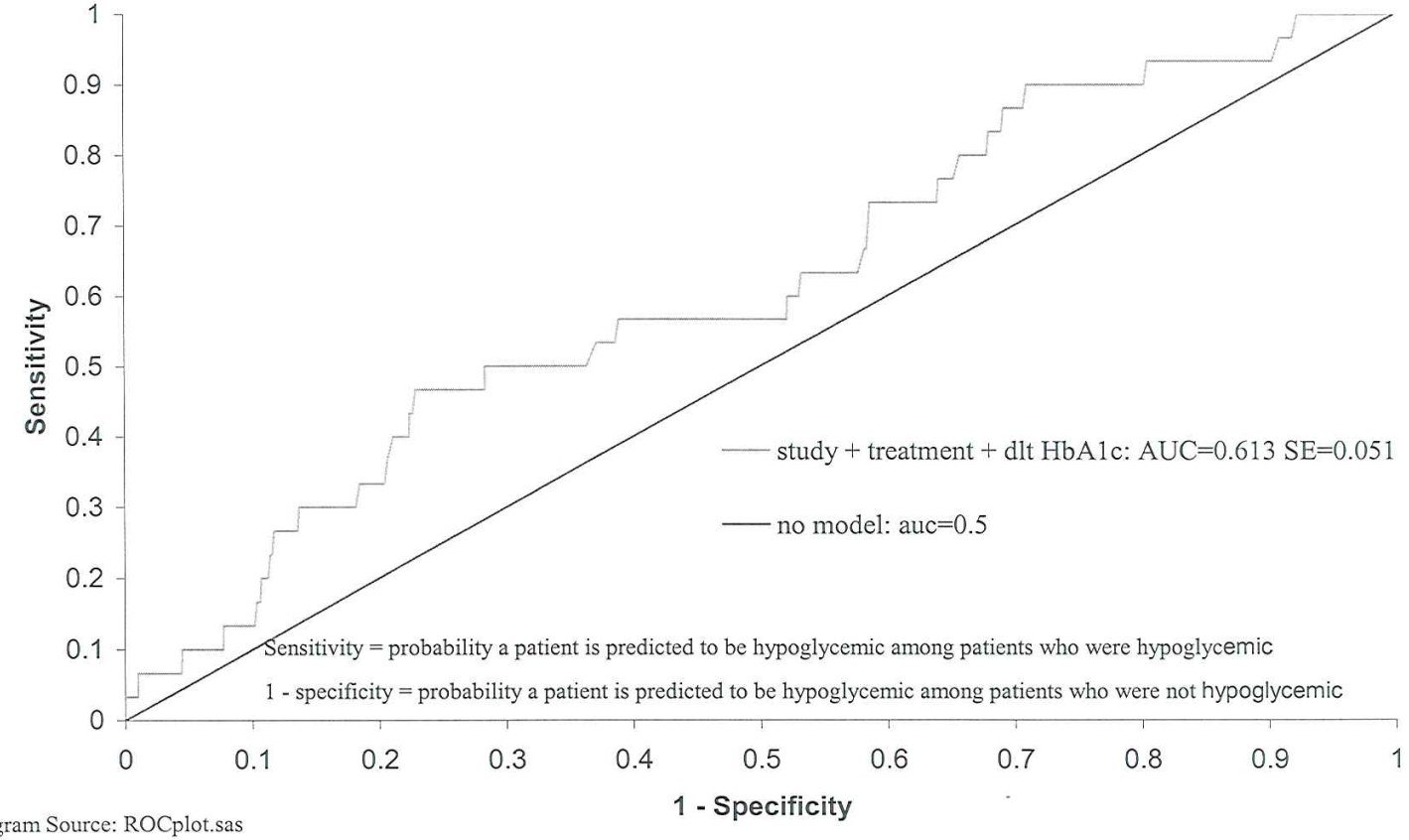
Percentage of subjects having at least one hypoglycemic event:
Daytime Severe
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



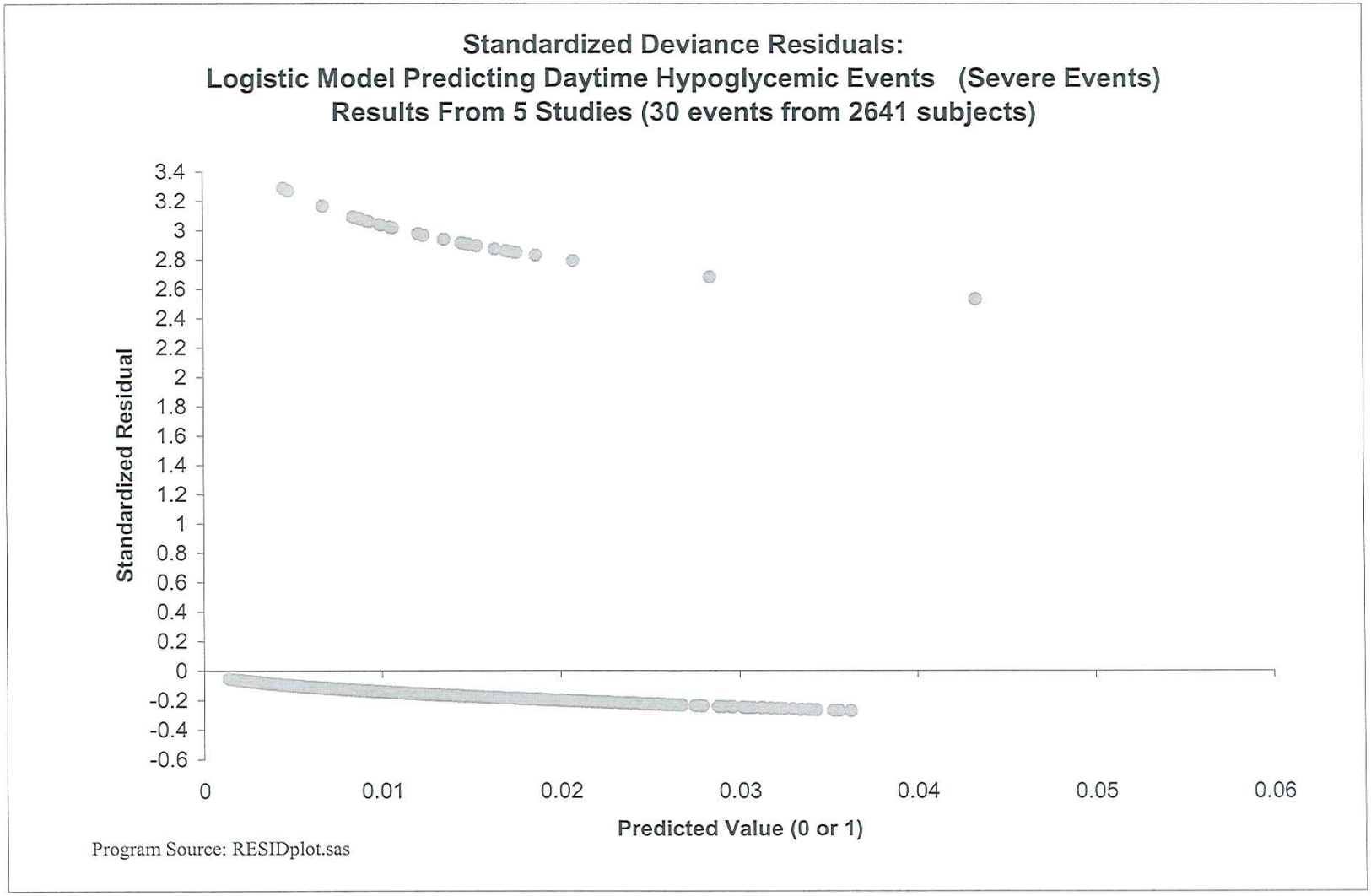
The results from 3002 with 0 events in the NPH arm was combined with 4002 because 4002 was the only other study with an NPH advatage.
70 subjects missing their HbA1c change from baseline were excluded: none had a severe daytime hypo event

Program Source: forestSVRd.sas

ROC Curves: Logistic Models Predicting Daytime Hypoglycemic Events (Severe Events)
Results From 5 Studies (30 events from 2641 subjects)



Program Source: ROCplot.sas

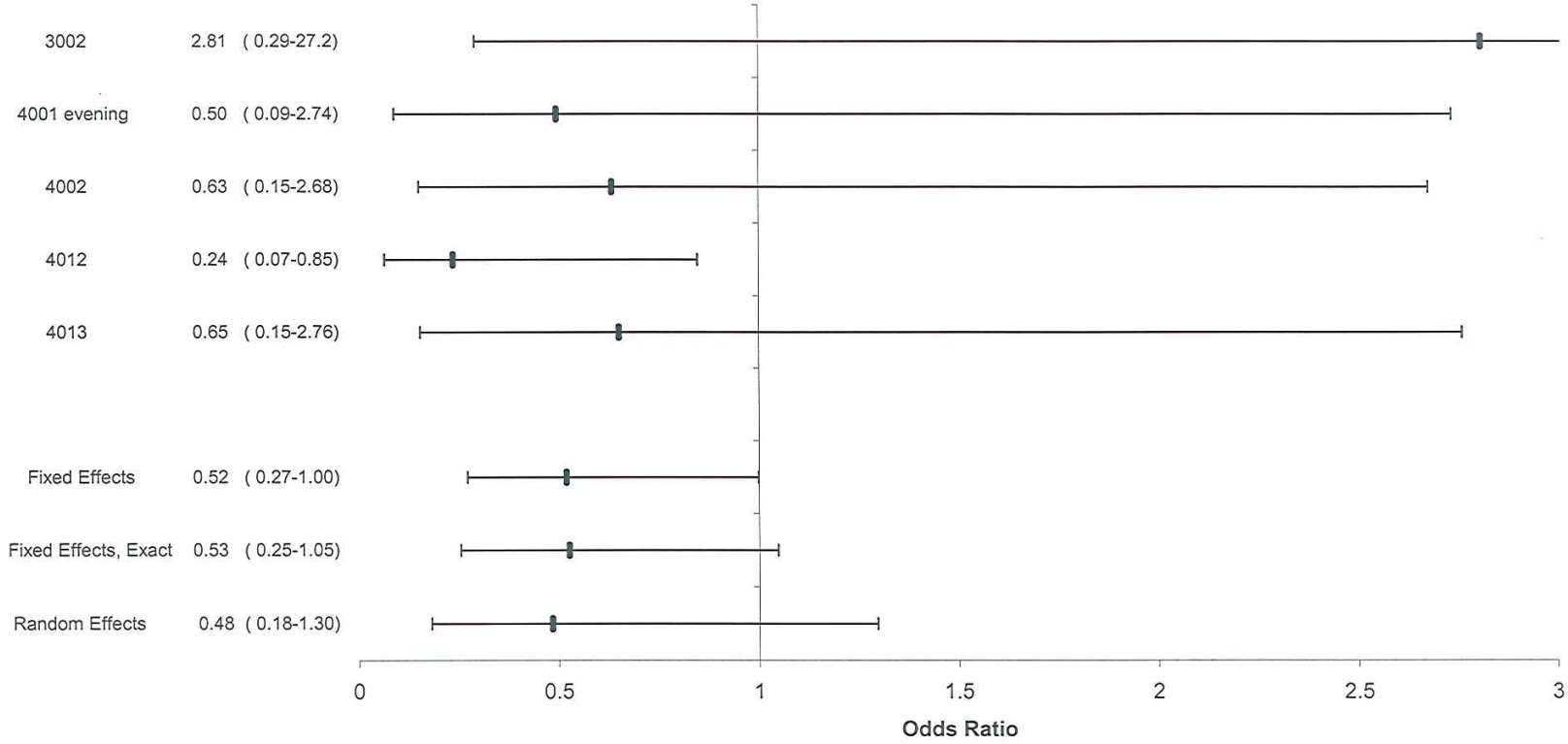

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (Severe Events)
Results From 5 Studies (30 events from 2641 subjects)
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
0.04
0.05
0.06
Predicted Value (0 or 1)
Program Source: RESIDplot.sas

Percentage of subjects having at least one severe hypoglycemic event during the nighttime hours: Severe defined by study
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)		P Value
3002	3/280 (1.1%)	1/266 (0.4%)	2.81	(0.29-27.2)	0.3723
4001 evening	2/224 (0.9%)	4/227 (1.8%)	0.50	(0.09-2.74)	0.4207
4002	3/355 (0.8%)	5/380 (1.3%)	0.63	(0.15-2.68)	0.5360
4012	3/218 (1.4%)	12/219 (5.5%)	0.24	(0.07-0.85)	0.0269
4013	3/226 (1.3%)	5/246 (2.0%)	0.65	(0.15-2.76)	0.5620
Total	14/1303 (1.1%)	27/1338 (2.0%)			
Results from subject level (fixed effects) model					
Slope: HbA1c Change			Non-significant change in hypo odds		0.3919
Odds Ratio Homogeneity					0.3543
Odds Ratio Summaries					
Fixed Effects*			0.52	(0.27-1.00)	0.0498
Fixed Effects, Exact Inference**			0.53	(0.25-1.05)	0.0702
Random Effects***			0.48	(0.18-1.30)	0.1105

70 subjects missing their HbA1c change from baseline were excluded, including 5 subjects with a hypo. event: 1 glargine and 4 NPH subjects.
* The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).
** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.
*** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).
. Log odds ratio standard errors took into account both the within and among study variances

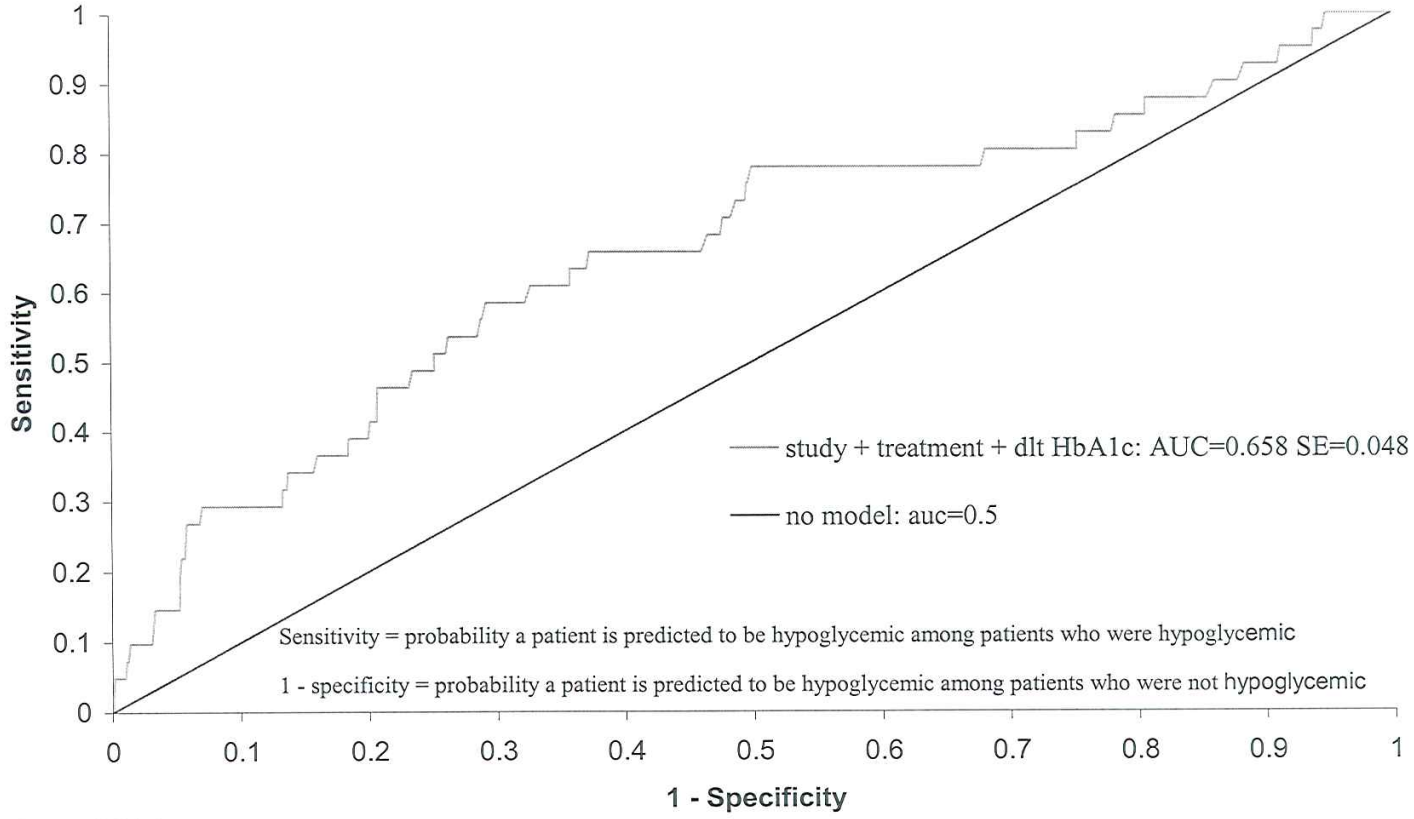
Percentage of subjects having at least one hypoglycemic event:
Nighttime Severe
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



70 subjects missing their HbA1c change from baseline were excluded: including 5 subjects with a hypo. event: 1 glargine and 4 NPH

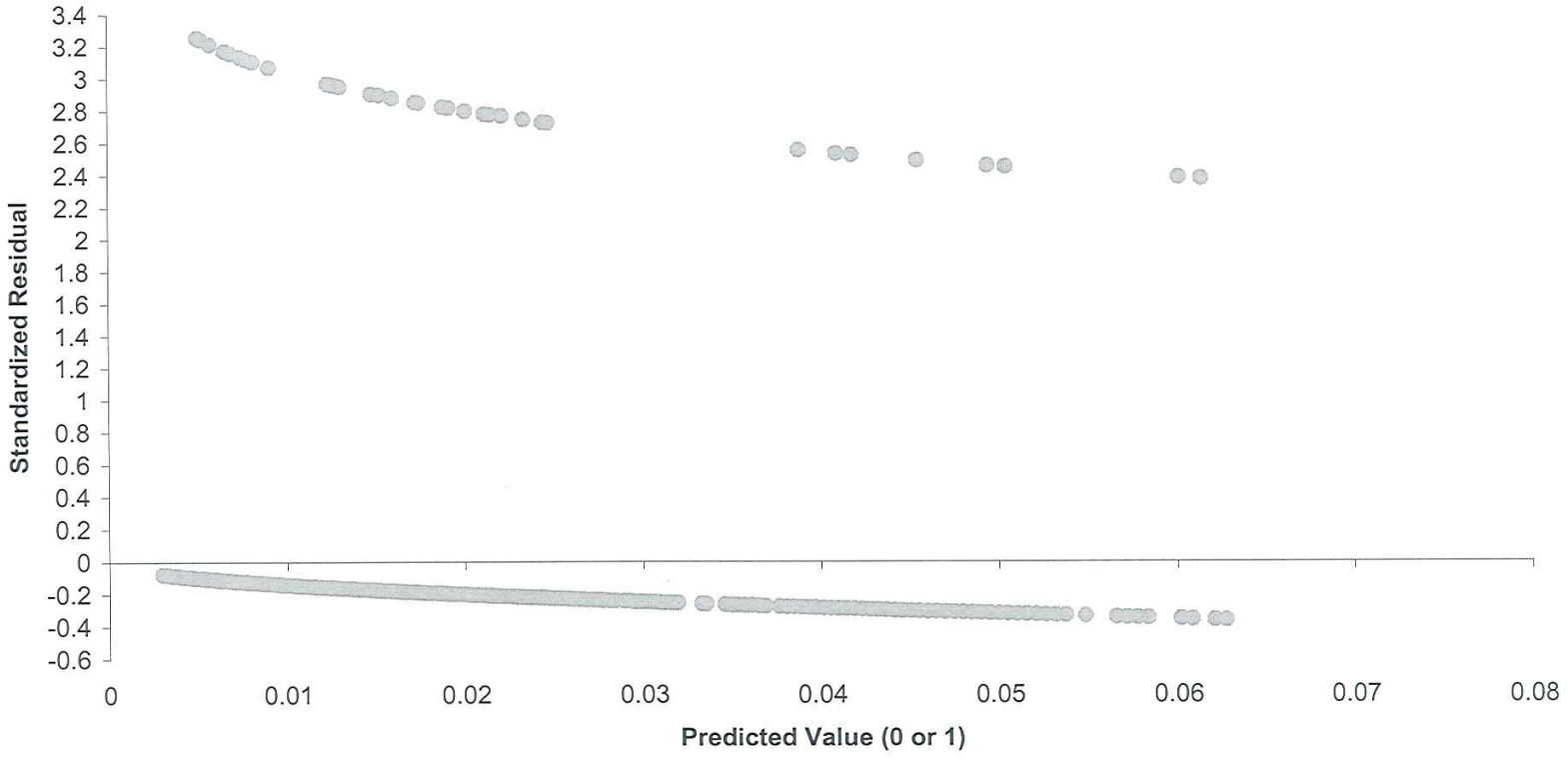
Program Source: forestSVRn.sas

ROC Curves: Logistic Models Predicting Evening Hypoglycemic Events (Severe Events)
Results From 5 Studies (41 events from 2641 subjects)



Program Source: ROCplot.sas

Standardized Deviance Residuals:
Logistic Model Predicting Evening Hypoglycemic Events (Severe Events)
Results From 5 Studies (41 events from 2641 subjects)



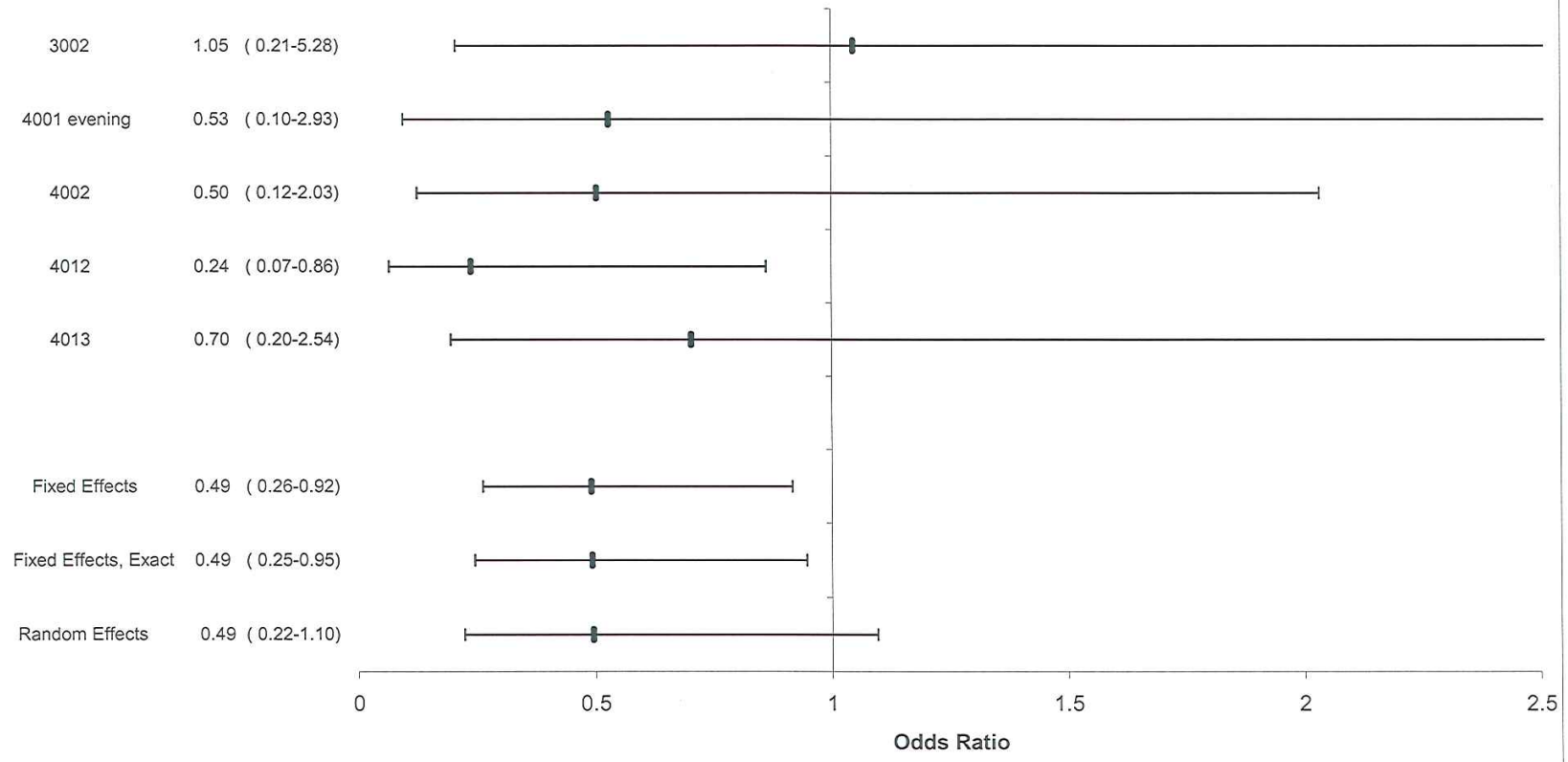
Program Source: RESIDplot.sas

Percentage of subjects having at least one severe hypoglycemic event during the nighttime hours: Severe defined by study
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study

Study or Parameter	Glargine n/N (%)	NPH n/N (%)	Adjusted* Odds Ratio (95% CB)	P Value
3002	3/289 (1.0%)	3/281 (1.1%)	1.05 (0.21-5.28)	0.9552
4001 evening	2/227 (0.9%)	4/233 (1.7%)	0.53 (0.10-2.93)	0.4658
4002	3/367 (0.8%)	6/389 (1.5%)	0.50 (0.12-2.03)	0.3346
4012	3/221 (1.4%)	12/223 (5.4%)	0.24 (0.07-0.86)	0.0288
4013	4/231 (1.7%)	6/250 (2.4%)	0.70 (0.20-2.54)	0.5911
Total	15/1335 (1.1%)	31/1376 (2.3%)		
Results from subject level (fixed effects) model				
Slope: HbA1c Change			Non-significant change in hypo odds	0.3868
HbA1c Change missing			485.5% increase in hypo odds if A1c change missing	0.0005
Odds Ratio Homogeneity				0.6446
Odds Ratio Summaries				
Fixed Effects*			0.49 (0.26-0.92)	0.0257
Fixed Effects, Exact Inference**			0.49 (0.25-0.95)	0.0321
Random Effects***			0.49 (0.22-1.10)	0.0698

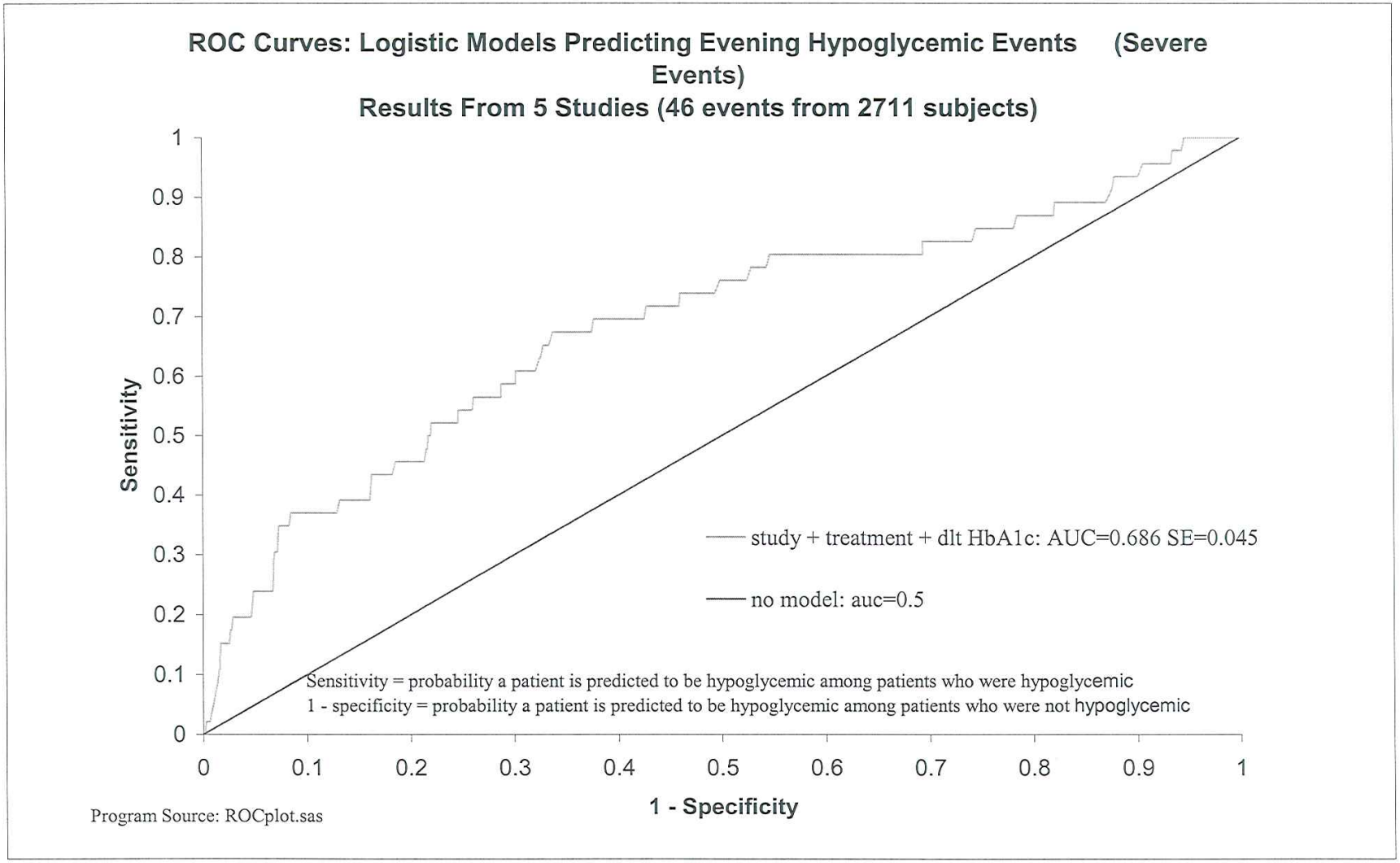
70 subjects missing their HbA1c change from baseline were included in this analysis , including 5 subjects with a severe nocturnal hypo. event: 1 glargine and 4 NPH subjects. Excluding these subjects would have removed just under 11% of the total severe nocturnal hypoglycemic events.
*** The log odds and log odds ratios were estimated at the same (mean) HbA1c change from baseline. The model also included study, treatment, study x treatment, indicator (HbA1c change missing).**
**** The exact confidence bounds and p values were estimated at the study level without an HbA1c adjustment.**
***** Log odds by study and treatment, estimated at the subject level by adjusting for HbA1c change were entered into a study by treatment mixed model (study as the random effect).**
. Log odds ratio standard errors took into account both the within and among study variances

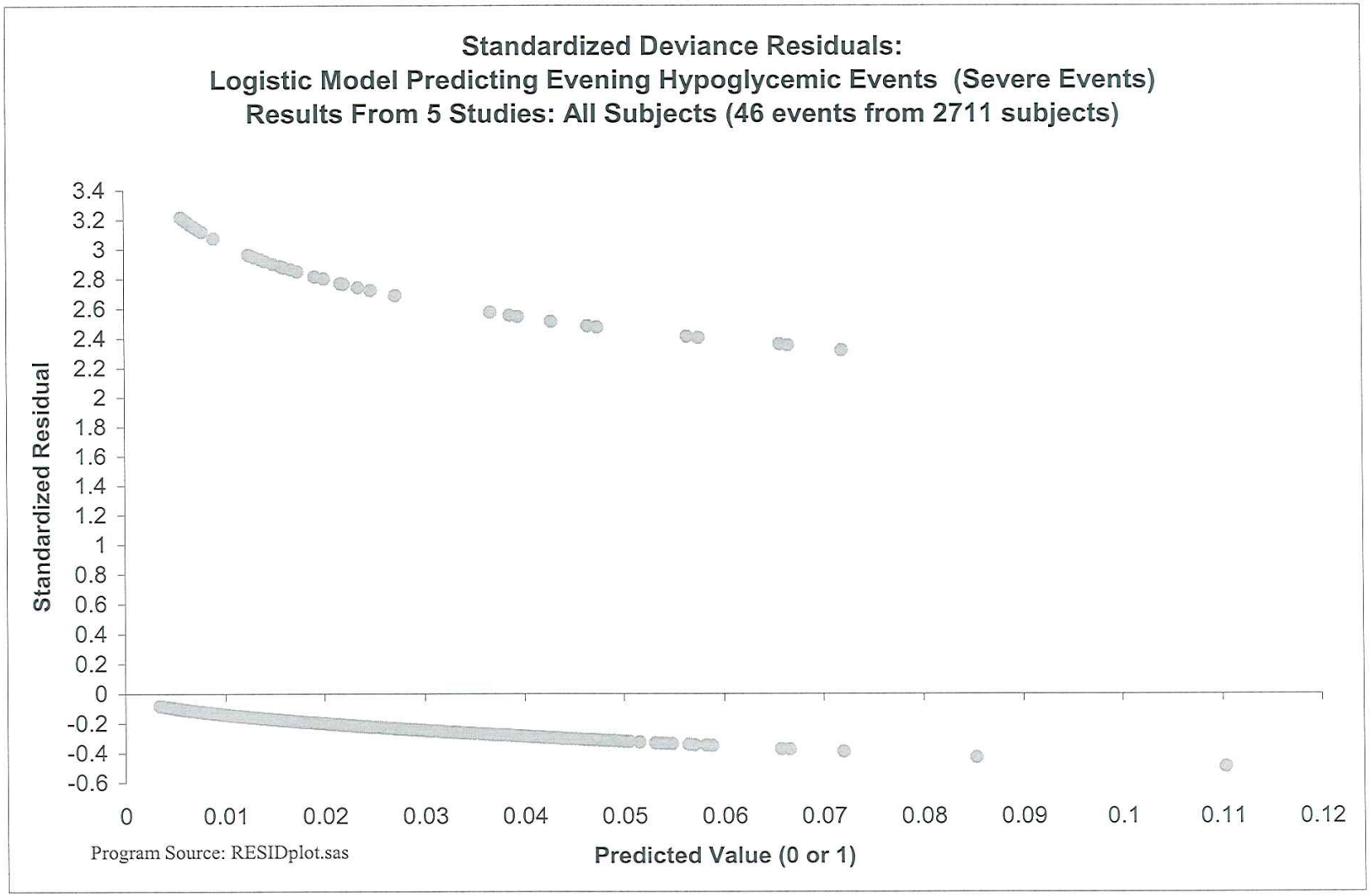
Percentage of subjects having at least one hypoglycemic event:
Nighttime Severe (Using All Subjects)
Odds Ratios: Glargine/NPH (95% Confidence Bounds) by Study



All subjects were included by using a missing HbA1c indicator in the subject level models

Program Source: forestSVRn-all.sas


ROC Curves: Logistic Models Predicting Evening Hypoglycemic Events (Severe Events)
Results From 5 Studies (46 events from 2711 subjects)
Sensitivity
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1 - Specificity
study + treatment + dlt HbA1c: AUC=0.686 SE=0.045
no model: auc=0.5
Sensitivity = probability a patient is predicted to be hypoglycemic among patients who were hypoglycemic
1 - specificity = probability a patient is predicted to be hypoglycemic among patients who were not hypoglycemic
Program Source: ROCplot.sas


Standardized Deviance Residuals:
Logistic Model Predicting Evening Hypoglycemic Events (Severe Events)
Results From 5 Studies: All Subjects (46 events from 2711 subjects)
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07
0.08
0.09
0.1
0.11
0.12
Program Source: RESIDplot.sas
Predicted Value (0 or 1)

Logistic Regression Analyses Studypools II - V

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (235)		NPH (227)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	5	2.1	3	1.3	1.40	0.32	6.05	0.653	0.225	0.653
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.3	4	1.8	0.75	0.16	3.48	0.715	0.752	0.546
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	95	40.4	72	31.7	1.28	0.86	1.89	0.224	<0.001	0.622
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	4	1.8	0.16	0.00	1.30	0.045	0.007	0.863
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	2	0.9	0.64	0.06	7.26	0.715	0.154	0.706
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	19	8.1	65	28.6	0.20	0.11	0.34	<0.001	0.027	0.711

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 237, NPH: 233. Patients who have values in change in HbA1C: Glargine: 235, NPH: 227
(4). 8 patients who have missing values in change in HbA1C were excluded, among them no one has at least one hypoglycemia
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C
(6). For "Nocturnal Severe Hypoglycemia" item, Adjusted odds ratio, 95% CI and P-values come from "Conditional Exact Tests" of Logistic regression

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (157)		NPH (157)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.6	0	0.0	0.33	0.01	Infinity	1.000	0.930	0.855
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.9	2	1.3	1.50	0.25	9.12	0.659	0.820	0.540
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	30	19.1	31	19.7	0.96	0.55	1.67	0.876	0.626	0.524
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	0	0.0	-	-	-	-	-	-
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	0	0.0	0	0.0	-	-	-	-	-	-
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	6	3.8	2	1.3	2.99	0.59	15.12	0.185	0.202	0.709

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 158, NPH: 159. Patients who have values in change in HbA1C: Glargine: 157, NPH: 157
(4). 3 patients who have missing values in change in HbA1C were excluded, among them no one has at least one hypoglycemia
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C
(6). For "Day Time Severe Hypoglycemia" item, Adjusted odds ratio, 95% CI and P-values come from "Conditional Exact Tests" of Logistic regression

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (498)		NPH (486)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	20	4.0	36	7.4	0.52	0.30	0.91	0.023	0.575	0.585
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	103	20.7	139	28.6	0.65	0.49	0.88	0.004	0.421	0.563
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	325	65.3	352	72.4	0.72	0.55	0.94	0.017	0.321	0.548
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	27	5.4	32	6.6	0.81	0.48	1.37	0.430	0.546	0.536
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	105	21.1	116	23.9	0.85	0.63	1.15	0.292	0.836	0.518
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	269	54.0	282	58.0	0.85	0.66	1.09	0.204	0.905	0.527

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 514, NPH: 503. Patients who have values in change in HbA1C: Glargine: 498, NPH: 486
(4). 33 patients who have missing values in change in HbA1C were excluded, including 3 patients who have at least one hypoglycemia: 2 Glargine and 1 NPH patients
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (246)		NPH (255)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	4	1.6	0.26	0.03	2.37	0.234	0.170	0.785
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	10	4.1	21	8.2	0.47	0.22	1.02	0.055	0.716	0.606
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	120	48.8	136	53.3	0.85	0.60	1.21	0.369	0.036	0.560
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	2	0.8	0.48	0.04	5.43	0.556	0.475	0.650
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	11	4.5	12	4.7	0.96	0.41	2.22	0.923	0.592	0.500
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	71	28.9	91	35.7	0.75	0.51	1.09	0.130	0.027	0.574

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 259, NPH: 259. Patients who have values in change in HbA1C: Glargine: 246, NPH: 255
(4). 17 patients who have missing values in change in HbA1C were excluded, including 5 patients who have at least one hypoglycemia: 3 Glargine and 2 NPH patients
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C

Appendices Studypool I

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=289)	NPH (N=281)	Total (N=570)
Age (yrs)	N	289	281	570
	Mean (SD)	59.6 (9.3)	59.4 (9.1)	59.5 (9.2)
	Median	60.0	59.0	59.5
	Min, Max	34.0, 79.0	39.0, 80.0	34.0, 80.0
Age Group, n (%)	Age <65 (yrs)	192 (66.4%)	190 (67.6%)	382 (67.0%)
	65 <= Age <=75 (yrs)	89 (30.8%)	85 (30.2%)	174 (30.5%)
	Age > 75 (yrs)	8 (2.8%)	6 (2.1%)	14 (2.5%)
	Age >= 65 (yrs)	97 (33.6%)	91 (32.4%)	188 (33.0%)
Sex, n (%)	Male	154 (53.3%)	152 (54.1%)	306 (53.7%)
	Female	135 (46.7%)	129 (45.9%)	264 (46.3%)
Race, n (%)	White	267 (92.4%)	262 (93.2%)	529 (92.8%)
	Black	13 (4.5%)	10 (3.6%)	23 (4.0%)
	Asian/Oriental	7 (2.4%)	5 (1.8%)	12 (2.1%)
	Multiracial	2 (0.7%)	4 (1.4%)	6 (1.1%)
Weight (kg) [b]	N	289	281	570
	Mean (SD)	83.5 (15.6)	82.1 (14.7)	82.8 (15.2)
	Median	83.0	80.8	82.0
	Min, Max	49.3, 134.1	48.3, 120.3	48.3, 134.1

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=289)	NPH (N=281)	Total (N=570)
Height (cm) [b]	N	289	281	570
	Mean (SD)	168.6 (9.8)	168.6 (9.4)	168.6 (9.6)
	Median	169.0	168.0	168.0
	Min, Max	145.0, 191.0	148.0, 191.0	145.0, 191.0
BMI (kg/m2)	N	289	281	570
	Mean (SD)	29.3 (4.3)	28.8 (4.3)	29.1 (4.3)
	Median	29.1	28.4	28.7
	Min, Max	19.0, 40.0	18.1, 44.2	18.1, 44.2
BMI Group, n (%)	BMI >=30 (kg/m2)	122 (42.2%)	102 (36.3%)	224 (39.3%)
Duration of Diabetes (yrs)	N	289	281	570
	Mean (SD)	10.2 (6.2)	10.5 (6.0)	10.3 (6.1)
	Median	9.0	9.0	9.0
	Min, Max	2.0, 41.0	2.0, 37.0	2.0, 41.0
Prior use of OAD, n (%)	Yes	288 (99.7%)	281 (100%)	569 (99.8%)
	No	1 (0.3%)	0 (0.0%)	1 (0.2%)
Prior Use of Insulin, n (%)	Yes	67 (23.2%)	77 (27.4%)	144 (25.3%)
	No	222 (76.8%)	204 (72.6%)	426 (74.7%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=289)	NPH (N=281)	Total (N=570)
HBA1C at Baseline (%)	N	281	271	552
	Mean (SD)	9.0 (1.2)	8.9 (1.1)	9.0 (1.2)
	Median	9.0	8.8	8.9
	Min, Max	6.5, 12.3	6.3, 11.8	6.3, 12.3
Country, n (%)	AUSTRIA	34 (11.8%)	34 (12.1%)	68 (11.9%)
	CROATIA	4 (1.4%)	3 (1.1%)	7 (1.2%)
	CZECH REPUBLIC	11 (3.8%)	10 (3.6%)	21 (3.7%)
	DENMARK	31 (10.7%)	31 (11.0%)	62 (10.9%)
	FINLAND	37 (12.8%)	35 (12.5%)	72 (12.6%)
	GERMANY	30 (10.4%)	28 (10.0%)	58 (10.2%)
	ITALY	12 (4.2%)	9 (3.2%)	21 (3.7%)
	NETHERLANDS	23 (8.0%)	24 (8.5%)	47 (8.2%)
	NORWAY	8 (2.8%)	8 (2.8%)	16 (2.8%)
	SLOVENIA	16 (5.5%)	18 (6.4%)	34 (6.0%)
	SOUTH AFRICA	47 (16.3%)	44 (15.7%)	91 (16.0%)
	SPAIN	4 (1.4%)	6 (2.1%)	10 (1.8%)
	SWEDEN	26 (9.0%)	23 (8.2%)	49 (8.6%)
	SWITZERLAND	2 (0.7%)	2 (0.7%)	4 (0.7%)
	UK	4 (1.4%)	6 (2.1%)	10 (1.8%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=289)	NPH (N=281)	Total (N=570)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	280	266	546		
	Mean (SD)	9.04 (1.2)	8.85 (1.1)	8.95 (1.2)		
	Median	9.00	8.80	8.90		
	Min, Max	6.5, 12.3	6.3, 11.8	6.3, 12.3		
HBA1C at Endpoint (%)	N	280	266	546		
	Mean (SD)	8.51 (1.4)	8.48 (1.3)	8.49 (1.4)		
	Median	8.30	8.30	8.30		
	Min, Max	5.8, 14.3	5.3, 13.0	5.3, 14.3		
Change in HBA1C (%)	N	280	266	546		
	Mean (SD)	-0.53 (1.4)	-0.38 (1.2)	-0.45 (1.3)		
	Median	-0.50	-0.30	-0.40		
	Min, Max	-4.4, 4.2	-3.6, 2.9	-4.4, 4.2		
HBA1C Change(ANCOVA) (%) [b]	N	280	266			
	LS Means (se)	-0.45 (0.08)	-0.37 (0.08)		-0.09 (-0.29, 0.11)	0.3953

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=289)		NPH (N=281)		Total (N=570)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	3	1.0	0	0.0	3	0.5	-	0.56	Infty	0.267
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	4	1.4	3	1.1	7	1.2	1.27	0.21	8.77	1.000
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	79	27.3	77	27.4	156	27.4	1.00	0.67	1.48	1.000
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	3	1.0	3	1.1	6	1.1	0.97	0.13	7.34	1.000
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	2	0.7	3	1.1	5	0.9	0.64	0.05	5.69	0.972
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	29	10.0	59	21.0	88	15.4	0.41	0.24	0.68	<0.001

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=227)	NPH (N=233)	Total (N=460)
Age (yrs)	N	227	233	460
	Mean (SD)	60.2 (9.5)	61.8 (8.5)	61.0 (9.0)
	Median	60.6	62.7	61.8
	Min, Max	19.6, 85.5	38.4, 85.6	19.6, 85.6
Age Group, n (%)	Age <65 (yrs)	150 (66.1%)	147 (63.1%)	297 (64.6%)
	65 <= Age <=75 (yrs)	69 (30.4%)	75 (32.2%)	144 (31.3%)
	Age > 75 (yrs)	8 (3.5%)	11 (4.7%)	19 (4.1%)
	Age >= 65 (yrs)	77 (33.9%)	86 (36.9%)	163 (35.4%)
Sex, n (%)	Male	130 (57.3%)	120 (51.5%)	250 (54.3%)
	Female	97 (42.7%)	113 (48.5%)	210 (45.7%)
Race, n (%)	White	219 (96.5%)	227 (97.4%)	446 (97.0%)
	Black	1 (0.4%)	1 (0.4%)	2 (0.4%)
	Asian/Oriental	6 (2.6%)	1 (0.4%)	7 (1.5%)
	Multiracial	0 (0.0%)	1 (0.4%)	1 (0.2%)
	Other	1 (0.4%)	3 (1.3%)	4 (0.9%)
Weight (kg) [b]	N	227	233	460
	Mean (SD)	81.4 (13.4)	80.4 (14.8)	80.9 (14.1)
	Median	80.0	80.0	80.0
	Min, Max	48.0, 120.8	46.0, 139.4	46.0, 139.4

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=227)	NPH (N=233)	Total (N=460)
Height (cm) [b]	N	226	233	459
	Mean (SD)	169.0 (9.0)	167.0 (9.8)	168.0 (9.5)
	Median	169.0	167.5	168.0
	Min, Max	145.0, 191.0	122.0, 190.0	122.0, 191.0
BMI (kg/m2)	N	226	233	459
	Mean (SD)	28.5 (3.9)	28.7 (3.9)	28.6 (3.9)
	Median	28.2	28.6	28.5
	Min, Max	16.8, 37.6	20.2, 42.5	16.8, 42.5
BMI Group, n (%)	BMI >=30 (kg/m2)	77 (34.1%)	86 (36.9%)	163 (35.5%)
Duration of Diabetes (yrs)	N	227	233	460
	Mean (SD)	9.4 (6.9)	9.8 (6.1)	9.6 (6.5)
	Median	8.0	9.0	8.2
	Min, Max	0.8, 50.5	0.7, 38.3	0.7, 50.5
Prior use of OAD, n (%)	Yes	225 (99.1%)	232 (99.6%)	457 (99.3%)
	No	2 (0.9%)	1 (0.4%)	3 (0.7%)
Prior Use of Insulin, n (%)	No	227 (100%)	233 (100%)	460 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=227)	NPH (N=233)	Total (N=460)
HBA1C at Baseline (%)	N	227	233	460
	Mean (SD)	9.1 (1.0)	9.1 (1.1)	9.1 (1.0)
	Median	9.0	9.0	9.0
	Min, Max	5.7, 11.9	6.9, 12.3	5.7, 12.3
Country, n (%)	AUSTRIA	16 (7.0%)	16 (6.9%)	32 (7.0%)
	BELGIUM	5 (2.2%)	7 (3.0%)	12 (2.6%)
	FINLAND	10 (4.4%)	9 (3.9%)	19 (4.1%)
	GERMANY	82 (36.1%)	81 (34.8%)	163 (35.4%)
	IRELAND	9 (4.0%)	11 (4.7%)	20 (4.3%)
	ITALY	8 (3.5%)	8 (3.4%)	16 (3.5%)
	NETHERLANDS	8 (3.5%)	6 (2.6%)	14 (3.0%)
	PORTUGAL	7 (3.1%)	7 (3.0%)	14 (3.0%)
	SPAIN	13 (5.7%)	13 (5.6%)	26 (5.7%)
	SWEDEN	3 (1.3%)	4 (1.7%)	7 (1.5%)
	SWITZERLAND	7 (3.1%)	7 (3.0%)	14 (3.0%)
	TURKEY	7 (3.1%)	9 (3.9%)	16 (3.5%)
	UK	52 (22.9%)	55 (23.6%)	107 (23.3%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=227)	NPH (N=233)	Total (N=460)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	224	227	451		
	Mean (SD)	9.07 (1.0)	9.09 (1.1)	9.08 (1.0)		
	Median	9.00	9.00	9.00		
	Min, Max	5.7, 11.9	6.9, 12.3	5.7, 12.3		
HBA1C at Endpoint (%)	N	224	227	451		
	Mean (SD)	8.12 (1.3)	8.26 (1.3)	8.19 (1.3)		
	Median	7.95	8.30	8.10		
	Min, Max	4.7, 13.5	4.9, 12.8	4.7, 13.5		
Change in HBA1C (%)	N	224	227	451		
	Mean (SD)	-0.95 (1.2)	-0.83 (1.3)	-0.89 (1.3)		
	Median	-0.95	-0.80	-0.90		
	Min, Max	-5.1, 2.7	-5.1, 2.2	-5.1, 2.7		
HBA1C Change(ANCOVA) (%) [b]	N	224	227			
	LS Means (se)	-0.90 (0.07)	-0.80 (0.07)		-0.10 (-0.30, 0.09)	0.3094

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=227)		NPH (N=233)		Total (N=460)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Porb
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	2	0.9	3	1.3	5	1.1	0.87	0.07	8.38	1.000
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	4	1.7	5	1.1	0.25	0.01	2.53	0.375
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	61	26.9	72	30.9	133	28.9	0.80	0.51	1.26	0.373
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	2	0.9	4	1.7	6	1.3	0.50	0.04	3.51	0.684
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	2	0.9	3	0.7	0.50	0.01	9.60	1.000
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	31	13.7	65	27.9	96	20.9	0.36	0.21	0.63	<0.001

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=367)	NPH (N=389)	Total (N=756)
Age (yrs)	N	367	389	756
	Mean (SD)	54.7 (9.5)	55.6 (8.9)	55.1 (9.2)
	Median	54.0	55.0	55.0
	Min, Max	29.0, 74.0	31.0, 75.0	29.0, 75.0
Age Group, n (%)	Age <65 (yrs)	308 (83.9%)	316 (81.2%)	624 (82.5%)
	65 <= Age <=75 (yrs)	59 (16.1%)	73 (18.8%)	132 (17.5%)
	Age >= 65 (yrs)	59 (16.1%)	73 (18.8%)	132 (17.5%)
Sex, n (%)	Male	202 (55.0%)	219 (56.3%)	421 (55.7%)
	Female	165 (45.0%)	170 (43.7%)	335 (44.3%)
Race, n (%)	White	308 (83.9%)	324 (83.3%)	632 (83.6%)
	Black	42 (11.4%)	51 (13.1%)	93 (12.3%)
	Asian/Oriental	12 (3.3%)	10 (2.6%)	22 (2.9%)
	Multiracial	5 (1.4%)	4 (1.0%)	9 (1.2%)
Weight (kg) [b]	N	367	389	756
	Mean (SD)	95.2 (18.2)	94.4 (16.8)	94.8 (17.5)
	Median	93.9	93.1	93.4
	Min, Max	52.5, 154.4	57.6, 151.5	52.5, 154.4

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=367)	NPH (N=389)	Total (N=756)
Height (cm) [b]	N	367	389	756
	Mean (SD)	170.8 (10.5)	171.1 (9.7)	171.0 (10.1)
	Median	170.2	171.5	171.0
	Min, Max	130.8, 200.7	147.3, 207.0	130.8, 207.0
BMI (kg/m2)	N	367	389	756
	Mean (SD)	32.5 (4.7)	32.2 (4.8)	32.3 (4.7)
	Median	32.2	31.5	31.8
	Min, Max	23.0, 45.1	19.3, 43.9	19.3, 45.1
BMI Group, n (%)	BMI >=30 (kg/m2)	251 (68.4%)	244 (62.7%)	495 (65.5%)
Duration of Diabetes (yrs)	N	367	389	756
	Mean (SD)	8.4 (5.5)	9.0 (5.6)	8.7 (5.6)
	Median	7.0	8.0	7.5
	Min, Max	1.0, 34.0	1.0, 37.0	1.0, 37.0
Prior use of OAD, n (%)	Yes	367 (100%)	389 (100%)	756 (100%)
Prior Use of Insulin, n (%)	No	367 (100%)	389 (100%)	756 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=367)	NPH (N=389)	Total (N=756)
HBA1C at Baseline (%)	N	367	389	756
	Mean (SD)	8.6 (0.9)	8.6 (0.9)	8.6 (0.9)
	Median	8.5	8.5	8.5
	Min, Max	6.7, 10.9	6.8, 12.2	6.7, 12.2
Country, n (%)	CANADA	57 (15.5%)	61 (15.7%)	118 (15.6%)
	USA	310 (84.5%)	328 (84.3%)	638 (84.4%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=367)	NPH (N=389)	Total (N=756)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	355	380	735		
	Mean (SD)	8.60 (0.9)	8.56 (0.9)	8.58 (0.9)		
	Median	8.50	8.50	8.50		
	Min, Max	6.7, 10.9	6.8, 12.2	6.7, 12.2		
HBA1C at Endpoint (%)	N	355	380	735		
	Mean (SD)	6.96 (0.8)	6.97 (0.8)	6.97 (0.8)		
	Median	6.90	6.90	6.90		
	Min, Max	5.2, 11.2	5.3, 10.5	5.2, 11.2		
Change in HBA1C (%)	N	355	380	735		
	Mean (SD)	-1.64 (0.9)	-1.58 (0.9)	-1.61 (0.9)		
	Median	-1.60	-1.50	-1.60		
	Min, Max	-4.4, 1.5	-5.6, 1.2	-5.6, 1.5		
HBA1C Change(ANCOVA) (%) [b]	N	355	380			
	LS Means (se)	-1.65 (0.04)	-1.62 (0.04)		-0.03 (-0.13, 0.08)	0.6118

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=367)		NPH (N=389)		Total (N=756)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	6	1.6	3	0.8	9	1.2	2.11	0.45	13.56	0.457
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	21	5.7	35	9.0	56	7.4	0.61	0.33	1.11	0.114
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	229	62.4	252	64.8	481	63.6	0.90	0.66	1.23	0.543
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	3	0.8	6	1.5	9	1.2	0.51	0.08	2.43	0.541
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	14	3.8	32	8.2	46	6.1	0.44	0.21	0.86	0.014
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	153	41.7	199	51.2	352	46.6	0.67	0.49	0.91	0.009

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=221)	NPH (N=223)	Total (N=444)
Age (yrs)	N	221	223	444
	Mean (SD)	56.1 (8.4)	56.2 (8.7)	56.1 (8.5)
	Median	56.0	55.6	55.8
	Min, Max	38.2, 77.0	31.5, 78.8	31.5, 78.8
Age Group, n (%)	Age <65 (yrs)	184 (83.3%)	187 (83.9%)	371 (83.6%)
	65 <= Age <=75 (yrs)	36 (16.3%)	32 (14.3%)	68 (15.3%)
	Age > 75 (yrs)	1 (0.5%)	4 (1.8%)	5 (1.1%)
	Age >= 65 (yrs)	37 (16.7%)	36 (16.1%)	73 (16.4%)
Sex, n (%)	Male	89 (40.3%)	99 (44.4%)	188 (42.3%)
	Female	132 (59.7%)	124 (55.6%)	256 (57.7%)
Race, n (%)	Asian/Oriental	221 (100%)	223 (100%)	444 (100%)
Weight (kg) [b]	N	221	223	444
	Mean (SD)	63.9 (10.2)	64.3 (10.8)	64.1 (10.5)
	Median	63.0	64.0	63.0
	Min, Max	44.0, 102.0	41.0, 99.0	41.0, 102.0

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=221)	NPH (N=223)	Total (N=444)
Height (cm) [b]	N	221	223	444
	Mean (SD)	160.6 (8.8)	159.9 (8.7)	160.3 (8.8)
	Median	160.0	160.0	160.0
	Min, Max	144.0, 185.0	142.0, 183.0	142.0, 185.0
BMI (kg/m2)	N	221	223	444
	Mean (SD)	24.7 (3.0)	25.1 (3.2)	24.9 (3.1)
	Median	24.6	24.8	24.7
	Min, Max	18.6, 35.3	19.5, 35.0	18.6, 35.3
BMI Group, n (%)	BMI >=30 (kg/m2)	14 (6.3%)	20 (9.0%)	34 (7.7%)
Duration of Diabetes (yrs)	N	221	223	444
	Mean (SD)	10.5 (6.2)	10.2 (5.2)	10.4 (5.7)
	Median	9.7	9.8	9.8
	Min, Max	0.4, 31.1	0.4, 24.8	0.4, 31.1
Prior use of OAD, n (%)	Yes	221 (100%)	223 (100%)	444 (100%)
Prior Use of Insulin, n (%)	No	221 (100%)	223 (100%)	444 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=221)	NPH (N=223)	Total (N=444)
HBA1C at Baseline (%)	N	199	204	403
	Mean (SD)	9.0 (0.9)	9.0 (0.8)	9.0 (0.9)
	Median	9.0	9.1	9.0
	Min, Max	7.5, 11.2	7.5, 10.8	7.5, 11.2
Country, n (%)	CHINA	63 (28.5%)	63 (28.3%)	126 (28.4%)
	HONG KONG	12 (5.4%)	14 (6.3%)	26 (5.9%)
	INDONESIA	8 (3.6%)	11 (4.9%)	19 (4.3%)
	MALAYSIA	8 (3.6%)	8 (3.6%)	16 (3.6%)
	PAKISTAN	18 (8.1%)	18 (8.1%)	36 (8.1%)
	PHILIPPINES	12 (5.4%)	12 (5.4%)	24 (5.4%)
	SINGAPORE	2 (0.9%)	2 (0.9%)	4 (0.9%)
	SOUTH KOREA	59 (26.7%)	53 (23.8%)	112 (25.2%)
	TAIWAN	15 (6.8%)	17 (7.6%)	32 (7.2%)
	THAILAND	24 (10.9%)	25 (11.2%)	49 (11.0%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=221)	NPH (N=223)	Total (N=444)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	218	219	437		
	Mean (SD)	9.01 (0.9)	9.06 (0.8)	9.04 (0.9)		
	Median	9.00	9.10	9.00		
	Min, Max	7.5, 11.2	7.5, 10.8	7.5, 11.2		
HBA1C at Endpoint (%)	N	218	219	437		
	Mean (SD)	7.91 (1.2)	8.15 (1.2)	8.03 (1.2)		
	Median	7.75	8.10	8.00		
	Min, Max	4.5, 12.2	5.0, 12.3	4.5, 12.3		
Change in HBA1C (%)	N	218	219	437		
	Mean (SD)	-1.11 (1.2)	-0.91 (1.3)	-1.01 (1.2)		
	Median	-1.20	-0.90	-1.00		
	Min, Max	-4.2, 2.9	-4.3, 3.5	-4.3, 3.5		
HBA1C Change(ANCOVA) (%) [b]	N	218	219			
	LS Means (se)	-1.03 (0.08)	-0.83 (0.08)		-0.21 (-0.41, -0.01)	0.0427

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=221)		NPH (N=223)		Total (N=444)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	2	0.9	4	1.8	6	1.4	0.53	0.05	3.76	0.736
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.5	1	0.4	2	0.5	0.90	0.01	71.92	1.000
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	51	23.1	67	30.0	118	26.6	0.68	0.42	1.08	0.103
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	3	1.4	12	5.4	15	3.4	0.25	0.04	0.96	0.042
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	0	0.0	2	0.9	2	0.5	0.00	0.00	3.89	0.548
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	45	20.4	84	37.7	129	29.1	0.40	0.25	0.64	<0.001

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=231)	NPH (N=250)	Total (N=481)
Age (yrs)	N	231	250	481
	Mean (SD)	56.3 (9.8)	57.1 (9.6)	56.7 (9.7)
	Median	56.2	57.4	56.8
	Min, Max	24.4, 75.7	23.2, 75.9	23.2, 75.9
Age Group, n (%)	Age <65 (yrs)	182 (78.8%)	198 (79.2%)	380 (79.0%)
	65 <= Age <=75 (yrs)	47 (20.3%)	48 (19.2%)	95 (19.8%)
	Age > 75 (yrs)	2 (0.9%)	4 (1.6%)	6 (1.2%)
	Age >= 65 (yrs)	49 (21.2%)	52 (20.8%)	101 (21.0%)
Sex, n (%)	Male	100 (43.3%)	95 (38.0%)	195 (40.5%)
	Female	131 (56.7%)	155 (62.0%)	286 (59.5%)
Race, n (%)	White	101 (43.7%)	121 (48.4%)	222 (46.2%)
	Black	12 (5.2%)	7 (2.8%)	19 (4.0%)
	Asian/Oriental	3 (1.3%)	0 (0.0%)	3 (0.6%)
	Multiracial	100 (43.3%)	108 (43.2%)	208 (43.2%)
	Hispanic	15 (6.5%)	13 (5.2%)	28 (5.8%)
	Other	0 (0.0%)	1 (0.4%)	1 (0.2%)
Weight (kg) [b]	N	231	250	481
	Mean (SD)	71.1 (12.7)	70.8 (13.7)	70.9 (13.2)
	Median	71.0	70.0	70.0
	Min, Max	33.0, 111.0	40.4, 121.0	33.0, 121.0

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=231)	NPH (N=250)	Total (N=481)
Height (cm) [b]	N	231	250	481
	Mean (SD)	161.0 (10.1)	160.9 (10.1)	161.0 (10.1)
	Median	161.0	160.0	160.0
	Min, Max	140.0, 184.0	128.0, 191.0	128.0, 191.0
BMI (kg/m2)	N	231	250	481
	Mean (SD)	27.3 (3.7)	27.2 (4.0)	27.3 (3.9)
	Median	27.1	26.7	26.9
	Min, Max	16.4, 35.1	14.3, 35.8	14.3, 35.8
BMI Group, n (%)	BMI >=30 (kg/m2)	62 (26.8%)	64 (25.6%)	126 (26.2%)
Duration of Diabetes (yrs)	N	231	250	481
	Mean (SD)	10.3 (6.4)	10.8 (6.4)	10.6 (6.4)
	Median	9.0	10.1	9.6
	Min, Max	0.7, 31.4	0.6, 38.1	0.6, 38.1
Prior use of OAD, n (%)	Yes	231 (100%)	250 (100%)	481 (100%)
Prior Use of Insulin, n (%)	No	231 (100%)	250 (100%)	481 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=231)	NPH (N=250)	Total (N=481)
HBA1C at Baseline (%)	N	230	250	480
	Mean (SD)	9.1 (1.1)	9.2 (1.0)	9.1 (1.1)
	Median	9.0	9.2	9.1
	Min, Max	6.0, 12.2	6.4, 12.1	6.0, 12.2
Country, n (%)	ARGENTINA	24 (10.4%)	30 (12.0%)	54 (11.2%)
	BRAZIL	76 (32.9%)	85 (34.0%)	161 (33.5%)
	CHILE	2 (0.9%)	3 (1.2%)	5 (1.0%)
	COLOMBIA	8 (3.5%)	10 (4.0%)	18 (3.7%)
	GUATEMALA	8 (3.5%)	8 (3.2%)	16 (3.3%)
	MEXICO	81 (35.1%)	85 (34.0%)	166 (34.5%)
	PARAGUAY	5 (2.2%)	3 (1.2%)	8 (1.7%)
	PERU	5 (2.2%)	7 (2.8%)	12 (2.5%)
	URUGUAY	3 (1.3%)	1 (0.4%)	4 (0.8%)
	VENEZUELA	19 (8.2%)	18 (7.2%)	37 (7.7%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=231)	NPH (N=250)	Total (N=481)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	226	246	472		
	Mean (SD)	9.06 (1.1)	9.21 (1.0)	9.14 (1.1)		
	Median	9.00	9.20	9.10		
	Min, Max	6.0, 12.2	7.2, 12.1	6.0, 12.2		
HBA1C at Endpoint (%)	N	226	246	472		
	Mean (SD)	7.68 (1.3)	7.78 (1.3)	7.74 (1.3)		
	Median	7.50	7.60	7.60		
	Min, Max	4.9, 12.3	5.0, 11.9	4.9, 12.3		
Change in HBA1C (%)	N	226	246	472		
	Mean (SD)	-1.37 (1.3)	-1.43 (1.4)	-1.40 (1.4)		
	Median	-1.30	-1.45	-1.40		
	Min, Max	-5.5, 3.4	-4.9, 3.5	-5.5, 3.5		
HBA1C Change(ANCOVA) (%) [b]	N	226	246			
	LS Means (se)	-1.46 (0.08)	-1.40 (0.08)		-0.05 (-0.27, 0.16)	0.6298

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=231)		NPH (N=250)		Total (N=481)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	2	0.9	5	2.0	7	1.5	0.30	0.03	2.18	0.344
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	2	0.8	3	0.6	0.43	0.01	9.07	0.950
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	88	38.1	98	39.2	186	38.7	0.94	0.62	1.43	0.844
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	4	1.7	6	2.4	10	2.1	0.77	0.16	3.37	0.959
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	2	0.9	7	2.8	9	1.9	0.26	0.03	1.43	0.152
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	37	16.0	73	29.2	110	22.9	0.37	0.23	0.64	<0.001

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=1335)	NPH (N=1376)	Total (N=2711)
Age (yrs)	N	1335	1376	2711
	Mean (SD)	57.2 (9.6)	57.8 (9.2)	57.5 (9.4)
	Median	57.0	57.6	57.2
	Min, Max	19.6, 85.5	23.2, 85.6	19.6, 85.6
Age Group, n (%)	Age <65 (yrs)	1016 (76.1%)	1038 (75.4%)	2054 (75.8%)
	65 <= Age <=75 (yrs)	300 (22.5%)	313 (22.7%)	613 (22.6%)
	Age > 75 (yrs)	19 (1.4%)	25 (1.8%)	44 (1.6%)
	Age >= 65 (yrs)	319 (23.9%)	338 (24.6%)	657 (24.2%)
Sex, n (%)	Male	675 (50.6%)	685 (49.8%)	1360 (50.2%)
	Female	660 (49.4%)	691 (50.2%)	1351 (49.8%)
Race, n (%)	White	895 (67.0%)	934 (67.9%)	1829 (67.5%)
	Black	68 (5.1%)	69 (5.0%)	137 (5.1%)
	Asian/Oriental	249 (18.7%)	239 (17.4%)	488 (18.0%)
	Multiracial	107 (8.0%)	117 (8.5%)	224 (8.3%)
	Hispanic	15 (1.1%)	13 (0.9%)	28 (1.0%)
	Other	1 (0.1%)	4 (0.3%)	5 (0.2%)
Weight (kg) [b]	N	1335	1376	2711
	Mean (SD)	81.0 (18.5)	80.3 (18.1)	80.6 (18.3)
	Median	79.0	78.9	79.0
	Min, Max	33.0, 154.4	40.4, 151.5	33.0, 154.4

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=1335)	NPH (N=1376)	Total (N=2711)
Height (cm) [b]	N	1334	1376	2710
	Mean (SD)	166.6 (10.6)	166.2 (10.6)	166.4 (10.6)
	Median	166.3	166.0	166.0
	Min, Max	130.8, 200.7	122.0, 207.0	122.0, 207.0
BMI (kg/m2)	N	1334	1376	2710
	Mean (SD)	28.9 (4.8)	28.9 (4.8)	28.9 (4.8)
	Median	28.4	28.3	28.3
	Min, Max	16.4, 45.1	14.3, 44.2	14.3, 45.1
BMI Group, n (%)	BMI >=30 (kg/m2)	526 (39.4%)	516 (37.5%)	1042 (38.5%)
Duration of Diabetes (yrs)	N	1335	1376	2711
	Mean (SD)	9.6 (6.3)	10.0 (5.9)	9.8 (6.1)
	Median	8.2	9.0	8.8
	Min, Max	0.4, 50.5	0.4, 38.3	0.4, 50.5
Prior use of OAD, n (%)	Yes	1332 (99.8%)	1375 (99.9%)	2707 (99.9%)
	No	3 (0.2%)	1 (0.1%)	4 (0.1%)
Prior Use of Insulin, n (%)	Yes	67 (5.0%)	77 (5.6%)	144 (5.3%)
	No	1268 (95.0%)	1299 (94.4%)	2567 (94.7%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=1335)	NPH (N=1376)	Total (N=2711)
HBA1C at Baseline (%)	N	1304	1347	2651
	Mean (SD)	8.9 (1.0)	8.9 (1.0)	8.9 (1.0)
	Median	8.9	8.9	8.9
	Min, Max	5.7, 12.3	6.3, 12.3	5.7, 12.3
Country, n (%)	ARGENTINA	24 (1.8%)	30 (2.2%)	54 (2.0%)
	AUSTRIA	50 (3.7%)	50 (3.6%)	100 (3.7%)
	BELGIUM	5 (0.4%)	7 (0.5%)	12 (0.4%)
	BRAZIL	76 (5.7%)	85 (6.2%)	161 (5.9%)
	CHILE	2 (0.1%)	3 (0.2%)	5 (0.2%)
	CHINA	63 (4.7%)	63 (4.6%)	126 (4.6%)
	COLOMBIA	8 (0.6%)	10 (0.7%)	18 (0.7%)
	CROATIA	4 (0.3%)	3 (0.2%)	7 (0.3%)
	CZECH REPUBLIC	11 (0.8%)	10 (0.7%)	21 (0.8%)
	DENMARK	31 (2.3%)	31 (2.3%)	62 (2.3%)
	FINLAND	47 (3.5%)	44 (3.2%)	91 (3.4%)
	GERMANY	112 (8.4%)	109 (7.9%)	221 (8.2%)
	GUATEMALA	8 (0.6%)	8 (0.6%)	16 (0.6%)
	HONG KONG	12 (0.9%)	14 (1.0%)	26 (1.0%)
	INDONESIA	8 (0.6%)	11 (0.8%)	19 (0.7%)
	IRELAND	9 (0.7%)	11 (0.8%)	20 (0.7%)
	ITALY	20 (1.5%)	17 (1.2%)	37 (1.4%)
	MALAYSIA	8 (0.6%)	8 (0.6%)	16 (0.6%)
	MEXICO	81 (6.1%)	85 (6.2%)	166 (6.1%)
	NETHERLANDS	31 (2.3%)	30 (2.2%)	61 (2.3%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

Statistics or Category [a]	Glargine (N=1335)	NPH (N=1376)	Total (N=2711)
NORWAY	8 (0.6%)	8 (0.6%)	16 (0.6%)
PAKISTAN	18 (1.3%)	18 (1.3%)	36 (1.3%)
PARAGUAY	5 (0.4%)	3 (0.2%)	8 (0.3%)
PERU	5 (0.4%)	7 (0.5%)	12 (0.4%)
PHILIPPINES	12 (0.9%)	12 (0.9%)	24 (0.9%)
PORTUGAL	7 (0.5%)	7 (0.5%)	14 (0.5%)
SINGAPORE	2 (0.1%)	2 (0.1%)	4 (0.1%)
SLOVENIA	16 (1.2%)	18 (1.3%)	34 (1.3%)
SOUTH AFRICA	47 (3.5%)	44 (3.2%)	91 (3.4%)
SOUTH KOREA	59 (4.4%)	53 (3.9%)	112 (4.1%)
SPAIN	17 (1.3%)	19 (1.4%)	36 (1.3%)
SWEDEN	29 (2.2%)	27 (2.0%)	56 (2.1%)
SWITZERLAND	9 (0.7%)	9 (0.7%)	18 (0.7%)
TAIWAN	15 (1.1%)	17 (1.2%)	32 (1.2%)
THAILAND	24 (1.8%)	25 (1.8%)	49 (1.8%)
TURKEY	7 (0.5%)	9 (0.7%)	16 (0.6%)
UK	56 (4.2%)	61 (4.4%)	117 (4.3%)
URUGUAY	3 (0.2%)	1 (0.1%)	4 (0.1%)
VENEZUELA	19 (1.4%)	18 (1.3%)	37 (1.4%)
CANADA	57 (4.3%)	61 (4.4%)	118 (4.4%)
USA	310 (23.2%)	328 (23.8%)	638 (23.5%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=1335)	NPH (N=1376)	Total (N=2711)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	1303	1338	2641		
	Mean (SD)	8.93 (1.0)	8.91 (1.0)	8.92 (1.0)		
	Median	8.90	8.90	8.90		
	Min, Max	5.7, 12.3	6.3, 12.3	5.7, 12.3		
HBA1C at Endpoint (%)	N	1303	1338	2641		
	Mean (SD)	7.78 (1.3)	7.83 (1.3)	7.81 (1.3)		
	Median	7.60	7.60	7.60		
	Min, Max	4.5, 14.3	4.9, 13.0	4.5, 14.3		
Change in HBA1C (%)	N	1303	1338	2641		
	Mean (SD)	-1.15 (1.3)	-1.08 (1.3)	-1.11 (1.3)		
	Median	-1.20	-1.10	-1.10		
	Min, Max	-5.5, 4.2	-5.6, 3.5	-5.6, 4.2		
HBA1C Change(ANCOVA: Fixed) (%) [b]	N	1303	1338			
	LS Means (se)	-1.10 (0.03)	-1.01 (0.03)		-0.09 (-0.18, -0.01)	0.0309
HBA1C Change(ANCOVA: Random) (%) [c]	N	1303	1338			
	LS Means (se)	-1.09 (0.23)	-1.01 (0.23)		-0.08 (-0.20, 0.03)	0.1181

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + study number + Actual treatment group * study number, with study as fixed effect
[c] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group, with study number + Actual treatment group * study number as random effect

Appendices Studypools II - V

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=237)	NPH (N=233)	Total (N=470)
Age (yrs)	N	237	233	470
	Mean (SD)	60.5 (9.0)	61.8 (8.5)	61.1 (8.8)
	Median	61.1	62.7	61.8
	Min, Max	34.0, 80.0	38.4, 85.6	34.0, 85.6
Age Group, n (%)	Age <65 (yrs)	153 (64.6%)	147 (63.1%)	300 (63.8%)
	65 <= Age <=75 (yrs)	77 (32.5%)	75 (32.2%)	152 (32.3%)
	Age > 75 (yrs)	7 (3.0%)	11 (4.7%)	18 (3.8%)
	Age >= 65 (yrs)	84 (35.4%)	86 (36.9%)	170 (36.2%)
Sex, n (%)	Male	124 (52.3%)	120 (51.5%)	244 (51.9%)
	Female	113 (47.7%)	113 (48.5%)	226 (48.1%)
Race, n (%)	White	227 (95.8%)	227 (97.4%)	454 (96.6%)
	Black	1 (0.4%)	1 (0.4%)	2 (0.4%)
	Asian/Oriental	8 (3.4%)	1 (0.4%)	9 (1.9%)
	Multiracial	0 (0.0%)	1 (0.4%)	1 (0.2%)
	Other	1 (0.4%)	3 (1.3%)	4 (0.9%)
Weight (kg) [b]	N	237	233	470
	Mean (SD)	80.2 (15.8)	80.4 (14.8)	80.3 (15.3)
	Median	79.3	80.0	79.6
	Min, Max	40.4, 150.0	46.0, 139.4	40.4, 150.0

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=237)	NPH (N=233)	Total (N=470)
Height (cm) [b]	N	237	233	470
	Mean (SD)	167.8 (9.6)	167.0 (9.8)	167.4 (9.7)
	Median	168.0	167.5	168.0
	Min, Max	143.0, 191.0	122.0, 190.0	122.0, 191.0
BMI (kg/m2)	N	237	233	470
	Mean (SD)	28.4 (4.5)	28.7 (3.9)	28.5 (4.2)
	Median	28.4	28.6	28.6
	Min, Max	17.9, 51.9	20.2, 42.5	17.9, 51.9
BMI Group, n (%)	BMI >=30 (kg/m2)	81 (34.2%)	86 (36.9%)	167 (35.5%)
Duration of Diabetes (yrs)	N	237	233	470
	Mean (SD)	10.5 (6.8)	9.8 (6.1)	10.2 (6.5)
	Median	8.9	9.0	8.9
	Min, Max	0.5, 37.2	0.7, 38.3	0.5, 38.3
Prior use of OAD, n (%)	Yes	236 (99.6%)	232 (99.6%)	468 (99.6%)
	No	1 (0.4%)	1 (0.4%)	2 (0.4%)
Prior Use of Insulin, n (%)	No	237 (100%)	233 (100%)	470 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=237)	NPH (N=233)	Total (N=470)
HBA1C at Baseline (%)	N	237	233	470
	Mean (SD)	9.1 (1.0)	9.1 (1.1)	9.1 (1.0)
	Median	9.1	9.0	9.1
	Min, Max	6.8, 11.5	6.9, 12.3	6.8, 12.3
Country, n (%)	AUSTRIA	14 (5.9%)	16 (6.9%)	30 (6.4%)
	BELGIUM	5 (2.1%)	7 (3.0%)	12 (2.6%)
	FINLAND	10 (4.2%)	9 (3.9%)	19 (4.0%)
	GERMANY	84 (35.4%)	81 (34.8%)	165 (35.1%)
	IRELAND	10 (4.2%)	11 (4.7%)	21 (4.5%)
	ITALY	8 (3.4%)	8 (3.4%)	16 (3.4%)
	NETHERLANDS	7 (3.0%)	6 (2.6%)	13 (2.8%)
	PORTUGAL	6 (2.5%)	7 (3.0%)	13 (2.8%)
	SPAIN	14 (5.9%)	13 (5.6%)	27 (5.7%)
	SWEDEN	4 (1.7%)	4 (1.7%)	8 (1.7%)
	SWITZERLAND	8 (3.4%)	7 (3.0%)	15 (3.2%)
	TURKEY	10 (4.2%)	9 (3.9%)	19 (4.0%)
	UK	57 (24.1%)	55 (23.6%)	112 (23.8%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=237)	NPH (N=233)	Total (N=470)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	235	227	462		
	Mean (SD)	9.14 (1.0)	9.09 (1.1)	9.12 (1.0)		
	Median	9.10	9.00	9.10		
	Min, Max	6.8, 11.5	6.9, 12.3	6.8, 12.3		
HBA1C at Endpoint (%)	N	235	227	462		
	Mean (SD)	7.86 (1.2)	8.26 (1.3)	8.06 (1.3)		
	Median	7.70	8.30	8.00		
	Min, Max	4.9, 11.6	4.9, 12.8	4.9, 12.8		
Change in HBA1C (%)	N	235	227	462		
	Mean (SD)	-1.27 (1.3)	-0.83 (1.3)	-1.06 (1.3)		
	Median	-1.30	-0.80	-1.10		
	Min, Max	-4.6, 2.4	-5.1, 2.2	-5.1, 2.4		
HBA1C Change(ANCOVA) (%) [b]	N	235	227			
	LS Means (se)	-1.27 (0.08)	-0.86 (0.08)		-0.41 (-0.62, -0.20)	0.0001

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=237)		NPH (N=233)		Total (N=470)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Porb
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	5	2.1	3	1.3	8	1.7	2.07	0.39	13.81	0.520
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.3	4	1.7	7	1.5	0.71	0.10	4.20	0.938
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	95	40.1	72	30.9	167	35.5	1.58	1.03	2.42	0.033
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	4	1.7	4	0.9	0.00	0.00	1.27	0.160
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	2	0.9	3	0.6	0.43	0.01	8.76	0.929
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	19	8.0	65	27.9	84	17.9	0.18	0.09	0.34	<0.001

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=158)	NPH (N=159)	Total (N=317)
Age (yrs)	N	158	159	317
	Mean (SD)	55.1 (9.6)	56.9 (7.9)	56.0 (8.8)
	Median	56.5	58.0	57.0
	Min, Max	20.0, 70.0	36.0, 70.0	20.0, 70.0
Age Group, n (%)	Age <65 (yrs)	135 (85.4%)	131 (82.4%)	266 (83.9%)
	65 <= Age <=75 (yrs)	23 (14.6%)	28 (17.6%)	51 (16.1%)
	Age >= 65 (yrs)	23 (14.6%)	28 (17.6%)	51 (16.1%)
Sex, n (%)	Male	98 (62.0%)	101 (63.5%)	199 (62.8%)
	Female	60 (38.0%)	58 (36.5%)	118 (37.2%)
Race, n (%)	Asian/Oriental	158 (100%)	159 (100%)	317 (100%)
Weight (kg) [b]	N	158	159	317
	Mean (SD)	61.6 (10.4)	61.7 (10.3)	61.7 (10.3)
	Median	60.3	62.5	62.0
	Min, Max	40.0, 93.4	36.7, 86.5	36.7, 93.4

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=158)	NPH (N=159)	Total (N=317)
Height (cm) [b]	N	158	159	317
	Mean (SD)	161.8 (8.7)	161.9 (8.8)	161.9 (8.7)
	Median	161.7	163.0	162.6
	Min, Max	141.5, 187.0	142.0, 182.0	141.5, 187.0
BMI (kg/m2)	N	158	159	317
	Mean (SD)	23.5 (3.0)	23.4 (2.8)	23.4 (2.9)
	Median	23.2	23.1	23.1
	Min, Max	16.6, 31.2	16.7, 29.9	16.6, 31.2
BMI Group, n (%)	BMI >=30 (kg/m2)	2 (1.3%)	0 (0.0%)	2 (0.6%)
Duration of Diabetes (yrs)	N	158	159	317
	Mean (SD)	12.0 (6.9)	12.6 (6.7)	12.3 (6.8)
	Median	10.5	11.1	11.1
	Min, Max	0.6, 37.2	0.5, 38.1	0.5, 38.1
Prior use of OAD, n (%)	Yes	158 (100%)	159 (100%)	317 (100%)
Prior Use of Insulin, n (%)	No	158 (100%)	159 (100%)	317 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=158)	NPH (N=159)	Total (N=317)
HBA1C at Baseline (%)	N	158	159	317
	Mean (SD)	9.1 (1.1)	9.1 (1.0)	9.1 (1.0)
	Median	8.9	8.9	8.9
	Min, Max	7.3, 12.0	7.3, 11.8	7.3, 12.0
Country, n (%)	JAPAN	158 (100%)	159 (100%)	317 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=158)	NPH (N=159)	Total (N=317)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	157	157	314		
	Mean (SD)	9.10 (1.1)	9.12 (1.0)	9.11 (1.0)		
	Median	8.90	8.90	8.90		
	Min, Max	7.3, 12.0	7.3, 11.8	7.3, 12.0		
HBA1C at Endpoint (%)	N	157	157	314		
	Mean (SD)	8.01 (1.0)	8.08 (1.1)	8.05 (1.0)		
	Median	7.90	7.90	7.90		
	Min, Max	5.9, 11.8	5.9, 12.0	5.9, 12.0		
Change in HBA1C (%)	N	157	157	314		
	Mean (SD)	-1.08 (1.0)	-1.03 (0.9)	-1.06 (0.9)		
	Median	-1.00	-0.90	-1.00		
	Min, Max	-4.4, 1.6	-4.5, 1.1	-4.5, 1.6		
HBA1C Change(ANCOVA) (%) [b]	N	157	157			
	LS Means (se)	-1.09 (0.07)	-1.03 (0.07)		-0.06 (-0.25, 0.13)	0.5346

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=158)		NPH (N=159)		Total (N=317)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.6	0	0.0	1	0.3	-	0.05	Infty	0.997
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.9	2	1.3	5	1.6	1.52	0.17	18.40	0.994
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	30	19.0	31	19.5	61	19.2	0.97	0.53	1.76	1.000
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	0	0.0	0	0.0	-			
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	0	0.0	0	0.0	0	0.0	-			
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	6	3.8	2	1.3	8	2.5	3.10	0.54	31.75	0.279

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=514)	NPH (N=503)	Total (N=1017)
Age (yrs)	N	514	503	1017
	Mean (SD)	54.9 (8.8)	55.3 (8.5)	55.1 (8.6)
	Median	55.0	56.0	56.0
	Min, Max	33.0, 74.0	29.0, 72.0	29.0, 74.0
Age Group, n (%)	Age <65 (yrs)	430 (83.7%)	426 (84.7%)	856 (84.2%)
	65 <= Age <=75 (yrs)	84 (16.3%)	77 (15.3%)	161 (15.8%)
	Age >= 65 (yrs)	84 (16.3%)	77 (15.3%)	161 (15.8%)
Sex, n (%)	Male	278 (54.1%)	270 (53.7%)	548 (53.9%)
	Female	236 (45.9%)	233 (46.3%)	469 (46.1%)
Race, n (%)	White	447 (87.0%)	421 (83.7%)	868 (85.3%)
	Black	49 (9.5%)	62 (12.3%)	111 (10.9%)
	Asian/Oriental	9 (1.8%)	12 (2.4%)	21 (2.1%)
	Multiracial	9 (1.8%)	8 (1.6%)	17 (1.7%)
Weight (kg) [b]	N	513	503	1016
	Mean (SD)	100.2 (22.7)	98.7 (22.3)	99.4 (22.5)
	Median	97.1	97.0	97.1
	Min, Max	49.5, 183.3	48.5, 180.1	48.5, 183.3

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=514)	NPH (N=503)	Total (N=1017)
Height (cm) [b]	N	511	503	1014
	Mean (SD)	170.1 (10.1)	170.1 (10.3)	170.1 (10.2)
	Median	170.2	170.0	170.2
	Min, Max	137.2, 198.1	127.0, 200.7	127.0, 200.7
BMI (kg/m2)	N	510	503	1013
	Mean (SD)	34.6 (7.2)	34.1 (7.2)	34.3 (7.2)
	Median	33.5	33.6	33.5
	Min, Max	20.9, 64.4	18.6, 62.7	18.6, 64.4
BMI Group, n (%)	BMI >=30 (kg/m2)	360 (70.6%)	341 (67.8%)	701 (69.2%)
Duration of Diabetes (yrs)	N	514	503	1017
	Mean (SD)	10.7 (6.9)	10.8 (6.8)	10.8 (6.8)
	Median	9.6	10.0	10.0
	Min, Max	1.1, 38.0	1.0, 51.0	1.0, 51.0
Prior use of OAD, n (%)	Yes	495 (96.3%)	475 (94.4%)	970 (95.4%)
	No	19 (3.7%)	28 (5.6%)	47 (4.6%)
Prior Use of Insulin, n (%)	Yes	345 (67.1%)	353 (70.2%)	698 (68.6%)
	No	169 (32.9%)	150 (29.8%)	319 (31.4%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=514)	NPH (N=503)	Total (N=1017)
HBA1C at Baseline (%)	N	513	503	1016
	Mean (SD)	8.4 (1.4)	8.3 (1.4)	8.4 (1.4)
	Median	8.2	8.0	8.1
	Min, Max	5.7, 12.7	5.8, 13.5	5.7, 13.5
Country, n (%)	CANADA	154 (30.0%)	156 (31.0%)	310 (30.5%)
	USA	360 (70.0%)	347 (69.0%)	707 (69.5%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=514)	NPH (N=503)	Total (N=1017)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	498	486	984		
	Mean (SD)	8.42 (1.4)	8.29 (1.4)	8.36 (1.4)		
	Median	8.20	8.00	8.10		
	Min, Max	5.7, 12.7	5.8, 13.5	5.7, 13.5		
HBA1C at Endpoint (%)	N	498	486	984		
	Mean (SD)	7.80 (1.3)	7.56 (1.3)	7.68 (1.3)		
	Median	7.60	7.40	7.50		
	Min, Max	5.1, 13.6	4.8, 14.2	4.8, 14.2		
Change in HBA1C (%)	N	498	486	984		
	Mean (SD)	-0.62 (1.5)	-0.73 (1.5)	-0.67 (1.5)		
	Median	-0.50	-0.70	-0.60		
	Min, Max	-5.6, 3.5	-5.3, 6.5	-5.6, 6.5		
HBA1C Change(ANCOVA) (%) [b]	N	498	486			
	LS Means (se)	-0.54 (0.06)	-0.73 (0.06)		0.19 (0.04, 0.33)	0.0122

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=514)		NPH (N=503)		Total (N=1017)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	20	3.9	36	7.2	56	5.5	0.51	0.27	0.93	0.027
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	104	20.2	139	27.6	243	23.9	0.65	0.47	0.88	0.005
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	327	63.6	353	70.2	680	66.9	0.72	0.53	0.96	0.027
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	27	5.3	32	6.4	59	5.8	0.81	0.46	1.43	0.532
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	106	20.6	116	23.1	222	21.8	0.87	0.64	1.20	0.429
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	270	52.5	282	56.1	552	54.3	0.87	0.66	1.13	0.303

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=259)	NPH (N=259)	Total (N=518)
Age (yrs)	N	259	259	518
	Mean (SD)	59.5 (9.7)	59.2 (9.9)	59.3 (9.8)
	Median	59.0	59.0	59.0
	Min, Max	35.0, 80.0	35.0, 79.0	35.0, 80.0
Age Group, n (%)	Age <65 (yrs)	170 (65.6%)	170 (65.6%)	340 (65.6%)
	65 <= Age <=75 (yrs)	82 (31.7%)	82 (31.7%)	164 (31.7%)
	Age > 75 (yrs)	7 (2.7%)	7 (2.7%)	14 (2.7%)
	Age >= 65 (yrs)	89 (34.4%)	89 (34.4%)	178 (34.4%)
Sex, n (%)	Male	150 (57.9%)	161 (62.2%)	311 (60.0%)
	Female	109 (42.1%)	98 (37.8%)	207 (40.0%)
Race, n (%)	White	208 (80.3%)	209 (80.7%)	417 (80.5%)
	Black	40 (15.4%)	36 (13.9%)	76 (14.7%)
	Asian/Oriental	8 (3.1%)	9 (3.5%)	17 (3.3%)
	Multiracial	2 (0.8%)	5 (1.9%)	7 (1.4%)
	Hispanic	1 (0.4%)	0 (0.0%)	1 (0.2%)
Weight (kg) [b]	N	259	259	518
	Mean (SD)	89.6 (17.3)	90.7 (17.8)	90.2 (17.6)
	Median	89.1	90.9	89.5
	Min, Max	43.6, 138.2	50.2, 150.0	43.6, 150.0

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=259)	NPH (N=259)	Total (N=518)
Height (cm) [b]	N	259	258	517
	Mean (SD)	170.7 (10.5)	172.6 (9.3)	171.6 (10.0)
	Median	171.5	172.7	172.7
	Min, Max	139.7, 194.3	149.9, 193.0	139.7, 194.3
BMI (kg/m2)	N	259	258	517
	Mean (SD)	30.7 (5.0)	30.4 (5.1)	30.5 (5.1)
	Median	30.3	30.2	30.2
	Min, Max	16.5, 45.1	19.6, 43.2	16.5, 45.1
BMI Group, n (%)	BMI >=30 (kg/m2)	137 (52.9%)	136 (52.7%)	273 (52.8%)
Duration of Diabetes (yrs)	N	259	259	518
	Mean (SD)	13.4 (8.3)	14.1 (9.0)	13.7 (8.7)
	Median	12.0	12.0	12.0
	Min, Max	1.0, 49.0	1.0, 56.0	1.0, 56.0
Prior use of OAD, n (%)	No	259 (100%)	259 (100%)	518 (100%)
Prior Use of Insulin, n (%)	Yes	259 (100%)	259 (100%)	518 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 1. Demographics and Baseline Characteristics
ITT Population (Actual Treatment)

	Statistics or Category [a]	Glargine (N=259)	NPH (N=259)	Total (N=518)
HBA1C at Baseline (%)	N	259	259	518
	Mean (SD)	8.6 (1.2)	8.5 (1.2)	8.6 (1.2)
	Median	8.3	8.4	8.3
	Min, Max	6.3, 12.3	6.2, 12.6	6.2, 12.6
Country, n (%)	USA	259 (100%)	259 (100%)	518 (100%)

[a] N = Number of patients, SD = Standard deviation
[b] Baseline weight and height are their last non missing values prior to the 1st dose of study medication

Table 2. HBA1C at Baseline and Endpoint
ITT Population, Patients with Non-missing HBA1C at Baseline and at Endpoint (Actual Treatment)

	Statistics	Glargine (N=259)	NPH (N=259)	Total (N=518)	Difference: Glargine - NPH (95% CI)	P-value
HBA1C at Baseline (%)	N	246	255	501		
	Mean (SD)	8.56 (1.2)	8.53 (1.2)	8.54 (1.2)		
	Median	8.30	8.40	8.30		
	Min, Max	6.3, 12.3	6.2, 12.6	6.2, 12.6		
HBA1C at Endpoint (%)	N	246	255	501		
	Mean (SD)	8.14 (1.3)	7.96 (1.2)	8.04 (1.2)		
	Median	7.95	7.80	7.90		
	Min, Max	5.5, 13.2	5.7, 13.6	5.5, 13.6		
Change in HBA1C (%)	N	246	255	501		
	Mean (SD)	-0.43 (1.3)	-0.57 (1.2)	-0.50 (1.2)		
	Median	-0.40	-0.50	-0.40		
	Min, Max	-4.6, 3.7	-4.0, 3.1	-4.6, 3.7		
HBA1C Change(ANCOVA) (%) [b]	N	246	255			
	LS Means (se)	-0.41 (0.07)	-0.59 (0.06)		0.17 (-0.00, 0.35)	0.0551

[a] N = Number of subjects who have values in HBA1C at both baseline and enopoint, SD = Standard Deviation, SE = Standard Error
[b] HBA1C Change(ANCOVA)(%) model: Change in HBA1C from baseline to endpoint = HBA1C at baseline + Actual treatment group + pooled center

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (N=259)		NPH (N=259)		Total (N=518)		Comparison (Glargine over NPH)			
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Odds Ratio	95% LCB	95% UCB	Prob
Incidence of Daytime Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	4	1.5	5	1.0	0.25	0.01	2.71	0.415
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	11	4.2	21	8.1	32	6.2	0.45	0.19	1.03	0.059
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	123	47.5	138	53.3	261	50.4	0.78	0.53	1.16	0.233
Incidence of Nocturnal Hypoglycemia, n/N(%)										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	2	0.8	3	0.6	0.43	0.01	8.16	0.878
Symptomatic Hypoglycemia with with SMBG<2.0(mmol/l) (36 mg/dl)	13	5.0	12	4.6	25	4.8	0.94	0.38	2.36	1.000
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	74	28.6	93	35.9	167	32.2	0.66	0.44	1.00	0.052

Note:
[a]. Odds ratio, 95%LCB, 95%UCB and P-value are from Common Odds Ratio (Exact Conf Limits) in Cochran-Mantel-Haenszel Test stratified by pooled center

sanofi aventis
Das Wichtigste ist die Gesundheit

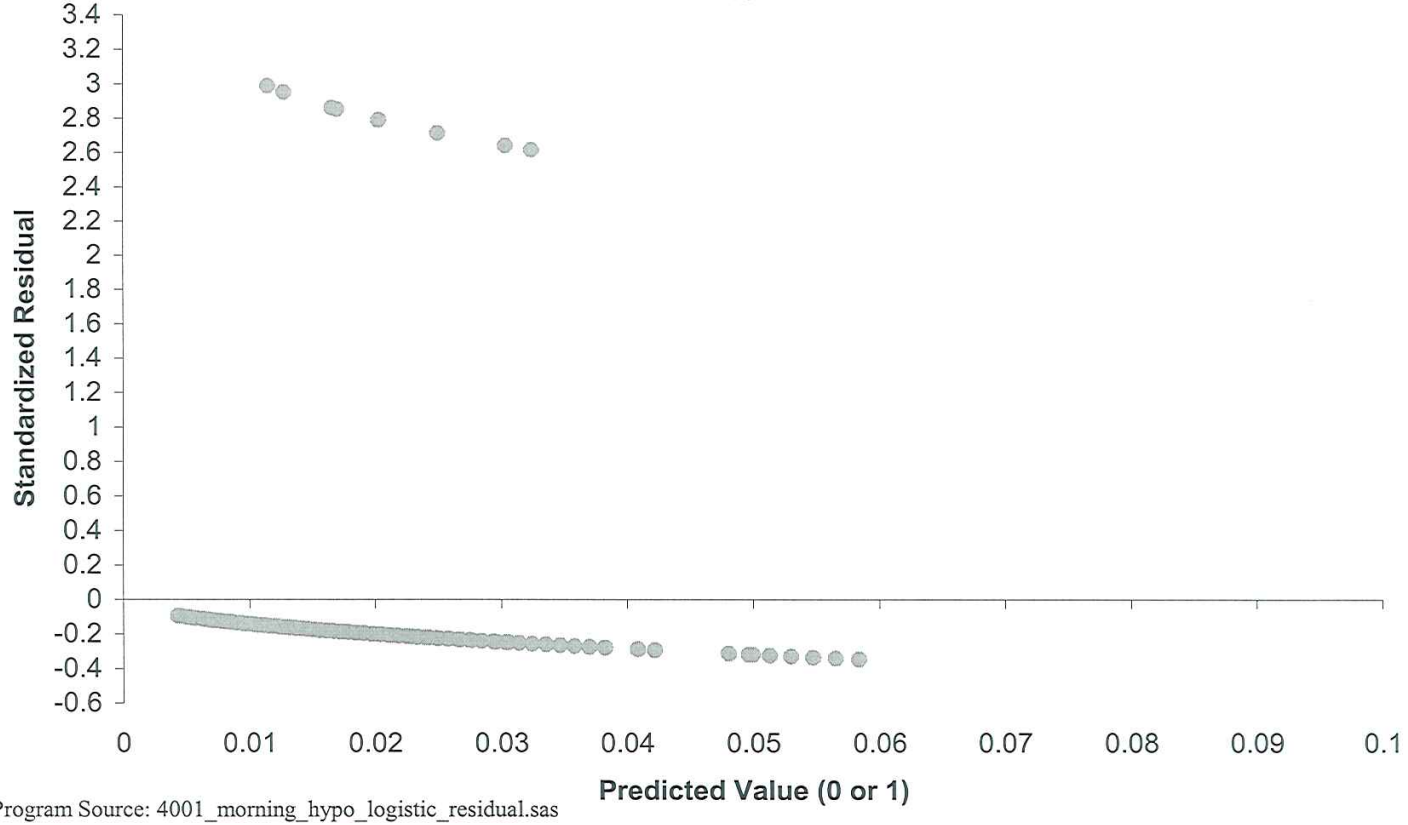
Logistic Regression Analyses Studypools II - V

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

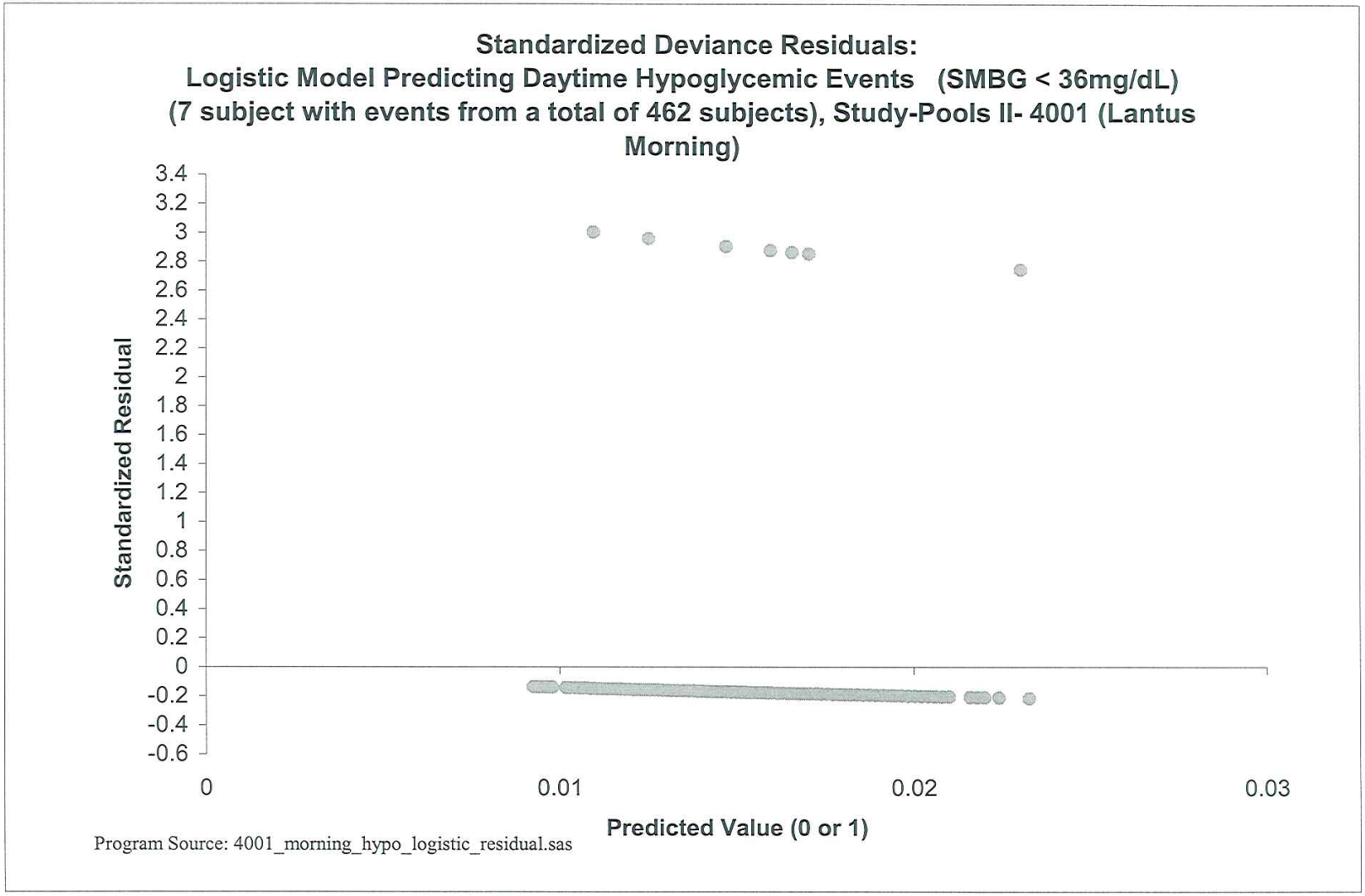
	Glargine (235)		NPH (227)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	5	2.1	3	1.3	1.40	0.32	6.05	0.653	0.225	0.653
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.3	4	1.8	0.75	0.16	3.48	0.715	0.752	0.546
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	95	40.4	72	31.7	1.28	0.86	1.89	0.224	<0.001	0.622
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	4	1.8	0.16	0.00	1.30	0.090	0.150	0.863
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	1	0.4	2	0.9	0.64	0.06	7.26	0.715	0.154	0.706
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	19	8.1	65	28.6	0.20	0.11	0.34	<0.001	0.027	0.711

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 237, NPH: 233. Patients who have values in change in HbA1C: Glargine: 235, NPH: 227
(4). 8 patients who have missing values in change in HbA1C were excluded, among them no one has at least one hypoglycemia
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C
(6). For "Nocturnal Severe Hypoglycemia" item, Adjusted odds ratio, 95% CI and P-values come from "Conditional Exact Tests" of Logistic regression

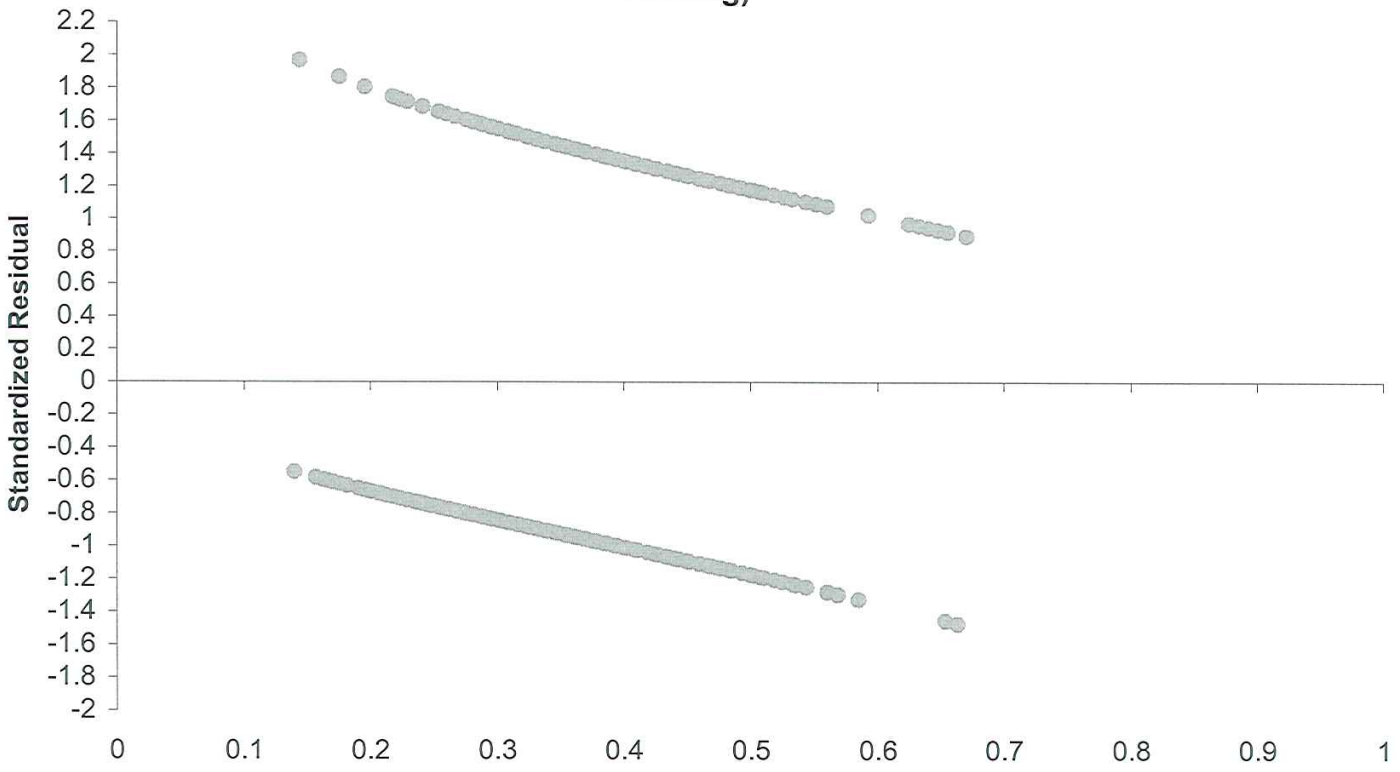
Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (Severe Events)
(8 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)



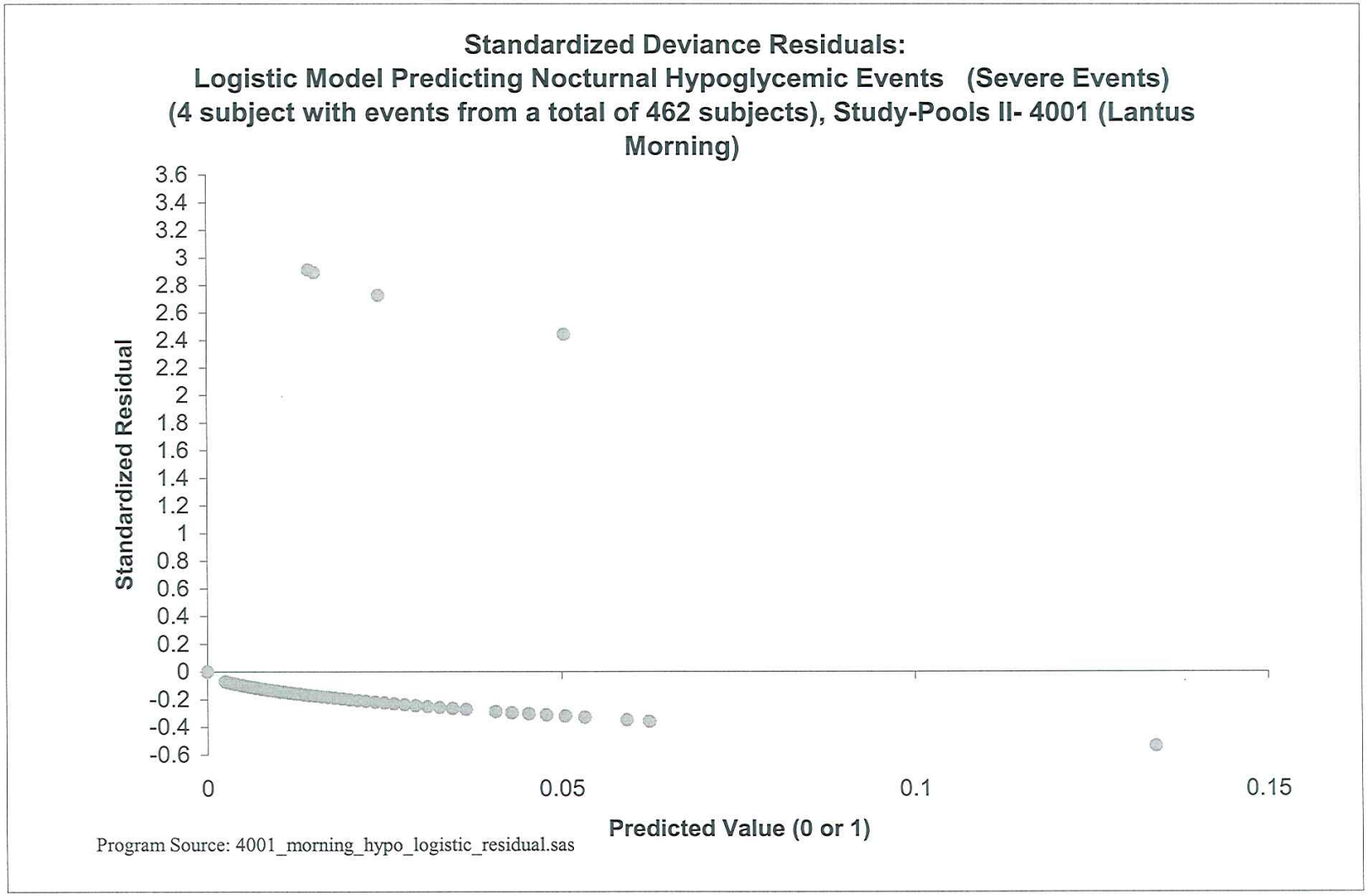
Program Source: 4001_morning_hypo_logistic_residual.sas

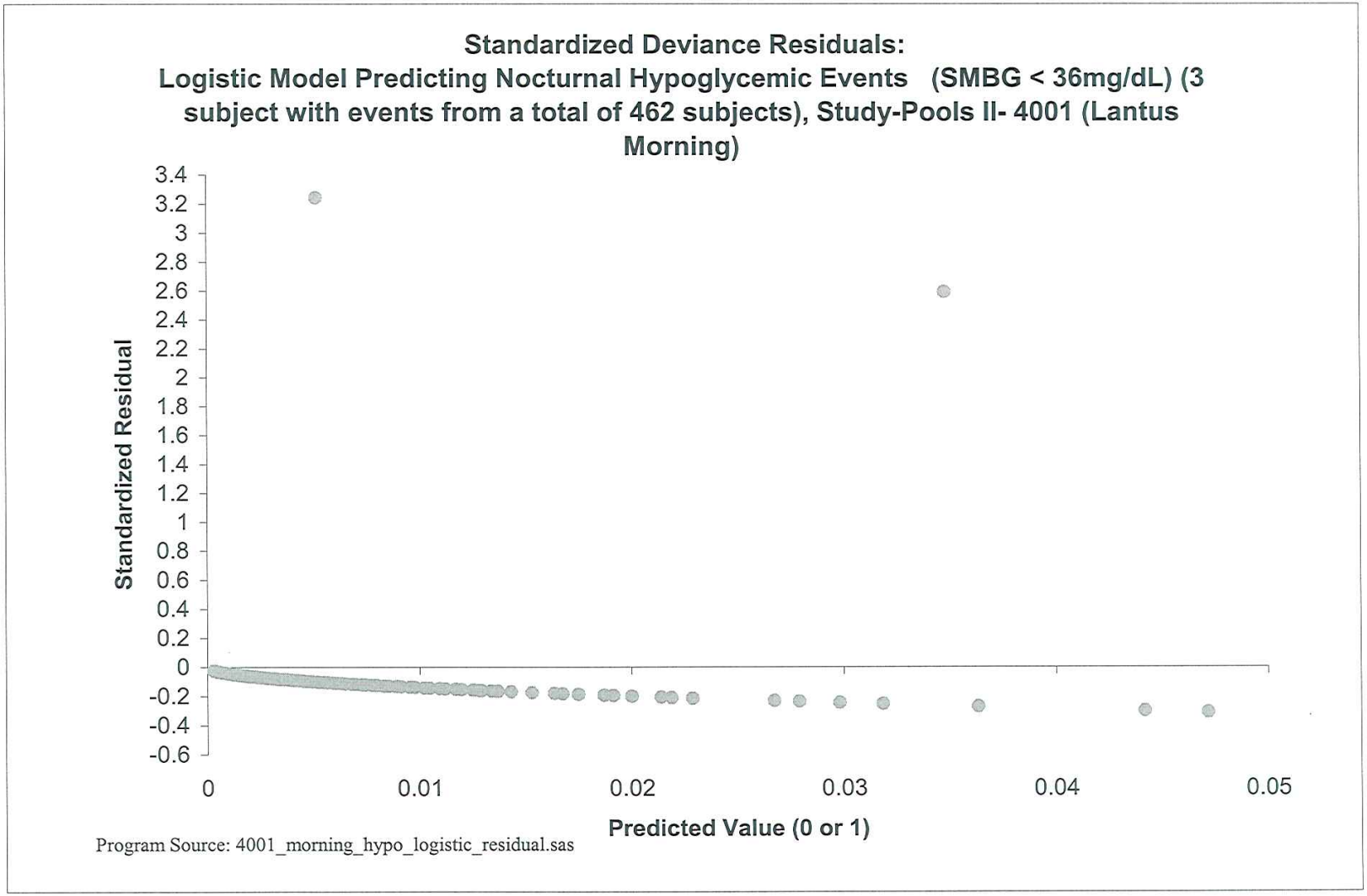

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
(7 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
Predicted Value (0 or 1)
Program Source: 4001_morning_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
(167 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)

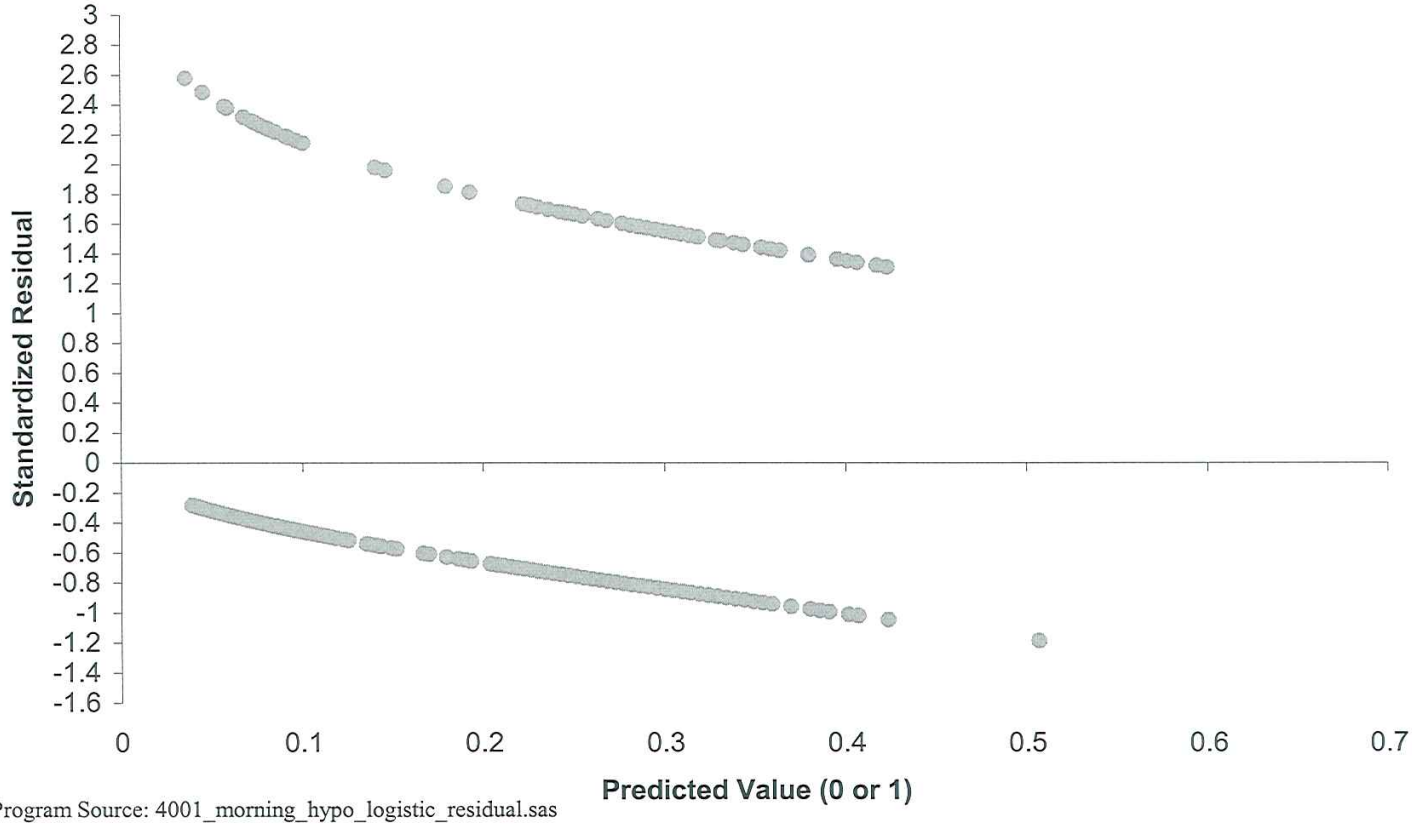


Program Source: 4001_morning_hypo_logistic_residual.sas


Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (Severe Events)
(4 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)
Standardized Residual
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.05
0.1
0.15
Predicted Value (0 or 1)
Program Source: 4001_morning_hypo_logistic_residual.sas


Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 36mg/dL) (3 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
0.04
0.05
Predicted Value (0 or 1)
Program Source: 4001_morning_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 70mg/dL)(84 subject with events from a total of 462 subjects), Study-Pools II- 4001 (Lantus Morning)

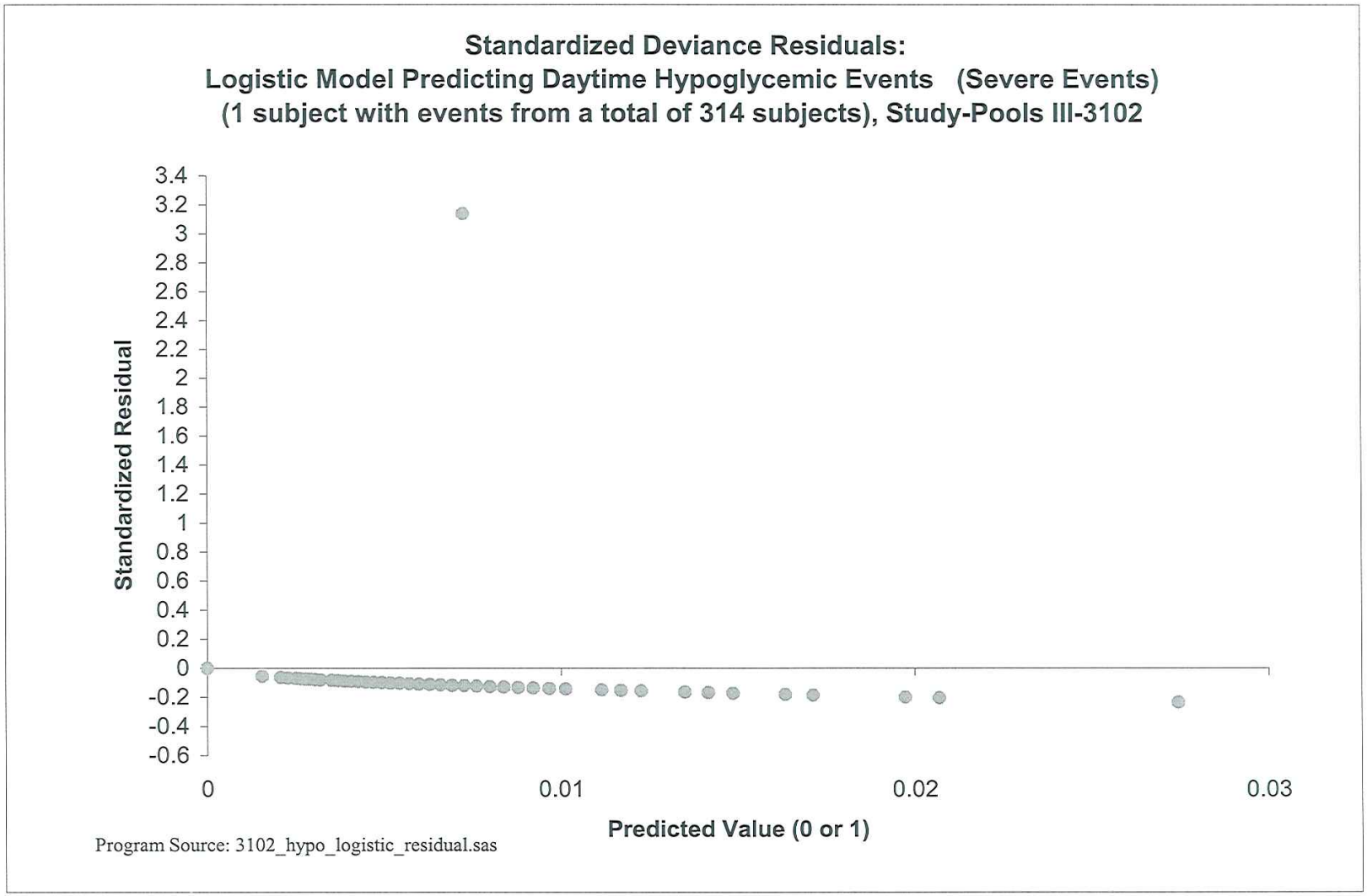


Program Source: 4001_morning_hypo_logistic_residual.sas

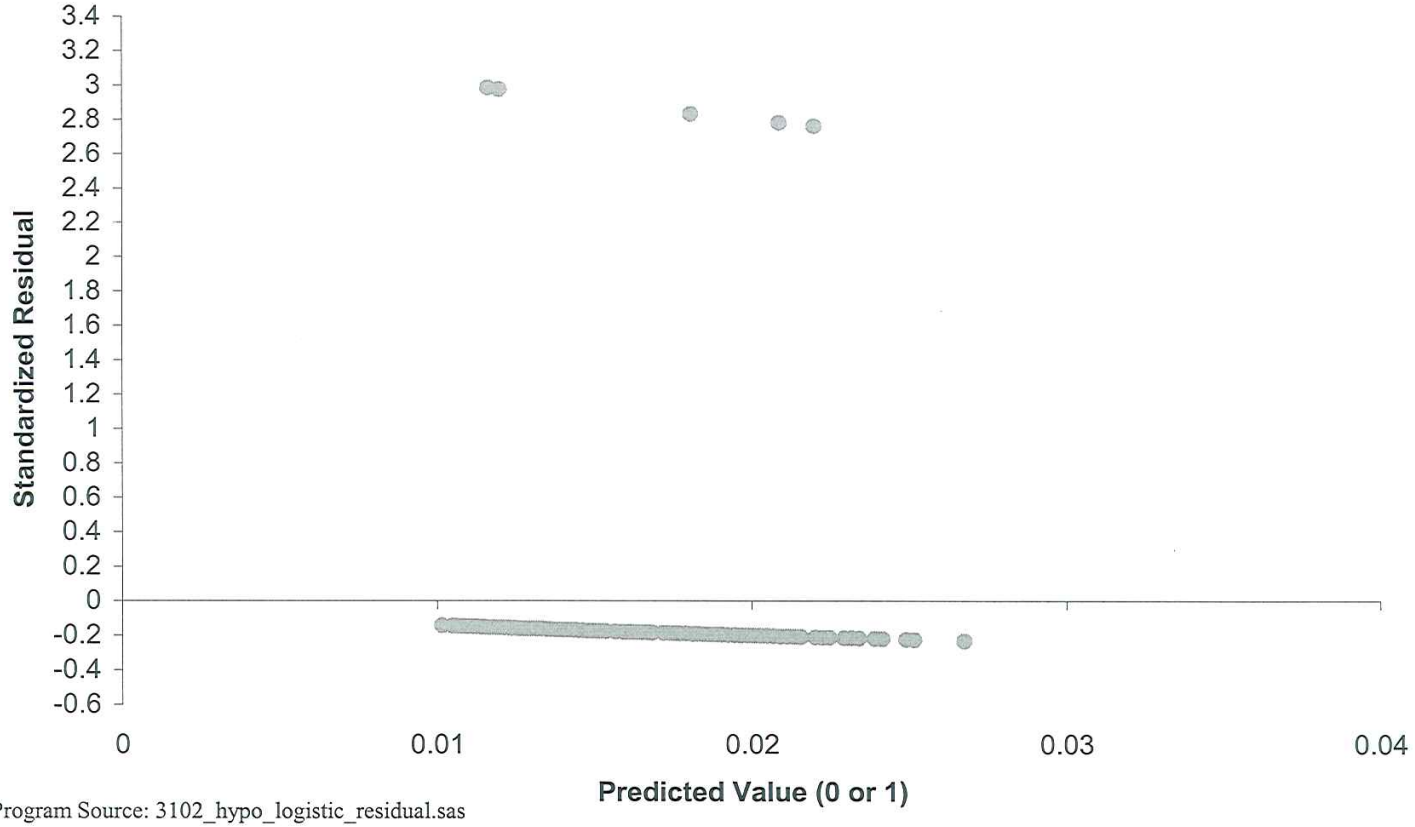
Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (157)		NPH (157)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.6	0	0.0	0.33	0.01	Infinity	1.000	0.573	0.855
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	3	1.9	2	1.3	1.50	0.25	9.12	0.659	0.820	0.540
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	30	19.1	31	19.7	0.96	0.55	1.67	0.876	0.626	0.524
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	0	0.0	0	0.0	-	-	-	-	-	-
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	0	0.0	0	0.0	-	-	-	-	-	-
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	6	3.8	2	1.3	2.99	0.59	15.12	0.185	0.202	0.709

Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 158, NPH: 159. Patients who have values in change in HbA1C: Glargine: 157, NPH: 157
(4). 3 patients who have missing values in change in HbA1C were excluded, among them no one has at least one hypoglycemia
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C
(6). For "Day Time Severe Hypoglycemia" item, Adjusted odds ratio, 95% CI and P-values come from "Conditional Exact Tests" of Logistic regression

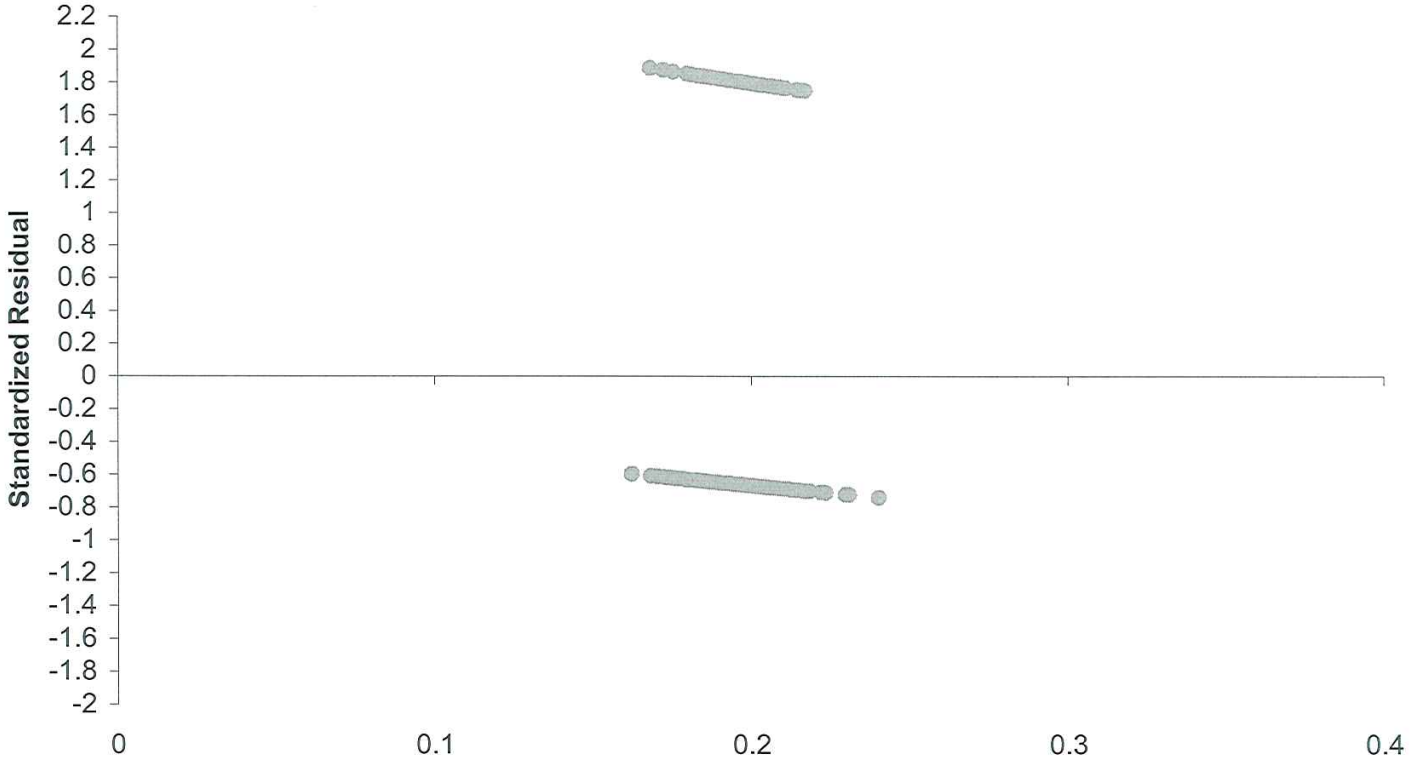

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (Severe Events)
(1 subject with events from a total of 314 subjects), Study-Pools III-3102
Standardized Residual
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
Predicted Value (0 or 1)
Program Source: 3102_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
(5 subjects with events from a total of 314 subjects), Study-Pools III-3102



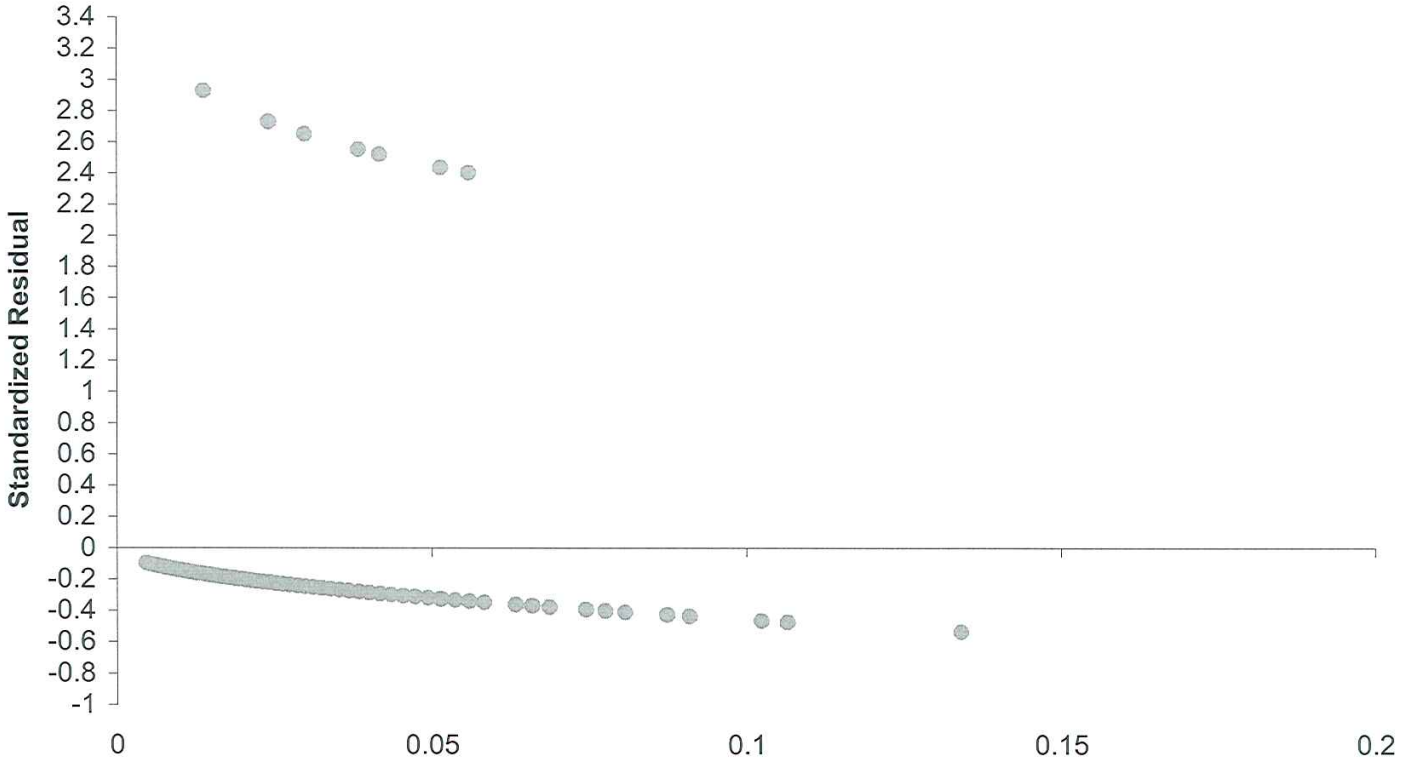
Program Source: 3102_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
(61 subjects with events from a total of 314 subjects), Study-Pools III-3102



Program Source: 3102_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 70mg/dL)(8 subjects with events from a total of 314 subjects), Study-Pools III-3102



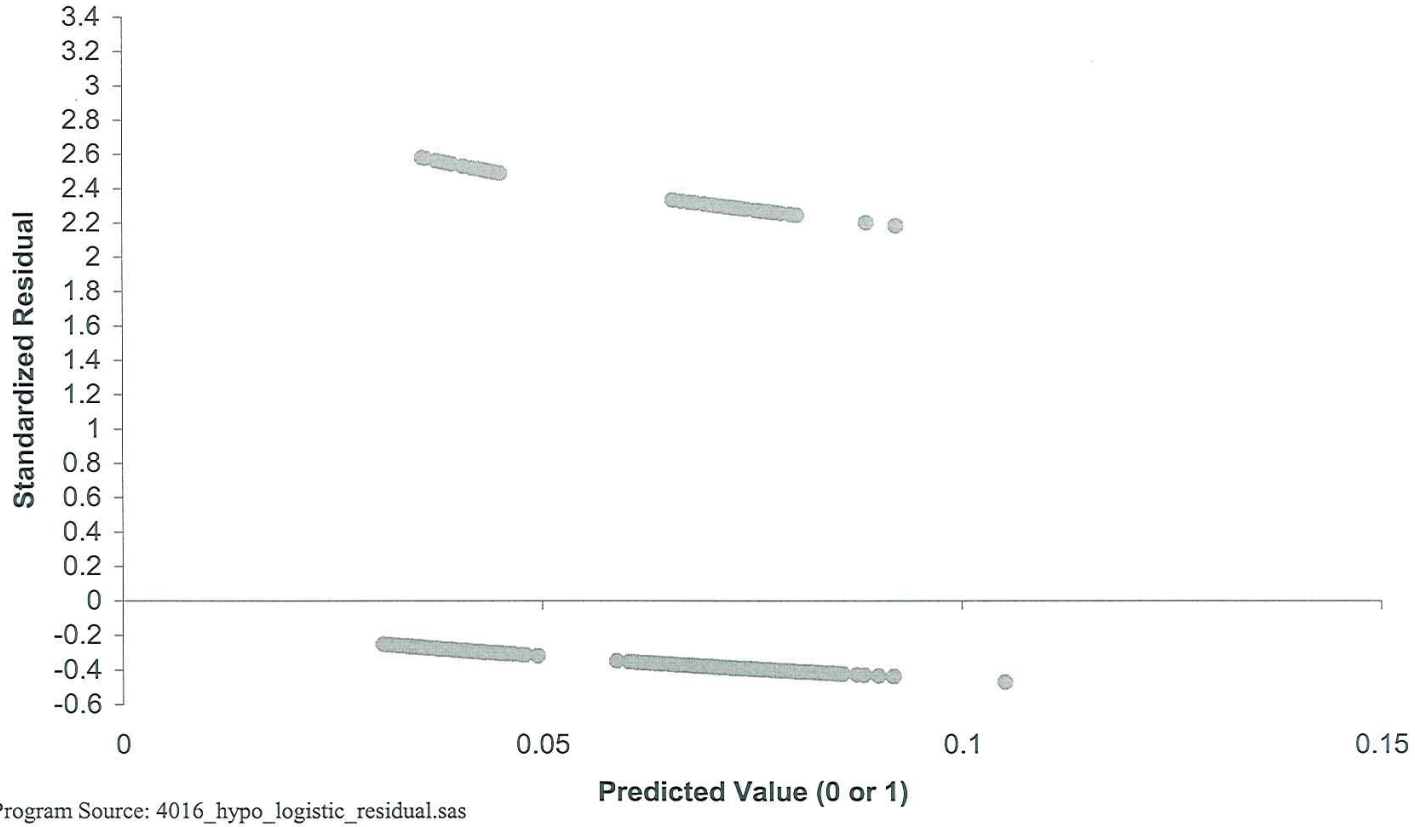
Program Source: 3102_hypo_logistic_residual.sas

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (498)		NPH (486)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	20	4.0	36	7.4	0.52	0.30	0.91	0.023	0.575	0.585
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	103	20.7	139	28.6	0.65	0.49	0.88	0.004	0.421	0.563
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	325	65.3	352	72.4	0.72	0.55	0.94	0.017	0.321	0.548
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	27	5.4	32	6.6	0.81	0.48	1.37	0.430	0.546	0.536
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	105	21.1	116	23.9	0.85	0.63	1.15	0.292	0.836	0.518
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	269	54.0	282	58.0	0.85	0.66	1.09	0.204	0.905	0.527

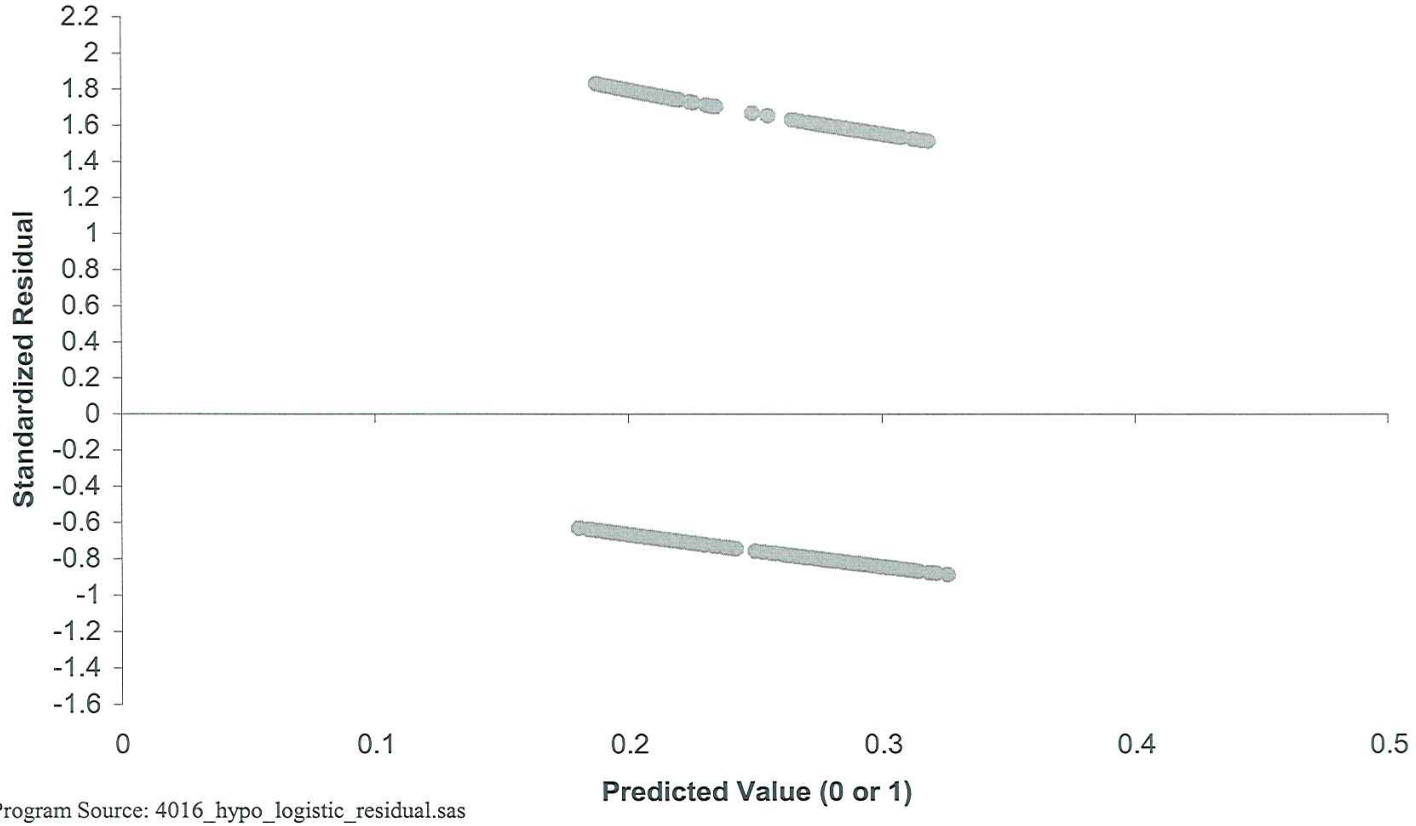
Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 514, NPH: 503. Patients who have values in change in HbA1C: Glargine: 498, NPH: 486
(4). 33 patients who have missing values in change in HbA1C were excluded, including 3 patients who have at least one hypoglycemia: 2 Glargine and 1 NPH patients
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (Severe Events)
(56 subject with events from a total of 984 subjects), Study-Pools IV-4016



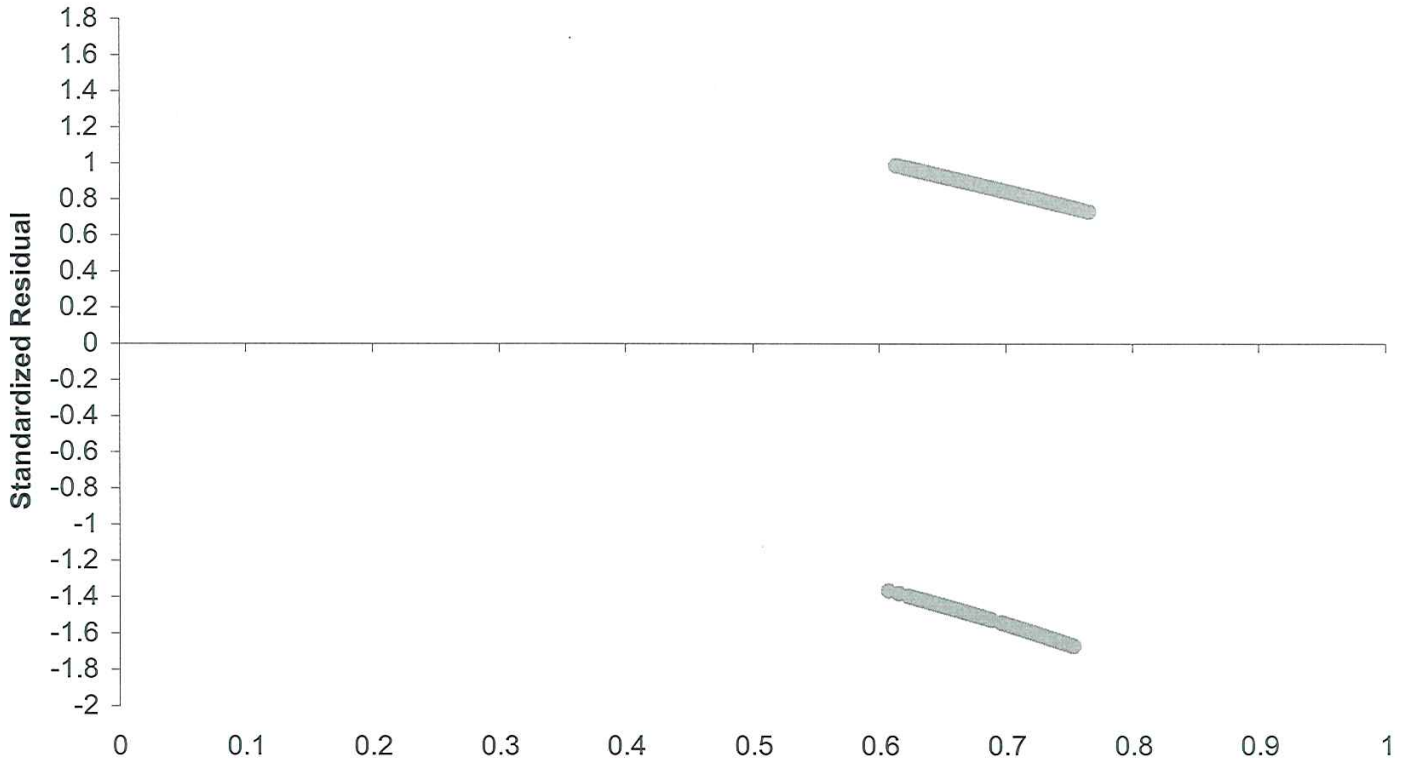
Program Source: 4016_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
(242 subject with events from a total of 984 subjects), Study-Pools IV-4016



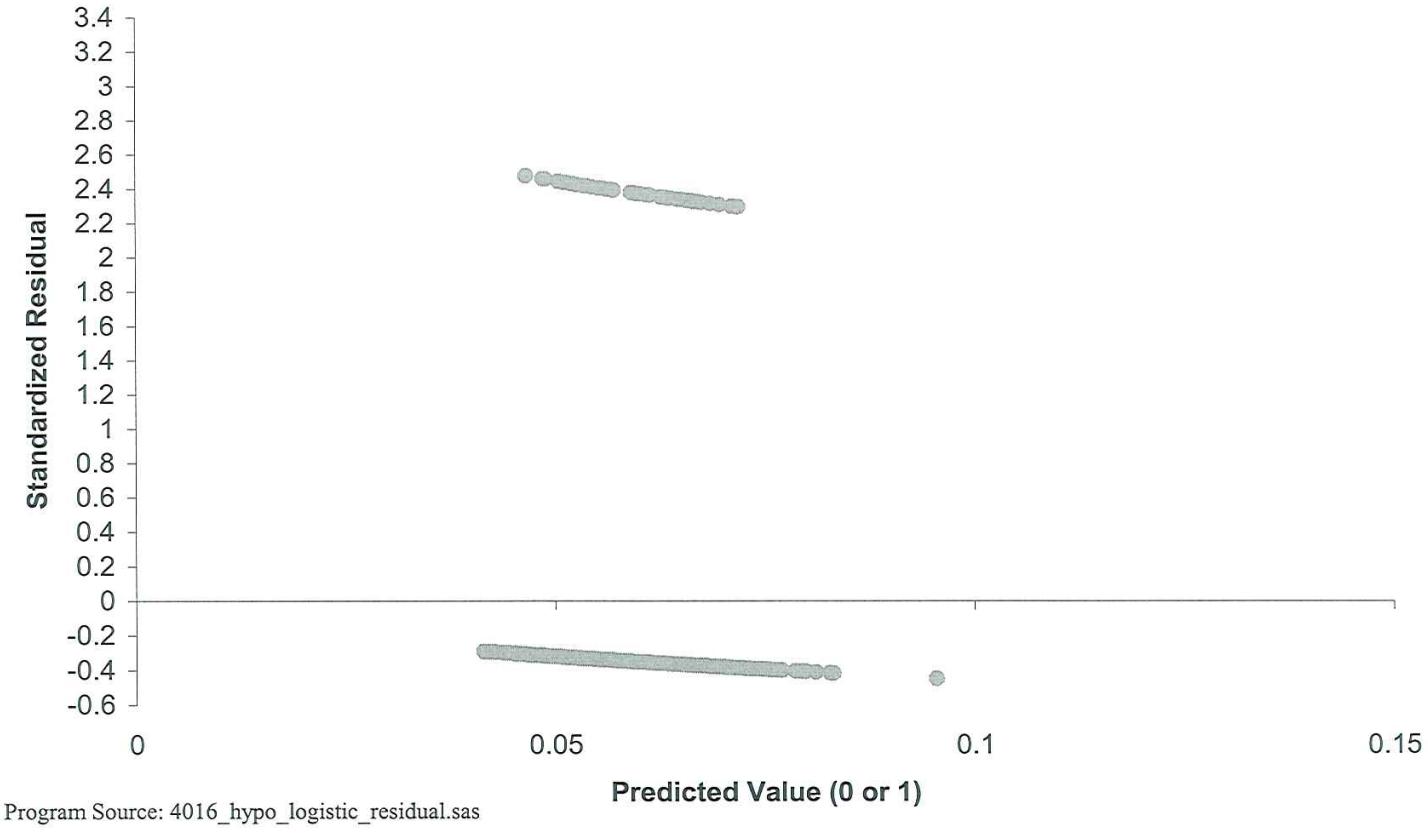
Program Source: 4016_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
(677 subject with events from a total of 984 subjects), Study-Pools IV-4016



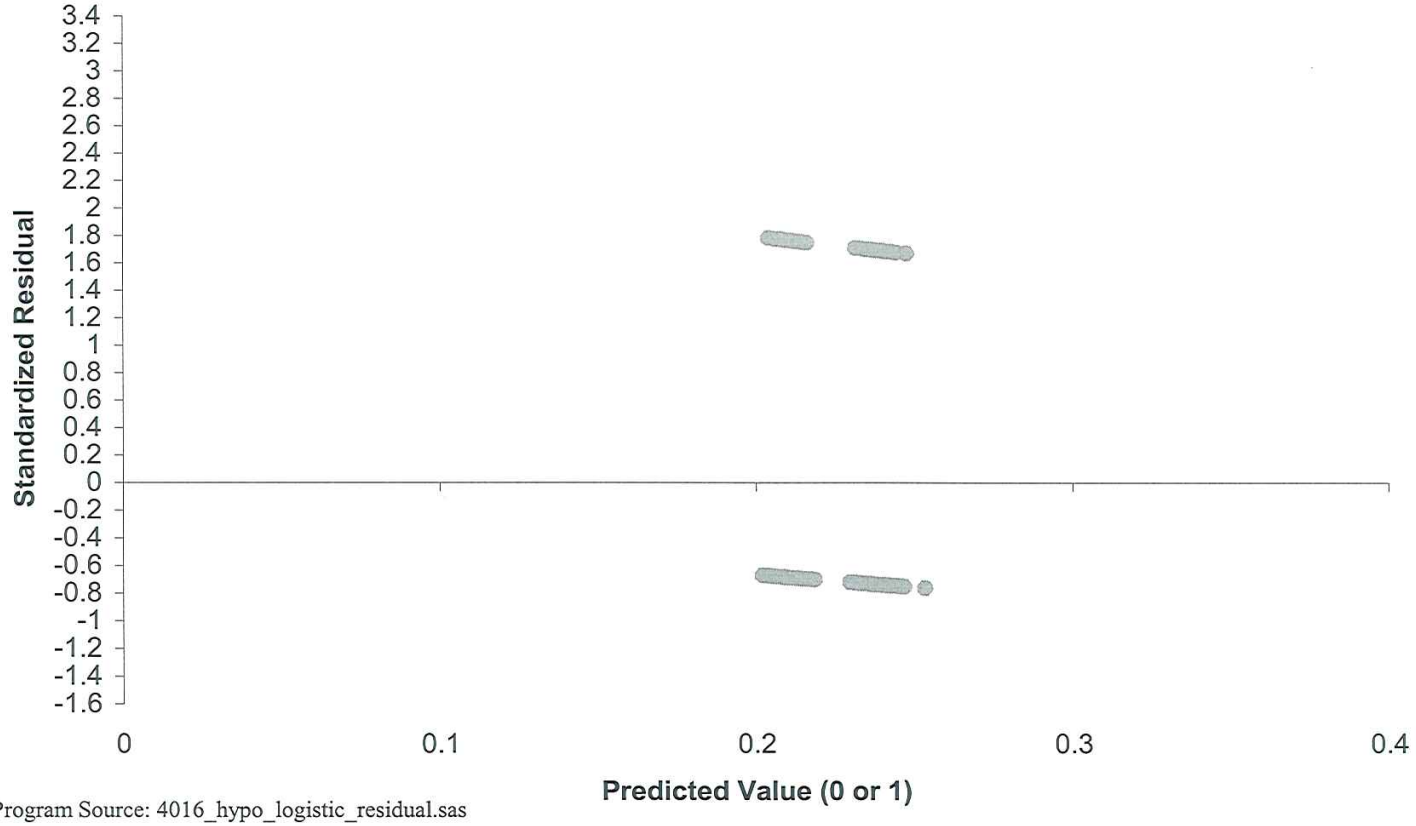
Program Source: 4016_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (Severe Events)
(59 subject with events from a total of 984 subjects), Study-Pools IV-4016



Program Source: 4016_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 36mg/dL) (221 subject with events from a total of 984 subjects), Study-Pools IV-4016



Program Source: 4016_hypo_logistic_residual.sas

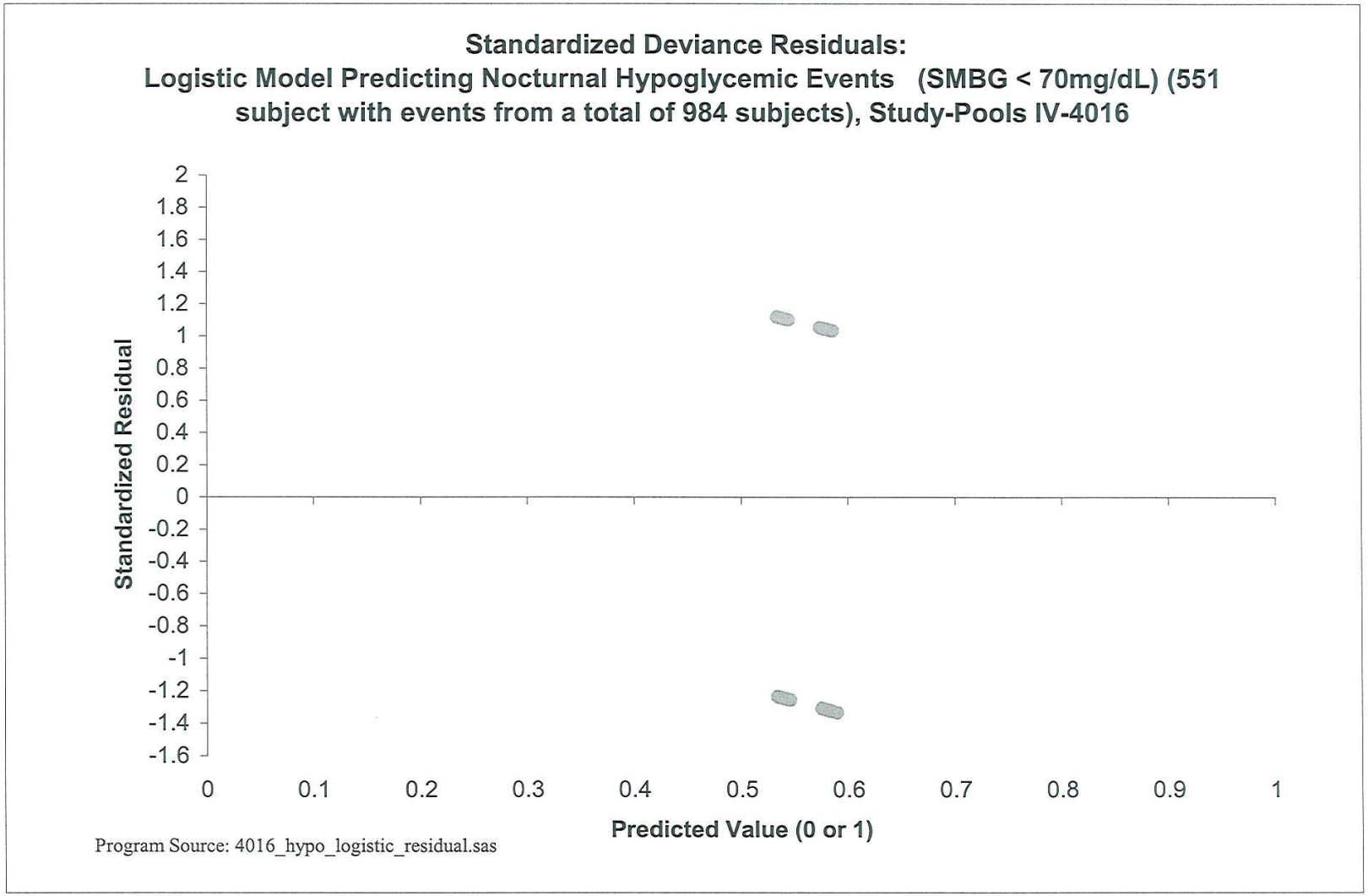
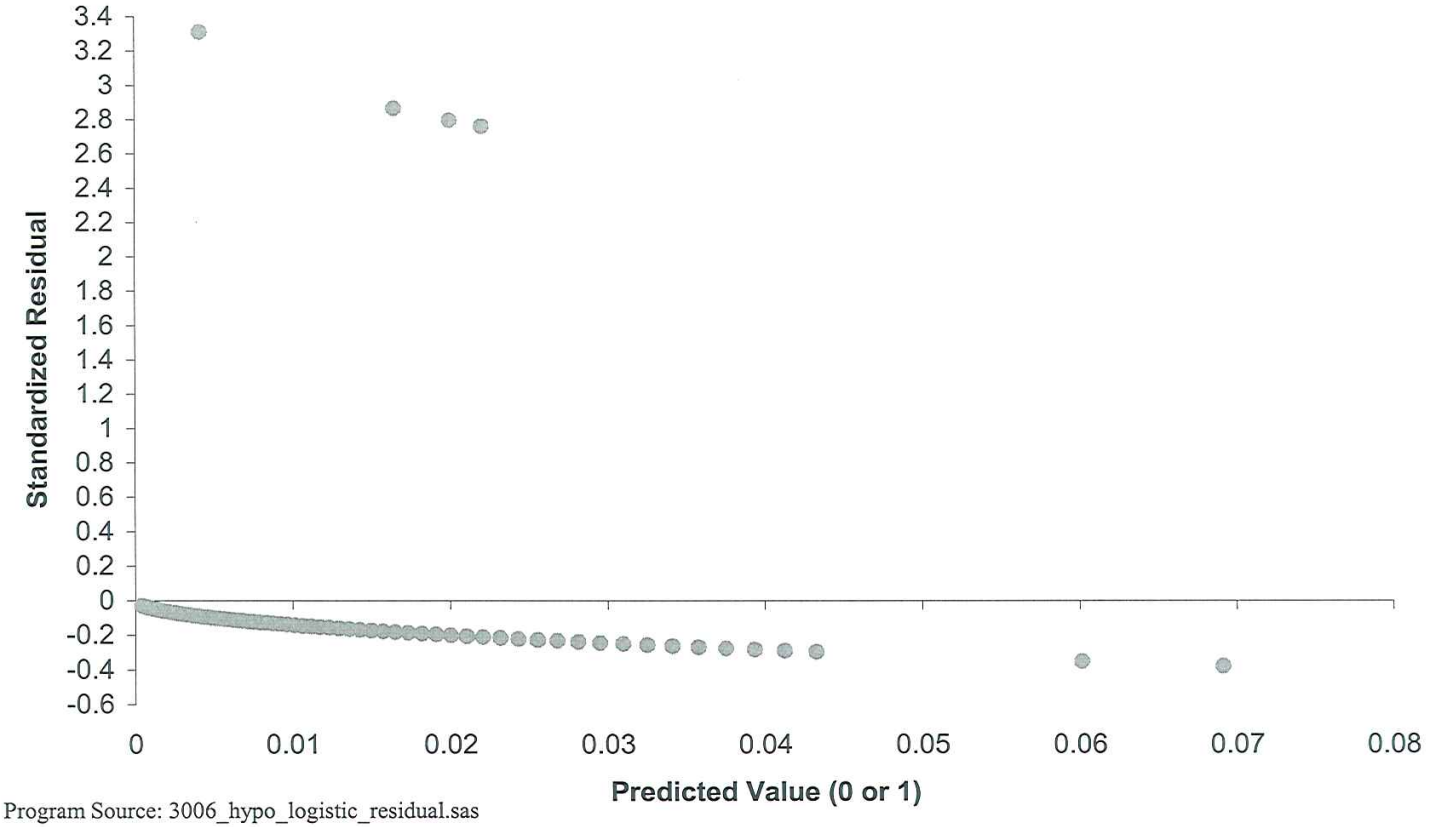

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 70mg/dL) (551 subject with events from a total of 984 subjects), Study-Pools IV-4016
Standardized Residual
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Predicted Value (0 or 1)
Program Source: 4016_hypo_logistic_residual.sas

Table 3. Percentage of patients with at least 1 Hypoglycemia Event over Treatment Period
ITT Population (Actual Treatment)

	Glargine (246)		NPH (255)		Comparison (Glargine over NPH)			P-value		C
Variable	Patients with at least 1 event	Incidence (%)	Patients with at least 1 event	Incidence (%)	Adjusted Odds Ratio	95% LCB	95% UCB	Treatment Effect	Change in HbA1C	Statistics
Day Time Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	4	1.6	0.26	0.03	2.37	0.234	0.170	0.785
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	10	4.1	21	8.2	0.47	0.22	1.02	0.055	0.716	0.606
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	120	48.8	136	53.3	0.85	0.60	1.21	0.369	0.036	0.560
Nocturnal Hypoglycemia										
Severe Hypoglycemia (Definition Given in the Studies)	1	0.4	2	0.8	0.48	0.04	5.43	0.556	0.475	0.650
Symptomatic Hypoglycemia with SMBG<2.0(mmol/l) (36 mg/dl)	11	4.5	12	4.7	0.96	0.41	2.22	0.923	0.592	0.500
Symptomatic Hypoglycemia with SMBG<3.888(mmol/l) (70 mg/dl)	71	28.9	91	35.7	0.75	0.51	1.09	0.130	0.027	0.574

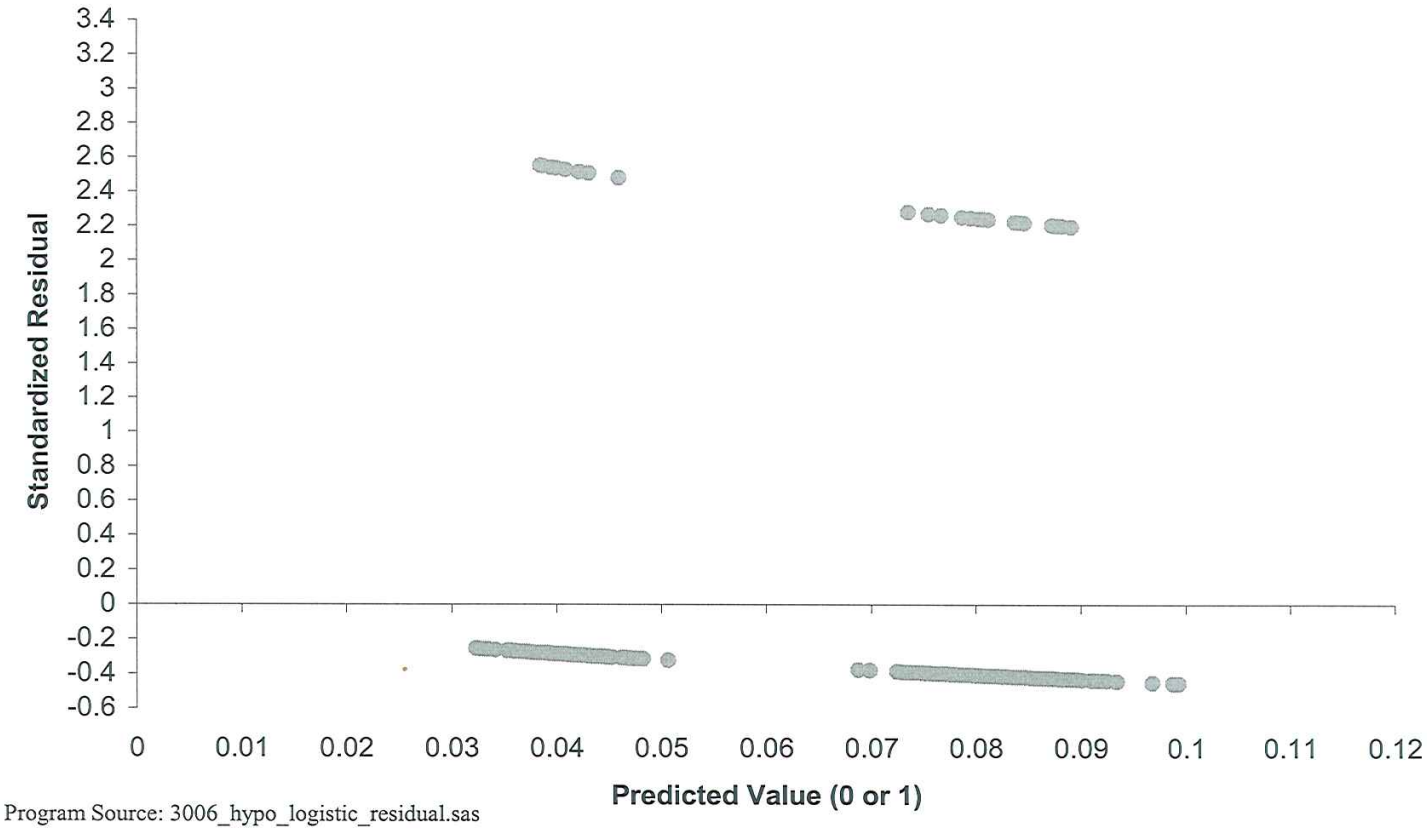
Note:
(1). HbA1C at baseline: last value collected before the first study drug intake
(2). HbA1C at endpoint: last value collected during treatment phase up to 14 days after last study drug intake (LOCF method)
(3). ITT population: Glargine: 259, NPH: 259. Patients who have values in change in HbA1C: Glargine: 246, NPH: 255
(4). 17 patients who have missing values in change in HbA1C were excluded, including 5 patients who have at least one hypoglycemia: 3 Glargine and 2 NPH patients
(5). Adjusted odds ratio, 95% CI, P-values and c-statistics come from Logistic regression adjusted for change in HbA1C

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (Severe Events)
(5 subjects with events from a total of 501 subjects), Study-Pools V-3006



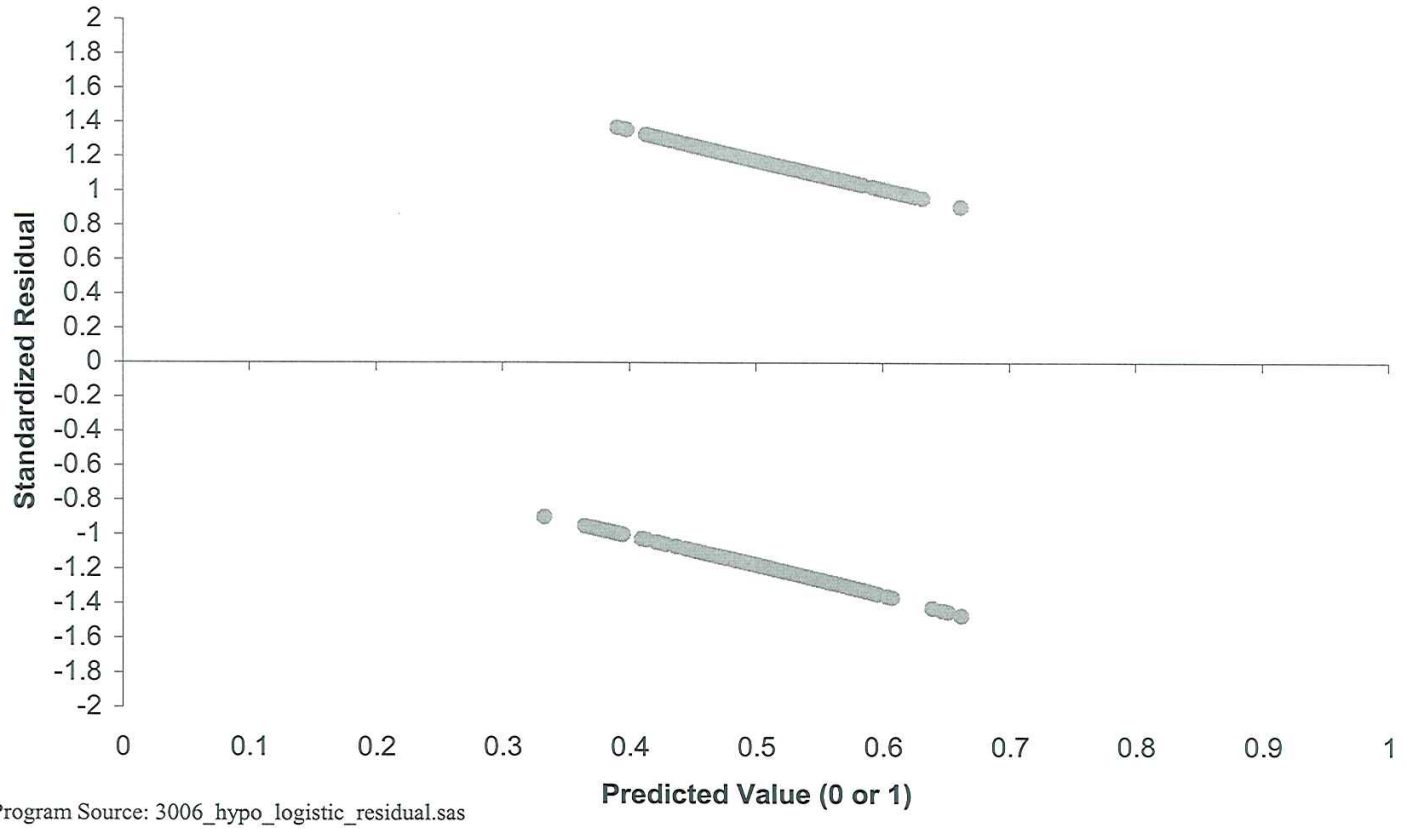
Program Source: 3006_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 36mg/dL)
(31 subjects with events from a total of 501 subjects), Study-Pools V-3006

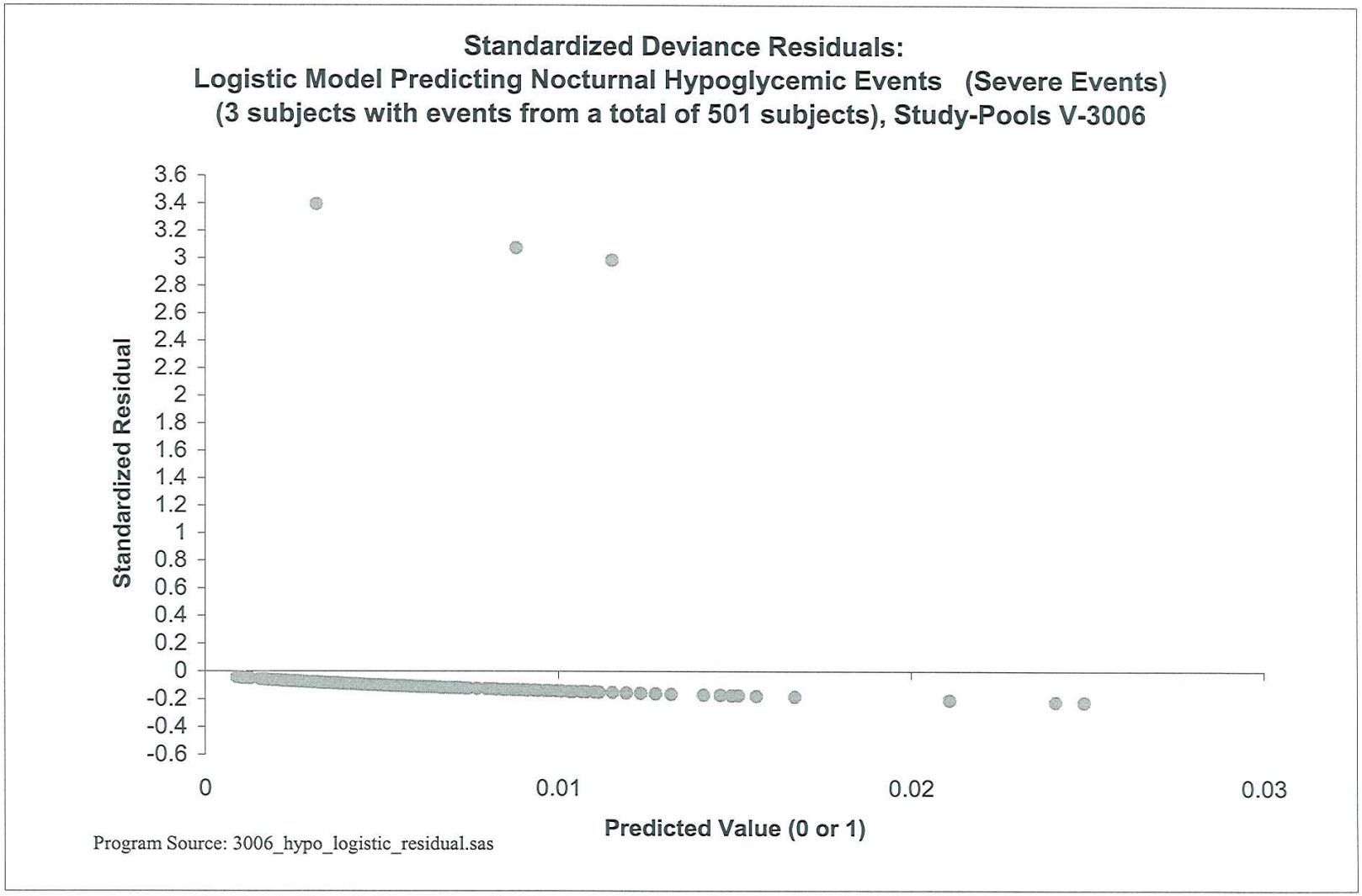


Program Source: 3006_hypo_logistic_residual.sas

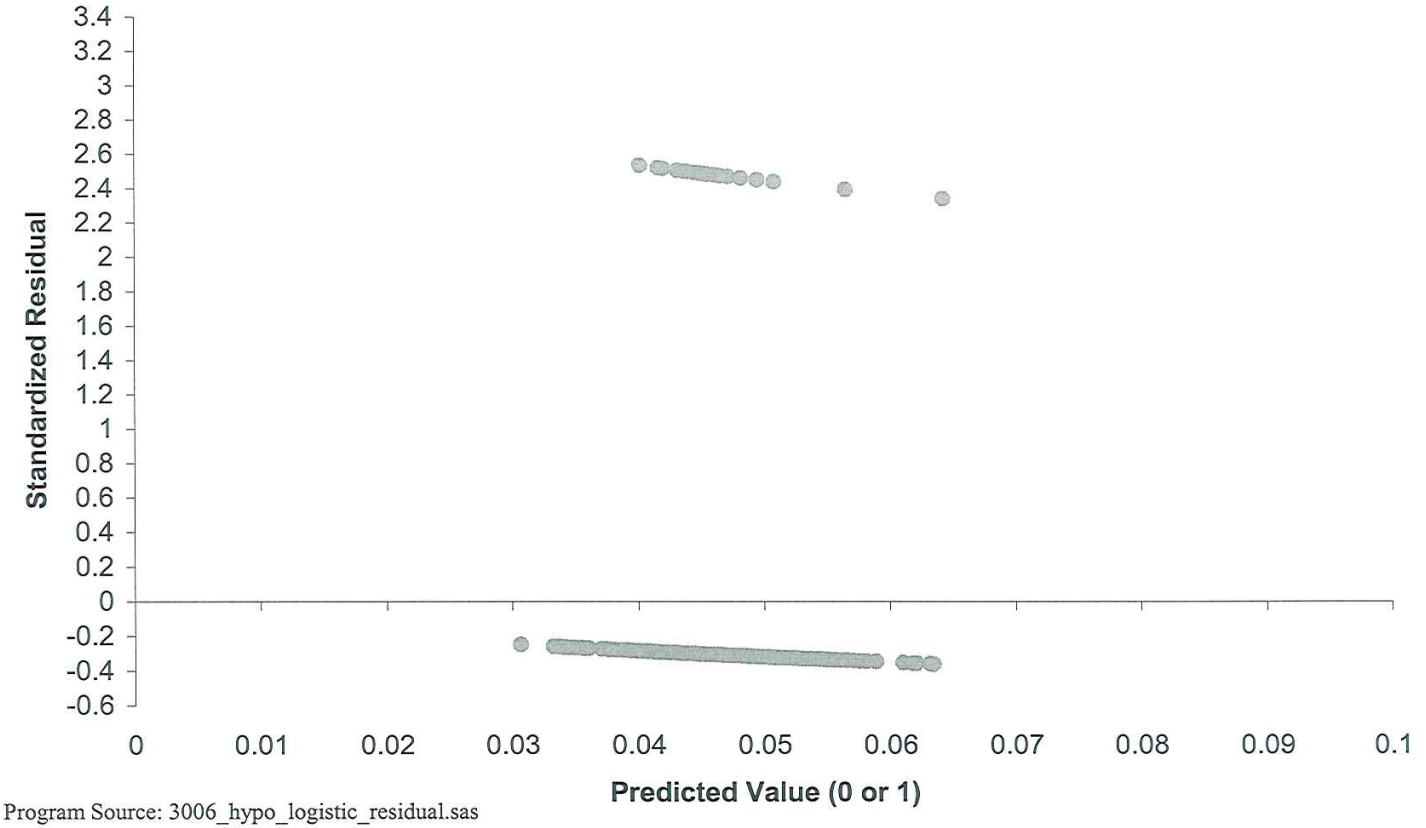
Standardized Deviance Residuals:
Logistic Model Predicting Daytime Hypoglycemic Events (SMBG < 70mg/dL)
(256 subjects with events from a total of 501 subjects), Study-Pools V-3006



Program Source: 3006_hypo_logistic_residual.sas

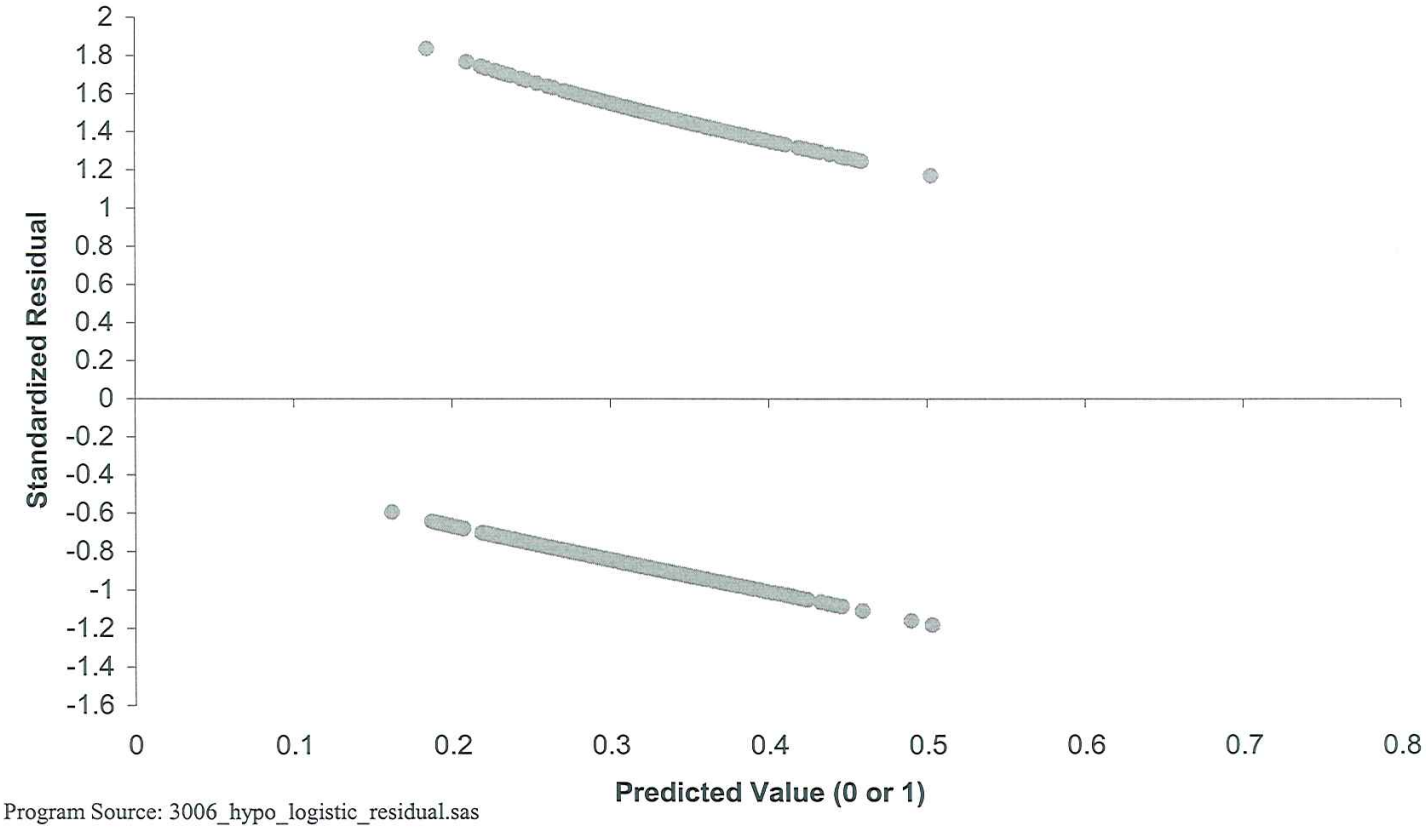

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (Severe Events)
(3 subjects with events from a total of 501 subjects), Study-Pools V-3006
Standardized Residual
3.6
3.4
3.2
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.01
0.02
0.03
Predicted Value (0 or 1)
Program Source: 3006_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 36mg/dL) (23 subjects with events from a total of 501 subjects), Study-Pools V - 3006



Program Source: 3006_hypo_logistic_residual.sas

Standardized Deviance Residuals:
Logistic Model Predicting Nocturnal Hypoglycemic Events (SMBG < 70mg/dL) (162 subjects with events from a total of 501 subjects), Study-Pools V-3006



Program Source: 3006_hypo_logistic_residual.sas

A 1.8 Verband forschender Arzneimittelhersteller e. V.

Autoren:
Stammberger, Ingo, Dr. med.
Thole, Henning
Wahler, Steffen, Dr. med.

Adresse:
Henning Thole
Verband forschender Arzneimittelhersteller e. V.
Hansvogteiplatz 13
10117 Berlin

Verband Forschender
Arzneimittelhersteller e.V.

**Stellungnahme zum Vorbericht
„Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2" - Auftrag A05-03**

Einleitung

Am 18.03.2008 wurde der Vorbericht „Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2" vom Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) veröffentlicht und zur Stellungnahme gestellt. Wir begrüßen die Möglichkeit, mit einer Stellungnahme den Vorbericht zu kommentieren, und sehen bereits jetzt den Bedarf von Änderungen am Vorbericht.

Der vorliegende Vorbericht macht deutlich, dass die Möglichkeit der Klärung durch ein Scoping zu Beginn des Verfahrens viele Fragen hätte klären können, die jetzt von Überarbeitung zu Überarbeitung anstehen und abgearbeitet werden müssen. Die Punkte, die hier vordringlich zur Klärung anstehen, werden im folgenden benannt.

Probleme, die aus der Anhörung des Berichtsplans hervorgehen

Aus Sicht des VFA ist es bei der zuvor durchgeführten Anhörung des Berichtsplans zu schweren formalen Mängeln gekommen, die in der Folge zu einer Verletzung gesetzlicher Forderungen geführt haben und unmittelbar in den Vorbericht wirken.

Nach § 35b Abs. 1 Satz 6 SGB V hat das IQWiG bei den Bewertungsverfahren *hohe Verfahrenstransparenz* und eine *angemessene Beteiligung* der in § 35 Abs. 2 und § 139a Abs. 5 Genannten zu gewährleisten. Ferner spricht § 139a Abs. 5 von einer *Beteiligung in allen wichtigen Abschnitten des Bewertungsverfahrens*. Dabei hatte der Gesetzgeber auch die Beteiligung der Fachkreise und Betroffenen bei der Erstellung der Berichtspläne im Blick.[1]

In der Anhörung des Berichtsplans war ein englischsprachiger Experte eingeladen worden. Nach mündlichen Informationen, die im zeitlichen Zusammenhang mit der Anhörung gegeben wurden, wurde diesem die Bitte, ggf. auch auf eigene Kosten einen Dolmetscher mitzubringen, vom IQWiG abschlägig beschieden. In der Anhörung selber wurde diesem Experten seitens des IQWiG teilweise auch auf deutsch geantwortet, so dass der Experte die Antworten auf seine Fragen nicht verstehen konnte. Überdies sind alle State-

Hausvogteiplatz 13
10117 Berlin
Telefon 030 20604-0
Telefax 030 20604-222
www.vfa.de

Hauptgeschäftsführerin
Cornelia Yzer

[1] Bundestagsdrucksache BT 16/3100, S. 151;
http://dip.bundestag.de/btd/16/031/1603100.pdf (Zugriff 23.04.08, 11:00 Uhr)

ments, die dieser Experte in der Anhörung abgegeben hatte, nicht im Protokoll enthalten, obwohl sie aufgezeichnet wurden.

Damit wurden alle Kommentare, die dieser Experte abgegeben hat, nicht zur Grundlage der Überarbeitung des Berichtes. Die Nachvollziehbarkeit der Kommentare ist überdies auch nicht gegeben.

Die Durchführung der Anhörung und der darauf aufgebaute Bericht muss in bezug auf diesen Punkt als grob Fehlerhaft bezeichnet werden. Die Überarbeitung des Berichtsplan muss unter Einbezug aller vorgebrachten Kommentare erfolgen. Die Statements des englischsprachigen Experten sollten veröffentlicht und in den Bericht mit einbezogen werden.

Die im Vorbericht enthaltene Darstellung, dass es bei der Bearbeitung des Projektes zu keinerlei Änderungsbedarf am Bericht gekommen sei, ist falsch und muss korrigiert werden. Die für die Notwendigkeit und die Nachvollziehbarkeit der vorgenommenen Änderungen abzugebenden Informationen sind offenzulegen.

Wertende Aussagen im Berichtsplan

In der mündlichen Anhörung zum Berichtsplan wurde (protokolliert auf S. 68-71 der Dokumentation der Anhörung) das Thema der wertenden Aussagen zur Mitogenität im Berichtsplan angesprochen.

Die im Vorbericht enthaltene Darstellung ist unverändert nicht vollständig und dadurch wertend und im Ergebnis irreführend. Die Ergebnisse sind, wie in der Anhörung diskutiert, von der Darstellung im Berichtsplan und Vorbericht abweichend.

Informationen zu diesem Thema sollten so dargestellt werden, dass die Informationen vollständig dargelegt werden, d.h. die gewählten Zitate sollte nicht gekürzt abgebildet werden. In den zitierten Dokumenten sind abschließende Wertungen zum Thema Mitogenität enthalten, die das IQWiG in stark wertender Weise bewusst nicht anführt. Dies muss als BIAS betrachtet werden.

Nichteinhaltung methodischer Standards

Der Vorbericht verstößt mit der Änderung im laufenden Nutzenbewertungsverfahren, die durch das Amendment zum Berichtsplan erfolgte, gegen methodische Standards[2].

2 Die hierzu anzuführende Literatur wurde bereits in der Stellungnahme zum Berichtsplan zitiert.

Besonders auffällig ist hier, dass im zeitlichen Zusammenhang zu diesem Verfahren unter Beteiligung von IQWiG-Mitarbeiter Arbeiten im gleichen Themenfeld in Form eines Cochrane-Reviews veröffentlicht wurden, die zu von diesem Bericht abweichenden Ergebnissen kommen. Dies stellt formal einen Interessenkonflikt dar, und führt zu der offensichtlichen Frage, in welchen Ausmaß die hier vorliegende Bewertung datengetrieben ist – die IQWiG-Mitarbeiter müssen ja Kenntnis von den Analysen und Ergebnissen des Cochrane-Reviews gehabt haben, und dies zu einem Zeitpunkt, als die eigene Analyse noch im Stadium Berichtsplan > Vorbericht war.

Dies ist insgesamt umso erstaunlicher, als das Cochrane-Review zu deutlich abweichenden Ergebnissen kommt.

Inkonsistenzen zwischen den verschiedenen Projekten im Themenbereich „Diabetes"

Es liegen relevante Inkonsistenzen zwischen den verschiedenen Berichten im Themenbereich „Diabetes" vor, die im Ergebnis eine ungleiche Bewertung zwischen den verschiedenen Indikationen darstellen und eine konsistente und gleichberechtigte Entscheidung des G-BA, der die Berichte des IQWiG als Empfehlung zu berücksichtigen hat, erschweren oder unmöglich machen.

In Bezug auf den vorliegenden Berichtsplan muss kritisiert werden, dass nur auf schwere Unterzuckerungen fokussiert wird, wohingegen im Verfahren A05-01 auch andere Formen der Unterzuckerung Berücksichtigung finden. Besonders verwirrend ist die Begründung des IQWiG, dass dies aus Gründen des Studiendesigns erfolgt sei, d.h. nicht schwerwiegende Hypoglykämien nicht sicher hätten ermittelt werden können.

Hier liegt ein offensichtlicher Widerspruch zu den Anforderungen vor, die das IQWiG selber in seinem Methodenpapier definiert hat.

Bestverfügbare Evidenz statt nur RCTs

Im Gegensatz zur bisherigen Vorgehensweise des IQWiG definiert das Konzept der Evidenzbasierten Medizin eine Betrachtung aller Evidenzklassen im Sinne der „best available Evidence" und damit ausdrücklich auch der Evidenzklassen Ia und II - V. Das IQWiG blendet aber sowohl existierende Systematische Übersichtsarbeiten als auch die Evidenz der Level II bis V regelhaft aus.

RCTs bergen systematische Probleme im Hinblick auf die externe Validität. In Bezug auf externe Validität sind RCTs nicht immer die „best available evidence".

Es ist nicht nachvollziehbar, warum nur RCTs berücksichtigt wurden. Selbst das noch gültige - wenn auch nicht gesetzeskonforme - Methodenpapier 2.0 verweist darauf, dass alle Evidenzklassen in die Bewertung einfließen. Dies ist aber von vornherein ausgeschlossen, wenn die Literatur nicht einmal identifiziert wird, weil die vom IQWiG vorgenommene theoretische Festlegung zum Ausschluss aller anderen Evidenz außer RCTs führt.

Ausdrücklich soll bei fehlenden Daten oder der fehlenden Realisierbarkeit entsprechende Daten in einer bestimmten Studienform zu erheben, immer nächstrangige Evidenz einbezogen werden. Auch Meta-Analysen wurden vom IQWiG nicht berücksichtigt – noch schwerer wirkt jedoch, dass auch Studien, die in anderen Systematischen Reviews berücksichtigt wurden, vom IQWiG nicht einbezogen wurden.

Die im Berichtsplan auf S. 11 unter Punkt 4.1.4 gegebene Erklärung, dass RCTs als einzig relevante Studienform einbezogen wurden, weil die Abbildung der Fragestellung in ihnen möglich gewesen wäre, ignoriert damit bewusst möglicherweise vorliegende Evidenz für genau den Fall, dass RCTs für die Frage oder für einen Aspekt der Frage nicht vorliegen.

Grundlage der Recherche und damit der Bewertung sollte keine theoretische Annahme sein, sondern ein praktischer Ansatz zur Ermittlung der tatsächlichen vorhandenen Evidenz.

Vor dem Hintergrund dieser methodischen Festlegungen ist es nicht nachvollziehbar, dass Studien ausgeschlossen wurden, erst Recht nicht unter unzutreffenden Gründen wie z.B. im Fall der Studie von Yokoyama. Die im Anhang B auf S. 279 des Vorberichts gegebene Begründung für den Ausschluss ist nicht korrekt, die Studie hat einen Vergleich Insulin Glargin vs. NPH-Insulin vorgenommen.

Die Ein- und Ausschlusskriterien sollten durch aktive Beteiligung der im GKV-WSG genannten Gruppen definiert werden. Die Auswahl der Studien muss sich daran orientieren, ob die entsprechende (Teil-) Frage aus der jeweiligen Evidenz beantwortet werden kann, und nicht daran, ob sie theoretisch aus Informationen einer bestimmten Evidenzstufe beantwortet werden könnte.

Verband Forschender
Arzneimittelhersteller e.V.

Zusammenfassende Wertung

Mit der Berücksichtigung einiger der Kommentare zum Berichtsplan wurden viele positive Änderungen in dieses Projekt eingeführt. An zentralen Stellen gibt es darüber hinaus jedoch unverändert Verbesserungsbedarf. Es wäre wünschenswert, wenn durch vollständige Berücksichtigung der Kommentare aus den Stellungnahmen die Qualität und Aussage der Berichte umfassend verbessert werden könnte.

Der Kernpunkt der notwendigen Veränderungen ist jedoch die Transparenz im Verfahren.

Es werden weiterhin keine Dokumente veröffentlicht, die von den extern beauftragten Gutachtern erstellt wurden.

Damit bleibt ein ganz wesentlicher Punkt im unklaren, nämlich, an welcher Stelle Ergebnisse der Analyse der beauftragten externen Experten vorliegen, und an welchen Stellen Festlegungen des IQWiG vorliegen. Im Ergebnis können die zur Kommentierung vorgelegten Dokumente Werturteile und Änderungen in den Aussagen enthalten, die als solche nicht kenntlich gemacht sind.

Der mögliche Grad einer Abweichung in den Aussagen bleibt somit vollständig unklar. Dass dies nicht eine Forderung oder Ansicht nur des VFA ist, kann daran bemessen werden, dass Defizite an der Transparenz und der Beteiligung im letzten Jahr von verschiedenen Autoren, so z.B. den BMG-Experten der beiden Fachtagungen, formuliert wurden.

Auffällig ist, dass der IQWiG-Vorbericht zu erheblich schlechteren Einschätzungen der Bedeutung der betreffenden Medikamente kommt als die international hierzu verfügbare Literatur. Die zuvor vorgebrachten Argumente weisen auf die Sachverhalte hin, die hierfür ursächlich sind.

Die zeitnahe Veröffentlichung aller notwendigen Hintergrundinformationen, u.a. die Berichte der extern beauftragten Experten könnte, zusammen mit eine verlängerten und damit angemessenen Kommentierungsfrist, die Möglichkeiten zur Abgabe einer Stellungnahme verbessern.

Berlin, 25.04.08

A 2 Stellungnahmen von Privatpersonen

A 2.1 Bradley Clare, Prof., PhD

Adresse:
Royal Holloway, University of London
Egham Hill
Egham, Surrey, TW20OEX

Vorbericht ‘Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2’

Auftragsnummer: A05-03

Stellungnahme von: Bradley, Clare; Professor of Health Psychology, Royal Holloway, University of London; Developer of the DTSQs and c, W-BQ22 discussed in IQWiG’s preliminary report and other patient reported outcome measures.

Kontaktdaten:
Clare Bradley, PhD
Professor of Health Psychology,
Department of Psychology,
Royal Holloway, University of London,
Egham Hill, Egham, Surrey,
TW20 0EX, UK
Tel: +44 1784 443708 (direct) 443714 (admin)
Fax: +44 1784 471168
Email: c.bradley@rhul.ac.uk

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien

Obwohl der Vorbericht ein Kapitel zur Entwicklung des 22-Item Well-Being Questionnaires / Fragebogens zum Wohlbefinden erwähnt (ref 68), sind die Angaben zu Verlag und Ort der Veröffentlichung nicht korrekt.
Der Vorbericht verweist nicht auf die Literatur zur Entwicklung des Diabetes Satisfaction Questionnaires (DTSQ, Fragebogen zur Zufriedenheit mit der Diabetes-Behandlung) in seiner Status- und Veränderungsversion.
Für eines der anderen in den beschriebenen/angegebenen Studien verwendeten Instrumente (SF-36) wird eine Quellenangabe angegeben, nicht jedoch für das andere Instrument (ITSQ). Ich gebe diese unten an.

DTSQs

Bradley C. The Diabetes Treatment Satisfaction Questionnaire: DTSQ. In Bradley C (Ed) (1994) *Handbook of Psychology and Diabetes: a guide to psychological measurement in diabetes research and practice*. Chur, Switzerland: Harwood Academic Publishers.

Bradley C and Lewis KS (1990) Measures of psychological well-being and treatment satisfaction developed from the responses of people with tablet-treated diabetes. *Diabetic Medicine,* **7**, 445-451.

DTSQc

Bradley C, Plowright R, Stewart J, Valentine J and Witthaus E (2007) The Diabetes Treatment Satisfaction Questionnaire change version (DTSQc) evaluated in insulin glargine trials shows greater responsiveness to improvements than the original DTSQ. *Health and Quality of Life Outcomes* **5** (5) 57, http://www.hqlo.com/content/5/1/57.

Howorka K, Pumprla J, Schlusche C, Wagner-Nosiska D, Schabmann A and Bradley C (2000) Dealing with ceiling baseline treatment satisfaction level in patients with diabetes under flexible, functional insulin treatment: assessment of improvements in treatment satisfaction with a new insulin analogue. *Quality of Life Research*, **9**: 915-930.

ITSQ

Anderson RT, Scovland SE, Marrero D et al. Development and validation of the Insulin Treatment Satisfaction Questionnaire. Clin Ther 2004; **26:** 565-578.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien.

Abschnitt 5.3.5 Gesundheitsbezogene Lebensqualität und Behandlungszufriedenheit

Der für gesundheitsbezogene Lebensqualität gegebenen Definition kann ich nicht zustimmen. Sie scheint aus einem deutschen Lexikon zu kommen, wird jedoch präsentiert, als wäre sie allgemein akzeptiert.
Der SF-36 ist der einzige der sechs verwendeten Fragebögen, der als die Definition erfüllend erachtet wird und als Instrument zur Messung der gesundheitsbezogenen Lebensqualität angesehen wird. Ich wende ein, dass der SF-36 nicht Lebensqualität, sondern vielmehr Gesundheitsstatus misst und dass es sehr wichtig ist, zwischen Instrumenten zur Messung der Lebensqualität und des Gesundheitsstatus' zu unterscheiden (Bradley, 2001), wenn wir irreführende Schlussfolgerungen in Bezug auf die Daten vermeiden möchten. Der Fragebogen zum Wohlbefinden (W-BQ22) wird in der englischen Übersetzung des Vorberichts als ‚illness-specific', (‚krankheitsspezifisch' im deutschen Original) bezeichnet, obwohl es tatsächlich ein generisches Instrument ist, das nicht diabetesspezifisch ist, auch wenn es an Personen mit Diabetes validiert worden ist.
Sowohl vom W-BQ22 als auch vom Diabetes Health Profile (DHP-18) wird behauptet, dass sie Aspekte krankheitsbezogener Lebensqualität messen, obwohl es wesentlich zutreffender wäre, alle diese Instrumente als Instrumente zur Messung von patientenberichteten Endpunkten (PRO) anzusehen und keines von ihnen als Instrument zur Messung von Lebensqualität.
In Tabelle 53 werden die Richtungen der Skalen fälschlicherweise als „niedrigere Werte bedeuten Verbesserung" oder „höhere Werte bedeuten Verbesserung" beschrieben, es wird also eine Veränderung impliziert, wo tatsächlich keines dieser Instrumente Veränderung misst. Wo es für den

Gesamtscore des W-BQs heißt, dass „höhere Werte Verbesserung bedeuten“, sollte es heißen, dass „höhere Werte besseres Wohlbefinden“ bedeuten. Soweit ich weiß, erscheint dieser Fehler ebenfalls im deutschen Original und ist nicht nur ein Problem der Übersetzung des Vorberichts ins Englische, welche ich benutze.
Ähnliche Probleme zeigen sich in Tabelle 58, welche die Instrumente zur Behandlungszufriedenheit beschreibt. Hier wird kein Unterschied zwischen der Bedeutung der Scores des DTSQs und DTSQc gemacht, da die Richtung der Scores des DTSQs beschrieben wurde, als wäre es ein Instrument zur Messung von Veränderung, was es nicht ist. Es ist möglich, dass die unangemessene Vermischung von Daten des DTSQs und DTSQc in von IQWiG durchgeführten Meta-Analysen teilweise auf fehlendem Verständnis hinsichtlich der Tatsache beruht, dass der DTSQs und DTSQc zwei unterschiedliche Fragebögen sind, die unterschiedliche Phänomene messen. Der DTSQs misst Zufriedenheit mit einer Behandlung während der letzten Wochen, während der DTSQc die Zufriedenheit mit einer gegenwärtigen Behandlung im Vergleich zur Zufriedenheit mit einer vorherigen Behandlung misst. Die Antwortmöglichkeiten sind verschieden und die Auswertung ist verschieden und ich konnte keine Rechtfertigung dafür erkennen, dass diese beiden Instrumente in einer einzigen Meta-Analyse zusammengefasst werden. Obwohl ich Vergleiche zwischen Scores des DTSQc und Scores des DTSQs, die durch Subtraktion der Baseline-Scores des DTSQs von den Endpunkt Scores des DTSQs entstehen, berichtet habe, ist es notwendig, zuerst die Skala so anzupassen, dass eine vergleichbare Spannweite der Werte entsteht (Bradley et al 2007). In unserer Veröffentlichung von 2007 haben wir dargelegt, dass der DTSQc Deckeneffekte überwinden kann, die mit dem DTSQs und anderen Instrumenten, die den Status von Behandlungszufriedenheit messen, häufig auftreten. Es wäre angemessen, das Niveau der Zufriedenheit bei Baseline in Betracht zu ziehen und anzuerkennen, dass es schwierig ist, Verbesserungen aufzuzeigen, wenn bei Statusmessungen der Zufriedenheit Deckeneffekte auftreten, wie etwa beim DTSQs und ITSQ. Die Ergebnisse des DTSQc sollten stärkere Berücksichtigung erhalten, wenn bei Baseline die Zufriedenheit bereits recht hoch ist.

Fallzahlenplanung
Durch den Vorbericht hinweg wird die Aufmerksamkeit auf einen Mangel an „Fallzahlenplanung“ gelenkt, als würde dies die Qualität der Ergebnisse in Zweifel ziehen. Meine Erfahrung ist, dass Studien, deren Stichproben geplant wurden, um signifikante Unterschiede in HbA1c zu ermitteln, bei weitem die Stichprobengröße überschreiten, die hinreichend ist, um signifikante Unterschiede im DTSQ zu ermitteln und gerade groß genug sind, um Unterschiede im W-BQ zu ermitteln. Obwohl ich verstehe, dass eine fehlende Fallzahlenplanung problematisch ist, wenn Ergebnisse insignifikant sind und wir womöglich einen tatsächlichen Effekt übersehen, ist unklar, warum fehlende Fallzahlenplanung IQWiG Sorgen bereiten, wenn die Ergebnisse signifikant sind.

Verblindung
Im Vorbericht wird durchgehen suggeriert, dass das offene Design die Stärke

der Ergebnisse bedeutend einschränkt. Dieser Annahme muss ich dringend widersprechen. Die ökologische Validität von Studien zu patientenberichteten Endpunkten, im besonderen Behandlungszufriedenheit, kann durch Verblindung ernsthaft beeinträchtigt werden, insbesondere, wenn zusätzlich zu aktiven Behandlungen Scheinbehandlungen durchgeführt werden und die vom Patienten erlebte Behandlung beschwerlicher oder lästiger wird als jede der Behandlungen in der normalen klinischen Realität außerhalb der Studie wäre. Einer der Vorteile von Glargin ist, dass es nur eine Injektion pro Tag umfasst, während NPH-Insulin häufiger zwei Injektionen erfordert. Es ist vollkommen angemessen, dass Patientenberichte zu ihrer Zufriedenheit mit einer Behandlung in Bezug auf die Realität der Behandlung abgegeben werden und nicht verzerrt werden durch Versuche, ihr Bewusstsein, welche Behandlung sie nutzen, zu verschleiern, was effektiv ihre Erfahrung der Behandlung verändert.

Patienten, die nicht in die Auswertung mit einbezogen werden

Fehlende Fragebögen
IQWiG war beunruhigt aufgrund der Anzahl nicht in die Analyse einbezogener Patienten. Ich kann bestätigen, dass manche Länder den DTSQs und/oder DTSQs nicht benutzten, da die benötigte Sprachversion zu dieser Zeit noch nicht linguistisch validiert war (linguistische Validierung des DTSQ in einer neuen Sprache dauert 4-5 Monate und es ist nicht immer möglich, dies vor Beginn einer Studie zu erreichen). Patienten, die aus solchen Gründen ausgeschlossen wurden, sollten kein Problem für die Validität der Daten darstellen. Studienabbrecher sind eine andere Sache und müssen sorgfältig geprüft werden.

Fehlende Patienten
Die Signifikanz unterschiedlich vieler Studienabbrechern in verschiedenen Behandlungsgruppen muss mit Blick auf die Behandlung, die vor Beginn der Studie verwendet wurde, betrachtet werden.
Falls Patienten vor Beginn der Studie NPH-Insulin benutzt haben und dann randomisiert werden, um weiterhin NPH-Insulin zu benutzen oder zu einem Insulin-Analogon zu wechseln, ist es nicht überraschend, wenn unter den Patienten, die auf das neue Insulin randomisiert wurden, eher mehr Patienten aufgrund ungünstiger Ereignisse abbrechen. Diejenigen, die auf NPH-Insulin randomisiert wurden, werden aufgrund von Enttäuschung, nicht auf die neue Behandlung randomisiert worden zu sein, mit einer höheren Wahrscheinlichkeit abbrechen.
Die Studien, die Detemir und Glargin vergleichen, sind besonders interessant, um die Anzahl der Abbrecher in jeder Gruppe zu untersuchen, da diese Studien wahrscheinlich frei von Enttäuschungseffekten sind (wenn jeder ein Insulinanalogon ausprobieren kann, welches sie vorher nicht benutzt haben) und jeder einen Wechsel der Behandlung erfährt, so dass Teilnehmer, die aufgrund ungünstiger Ereignisse abbrechen, zusammen mit Zufriedenheitsscores sinnvoll zwischen den Behandlungsgruppen verglichen werden können. Die unpublizierten Studien, die Detemir und Glargin vergleichen, sind mir nicht bekannt, ich möchte jedoch anmerken, dass es wichtig ist zu prüfen, ob beide Gruppen den Vorteil von Insulin-Pens hatten.

An anderer Stelle wurde darauf hingewiesen (Bradley und Gilbride 2008 in Druck), dass, wenn Insulin einmal mit einem Pen gegeben wird, was Patienten üblicherweise bevorzugen, und einmal mit Spritze und Ampulle, das mit Spritze und Ampulle gegebene benachteiligt ist, besonders, wenn der ITSQ benutzt wird, der eine Subskala zur Messung der Zufriedenheit mit der Methode der Zufuhr des Insulins hat.

Minimal Important Difference (MID)

Obwohl die (problematische) Meta-Analyse einen statistisch signifikanten Unterschied zugunsten Insulin Glargin im Vergleich zu NPH-Insulin aufgezeigt hat, wurde der Effekt als „klinisch nicht relevant" abgetan, anscheinend da für den DTSQs und DTSQc kein MID definiert wurde. Ich habe an anderer Stelle erörtert (Bradley, 2006), dass es nicht angebracht ist, für den DTSQs oder c eine MID zu spezifizieren, da diese Endpunkte für sich genommen wichtig sind: Sie sind keine intermediären Endpunkte, die nur aufgrund ihrer Auswirkung auf andere Endpunkte wichtig sind (z.B. HbA1c ist vor allem wegen seiner Assoziation mit mikrovaskulären Komplikationen des Diabetes von klinischer Bedeutung. Für HbA1c ist es daher sinnvoller, eine MID festzulegen). Für den DTSQ spiegelt eine statistisch signifikante Veränderung eine klinisch signifikante Veränderung wider. Es ist Sache des Patienten und nicht des Instrumentenentwicklers oder des Forschers, festzulegen, was eine klinisch signifikante Verbesserung ist. Wenn Patienten verbesserte Zufriedenheit berichten und diese Verbesserung statistisch signifikant ist, reicht dies aus, um zu akzeptieren, dass die Zufriedenheit sich systematisch in einer Art und Weise verbessert hat, die nicht auf zufällige Faktoren zurückgeführt werden kann. Wie für jede Variable werden Kennwerte der Effektgrößen uns zeigen, wie groß der Effekt ist, aber auch ein kleiner Effekt kann wertvoll sein und sollte nicht von der Hand gewiesen werden. Für den DTSQ haben wir sowohl den DTSQc, um vom Patienten eine direkte Schätzung der Veränderung zu erfahren, als auch den DTSQc, der Messungen der Zufriedenheit zu bestimmten Zeitpunkten ermöglicht. Die Tatsache, dass der DTSQc, der nicht durch Deckeneffekte eingeschränkt ist, üblicherweise stärkere Effekte zeigt als die Differenz von Endpunkt und Baseline Scores des DTSQs, bestätigt, dass die vom DTSQs gezeigten Verbesserungen der Zufriedenheit die von Patienten erlebten Verbesserungen unterschätzen und nicht überschätzen.

Zusammenfassend bin ich besorgt, dass IQWiG auf der Basis fehlerhafter Annahmen hinsichtlich der Notwendigkeit, MID zu definieren, tatsächlich vorhandene Vorteile von Glargin im Vergleich zu NPH-Insulin zurückweisen könnte und dass ihre Meta-Analyse von Glargin-Studien, die den DTSQs und/oder c benutzten, auf unangemessene Weise Daten der beiden Fragebögen kombiniert haben könnte. Während es sehr wichtig ist, die Anzahl der Studienabbrecher zu beachten und Gründe für den Abbruch zu erörtern, und dies ist besonders wichtig im Hinblick auf die Interpretation der Ergebnisse zur Behandlungszufriedenheit, gibt es keinen Grund zur Sorge wegen fehlender Daten zur Behandlungszufriedenheit von bestimmten Ländern, die nur entstehen, da der Fragebogen nicht in der entsprechenden Sprache zur Verfügung stand.

Fehlende Verblindung ist ein Vorteil in Studien zur Behandlungszufriedenheit und es ist nicht nötig, sich über die Angemessenheit der Stichprobengrößen zur Ermittlung von Veränderungen im DTSQ Sorgen zu machen. Ich hoffe, dass IQWiG sich in der Lage sieht, eine positivere Sicht auf diese Daten anzunehmen und ich freue mich, diese Aspekte mit den Beteiligten zu diskutieren.

References

Bradley C (2001) Importance of differentiating health status from quality of life. *The Lancet*, **357**, 7-8. http://eprints.rhul.ac.uk/502

Bradley C (2006) Feedback on the FDA's February 2006 draft guidance on Patient Reported Outcome (PRO) measures from a developer of PRO measures. *Health and Quality of Life Outcomes* **4** (1) 78, http://www.hqlo.com/content/4/1/78.

Bradley C and Gilbride CJB (2008) Improving treatment satisfaction and other patient-reported outcomes in people with Type 2 diabetes: the role of once-daily insulin glargine *Diabetes, Obesity and Metabolism;* **10:** (Suppl 1), 50-65 (in press).

Bradley C, Plowright R, Stewart J, Valentine J and Witthaus E (2007) The Diabetes Treatment Satisfaction Questionnaire change version (DTSQc) evaluated in insulin glargine trials shows greater responsiveness to improvements than the original DTSQ. *Health and Quality of Life Outcomes* **5** (5) 57, http://www.hqlo.com/content/5/1/57.

I have translated the comments of my supervisor, Professor Clare Bradley, from English to German.
Ich habe die Stellungnahme meiner Supervisorin, Professor Clare Bradley aus dem Englischen ins Deutsche übersetzt.

Leonie Brose, PhD student

I understand that this German version faithfully translates my comments originally made in English.
Ich nehme an, dass diese deutsche Version getreu meiner im Original auf Englisch gemachten Stellungnahme ist.

Professor Clare Bradley

A 2.2 Bramlage, Peter, PD Dr.

Adresse:
Menzelstraße 21
15831 Mahlow

PRIV.–DOZ. DR. PETER BRAMLAGE

MENZELSTR. 21
15831 MAHLOW
GERMANY

PHONE: +49–3379–3147890
FAX: +49–3379–3147892
E–MAIL: peter.bramlage@gmx.de

Institut für Qualität und Wirtschaftlichkeit
im Gesundheitswesen (IQWiG)

Stellungnahme zum Vorbericht A05-03

Herr Prof. Dr. med. Sawicki

Dillenburger Strasse 27

D-51105 Köln

Vorab per email: berichte@iqwig.de

Mahlow, 21. April 2008

Stellungnahmeverfahren A05-03

Sehr geehrter Herr Professor Sawicki,

sehr geehrter Herr Dr. Kaiser,

Bitte finden Sie in der Anlage meine Stellungnahme zum Vorbericht *„Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2“*. Gerne kann ich die angesprochenen Punkte in einem eventuellen Anhörungsverfahren näher erläutern.

Mit freundlichen Grüßen,

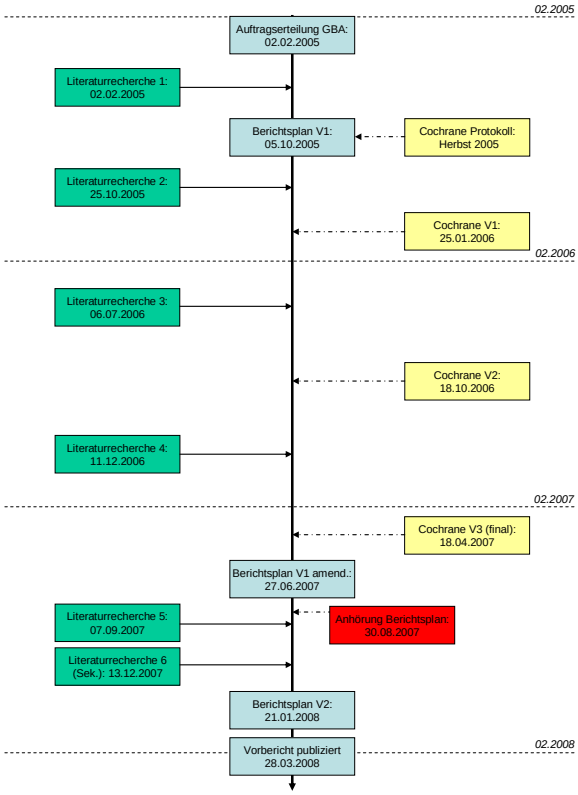
Priv.-Doz. Dr. Peter Bramlage

Stellungnahme zum Vorbericht „Langwirksame Insulinanaloga zur Behandlung des Typ-2 Diabetes mellitus

Trennung von Berichtsplanerstellung und Ergebnissichtung

Meta-Analysen sind definiert als eine systematische, zusammenfassende Analyse einer größeren Zahl einzelner Studienergebnisse mit dem Ziel eines integrierten Gesamtergebnisses. Meta-Analysen sind Grundlage für die Nutzenbewertung des IQWiG. Prinzipiell werden in die vom IQWiG durchgeführten Meta-Analysen ausschließlich prospektive randomisierte, kontrollierte Studien eingeschlossen. Zwar wird diese Restriktion zu Recht häufig kontrovers diskutiert, RCTs bieten aber mit der größten Ergebnissicherheit im Sinne einer hohen internen Validität. Gründe dafür liegen im prospektiven Design, der Randomisierung zur Gleichverteilung von bekannten und unbekannten Einflussfaktoren, in der Kontrolle (Vergleichsarm), der Verblindung (fakultativ) und nicht zuletzt auch im strikten Studiendesign. RCTs genügen dabei besonderen Qualitätsanforderungen: Die Protokolle werden im Vorfeld festgelegt, der erste Patient erst nach Abschluss des Studienprotokolls eingeschlossen und Amendments sind nur möglich, wenn sie gut begründet sind und solange die Verblindung nicht aufgehoben ist.

Kriterien guten Studiendesigns aus RCTs lassen sich auch auf Meta-Analysen übertragen. So sind ein im Vorfeld festgelegtes alle Auswertungsschritte eindeutig definierendes Studienprotokoll, eine zeitlich danach erfolgende Literaturrecherche, eine von Ergebnissen unabhängige Bewertung und eine Sensitivitätsanalyse für die ein- und ausgeschlossenen Studien wichtiges Qualitätsmerkmal für eine objektive Meta-Analyse. Veränderungen des Protokolls erhöhen das Verzerrungspotential und sollten restriktiv und nur in begründeten Ausnahmefällen erfolgen und deren Auswirkungen sind in Sensitivitätsanalysen zu prüfen.



Der zeitliche Ablauf bei der Erstellung des vorliegenden Vorberichts zu langwirksamen Insulinanaloga beim Typ-2 Diabetes ist in der nebenstehenden Abbildung skizziert. Wie aus dem von der Grazer Arbeitsgruppe maßgeblich erstellten Vorbericht und dem unter finanzieller Unterstützung des IQWiG zeitgleich in der gleichen Arbeitsgruppe erstellten Cochrane Review (Horvath, (10)) hervorgeht, ist die Erstellung des Berichtsplans zeitlich eng mit einer Sichtung und Bewertung der zur Verfügung stehenden Literaturdaten verbunden.

Wie aus der Projektübersicht auf der Internetseite der Grazer Arbeitsgruppe hervorgeht (13), wurde der vorläufige Vorbericht bereits vor dem 27.02.2007 (Datum des Copyright der Seite) fertig gestellt. Hinweisend darauf ist auch, dass auf der Seite der Cochrane Review als kurz vor der Veröffentlichung stehend bezeichnet wird, während der Ergebnisbericht der Arbeit

(vorläufiger Vorbericht) als an das IQWiG übermittelt benannt war. Diese Übermittlung hat offensichtlich noch vor dem Amendment 1 des Berichtsplans stattgefunden.

Die Kritikpunkte an diesem Vorgehen sind wie folgt:

1) Literaturrecherche 1 erfolgt schon vor Vorlage der ersten Version des Berichtsplans.

2) Zwischen der Version 1 des Berichtsplans und dem Amendment zu dieser Version wurden insgesamt 3 Literaturrecherchen durchgeführt und insgesamt 3 Versionen des Cochrane Reviews (inklusive der finalen Version) vorgelegt.

3) Literaturrecherchen 5 (Primärliteratur) und 6 (Sekundärliteratur) werden nach Anhörung der Betroffenen durch das IQWiG und vor der Publikation der 2. Version des Berichtsplans durchgeführt.

4) Zwischen Berichtsplan in der Version 2 und der Publikation des Vorberichts gibt es keine weitere Literaturrecherche.

5) Der Vorbericht erscheint schon etwa 2 Monate nach Vorliegen des „endgültigen“ Berichtsplans (Version 2).

Angeregt durch die zeitliche Verzahnung von Berichtsplanerstellung und der Erstellung des Cochrane Reviews erscheint es notwendig zu erfahren, wie sich der Berichtsplan und auch der Vorbericht in diesem Prozess verändert haben und welche Studien ein- bzw. ausgeschlossen wurden. Die Tabelle auf der rechten Seite gibt eine Übersicht über die Unterschiede und Gemeinsamkeiten der beiden Berichte und vergleicht diese mit anderen vorliegenden Bewertungen bzw. Meta-Analysen. Auffällig ist, dass die Studien von Pan (16), Fonseca (4), Yki-Jarvinen 2000 (24)und Kawamori (11) im Vorbericht berücksichtigt werden, während sie nicht Gegenstand der Bewertung im Cochrane Review sind. Auf der anderen Seite findet die Studie von Yokoyama (26) im IQWiG Vorbericht keine Berücksichtigung, während das Cochrane Review sie einbezieht. Wichtig wären Sensitivitätsanalysen für die vorgenommenen Veränderungen, diese wurden aber entweder nicht durchgeführt oder im Vorbericht nicht berichtet.

	NICE (15)	Tran (22)	Horvath (10)	Mullins (14)	Rosenstock (20)	Garber (6)	IQWIG
Pan (16)				X			X
Eliaschewitz (3)			X	X			X
Yokoyama (26)			X				
Yki-Jarvinen (25)		X	X	X			X
Fonseca (4)		X					X
Riddle (19)		X	X	X	X		X
HOE 901 (1)		X		X			
Fritsche (5)		X	X	X	X		X
Massi-Benedetti (12)		X	X		X		X
Rosenstock (21)	X	X	X	X	X		X
Yki-Jarvinen (24)	X	X		X			X
Kawamori (11)							X
Hermansen (8)			X			X	X
Haak (7)		X	X			X	X
Raslova (18)						X	

Beispiel Hypoglykämien

Beim Vergleich der Ergebnisse des Cochrane Reviews mit den Ergebnissen des Vorberichts fallen erheblich Unterschiede auf. So heißt es im Abstract des Cochrane Reviews: *"while no statistically significant difference for severe hypoglycemia rates was shown in any of the trials, the rate of symptomatic, overall and nocturnal hypoglycemia was statistically*

significantly lower in patients treated with either insulin glargine or detemir." Dagegen kommt der Vorbericht (S. 240 Mitte) zu folgender Einschätzung: *"Hypoglykämien wurden zwar in allen Studien erfasst, aufgrund des durchweg offenen Studiendesigns einschließlich fehlender oder unklarer Verblindung der Erhebung der Endpunkte waren alle Studien jedoch anfällig für systematische Verzerrungen. Unter diesen Bedingungen wurden nur die Angaben zu schweren und schwerwiegenden Hypoglykämien als ausreichend ergebnissicher angesehen".* Die unterschiedliche Herangehensweise erscheint auf den ersten Blick unverständlich und bedarf sicher einer Erklärung.

Interessanterweise weicht der Vorbericht auch von den Ergebnissen der meisten bislang zur Frage von Hypoglykämien durchgeführten Meta-Analysen ab: Es findet sich mehrheitlich eine Reduktion der nächtlichen Hypoglykämien und in der Regel der symptomatischen und gesamten Hypoglykämien (6,10,14,15,20,22). Schwere Hypoglykämien werden zum Teil günstig beeinflusst, waren aber im ungünstigsten Fall immer noch mit NPH-Insulinen vergleichbar.

Die zu dieser Fragestellung durchgeführte Meta-Regressions Analyse von Mullins berücksichtigt in der Modellierung den HbA1c Wert als konfundierende Variable und kommt zu dem Ergebnis, dass die Vorteile für das untersuchte IGlar größer sind, wenn für diese Variable als Confounder adjustiert wird (14). 4) Dass die Mullins Analyse ausgeschlossen wurde, erscheint unverständlich. Obwohl inhaltlich hochrelevant werden formale Kriterien für den Ausschluss vorgebracht. Sinngemäß heißt es, dass *"Meta-Analysen auf Basis individueller Patientendaten nur in die Bewertung eingehen, wenn z.B. eine nicht zu erklärende Heterogenität zwischen den einzelnen Studienergebnissen besteht“* (S. 19). Im Detail heißt es weiter (Vorbericht S. 237): *"Der Pool der in diese Publikation eingeschlossenen Studien entsprach jedoch nicht der vorliegenden Nutzenbewertung". ... "Darüber hinaus wurden formale Kriterien für die Erstellung von Meta-Analysen nicht erfüllt. So liegt z.B. weder ein veröffentlichtes Protokoll vor, noch wurde in der Publikation auf ein Protokoll verwiesen, mit dem die angewandte Methodik und die verwendete Suchstrategie im Voraus geplant wurden.“* Dieses Vorgehen und die vorgebrachte Argumentation erscheinen nicht statthaft, da der Vorbericht selbst gegen einen Teil dieser Argumente (prädefiniertes Protokoll ohne Kenntnis der Ergebnisse etc.) verstößt.

Die Bewertung des Themas Hypoglykämien im Vorbericht weicht deutlich von anderen Bewertungen inklusive den Ergebnissen des Cochrane Reviews ab. Dieses Vorgehen erscheint unverständlich und ist aus klinischer Sicht nicht nachvollziehbar.

Beispiel Yokoyama (26)

In diesem Kontext ist auch der Ausschluss der Studie von Yokoyama (26) problematisch. Auch im Cochrane Review wird die Studie kontrovers diskutiert – sie kann nach Aussage der Autoren aber aufgrund des fehlenden Verstoßes gegen die Ein- und Ausschlusskriterien nicht von der Bewertung ausgeschlossen werden. Sie wird in den Meta-Analysen allerdings nicht berücksichtigt. Der Vorbericht geht hier einen Schritt weiter und schließt die Studie über das Kriterium „sonstige Blutzucker senkende Behandlung unterschiedlich“ (siehe Seite 279 Vorbericht) von der Bewertung aus. Wie es zu diesem Ausschluss kommt ist nicht einfach nachvollziehbar:

Tabelle 1 der Arbeit von Yokoyama (S. 37, (26)) zeigt, dass 12 bzw. 14 von jeweils 31 Patienten mit Glimepiride und 20 bzw. 19 von 31 Patienten mit Metformin im Sinne einer

„sonstige Blutzucker senkende Behandlung“ erhielten. Darüber hinaus erhielten 26 bzw. 27 von 31 Patienten Aspartat, 5 bzw. 4 von 31 Patienten Lispro. Dass die Dosis des kurzwirksamen Insulins bei gleichzeitiger Steigerung der IGlar Dosis reduziert wurde, während das Verhältnis beim NPH Insulin gleich blieb (S. 38, Tabelle 2, (26)) reicht als Argument nicht aus, wird jedoch auf Seite 22 des Vorberichts als Argument bemüht. Die Anfrage an den Autor vom 14.09.2006 ergab, dass die Titrierung der NPH Dosis durchgeführt wurde, aber aufgrund der Befürchtung nächtlicher Hypoglykämien nicht fortgesetzt wurde. Diese Erkenntnis wird in den Ergebnissen und in der Diskussion zu wenig gewürdigt. Die Möglichkeit der morgendlichen Gabe in Kombination mit der Titration des Analogs im Hinblick auf die Erreichung der HbA1c Werte bei fehlendem Anstieg der Hypoglykämierate scheint hier der entscheidende Vorteil gegenüber dem NPH-Insulin zu sein.

Die Datenlage spiegelt nicht die klinische Realität

Die mangelnde Praxisrelevanz und die fehlende Einbeziehung der besten verfügbaren Evidenz wurden an den Berichten des Instituts wiederholt bemängelt. Strikte Ein- und Ausschlusskriterien führen zu einer massiven Datenreduktion, die zum Teil zu einer erheblichen Verfälschung der Bewertung führen kann (illustriert am Beispiel der kurzwirksamen Insulinanaloga (23)). Dass der Vorbericht jetzt die fehlende Praxisrelevanz der zu langwirksamen Insulinanaloga auf Datenbasis der DETECT Studie (17) bemängelt, (siehe S. 252 Vorbericht) verwundert, zumal die Datenreduktion ja artifiziell und damit behebbar ist. Beispiele für Studien, die für die klinische Praxis von höchster Relevanz sind aber im Vorbericht nicht berücksichtigt wurden, sind die Studien von Holman (9) und Bretzel (2). Ursachen dafür sind die „falsche“ Vergleichsbehandlung bzw. auch das Publikationsdatum (Ende 2007 bzw. Anfang 2008).

Zusammenfassung und Fazit

Der wesentliche Vorteil der langwirksamen Insulinanaloga ist ihre Eigenschaft einer stabilen, berechenbaren Insulinbasis über 24 Stunden nach einmaliger (IGlar) bzw. zweimaliger Injektion (IDet). Die NPH Insuline weisen in ihrer Pharmakokinetik Defizite als Basisinsulin auf, die durch mehrmalige Applikation am Tag nur teilweise kompensiert werden können. Die bessere Verfügbarkeit und Handhabung der langwirksamen Insulinanaloga als Basisinsulin hat in den bis dato veröffentlichten Studien in der Regel zu niedrigeren Nüchternblutglukosekonzentrationen und geringeren Hypoglykämieraten (insbesondere nachts) geführt.

Der Vorbericht unternimmt keinen Versuch, diese Vorteile der langwirksamen Insulinanaloga zu objektivieren, sondern umgeht diese Fragen durch die entsprechende Formulierung des Berichtsplans. Er ist darüber hinaus methodisch angreifbar und präsentiert für wichtige Entscheidungen zum Ein- und Ausschluss von Studien keine Sensitivitätsanalysen. Die enge Verzahnung mit einem zeitgleich erstellten Cochrane Review ist unglücklich und der Abgleich der beiden Berichte dokumentiert wesentlichen Unterschiede im Ergebnis.

Mahlow, den 21. April 2008 Priv.-Doz. Dr. Peter Bramlage

Literatur

1. Safety and efficacy of insulin glargine (HOE 901) versus NPH insulin in combination with oral treatment in Type 2 diabetic patients. Diabet Med 2003; 20: 545-51.
2. Bretzel RG, Nuber U, Landgraf W, Owens DR, Bradley C, Linn T. Once-daily basal insulin glargine versus thrice-daily prandial insulin lispro in people with type 2 diabetes on oral hypoglycaemic agents (APOLLO): an open randomised controlled trial. Lancet 2008; 371: 1073-84.
3. Eliaschewitz FG, Calvo C, Valbuena H, et al. Therapy in type 2 diabetes: insulin glargine vs. NPH insulin both in combination with glimepiride. Arch Med Res 2006; 37: 495-501.
4. Fonseca V, Bell DS, Berger S, Thomson S, Mecca TE. A comparison of bedtime insulin glargine with bedtime neutral protamine hagedorn insulin in patients with type 2 diabetes: subgroup analysis of patients taking once-daily insulin in a multicenter, randomized, parallel group study. Am J Med Sci 2004; 328: 274-80.
5. Fritsche A, Schweitzer MA, Haring HU. Glimepiride combined with morning insulin glargine, bedtime neutral protamine hagedorn insulin, or bedtime insulin glargine in patients with type 2 diabetes. A randomized, controlled trial. Ann Intern Med 2003; 138: 952-9.
6. Garber AJ, Clauson P, Pedersen CB, Kolendorf K. Lower risk of hypoglycemia with insulin detemir than with neutral protamine hagedorn insulin in older persons with type 2 diabetes: a pooled analysis of phase III trials. J Am Geriatr Soc 2007; 55: 1735-40.
7. Haak T, Tiengo A, Draeger E, Suntum M, Waldhausl W. Lower within-subject variability of fasting blood glucose and reduced weight gain with insulin detemir compared to NPH insulin in patients with type 2 diabetes. Diabetes Obes Metab 2005; 7: 56-64.
8. Hermansen K, Davies M, Derezinski T, Martinez Ravn G, Clauson P, Home P. A 26-week, randomized, parallel, treat-to-target trial comparing insulin detemir with NPH insulin as add-on therapy to oral glucose-lowering drugs in insulin-naive people with type 2 diabetes. Diabetes Care 2006; 29: 1269-74.
9. Holman RR, Thorne KI, Farmer AJ, et al. Addition of biphasic, prandial, or basal insulin to oral therapy in type 2 diabetes. N Engl J Med 2007; 357: 1716-30.

10. Horvath K, Jeitler K, Berghold A, et al. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Database Syst Rev 2007: CD005613.
11. Kawamori R, Iwamoto Y, Kadowaki T, Iwasaki M. Efficacy and safety of insulin glargine in concurrent use with oral hypoglycemic agents for the treatment of type 2 diabetic patients. J Clin Ther Med 2003; 19: 445-464.
12. Massi Benedetti M, Humburg E, Dressler A, Ziemen M. A one-year, randomised, multicentre trial comparing insulin glargine with NPH insulin in combination with oral agents in patients with type 2 diabetes. Horm Metab Res 2003; 35: 189-96.
13. Medizinische Universität Graz ERC, Aktuelle und Abgeschlossene Projekte. http://www.meduni-graz.at/ebm/projekte/, 2008.
14. Mullins P, Sharplin P, Yki-Jarvinen H, Riddle MC, Haring HU. Negative binomial meta-regression analysis of combined glycosylated hemoglobin and hypoglycemia outcomes across eleven Phase III and IV studies of insulin glargine compared with neutral protamine Hagedorn insulin in type 1 and type 2 diabetes mellitus. Clin Ther 2007; 29: 1607-19.
15. NICE. Guidance on the use of long-acting insulin analogues for the treatment of diabetes - insulin glargine. London: NICE, 2002:http://www.nice.org.uk/nicemedia/pdf/53_Insulin_analogues_full_guidance.pdf.
16. Pan CY, Sinnassamy P, Chung KD, Kim KW. Insulin glargine versus NPH insulin therapy in Asian Type 2 diabetes patients. Diabetes Res Clin Pract 2007; 76: 111-8.
17. Pittrow D, Stalla GK, Zeiher AM, et al. Prävalenz, medikamentose Behandlung und Einstellung des Diabetes mellitus in der Hausarztpraxis. Med Klin (München) 2006; 101: 635-44.
18. Raslova K, Bogoev M, Raz I, Leth G, Gall MA, Hancu N. Insulin detemir and insulin aspart: a promising basal-bolus regimen for type 2 diabetes. Diabetes Res Clin Pract 2004; 66: 193-201.
19. Riddle MC, Rosenstock J, Gerich J. The treat-to-target trial: randomized addition of glargine or human NPH insulin to oral therapy of type 2 diabetic patients. Diabetes Care 2003; 26: 3080-6.
20. Rosenstock J, Dailey G, Massi-Benedetti M, Fritsche A, Lin Z, Salzman A. Reduced hypoglycemia risk with insulin glargine: a meta-analysis comparing insulin glargine with human NPH insulin in type 2 diabetes. Diabetes Care 2005; 28: 950-5.

21. Rosenstock J, Schwartz SL, Clark CM, Jr., Park GD, Donley DW, Edwards MB. Basal insulin therapy in type 2 diabetes: 28-week comparison of insulin glargine (HOE 901) and NPH insulin. Diabetes Care 2001; 24: 631-6.
22. Tran K, Banerjee S, Li H, et al. Long-acting insulin analogues for diabetes mellitus: meta-analysis of clinical outcomes and assessment of cost-effectiveness [Technology Report number 92]. Ottawa: Candian Agency for Drugs and Technologies in Health, 2007.
23. Wehling M. Kommentar zur Bewertung des Einsatzes kurz wirksamer Insulinanaloga bei Typ-I-Diabetikern durch das IQWiG. Diabetologe 2007; 3: 372-9.
24. Yki-Jarvinen H, Dressler A, Ziemen M. Less nocturnal hypoglycemia and better post-dinner glucose control with bedtime insulin glargine compared with bedtime NPH insulin during insulin combination therapy in type 2 diabetes. HOE 901/3002 Study Group. Diabetes Care 2000; 23: 1130-6.
25. Yki-Jarvinen H, Kauppinen-Makelin R, Tiikkainen M, et al. Insulin glargine or NPH combined with metformin in type 2 diabetes: the LANMET study. Diabetologia 2006; 49: 442-51.
26. Yokoyama H, Tada J, Kamikawa F, Kanno S, Yokota Y, Kuramitsu M. Efficacy of conversion from bedtime NPH insulin to morning insulin glargine in type 2 diabetic patients on basal-prandial insulin therapy. Diabetes Res Clin Pract 2006; 73: 35-40.

A 2.3 Bretzel, Reinhard, Prof. Dr. med.

Adresse:
Medizinische Klinik und Poliklinik III
Rodthohl 6
35392 Gießen

Stellungnahme zur vorläufigen Nutzenbewertung langwirksamer Insulinanaloga des IQWiG (Vorbericht A05-03)

Mit großem Interesse haben wir den Vorbericht A05-03, am 28. März 2008 im Internet veröffentlicht, gelesen.

Wir sind seit vielen Jahren sowohl wissenschaftlich als klinisch auf dem Gebiet der Diabetologie tätig und haben u.a. große Multicenter-Studien geleitet. Das vorläufige Ergebnis der Nutzenbewertung hat uns verwundert. Wir hoffen, dass sich im Laufe des Stellungnahmeprozesses noch relevante Veränderungen ergeben, die wir für zwingend erforderlich halten. Nachfolgend möchten wir im Detail zum Vorbericht eine Stellungnahme abgeben.

Nicht genannte bzw. nicht vollständige anerkannte Studien

Morgenarm der Studie 4001 und Bewertung der Ergebnissicherheit der Unterzuckerungen

Der teilweise Ausschluss der Ergebnisdaten des Morgenarms der Studie 4001 ist ungerechtfertigt und ignoriert die Festlegungen aus dem Berichtsplan. Die Ergebnisse der morgendlichen Gabe von Insulin Glargin im Vergleich zur abendlichen Gabe von NPH-Insulin waren im Hinblick auf die Blutzuckereinstellung statistisch signifikant und mit einer HBA1c Absenkung oberhalb von 0,4 Prozent auch klinisch relevant. Auch die Häufigkeiten von Unterzuckerungen waren im Morgenarm der Studie 4001 unter Insulin Glargin seltener als unter NPH-Insulin. Die vom IQWiG angesprochenen sogenannten Tageseffekte werden in einer weiteren Studie, die die morgendliche Gabe von Insulin Glargin mit der abendlichen Gabe verglichen haben, nicht bestätigt (Standl et al., Horm Metab Res 2006; 38: 172-77). Durch eine zu enge Studienauswahl des IQWiG konnten diese Erkenntnisse bei der Abfassung des Vorberichtes nicht erkannt werden.

In einem weiteren Gesichtspunkt ist die Studie 4001 fehlerhaft beurteilt worden. Die Studie 4001 wurde mit besonderen Maßnahmen der verblindeten Endpunkterhebung durchgeführt. Alle Unterzuckerungen wurden einer nachträglichen, verblindeten Bewertung unterzogen. Für die Nutzenbewertung der kurzwirksamen Insulinanaloga war diese qualitätssichernde Maßnahme vom IQWiG gefordert worden, um die Unterzuckerungen als im Ergebnis sicher einstufen zu können. Diese Maßnahme wurde jetzt nicht vom IQWiG überraschend nicht besonders positiv im Hinblick auf die Ergebnissicherheit interpretiert.

Auswirkungen von Insulin Detemir auf die Entwicklungen des Körpergewichts

Es wurden umfassende präklinische Untersuchungen durchgeführt und durch Ergebnisse klinischer Studien bestätigt, dass sich unter der Behandlung mit Insulin Detemir eine günstigere Entwicklung des Körpergewichts nachweisen lässt. Da Typ 2 Diabetiker in der Regel übergewichtig sind, stellt die geringere Gewichtszunahme im Laufe der Therapie einen patientenrelevanten Vorteil dar.

Wir dürfen auf diese am 29.03.2008 in Lancet publizierte Studie hinweisen (Bretzel RG et al., Lancet 2008; 371:1073-1084).

Die Ergebnisse dieser Studie wurden 2006 in Teilen bereits in Abstractform publiziert (Diabetes [ADA Meeting 2006]; Diabetologia [EASD Meeting 2006]).

Die Fragestellung widmet sich dem Vergleich des in Deutschland weit verbreiteten Behandlungskonzepts einer supplementären, prandialen Insulintherapie im Vergleich zu einer einmaligen basalen Insulintherapie, beide mit der Gabe von oralen Antidiabetika kombiniert bei Typ 2 Diabetikern. Es konnte gezeigt werden, dass mit der einmal täglichen Gabe von Insulin Glargin die Blutzuckereinstellung gemessen am HbA1c ebenso gelingt, wie mit einer dreimal täglichen prandialen Gabe von Insulin Lispro. Die Häufigkeiten von gesamten und symptomatischen auch durch die Messung von Blutzuckerwerten gesicherten Unterzuckerungen waren für das Glargin gestützte Behandlungskonzept überlegen. Durch die Bestimmung der Blutzuckerwerte konnte einer vom IQWiG befürchteten systematischen Ergebnisverzerrung bei der Kategorisierung von Unterzuckerungen begegnet werden. Für nächtliche und schwere Unterzuckerungen wurde das Signifikanzniveau knapp verfehlt, wegen zu geringer Fallzahlen. Darüber hinaus sind signifikante und relevante Vorteile der Behandlungszufriedenheit unter Insulin Glargin im Vergleich zu Insulin Lispro gemessen worden.

Methodische Anforderungen an Meta-Analysen

Für bereits veröffentlichte Meta-Analysen wurden die Anforderungen post-hoc so hoch angelegt, dass keine der 33 identifizierten Meta-Analysen und HTA-Berichte vom IQWiG eingeschlossen worden ist. Wenn vom IQWiG die vollständige Datenanalyse einer präspezifizierten Fragestellung gefordert wird, dann muss dieses Kriterium zum Stand der Festlegung der jeweiligen Analysepläne geprüft werden. Eine im Jahre 2003 veröffentlichte Meta-Analyse kann nur die im Jahre 2002 verfügbaren Studiendaten beinhalten. Dieser Erfordernis hat das IQWiG bei der Fixierung der Anforderungen an Meta-Analysen bisher nicht Rechnung getragen.

Ein weiteres praktisches Problem ergibt sich aus der Strategie des IQWiG, meta-analytische Auswertungen nach Behandlungskonzepten zu trennen. Diese Stratifizierung wurde erst mit der Veröffentlichung des Vorberichtes erkennbar. Ein adäquater Diskurs über die berichtsspezifische Methodik hat dazu nicht stattgefunden. Es wurde versäumt, diese Anforderungen an Meta-Analysen bereits mit dem Berichtsplan zu veröffentlichen.

Vorbericht des IQWiG und Internationale Standards

Den einschlägigen Quellen ist zu entnehmen, dass das IQWiG bei seiner Nutzenbewertung an internationale Standards gebunden ist. Dies sollte im Ergebnis bedeuten, dass die Nutzenbewertung des IQWiG in der Arbeitsmethodik und im Ergebnis nicht unbegründet im Widerspruch zu international durchgeführten HTA´s stehen darf. Dies ist allerdings in mehreren Aspekten und offenkundig in mehreren Aspekten der Fall.

Das britische National Institute for Clinical Excellence (NICE) kommt in seiner Guidance No. 53 zu dem Ergebnis, dass folgende Patientengruppen besonders von einer Therapie mit langwirksamen Insulinen profitieren:

- Patienten die Unterstützung bei der Injektion benötigen
- Patienten deren Lebensführung stark durch symptomatische Hypoglykämien eingeschränkt wird
- Patienten die andernfalls (d.h. bei Nichteinsatz langwirksamer Analoga) zwei Insulin Injektionen pro Tag in Kombination mit oralen Antidiabetika benötigen würden

Die Vorteile langwirksamer Insulinanaloga liegen vor allem in der Möglichkeit der Einmalapplikation, der flexiblen Wahl des Tageszeitpunktes für die Insulininjektion und der niedrigeren Hypoglykämierate. Wir können uns dieser Einschätzung des NICE auch aus unserer eigenen klinischen Erfahrung heraus eindeutig anschließen.

Das IQWiG kommt in seiner vorläufigen Nutzenbewertung zu einem deutlich weniger positiven Ergebnis hinsichtlich des patientenrelevanten Nutzens langwirksamer Insulinanaloga. Besonders eklatant stellt sich der Unterschied der Beurteilung der Beschreibung der Hypoglykämiefrequenzen dar.

Dies hat nach unserer Einschätzung verschiedene Gründe, die thematisiert und im Endbericht korrigiert werden sollten.

Datenquellen

Das bisherige Verfahren nutzt die vorhandenen Datenquellen stark eingeschränkt. Es besteht kein Zweifel daran, dass randomisierte kontrollierte Studien eine sehr zuverlässige Quelle für die evidenzbasierte Medizin darstellen, dennoch müssen auch Studien aus dem Versorgungsalltag in der Diabetologie einbezogen werden.

Die vom IQWiG gewählten Einschlusskriterien für Datenquellen sind in Bezug auf das Studiendesign zu eng. Die verschiedenen Behandlungskonzepte beim Diabetes mellitus werden in Studien gegeneinander verglichen. Häufig besteht dabei begründbar keine genaue Übereinstimmung der Begleittherapien in den beiden Therapiearmen. Der daraus resultierende generelle Ausschluss solcher Studien von der Nutzenbewertung ist nicht gerechtfertigt. Es erscheint uns vielmehr erforderlich, den Vergleich zur prandialen Insulintherapie und auch zu biphasischen Insulinen zuzulassen. Hierdurch ergibt sich eine breitere, solide Datenbasis.

Gerne sind wir bereit, eine etwaige Einladung zur mündlichen Anhörung in Ihrem Institut wahrzunehmen.

Prof. Dr. med. Reinhard G. Bretzel
Arzt für Innere Medizin, Endokrinologie und Diabetologie
Lehrstuhl für Innere Medizin, Endokrinologie und Diabetologie
Justus-Liebig-Universität Giessen
Direktor der Medizinischen Klinik und Poliklinik III
Universitätsklinikum Giessen und Marburg GmbH
Standort Giessen

i. V.

Prof. Dr. med. Hans-Ulrich Häring
Arzt für Innere Medizin, Endokrinologie und Diabetologie
Direktor der Abteilung Innere Medizin IV
Universitätsklinikum Tübingen

A 2.4 Chantelau, Ernst, Prof. Dr. med.

Adresse:
Karolinger Straße 69 III
40223 Düsseldorf

Prof. Dr.med. E.Chantelau, Düsseldorf: Stellungnahme zur Bewertung von langwirkenden Insulinanaloga bei Typ-2 Diabetes mellitus durch das IQWiG 2008

Angesichts des vermehrten Auftretens von Lungenkrebs bei Anwendern des von Sanofi-Aventis hergestellten Insulinpulvers zur inhalativen Insulintherapie stellt sich erneut die Frage nach den Mitogenitätsrisiken der zur Injektionstherapie zugelassenen Insulinderivate. Grundlagenforscher melden erhebliche Bedenken an (1,2). Firmen-gestützte in-vitro Untersuchungsergebnisse (3-5), denen zufolge die Insulinderivate Lantus® bzw. Levemir® nicht stärker mitogen als „natürliches“ unmodifiziertes Insulin seien, konnten von Firmen-unabhängigen Untersuchern nicht reproduziert werden (6-8). Unbestritten ist außerdem für Lantus (entsprechende Untersuchungsergebnisse für Levemir liegen mir nicht vor) , dass es bei Typ-1 Diabetikern - bereits bei einmaliger Anwendung- den IGF-1 Spiegel im Serum erhöht (9-11). IGF-1 fördert das Wachstum vieler Tumoren (1). Diese Befunde dürfen bei der Bewertung von Lantus bzw. Levemir nicht ignoriert werden.

1)Frasca F, et al. The role of insulin receptors and IGF-1 receptors in cancer and other diseases. Arch Physiol Biochem 2008;114:DOI 10.10180/13813450801969715(ahead of print)
2)Eckardt K, Eckel J. Insulin analogues: action profiles beyond glycaemic control. Arch Physiol Biochem 2008;114: DOI 10.1080/13813450802983369(ahead of print)
3)Kurtzhals et al., Correlation of receptor binding and metabolic and mitogenic properties of insulin analogs designed for clinical use. Diabetes 2000;49:999-1005
4)Staiger K et al., Insulin and its analogue glargine do not affect viability and proliferation of human coronary artery endothelial and smooth muscle cells. Diabetologia 2005;48:1989-1905
5)Staiger K, et al. Effect of insulin glargine versus regular human insulin on proliferation of human breast epithelial cells, Abstract. Diabetes 2005;54/Suppl.1: A 111
6)Eckardt K, et al. IGF-1 receptor signalling determines the mitogenic potency of insulin analogues in human smooth muscle cells and fibroblasts. Diabetologia 2007; 50:2534-2543
7)Mayer D, et al. Proliferative effects of insulin analogues on mammary epithelial cells. Arch Physiol Biochem 2008;114: DOI 10.1080/13813450801900645(ahead of print)
8) Weinstein D, et al. Insulin analogues display IGF-1-like mitogenic and antiapoptotic activities in cultured cancer cells. Diabetes/Metabolism Research and Reviews,2008 (in press)
9)Horn T, et al. Plasma insulin and the IGF-1 system after Glargine. Abstract. Diabetes & Stoffwechsel 2001;10/Suppl.1:58

10)Ekström K, et al. Normalization of the IGF- IGFBP axis by sustained nightly insulinization in type-1 diabetes. Diabetes Care 2007; 30:1357-1363
11)Slawik M, et al. Treatment with insulin glargine does not suppress serum IGF-1. Diabetic Med 2006; 23:814-817

A 2.5 Debusmann, Edeltraud und Horst

Adresse:
Otto-Hahn-Straße 7
63150 Heusenstamm

Name, Vorname Edeltraud und Horst Debusmann

Straße Otto-Hahn-Straße 7

PLZ / Ort 63150 Heusenstamm

Auftragsnummer: A05-03

Langwirksame Insulinanaloga zur Behandlung des Diabetes mellitus Typ 2

Schriftliche Stellungnahme zum Vorbericht

Sehr geehrter Herr Prof. Sawicki,

ich bin kein Arzt, sondern Patient. Insofern bitte ich um Nachsicht, daß ich keine wissenschaftliche Stellungnahme zu den 307 Seiten Vorbericht abgeben kann.

Von Ihrem Institut wird behauptet, daß die Analoginsuline gegenüber dem Humaninsulin keinen Mehrnutzen hätten.

Soweit mir bekannt, hatten Sie als Mitarbeiter und Ko-Autor einer Studie "Erfahrungen mit Insulin Lispro (Erstautor war Lutz Heineman) noch vor Jahren die großen Vorteile des Analoginsulins gegenüber dem Humaninsulin herausgestellt und sprachen von einem Segen für die Betroffenen.

Auch ich habe die großen Vorteile der Analoginsuline kennengelernt und bin meinem Diabetologen dankbar, daß nach langen Versuchen endlich eine Therapie gefunden wurde, mit der ich mein Leben als Diabetiker entsprechend den Tagesanforderungen spontan gestalten kann.

Mit Bestürzung nehme ich jetzt zur Kenntnis, daß der Gemeinsame Bundesausschuss (G-BA) aufgrund Ihrer Empfehlung den Beschluss gefaßt hat, die Behandlung mit Insulinanaloga aus dem Leistungskatalog der gesetzlichen Krankenversicherung zu streichen. Die Behandlung mit modernen Insulinen stellt jedoch einen wesentlichen Fortschritt in der diabetologischen Versorgung und einen erheblichen Zugewinn bei der flexiblen Gestaltung des täglichen Lebens dar.

Da privat Versicherte weiterhin die Insulinanaloga von ihrer privaten Krankenkasse bezahlt bekommen, wird die Zweiklassenmedizin endgültig zum Tatbestand.

Es ist nicht verständlich, die durch ihre Krankheit bereits lebenslang erheblich in ihrer Lebensführung eingeschränkten Diabetiker jetzt noch mit einer weiteren Strafe zu belegen, indem sie vorrangig nur noch mit Insulinen behandelt werden sollen, die durch die Anwendung im Alltag erhebliche Einschränkungen auslösen.

Ist Ihnen bewußt, daß bisher mit Analoginsulin gut eingestellte Diabetiker mit der Entscheidung des Gem-BA jetzt erneut zum Versuchskaninchen werden? Jeder Diabetologe war für seine Patienten froh, nach langen Versuchsreihen mit all den Höhen und Tiefen für die Betroffenen endlich eine gute Blutzuckereinstellung gefunden zu haben, dies wird jetzt mit der Entscheidung des Gem-BA wieder fahrlässig zunichte gemacht.

Wer aber von Diabetes betroffen ist, weiß um die Probleme und deren Langwierigkeit einer guten Glucoseeinstellung. Jetzt sollen per Verordnung "von oben" die Betroffenen auf Humaninsulin (entgegen ärztlicher Empfehlung mit all den bekannten Nachteilen, insbesondere für Kinder, Jugendliche und im Beruf stehende Menschen) umgestellt werden. Erste Erfahrungen beim Treffen in den Selbsthilfegruppen zeigten, daß Patienten nach der "Zwangsumstellung" mehrfach in ein Koma verfielen und per Notarztwagen in die Klinik mußten. Diese Folgekosten plus der erneuten langwierigen Rückumstellung auf das vorheri-

ge Medikament übersteigen ein Mehrfaches dessen, was man einzusparen glaubte.
Was für eine Logik, welch ein Irrsinn!

Haben Sie bei Ihrer Überlegung auch die enormen Folgekosten mit einbezogen, die der Versichertengemeinschaft infolge schlechter Blutzuckereinstellung zukünftig entstehen werden (von Augen- über Nervenschäden bis hin zur Dialyse)? Ich denke, man will Kosten sparen. Mit der jetzigen Entscheidung werden die erheblichen Folgekosten aber in die Zukunft verlagert.

Man hat den Eindruck, IQWiG und Gem-BA kennen nicht das wirkliche Leben, kennen nicht die Sorgen und Nöte der betroffenen Menschen, die sich Ihre Krankheit nicht selber ausgesucht haben, sondern stellen lapidar fest ...

> Insulinanaloga haben keinen Mehrnutzen gegen über dem Humaninsulin.
> Welch eine Anmaßung!
> Ein Affront gegen alle verantwortungsvoll handelnden Ärzte und Diabetologen;
> eine Mißachtung international anerkannter Studien,
> ein Rückfall in die Vergangenheit.

In unserem Grundgesetz (GG) heißt es:

Artikel 1 (1)
Die Würde des Menschen ist unantastbar,
sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt.

Artikel 2 (2)
Jeder hat das Recht auf Leben und körperliche Unversehrtheit...

Mit der angedachten Entscheidung sehe ich einen erheblichen Konflikt zu den im GG verankerten Rechten. Die Politik maßt sich an, über mein Leben zu bestimmen.
Mit welchem Recht?

Gesundheit darf nicht zum Luxusgut werden.

Jeder insulinpflichtige Diabetiker muß das Recht haben, das für seinen Körper und seine therapeutischen Bedürfnisse erforderliche Insulin zu erhalten.

Heusenstamm, den 2.04.2008 Edelhard Debusmann
Datum/Unterschrift
Horst Debusmann

zur Kenntnis an:
- Bundesministerin Frau Ulla Schmidt, Berlin
- Frau Uta Zapf, SPD-Abgeordnete im Deutschen Bundestag für den Kreis Offenbach,
- Herrn Dr. Lippold, CDU-Abgeordneter im Deutschen Bundestag für den Kreis Offenbach,