

Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie

**Dokumentation und Würdigung der Stellung-
nahmen zum Berichtsplan**

Auftrag N04-01
Version 1.0
Stand: 10.09.2007

Impressum

Herausgeber:

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Thema:

Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum des Auftrags:

16.11.2004

Interne Auftragsnummer:

N04-01

Anschrift des Herausgebers:

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

Tel: 0221/35685-0

Fax: 0221/35685-1

Berichte@iqwig.de

www.iqwig.de

Inhaltsverzeichnis

Tabellenverzeichnis.....	3
Abkürzungsverzeichnis.....	4
1 Dokumentation des Stellungnahmeprozesses	6
2 Würdigung der Stellungnahmen.....	7
2.1 Auswahl der eingeschlossenen Studientypen	7
2.2 Bewertung der Studienqualität	8
2.3 Auswahl patientenrelevanter Endpunkte	8
2.4 Nutzenbewertung.....	9
2.5 Einteilung bzw. Klassifizierung von lokalen Behandlungsverfahren.....	10
2.6 Wirtschaftlichkeit.....	12
3 Literatur	13
4 Dokumentation der Stellungnahmen.....	15
4.1 Darlegung potenzieller Interessenkonflikte der Stellungnehmenden.....	15
4.2 Stellungnahmen von Fachgesellschaften, Firmen und Vereinigungen.....	20
4.2.1 Arbeitskreis BPH der Akademie der deutschen Urologen.....	21
4.2.2 Lumenis GmbH.....	40
4.2.3 Arbeitskreis BPH der Akademie der deutschen Urologen.....	42
4.2.4 Österreichische Gesellschaft für Urologie und Andrologie (ÖGU).....	53
4.2.5 Deutsche Krankenhausgesellschaft e. V.	58
4.2.6 Kassenärztliche Bundesvereinigung	60
4.2.7 Deutsche Krankenhausgesellschaft e. V.	62
4.2.8 Medtronic GmbH	67
4.3 Stellungnahmen von Einzelpersonen.....	76
4.3.1 Prof. Dr. Klaus Höfner	77
4.3.2 Prof. Dr. Rolf Muschter.....	85
4.3.3 Prof. Dr. J. J. M. C. H. de la Rosette	96
4.3.4 Dr. Christoph von Zastrow	100
5 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll.....	103

Tabellenverzeichnis

Tabelle 1: Stellungnahmen von Fachgesellschaften, Firmen und Organisationen	16
Tabelle 2: Stellungnahmen von Einzelpersonen	17

Abkürzungsverzeichnis

Abkürzung	Bedeutung
AHRQ	Agency for Healthcare Research and Quality
BPS	Benignes Prostatasyndrom
BQS	Bundesgeschäftsstelle Qualitätssicherung gGmbH
CCT	Controlled Clinical Trial
CLAP	Kontakt-Laserablation der Prostata
FDA	Food and Drug Administration
G-BA	Gemeinsamer Bundesausschuss
GKV-WSG	Gesetz zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung
HAS	Haute Autorité de Santé
HE-TUMT	Hochenergie Transurethrale Mikrowellentherapie
HoLEP	Holmiun-Laserenukleation der Prostata
HTA	Health Technology Assessment
IPSS	International Prostatic Symptom Score
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
LAP	Laser Ablation der Prostata
NE-TUMT	Niedrigenergie Transurethrale Mikrowellentherapie
PLFT	ProstaLund Feedbacktherapie
RCT	Randomized Controlled Trial
TEAP	Transurethrale Ethanolablation der Prostata
TUEP	Transurethral Evaporation of the Prostate
TUVRP/TUEVRP	Transurethrale (Elektro)vaporesektion der Prostata
TUIP	Transurethrale Inzision der Prostata
TULAP	Transurethrale Laserablation der Prostata
TUMT	Transurethrale Mikrowellentherapie
TUNA	Transurethrale Nadelablation
TUR	Transurethrale Resektion
TURP	Transurethrale Resektion der Prostata
TVP	Transurethrale Vaporisation der Prostata

Abkürzung	Bedeutung
UE	Unerwünschtes Ereignis
SGB	Sozialgesetzbuch
VLAP	Visuelle Laserablation der Prostata
WHO	World Health Organization

1 Dokumentation des Stellungnahmeprozesses

Der Berichtsplan „N04-01: Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“ wurde in der Version 1.0 vom 13.05.2005 am 07.06.2005 auf den Internetseiten des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) veröffentlicht [1].

Im Zuge des Gesetzes zur Stärkung des Wettbewerbs in der gesetzlichen Krankenversicherung (GKV-WSG) wurde die Verpflichtung zur Transparenz von Verfahrensschritten neu gefasst. Somit hat das Institut nach § 139a Abs. 5 SGB V in allen wichtigen Abschnitten des Bewertungsverfahrens Gelegenheit zur Stellungnahme zu geben (Berichtsplan und Vorbericht) und diese Stellungnahmen in die Entscheidung einzubeziehen. Aus diesem Grund wurde für Projekte, die vor Inkrafttreten dieses Gesetzes bereits in Bearbeitung waren, eine Übergangsregelung geschaffen, die zunächst ein kombiniertes Stellungnahmeverfahren vorsah, in dem zum Berichtsplan und zum Vorbericht Stellung genommen werden konnte. Dieses Vorgehen wurde in der Folge in mehrere Schritte gespalten, um Interessierten in allen Verfahrensschritten die Möglichkeit zur Beteiligung zu geben. Für den Auftrag N04-01 bedeutet dies konkret, dass zum Berichtsplan Version 1.0 [1] vom 13.05.2005 und zum Methodikteil des Vorberichts Version 1.0 [2] vom 26.03.2007 im Rahmen des kombinierten Stellungnahmeverfahrens vom 03.04.2007 bis 09.05.2007 schriftliche Stellungnahmen eingereicht werden konnten und am 03.07.2007 eine mündliche Erörterung hierzu stattfand.

Insgesamt wurden 12 Stellungnahmen verschiedener Fachgesellschaften, Firmen oder Privatpersonen (siehe Tabelle 1 und Tabelle 2) form- und fristgerecht abgegeben.

Aufgrund der Stellungnahmen kam es zu Änderungen der Version 1.0 des Berichtsplans.

2 Würdigung der Stellungnahmen

Folgende methodische Aspekte wurden in den 12 Stellungnahmen diskutiert: (1) Auswahl der eingeschlossenen Studientypen, (2) Bewertung der Studienqualität, (3) Auswahl patientenrelevanter Endpunkte, (4) Nutzenbewertung, (5) Einteilung bzw. Klassifizierung von Verfahren und (6) Wirtschaftlichkeit.

2.1 Auswahl der eingeschlossenen Studientypen

In den Stellungnahmen wurde kritisiert, dass hauptsächlich randomisierte kontrollierte Studien (RCTs) in die Nutzenbewertung eingeschlossen werden sollen und nichtrandomisierte kontrollierte Studien (CCTs) erst bei mangelnder Anzahl von RCTs Berücksichtigung fänden. Die Randomisierung stellt (im Mittel) die Strukturgleichheit von zu vergleichenden Gruppen her und ist ein wesentlicher Aspekt zur Gewinnung von unverzerrten Ergebnissen. Dagegen sind bei CCTs spezielle Maßnahmen zur adäquaten Berücksichtigung prognostischer Faktoren notwendig, um das Verzerrungspotential der Ergebnisse nicht ausufern zu lassen. Dabei ist das Verzerrungspotential bei nichtrandomisierten Studien immanent höher, als bei randomisierten Vergleichen, so dass die Ergebnisse i. d. R. lediglich Hinweise auf einen Vorteil geben können.

Um abschätzen zu können, inwieweit die Berücksichtigung qualitativ hochwertiger CCTs (eine Berücksichtigung prognostischer Faktoren muss erkennbar sein) ggf. die aus den RCTs gewonnenen Erkenntnisse beeinflussen, werden solche CCTs im Rahmen einer Sensitivitätsanalyse einbezogen, auch dann, wenn eine ausreichende Anzahl von RCTs vorliegen sollte. Eventuelle Unterschiede zu den Resultaten der RCTs werden dargestellt und diskutiert.

Die Qualitätsklassifizierung der Studien dient der Diskussion heterogener Ergebnisse und kann die Schlussfolgerungen beeinflussen, aber nicht zum Ausschluss von Ergebnissen aus der Nutzenbewertung oder dem Heranziehen eines Studientyps niedrigerer Evidenzstufen führen.

Im Rahmen der wissenschaftlichen Erörterung wurden die Anwesenden gebeten, nicht randomisierte Studien zu nennen, aus denen sich Hinweise auf einen so genannten dramatischen Effekt, dessen Entstehung allein durch Verzerrungen unwahrscheinlich wäre, durch die Anwendung der zu prüfenden Verfahren ableiten ließen. Es wurde weiterhin darum gebeten, diese Benennung innerhalb von 3 Wochen nach dem Erörterungstermin vorzunehmen. Eine solche Benennung von nicht randomisierten Studien mit dramatischen Effekten erfolgte jedoch nicht.

2.2 Bewertung der Studienqualität

Hinsichtlich des festgelegten Vorgehens, in die Nutzenbewertung nur die Studienergebnisse einzubeziehen, die auf mindestens 70 % der randomisierten bzw. zugeteilten Patienten in der Analyse beruhen, und für die der Unterschied der in die Analyse einbezogenen Patienten zwischen den Gruppen nicht größer als 15 % beträgt, wurde Stellung genommen. Vor allem bei Langzeitstudien mit mehrjährigen Beobachtungszeiten ließen sich bei diesem Patientenkontinuität hohe Ausfallraten im Sinne von Studienabbruchern nicht vermeiden und die festgelegten Grenzen zögen eine Nichtberücksichtigung wichtiger Ergebnisse nach sich. Diese Kritik beruht insofern auf einem Missverständnis, als dass sich - wie im Methodenteil beschrieben - die beiden oben genannten Grenzen auf den Anteil der Patienten beziehen, die in den Analysen nicht berücksichtigt wurden. Dieser Anteil ist nicht mit dem der Studienabbrucher gleichzusetzen. Hier gilt, wie für andere Studienpopulationen, dass für eine adäquate Interpretation berichteter Ergebnisse zum einen die Gründe für das Ausscheiden angegeben werden und zum anderen Studienabbrucher im Rahmen einer adäquaten Intention-to-Treat-Analyse in die Auswertungen einbezogen werden müssen. Darüber hinaus ist der Anteil von mindestens 70 % der randomisierten bzw. zugeteilten Patienten, die in der Auswertung verbleiben müssen, um deren Ergebnisse noch in die Bewertung aufzunehmen, vergleichsweise niedrig. International wird bereits ein Anteil von 20 % nicht in der Auswertung berücksichtigter Patienten als für eine adäquate Interpretation unzureichend angesehen [3].

2.3 Auswahl patientenrelevanter Endpunkte

Bezüglich der in der Methodik festgelegten patientenrelevanten Zielkriterien wurde in der Anhörung darauf hingewiesen, dass der Wunsch der Patienten nach geringerer Invasivität und Morbidität sowie nach Sicherheit der Anwendung und die Lebensqualität zu erfassen seien, um in die Nutzenbewertung einbezogen zu werden. Diese Zielkriterien werden mit jenen für den Berichtsplan Version 1.0 [1] definierten Zielkriterien erfasst. So wird die Invasivität einerseits mit der Notwendigkeit bzw. mit der Dauer eines Krankenhausaufenthalts und andererseits mittels der Darstellung unerwünschter Ereignisse erfasst. Die Morbidität und die Sicherheit der Anwendung werden ebenfalls im Rahmen der unerwünschten Ereignisse betrachtet. Die Lebensqualität wurde als Zielkriterium in der Methodik definiert (s. hierzu auch nächster Abschnitt). Daher werden die in der Anhörung genannten Zielkriterien in die Nutzenbewertung einbezogen, sofern die Studien hierzu Informationen liefern. Es erfolgte auch der Hinweis, dass die Betrachtung subjektiver Parameter unzureichend sei und sog. klinische Parameter wie Restharnfreiheit, Normalisierung der Miktions, Reduktion der Blasenwandhypertrophie und die Beseitigung einer etwaigen Nierenstauung zu berücksichtigen seien. Die Normalisierung der Miktions wird im Rahmen der unerwünschten Ereignisse dargestellt und vom Institut als patientenrelevant erachtet. Die Reduktion der Blasenwandhypertrophie ist ein Surrogat, und es bleibt unklar, wofür diese stehen soll und wie sie zu messen ist. Die Blasenwandhypertrophie (auch Trabekulierung der Blase) gilt als klassisches Merkmal der Obstruktion, jedoch weisen wiederum nicht alle Patienten mit einer intravesikalen Obstruktion eine

Trabekulierung auf [4]. Nach Grünewald et al. ist die Wertigkeit dieses Befunds daher umstritten [4]. Auch eine etwaige Nierenstauung stellt aus Sicht des Instituts ein Surrogat dar, da auch hier unklar ist, was sie messen soll. In der Literatur fanden sich uneinheitliche Angaben zur Diagnostik der Nierenfunktion. Die Leitlinie der deutschen Urologen [5] empfiehlt bei erhöhtem Serum-Kreatinin eine Untersuchung der Nierenfunktion, da bei Patienten mit eingeschränkter Nierenfunktion mit einer erhöhten Komplikationsrate gerechnet werden müsse. Roehrborn et al. weisen allerdings daraufhin, dass zunehmend weniger Patienten eine durch ein benignes Prostatasyndrom (BPS) eingeschränkte Nierenfunktion zum Zeitpunkt der ersten Konsultation aufweisen [6].

Die Erhebung der Lebensqualität kann nur umfassend sein, wenn sie mit validierten Instrumenten vorgenommen wird, die den Anforderungen der Multidimensionalität gerecht werden. Die Diskussion im Rahmen der wissenschaftlichen Erörterung ließ erkennen, dass die häufig in den Studien eingesetzte sog. 1-Item-Frage des International Prostatic Symptom Score (IPSS) sich auf die symptombezogene Zufriedenheit des Patienten bezieht. Dies wird durch eine Arbeit von van Venrooij [7] gestützt, die im Rahmen der Erörterung von den Stellungnehmenden zitiert wurde. Das Ergebnis einer Untersuchung von van Venrooij et al. zeigte, dass der IPSS und die Ergebnisse der 1-Item-Frage gut miteinander korrelieren. Daher spiegelt sie möglicherweise einen Teil der Lebensqualität wider, aber nicht die Multidimensionalität der Lebensqualität, die mindestens eine psychische, eine körperliche und eine soziale Komponente umfasst [8]. Ungeachtet dessen werden die Ergebnisse zu dieser 1-Item-Frage in die Nutzenbewertung einbezogen, da sie – auch von den Stellungnehmenden – als patientenrelevante Größe angesehen wird.

2.4 Nutzenbewertung

Im Rahmen der Anhörung entstand die Diskussion um die Definition eines Zusatznutzens von Alternativverfahren im Vergleich zur Standardbehandlung. Es kristallisierten sich 2 mögliche Herangehensweisen zur Beantwortung dieser Frage heraus. Da Patienten aufgrund ihres Leidensdrucks, der durch die Symptomatik maßgeblich bestimmt sei, einen Arzt aufsuchten und eine Behandlung initiiert werde, könne ein primäres Ziel für Patienten die Verbesserung der Symptomatik und ggf. damit verbunden der Lebensqualität sein. Ein Zusatznutzen gegenüber der Standardbehandlung sei demnach dann gegeben, wenn das entsprechende weniger invasive Verfahren zu einer stärkeren positiven Beeinflussung der Symptomatik führen werde. Von den Stellungnehmenden wurde die folgende Herangehensweise vorgeschlagen: Die Rationale zur Entwicklung und Einführung der weniger invasiven Alternativverfahren sei, dass diese bei vergleichbaren oder nur geringfügig schlechteren Ergebnissen hinsichtlich der Symptomatik zu geringeren Raten von unerwünschten Ereignissen bzw. Komplikationen führten. Ein Zusatznutzen liege daher in der Kombination von höchstens irrelevanter Unterlegenheit des Alternativverfahrens bzgl. der Symptomatik und gleichzeitig nachgewiesenen Vorteilen hinsichtlich des Auftretens von Komplikationen. Dieser Vorschlag führte die Diskussion zu der Frage, was ein irrelevanter Gruppenunterschied hinsichtlich des IPSS sei. Es stellte sich

heraus, dass es zur Beantwortung dieser Frage bisher keine Evidenz gebe. Die Definition eines Irrelevanzbereiches ist jedoch Voraussetzung, um diesen Vorschlag umzusetzen. Dies insbesondere dann, wenn Studien nicht darauf ausgerichtet sind, genau diese Frage zu untersuchen, d. h. also nicht als Äquivalenzstudien geplant wurden und eine solche Beurteilung post hoc auf Grundlage der Studienergebnisse erfolgen soll. In einer Arbeit von Barry et al. [9], die im Zusammenhang mit klinisch irrelevanten Unterschieden von den Stellungnehmenden zitiert wurde, untersuchten die Autoren den spürbaren intraindividuellen Unterschied vor und nach der Behandlung. Es wurde im Rahmen der Erörterung in diesem Zusammenhang auch die WHO-Leitlinie [10] zitiert, die auf weitere Untersuchungen zur Erfahrbarkeit von Veränderungen beim IPSS verweise. Es wurden über die Arbeit von Barry et al. hinaus jedoch keine Arbeiten indentifiziert. Es muss prinzipiell zwischen Grenzen für eine irrelevante Veränderung eines Patienten im Beobachtungsverlauf und Irrelevanzgrenzen für Gruppenunterschiede unterschieden werden. Da eine Umrechnung von intraindividuellen Unterschieden zu Gruppenunterschieden von diversen Annahmen und Spezifikationen (u. a. der Verteilung der Unterschiede) abhängt, sind Grenzwerte für individuell spürbare Veränderungen nicht direkt auf Gruppenunterschiede übertragbar. Um sich einer Definition eines solchen Irrelevanzbereiches für einen Gruppenunterschied bezüglich des IPSS zu nähern, wird das Institut vorhandene Hinweise in der Literatur heranziehen. Die daraus resultierenden Überlegungen werden im Berichtsplan in seiner (abschließenden) Version 2.0 dargestellt.

Darüber hinaus wurde im Rahmen dieser Diskussion von den Stellungnehmenden darauf hingewiesen, dass zur Interpretation der Ergebnisse eine Hierarchisierung der unerwünschten Ereignisse vorgenommen werden müsse, um eine Nutzen-Schaden-Abwägung vornehmen zu können. Die Diskussion zeigte, dass sich eine Hierarchisierung der einzelnen unerwünschten Ereignisse als schwierig erweisen könnte, insbesondere wenn es darum geht, mehrere nicht schwerwiegende unerwünschte Ereignisse in eine Reihenfolge zu bringen. Das Institut wird, sofern in den Publikationen beschrieben, eine Einteilung der unerwünschten Ereignisse in bspw. schwerwiegend/nicht schwerwiegend vornehmen.

2.5 Einteilung bzw. Klassifizierung von lokalen Behandlungsverfahren

Als sog. Standardverfahren werden für den Auftrag N04-01 die Transurethrale Resektion der Prostata (TURP), die Transurethrale Inzision der Prostata (TUIP) und die Adenomektomie zusammengefasst. Der sich unterscheidende Indikationsbereich dieser Verfahren – abhängig von der Prostatagröße – wurde bereits bei der Beschreibung der Verfahren im Berichtsplan der Version 1.0 berücksichtigt.

Die Transurethrale Vaporisation der Prostata (TVP) und die Transurethrale Vaporesektion der Prostata (TUVRP) wurden im Berichtsplan (Version 1.0) als Modifikationen der TURP betrachtet und daher ebenfalls den Standardverfahren zugeordnet. Das Vorgehen wurde vom Institut gewählt, da diese Verfahren aus geringfügigen Weiterentwicklungen der TURP entstanden sind. Im Rahmen der Anhörung wurde bestätigt, dass die TUVRP durchaus als Modi-

fikation der TURP betrachtet werden könne. Jedoch sei dies für die TVP zu hinterfragen, da hier das resezierende Element fehle und dieses Verfahren daher zu unterschiedlichen Gewebefeffekten führe. Das Institut ist nicht die einzige Institution, die sich für ein solches Vorgehen entschieden hat. So dokumentiert bspw. die Agency for Healthcare Research and Quality (AHRQ) in ihrem Bericht aus dem Jahr 2004 [11] die gleiche Einteilung, und auch ein kanadisches Health Technology Assessment (HTA) [12] beschreibt die TVP als Modifikation der TURP. Trotzdem wird dem nachvollziehbaren Hinweis aus der Anhörung nachgegangen. Dies wird mittels Sensitivitätsanalysen dort geschehen, wo die TURP und TVP als gleichwertige Verfahren betrachtet werden.

Darüber hinaus erfolgte in 2 Stellungnahmen die Anmerkung, dass im Zusammenhang mit der TURP in jüngster Zeit nicht mehr vom „Goldstandard“ gesprochen werde, sondern von der „Referenzmethode“, was auch in der aktuellen Leitlinie der World Health Organization (WHO) [13] zu finden ist. Dies wird für den Berichtsplan der Version 2.0 und die folgenden Berichtsversionen (Vorbericht [Version 2.0] und Abschlussbericht) übernommen.

Als eine Änderung des vorläufigen Berichtsplans (ohne wesentliche Konsequenz) wurde der Verzicht auf eine Klassifizierung der TUMT beschrieben. Dieses Vorgehen wurde gewählt, da eine häufig angewendete Klassifizierung in Hoch- und Niedrigenergie TUMT (HE- und NE-TUMT) lediglich für Geräte eines Vertreibers eindeutig operationalisiert ist. Für andere Geräte finden sich teils widersprüchliche Angaben in der Literatur. Der Versuch, eine Einteilung nach anderen technischen Merkmalen vorzunehmen, wie von den Autoren des Berichts der AHRQ 2004 [11] vorgenommen, zeigte sich als ebenfalls nicht zielführend (s. Methodik Vorbericht Version 1.0). Aus diesem Grund wurde für die Bearbeitung des Auftrags N04-01 auf eine Klassifizierung ganz verzichtet. In der Anhörung erfolgte der Hinweis, dies sei nicht adäquat, da unterschiedlich starke Gewebefeffekte nach HE- oder NE-TUMT stattfänden und daher auch unterschiedliche Effekte hinsichtlich der patientenrelevanten Therapieziele. Dies führe zu einer nicht angemessenen Betrachtung der Ergebnisse der HE-TUMT im Vergleich zu den Standardverfahren. Um diesem Hinweis zu begegnen, können auch hier Sensitivitätsanalysen durchgeführt werden, in denen klar klassifizierbare (HE- und NE-TUMT), aber auch nicht klassifizierbare Geräte voneinander getrennt betrachtet werden.

Die TUMT betreffend wurde von den Stellungnehmenden eine Publikation von Mattiasson et al. [14] eingebracht, die 5-Jahres-Ergebnisse einer im Vorbericht Version 1.0 einbezogenen Studie dokumentiert. Die Arbeit von Mattiasson et al. wurde nach der letzten Nachrecherche des Institut veröffentlicht und wird in den Vorbericht Version 2.0 aufgenommen.

Darüber hinaus erfolgte in der Anhörung der Hinweis auf den methodischen Umgang hinsichtlich der Zusammenfassung der Hybrid-Verfahren. Dies sei nicht angemessen, da unter „Hybrid“ eine Vielzahl unterschiedlicher Verfahren zusammengefasst werde. Für die weitere Bearbeitung des Auftrags wird geprüft, ob eine Zusammenfassung von in Publikationen vergleichbar beschriebenen Kombinationen möglich ist.

Hinsichtlich der Einteilung der Kontakt-Laserablation der Prostata (CLAP) und Visuelle Laserablation der Prostata (VLAP) erfolgte der Hinweis, dass nur jene Publikationen der CLAP zugeordnet werden könnten, in denen das eingesetzte Verfahren eindeutig mit dem Begriff „Kontakt“ beschrieben werde. Dies entspricht der Vorgehensweise des Instituts und wird bei der Erstellung des Vorberichts der Version 2.0 geprüft.

2.6 Wirtschaftlichkeit

Mehrfach wurde im Rahmen der Anhörung darauf hingewiesen, dass die Wirtschaftlichkeit als wichtiger Teilaspekt berücksichtigt werden müsse. Für eine ökonomische Analyse hat das Institut keinen Auftrag erhalten. Sie ist demzufolge nicht Gegenstand des Berichts.

3 Literatur

1. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Nichtmedikamentöse, lokale Verfahren zur Behandlung der benignen Prostatahyperplasie: Auftrag N04-01. Berichtsplan Version 1.0. Köln: IQWiG; 2005.
2. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Nichtmedikamentöse, lokale Verfahren zur Behandlung der benignen Prostatahyperplasie: Auftrag N04-01. Dokumentation und Würdigung der Stellungnahmen zum Berichtsplan Version 1.0. Köln: IQWiG; 2007.
3. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002; 359(9308): 781-785.
4. Grünewald U, Jonas U. Blasenentleerungsstörung. In: Höfner K, Stief C, Jonas U (Ed). *Benigne Prostatahyperplasie: ein Leitfaden für die Praxis*. Berlin: Springer; 2000. S. 110-136.
5. Berges R, Dreikorn K, Höfner K, Jonas U, Laval K, Madersbacher S, Michel M et al. Leitlinien der Deutschen Urologen zur Diagnostik des benignen Prostatasyndroms (BPS). *Urologe A* 2003; 42(4): 584-590.
6. Roehrborn C, McConnell J. Internationale- und USA-Leitlinien. In: Höfner K, Stief C, Jonas U (Ed). *Benigne Prostatahyperplasie: ein Leitfaden für die Praxis*. Berlin: Springer; 2000. S. 530-556.
7. van Venrooij GE, Boon TA, de Gier RP. International prostate symptom score and quality of life assessment versus urodynamic parameters in men with benign prostatic hyperplasia symptoms. *J Urol* 1995; 153(5): 1516-1519.
8. Schwarz R, Bernhard J, Flechtner H, Küchler T. Stellenwert des Begriffes „Lebensqualität“ in der Onkologie. In: Schwarz R (Ed). *Lebensqualität der Onkologie*. München: Zuckschwerdt; 1990. S. 3-10.
9. Barry MJ, Williford WO, Chang Y, Machi M, Jones KM, Walker-Corkery E, Lepor H. Benign prostatic hyperplasia specific health status measures in clinical research: How much change in the American Urological Association symptom index and the benign prostatic hyperplasia impact index is perceptible to patients? *J Urol* 1995; 154(5): 1770-1774.
10. Abrams PDC, Griffiths D, Van Kerrebroeck P, Nishizawa O, Nitti N, Keong Tatt F et al. Lower urinary tract symptom: Etiology, patient assessment and predicting, outcome from therapy. In: McConnell J, Abrams P, Denis L, Khoury S, Roehrborn C (Ed). *Proceedings of the 6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases: Male lower urinary tract dysfunction. Evaluation and management*. Paris: Editions 21; 2006. S. 69-142.

11. Agency for Healthcare Research and Quality. Treatments for benign prostatic hyperplasia. Rockville (MD): AHRQ; 2004.
12. The Medical Advisory Secretariat, Ontario Ministry of Health and Long-Term Care. Energy delivery systems for treatment of benign prostatic hyperplasia: Health technology policy assessment. Toronto: MAS; 2006.
13. De La Rosette JJ, Baba S, Badlani G, Elhilali M, Gravas S, Muschter R et al. New minimally Invasive and Surgical developments in the management of BPO. In: McConnell J, Abrams P, Denis L, Khoury S, Roehrborn C (Ed). Proceedings of the 6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases: Male lower urinary tract dysfunction evaluation and management. 2006. S. 195-233.
14. Mattiasson A, Wagrell L, Schelin S, Nordling J, Richthoff J, Magnusson B, Schain M et al. Five-year follow-up of feedback microwave thermotherapy versus TURP for clinical BPH: a prospective randomized multicenter study. Urology 2007; 69(1): 91-96.

4 Dokumentation der Stellungnahmen

4.1 Darlegung potenzieller Interessenkonflikte der Stellungnehmenden

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

Tabelle 1: Stellungnahmen von Fachgesellschaften, Firmen und Organisationen

Institution / Firma	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Arbeitskreis BPH der Akademie der Deutschen Urologen	Dr. R. Berges	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
Lumenis GmbH	H. A. Harms	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	M. Martin	Ja	Nein	Nein	Nein	Nein	Ja	Ja	Nein
Arbeitskreis BPH der Akademie der Deutschen Urologen	Prof. Dr. K. Höfner	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Nein
	Prof. Dr. K. Dreikorn	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	Dr. R. Berges	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	PD Dr. O. Reich	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	Prof. Dr. R. Muschter	Nein	Nein	Ja	Ja	Nein	Nein	Nein	Nein
	Dr. M. Oelke	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	Dr. W. Rulf	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	Dr. C. Tschuschke	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
Österreichische Gesellschaft für Urologie und Andrologie (ÖGU)	Doz. Dr. S. Madersbacher	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	Prof. Dr. G. Schatzl	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	Prim. Doz. Dr. M. Rauchenwald	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
Deutsche Krankenhausgesellschaft e. V.	Joachim Manz	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
Kassenärztliche Bundesvereinigung	Dr. P. Rheinberger	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	Dirk Horenkamp	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	Dr. R. Schiffner*	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein

Institution / Firma	Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Deutsche Krankenhausgesellschaft e. V.	Dr. N. Schlottmann	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	Dr. C. Simon	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	Dr. K. Schorn	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
Medtronic GmbH	C. Weinrank	Nein	Nein	Nein	Nein	Nein	Nein	Nein	Nein
	W. Frisch	Darlegung der potenziellen Interessenkonflikte liegt nicht vor.							
	S. Zenk	Ja	Nein	Nein	Nein	Nein	Ja	Nein	Nein
Erläuterung: „*“ = Teilnahme an der wissenschaftlichen Erörterung als Vertretung der Stellungnehmenden.									

Tabelle 2: Stellungnahmen von Einzelpersonen

Name	Frage 1	Frage 2	Frage 3	Frage 4	Frage 5	Frage 6	Frage 7	Frage 8
Prof. Dr. K. Höfner	Nein	Ja	Nein	Nein	Nein	Nein	Nein	Nein
Prof. Dr. R. Muschter	Nein	Nein	Ja	Ja	Nein	Nein	Nein	Nein
Prof. Dr. J. J. M. C. H. de la Rosette	Nein							
Dr. C. von Zastrow	Nein	Ja	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 Firma¹ abhängig (angestellt) beschäftigt, die von den Ergebnissen Ihrer wissenschaftlichen Arbeit für das Institut² finanziell profitieren könnte?

Frage 2: Beraten Sie oder haben Sie innerhalb der letzten 3 Jahre eine Person, Institution oder Firma direkt oder indirekt³ 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?

4.2 Stellungnahmen von Fachgesellschaften, Firmen und Vereinigungen

4.2.1 Arbeitskreis BPH der Akademie der deutschen Urologen

Autoren:

Dr. Richard Berges

Adresse:

Pan-Klinik Köln

Zeppelinstr. 1

50667 Köln

Dieser Kommentar zum Report “Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie” des Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) bezieht sich auf das Kapitel Transurethrale Mikrowellen Thermotherapie (TUMT; Seiten 258-332).

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien

Der Vorbericht berücksichtigt kürzlich publizierte 5-Jahres Langzeitdaten zur HE-TUMT im direkten Vergleich zur TUR-P nicht, welche mit einer technischen Weiterentwicklung herkömmlicher HE-TUMT, der PLFT-TUMT durchgeführt wurden. Die technische Innovation dieser TUMT-Variante besteht in der Möglichkeit, die in der Prostata während der Behandlung erzielten Temperaturen direkt zu messen und auf den Situs bedingten Hitzeabtransport (Patientenindividuelle Durchblutung der Prostata, eine höchst variable Größe) durch Anpassung von Mikrowellenenergie oder Behandlungsdauer zu reagieren. Durch dieses System wird erreicht, dass die angestrebten und zur Ablation von Prostatagewebe notwendigen Temperaturen im Individualfall auch erreicht werden. Daten dieser Untersuchung wurden bei der erneuten Analyse der HE-TUMT Daten in diesem Kommentar berücksichtigt[1].

Weitere Studien s. Literaturverzeichnis.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien

Korrektur der Analyse von HE-TUMT mit der Referenzmethode: Methodisch erfolgt im Vorbericht eine unzulässige Vermischung eines nichtablativen Behandlungsverfahrens mit einem Ablativen Behandlungsverfahren in einer gemeinsamen Analyse und Bewertung mit dem Referenzverfahren: Um Möglichkeiten und Grenzen der TUMT wirklichkeitsnäher abzubilden, haben wir die vorhandenen Daten erneut nach denselben statistischen Methoden wie im Vorbericht analysiert. Allerdings haben wir aus den vorgenannten Gründen nur solche Studien bei der Analyse berücksichtigt, die HE-TUMT mit der TUR-P vergleichen. Zusätzlich haben wir eine separate Analyse von PLFT-TUMT und TUR-P unter Verwendung der kürzlich publizierten 5-Jahres Daten durchgeführt[1]. Diese wurden durchgeführt mit der Cochrane Collaboration RevMan 4.2 Software (wenn in der Studie sowohl Mittelwert als auch Standardabweichung (SD) publiziert wurden).

Basierend auf dem Vorbericht sowie einer eigenen Medline-Suche wurden 10 Publikationen über Vergleiche von HE-TUMT mit TURP identifiziert (s. Tab. 1-3). Sieben dieser Publikationen wurden mit Prostatron (SoftwareVersion 2.5)[3-9], 3 mit PLFT durchgeführt [1, 10, 11]. Von den 7 Publikationen mit Prostatron, wurden 4 durch D’Ancona[5], Francisca[8] und Floratos[9] /De la Rosette[3] veröffentlicht (alle aus Nijmegen, NL). Da die Patientenrekrutierung von Francisca und Floratos/ De la Rosette gleich, und die Ergebnisse dieser Studien beinahe identisch sind, ist davon auszugehen, dass es sich hier um die selbe Patientenpopulation handelt. Gleichwohl existieren kleinere Unterschiede hinsichtlich behandelte Patientenzahlen und weiterer klinischer Daten in den 3 Artikeln von Francisca, Floratos and De la Rosette. Da die Publikation von De la Rosette sowohl Mittelwerte als auch Standardabweichungen beinhaltet und gleichzeitig als Basis für eine Kostenanalyse diente, wurde diese Publikation für die Meta-Analyse bei „Baseline“, 12, 24 und 36 Monate benutzt. Für die 3 Monatsdaten haben wir die Publikation von Francisca verwendet.. So wurden insgesamt 5 direkte Vergleichstudien zwischen HE-TUMT und TURP für die Analyse benutzt (4 x Prostatron 2.5, 1x PLFT).

Diese Studien vergleichen HE-TUMT mit der TUR-P vornehmlich bei Männern ≥ 45 Lebensjahren mit einem IPSS Score ≥ 7 oder $r \geq 12-13$ bzw. einem Madsen Score ≥ 8 , einem $Q_{\max} \leq 12-15$ mL/s und einem Prostatavolumen zwischen 30 und 100 mL; Patienten mit Harnverhaltung wurden ausgeschlossen aus den meisten dieser Untersuchungen. Die Zahl der Behandlungsfälle in diesen Studien schwankte zwischen 50 und 150 bei einer Follow-up Periode zwischen 6 und 60 Monaten.

Klinische Effektivitätsparameter in Abhängigkeit des Energie/Hitzeintrags ins Gewebe: Im Vorbericht wird auf Seite 11 richtig erkannt, dass Niedrigenergie-TUMT (NE-TUMT, Gerät: Prostatron 2.0) und Hochenergie-TUMT (HE-TUMT, Geräte: Prostatron 2.5, Targis™ oder ProstaLund® Feedback Therapy (PLFT)/CoreTherm™) unterschiedliche erzielen. Trotzdem werden gleichwohl Studien, die mit Niedrigenergie-TUMT durchgeführt wurden, in einer gemeinsamen Meta-Analyse hinsichtlich Symptomenreduktion (IPSS-Score), Lebensqualität (QoL) und maximalem Harnfluss (Q_{\max}) analysiert. Zitiert werden auf Seite 317 im Vorbericht Ergebnisse der Studie von Nørby et al. (2002)[4], in der sowohl HE- als auch NE-TUMT (Prostatron 2.0 und 2.5) angewandt wurden, für die Gesamtgruppe, obwohl im Artikel diskutiert wird, dass die Behandlungsergebnisse unter Verwendung von HE-TUMT besser ausfielen als mit NE-TUMT. Auch die EAU erkennt an, dass Behandlungsergebnisse unter LE- und HE-TUMT zugunsten letzterer unterschiedlich ausfallen[12]. In den deutschen S2 Leitlinien wird dem unterschiedlichen Energieeintrag in der Weise Rechnung getragen, dass zwischen ablativen und nichtablativen operativen Techniken unterschieden wird[13]. Es besteht ein unmittelbarer Zusammenhang zwischen Ablation und Symptomenverbesserung, dies wird auch im Vorbericht erkannt. Nicht erkannt wird der Zusammenhang zwischen Ablation und Invasivität respektive Morbidität, aus dem sich unmittelbar die Patientenreferenz ergibt und aus dessen Grunde alle Alternativverfahren überhaupt entwickelt wurden.

Behandlungseffizienz:

Symptome: Tab. 1 fasst die Auswirkungen von HE-TUMT und TURP auf die Symptomenlinderung zusammen. Der I-PSS (nahezu identisch mit dem AUA Symptom Index), ist ein in vielen Sprachen validierter Symptomen-Score welcher als Standardinstrument von EAU, AUA, DGU und der „International Consultation on BPH“ empfohlen wird. Für die Analyse wurden daher nur Daten berücksichtigt, die mit diesem Score erhoben wurden. Aus der Tabelle ist zu entnehmen, dass die meisten Patienten in diesen randomisierten Studien unter mittelschweren bis schweren Symptomen litten (I-PSS um die 20 Punkte). HE-TUMT reduzierte den mittleren I-PSS statistisch signifikant um mindestens 50% und bis zu 65% gegenüber den Ausgangsbefunden. D.h., Symptome wurden von mittel/schwer (=international anerkannte Behandlungsindikation) auf milde Symptomatik (I-PSS 0-7) herabgesetzt (= keine Behandlungsindikation). Dieser Effekt bleibt über 5 Jahre bestehen. Abb. 1 zeigt die Ergebnisse der Meta-Analyse dieser Daten. Daraus ist abzulesen, dass TURP der HE-TUMT leicht überlegen ist, statistisch signifikant nach 3, 24 und 30-36 Monaten, nicht aber nach 6 und 12 Monaten ($P < 0.05$). Wurden die Kurzzeitdaten (3 Monate), aus der Analyse herausgerechnet, beträgt die gewichtete mittlere Differenz (WMD) zwischen beiden Behandlungsgruppen 1.3 bis 4.7 Punkte. Der Unterschied in der Arbeit von Mattiasson (PLFT) lag bei 0.10-3.2 Punkten und erscheint damit kleiner als unter der Verwendung des Prostatron-Gerätes. Um eine leichte Symptomenverbesserung überhaupt wahrnehmen zu können, muss sich ein medianer IPSS von 16 Punkten um wenigstens 3 Punkte verringern[14]. Bei Patienten mit schweren Symptomen (I-PSS 20-35), muss eine Therapie den Score sogar um wenigstens 6.1 Punkte senken, bevor diese Verbesserung erfahrbar wird. Um eine moderate Verbesserung festzustellen, muss durch eine Therapie sogar eine Scoreverbesserung gegenüber dem Ausgangswert von 5.1 (mittelschwere Symptomatik) bzw. 8.7 Punkten (schwere Symptomatik) erfolgen. Es ist also fragwürdig, ob die in diesen Studien behandelten Patienten den Unterschied von 1.3-4.7 Punkten bei HE-TUMT verglichen mit der TURP wirklich hätten bemerken können. Es stellt sich somit die Frage der klinischen Relevanz des Unterschiedes zur Referenzmethode.

Lebensqualität: Tab. 2 zeigt die Effekte von HE-TUMT und TURP auf die Lebensqualität (QoL) gemessen mit dem krankheitsspezifischen validierten Fragebogen IPSS, respektive der darauf bezogenen Frage zur LUTS-bezogenen Lebensqualität und einem mehr generischen Instrument zur Erfassung von QoL (einer der Kritikpunkte im Vorbericht des IQWiG ist, dass QoL nur mit multidimensionalen Instrumenten erfasst werden kann. Die Autoren machen den Fehler, dass es hier nicht um die multidimensionale Erfassung von QoL geht, sondern um die Erfassung von LUTS-bezogener QoL, für die eben dieses Instrument auch umfangreich validiert wurde, der IPSS ist *das*

empfohlene Instrument zur Erfassung LUTS-bezogener QoL der „International Consultation on BPH [15]).

Sowohl HE-TUMT als auch TURP haben einen positiven Effekt auf QoL hinsichtlich der Verbesserung von LUTS bezogener QoL als auch hinsichtlich der Verbesserung von QoL gemessen mit multidimensionalen Instrumenten. Die statistisch signifikante Verbesserung von I-PSS-QoL betrug 50-70% vom Ausgangswert. Diese Wirkung hält auch in der Langzeitbeobachtung über 5 Jahre an.

Abb. 2 zeigt die Daten der Meta-Analyse für I-PSS-QoL. Diese Daten zeigen dass die geringgradig schlechtere Verbesserung von QoL durch HE-TUMT verglichen mit TURP (0.6-1.0 Punkte Unterschied) statistisch nicht signifikant ist. Der Unterschied zwischen PLFT-TUMT und TURP ist nochmals geringer (-0.1 to 0.3 Punkte).

Harnflussrate: Die Auswirkungen von TUMT und TURP auf Q_{max} werden in Tab. 3 dargestellt. In allen untersuchten Studien außer der von Ahmed zeigte sich eine statistisch signifikante Verbesserung von Q_{max} durch HE-TUMT im Vergleich zum Ausgangswert von mindestens 50% bis hin zu 65%. Auch dieser Effekt bleibt über 5 Jahre erhalten. Der Unterschied in der Harnstrahlverbesserung ist am ehesten mit dem Ausmaß des Hitzeeintrags korreliert, zum Beispiel sind dies nur 81 kJ in der Studie von Ahmed, aber 152 kJ in der von D'Ancona.

Abb. 3 zeigt die Daten der Meta-Analyse. Der Unterschied zwischen HE-TUMT und TURP fällt scheinbar zwar zugunsten der TURP aus, aber dieser Unterschied war an allen Datenpunkten entweder statistisch nicht signifikant ($P > 0.05$) oder erreichte gerade Signifikanz. Wurde Signifikanzniveau erreicht, muss dies der Untersuchung Francisca/De la Rosette oder Ahmed zugeschrieben werden. Der MWD lag zwischen 1.6 und 5.9 mL/s, bei der PLFT-TUMT zwischen 0.3 und 3.2 mL/s. Dieser geringe Unterschied ist wiederum vom Patienten nicht erfahbar, somit stellt sich die Frage nach der klinischen Relevanz. In den meisten Studien mit HE-TUMT lag der Q_{max} nach Behandlung zwischen 12 und 15 mL/s. Dies entspricht der physiologischen Spannweite für asymptomatische Männer dieser Altersklasse bei einem Miktionsvolumen von 250ml, deren Q_{max} etwa 14 mL/s beträgt[16]. Die Q_{max} Werte nach TURP von 20-25 mL/s sind in dieser Hinsicht als supranormal anzusehen und damit nicht Therapieziel.

Re-Interventionen: Abb. 4 zeigt die jährlichen Re-Interventionsraten in Langzeituntersuchungen. Diese zeigen mit Ausnahme der kürzeren D'Ancona Studie[5] ungefähr 2-fach höhere Re-Interventionsraten nach HE-TUMT verglichen zur TURP, allerdings auf einem akzeptablen Level. Dies war in der Untersuchung von De la Rosette[3] statistisch nicht signifikant.

Komplikationen: Tab. 4 listet die schweren und andere unerwünschte Ereignisse (Adverse Events = AEs) aus den direkten Vergleichsstudien. Daraus ergibt sich, dass die TURP mit häufigeren schweren AEs begleitet ist als die HE-TUMT (ca. 5% vs. 17%). Insbesondere schwere Blutungen (1% vs. 10%) mit resultierender Hospitalisation und Bluttransfusion sowie das gefürchtete TUR-Syndrom. Ebenso werden nach TURP Blasenhalssengen und Harnröhrenstrikturen beobachtet, die ebenfalls eine erneute stationäre Aufnahme und Re-Operation notwendig machen.

Die häufigsten AEs nach HE-TUMT sind initiale Urgency (30-40%), Harnwegsinfekte (20-30%) und akute Harnverhaltungen, die eine erneute Katheterisierung erforderlich machen (10-20%). Häufigste AEs nach TURP sind die retrograde Ejakulation (50-60%), Erektile Dysfunktion (10-20%), Haematurie, Harnwegsinfekte (5-12%), Harnverhaltungen (5-10%), Inkontinenz (5-10%) und persistierende Urgency.

Das Risiko einer akuten Harnverhaltung mit Katheterisierung von bis zu 14 Tagen (Abb. 5) ist das unbequemste AE nach HE-TUMT (und kann selbstverständlich auch nach TURP auftreten, allerdings in geringerer Frequenz). Da es sich bei der HE-TUMT um ein sekundär ablatives Behandlungsverfahren handelt, ist eine temporäre Harnableitung systemimmanent für ca. 14 Tage. Dies wird mit dem Pt. zuvor ausgiebig besprochen. Im übrigen werden heute zumindest nach PLFT-TUMT in der Regel temporäre Stents im Behandlungspaket mit integriert, die erstens eine infektfreie Ableitung postoperativ garantieren (Senkung der katheterbedingten Rate von

Harnwegsinfekten) und für den Patienten angenehmer zu tragen sind. Auch werden selbstauflösende Urethrastents eingesetzt mit gleichem Effekt[17].

Abhängigkeit der Anwendung vom Operateur: Es ist bekannt, dass der Ausgang einer TURP in hohem Maße Abhängig ist vom Ausbildungsstand und der Fähigkeit des Anwenders (lange Lernkurve, regelmäßiges Training). Im Gegensatz dazu ist die HE-TUMT oder PLFT-TUMT als einziges minimal-invasives Behandlungsverfahren weitgehend Anwenderunabhängig[9] und kann daher leicht erlernt werden und führt dann zu konstant guten Ergebnissen.

Invasivität: Für den Patienten ist gerade die Invasivität des Eingriffs und die damit verbundene potentielle Morbidität sowie die Notwendigkeit einer Narkose und eines stationären Aufenthalts für die Gesamteinschätzung seiner QoL wichtig. HE-TUMT ist im Gegensatz zur TURP und im Gegensatz zu allen anderen im Vorbericht des IQWiG untersuchten Verfahren ohne Narkose, und im Gegensatz zur TURP und nahezu allen anderen Verfahren immer ambulant durchführbar. Demgegenüber benötigt die TURP in jedem Fall eine Narkose und einen Krankenhausaufenthalt von 4-5 Tagen (Abb. 6)[3, 5, 18]. Die ambulante Durchführbarkeit des Verfahrens hat unmittelbar eine Kostenreduktion zur Folge, arbeitsfähigen Patienten ermöglicht sie zudem eine schnellere Wiedereingliederung in ihr Berufsleben.

Wegen der geringen Invasivität kann die HE-TUMT im Gegensatz zur TURP auch bei Hochrisikopatienten angewandt werden, bei denen eine Narkose nicht möglich ist oder eine Antikoagulantientherapie aufrecht erhalten bleiben muss. Dies erspart gerade diesen Patienten die ansonsten lebenslange Harnableitung bei BPS-bedingter Obstruktion.

Therapiekosten: International wurde in 2 Studien mit Langzeitvergleich (3-Jahre) eine Kostenanalysen zwischen HE-TUMT und TURP durchgeführt: Dies waren die niederländische Studie durchgeführt mit dem Prostatron-Gerät in Version 2.5, publiziert durch Floratos[9] und Francisca[8], sowie eine schwedisch-amerikanische Studie [2] mit PLFT-TUMT publiziert durch Wagrell[10, 11]. Die Vergleichsstudie LE-TUMT bzw. HE-TUMT (Prostatron-Gerät) mit TURP publiziert durch Nørby[4] beinhaltet ebenfalls eine Kostenanalyse[18]. Letztere hatte allerdings nur einen Beobachtungszeitraum von 6 Monaten, so dass Re-Interventionsraten möglicherweise unterschätzt wurden. Daher wurde diese Untersuchung aus unserer Analyse ausgeschlossen.

In der niederländischen Studie[3] wurden in die Kostenanalyse direkte medizinische Kosten einschließlich der Kosten für die Behandlungsprozedur selbst, die Hospitalisierungsdauer, die Kosten für Nachsorgeuntersuchungen, Antibiotika und medikamentöse Re-Interventionen über einen Zeitraum von 3 Jahren eingeschlossen. Die Gesamtkosten für den Beobachtungszeitraum von 3 Jahren wurden pro Fall (95% Konfidenzintervall) basierend auf Preisen in den Niederlanden von 1995 angegeben und betragen 2.385 (2.111-2.659) Euro für die HE-TUMT und 3.510 (3.203-3.817) Euro für die TURP. Die beobachtete Kostenreduktion von 32% durch die Anwendung der HE-TUMT (Abb. 7) wurde vornehmlich durch die ambulante Durchführbarkeit dieses Verfahrens erreicht (bei einer mittleren Hospitalisierungsdauer von 5.3 Tagen für die TURP).

Die schwedische Untersuchung[2] war primär eine “Resource Utilisation Study” über die eigentlichen Kosten der Behandlungsprozeduren, Kosten für mögliche und zu erwartende AEs mit Hospitalisierungsfolge und Kosten für Sekundärinterventionen über einen Zeitraum von 3 Jahren. Im Datensatz wurden im Einzelnen eingeschlossen: Daten aus klinischen Studien [11], Daten einer Registry von 52.010 stationären Behandlungen wegen einer TURP zwischen 1990 und 1995 in Schweden, sowie Daten aus einer retrospektiven Beobachtungsstudie über konsekutive Patienten in ambulanten Praxen, die zuvor eine PLFT-TUMT oder eine TURP erhalten hatten. Im ersten Jahr wurden Kosten von 1.763 Euro für die HE-TUMT errechnet und 3.209 Euro für die TURP. Über 3 Jahre kalkulierte man so Behandlungskosten von 2.151 Euro für eine PLFT-TUMT und 3.185 Euro

für die TURP. Auch in dieser Analyse errechnete sich somit wieder ein Kostenvorteil von 32% für die HE-TUMT gegenüber einer TURP in einem Zeitraum von 3 Jahren (Abb. 7).

In einer amerikanischen Modellberechnung (Decision-Analytic Markov Cost-Utility and Cost-Effectiveness-Model) wurde für einen 65-jährigen Mann mit mittelschwerer Symptomatik bei BPS sowohl bei Hochrisiko als auch bei nicht Hochrisikopatienten HE-TUMT als Behandlungsoption der Wahl gegenüber TURP errechnet, ebenfalls und zudem mit geringeren Behandlungskosten[19, 20].

Unterschiede einzelner HE-TUMT Geräte: Außer einer direkten Vergleichsstudie zwischen PLFT-TUMT und dem Prostatron Gerät in der Version 2.5 gibt es kaum Daten zu Unterschieden in der Effektivität einzelner TUMT-Systeme. Dennoch scheint es einen Vorteil der PLFT-TUMT gegenüber anderen Geräten zu geben, vergleicht man die Daten der Effektivitätsparameter in Tab. 1-3 und Abb. 1-3. Dieser Vorteil wird dem Umstand zugeschrieben, dass PLFT-TUMT als einziges Verfahren die tatsächlich in der Prostata erzeugte Hitze kontinuierlich während der Behandlung messen kann und so Behandlungsdauer und Mikrowellenenergie individuell an jeden Patienten angepasst werden können[21].

Einsatz der HE-TUMT bei speziellen Patientengruppen: Wie bereits erwähnt, beinhalten die 5 direkten Vergleichsstudien in der Hauptsache Patienten mit einem Prostatavolumen zwischen 30-100 mL und ohne vorherige Episode einer akuten Harnverhaltung. Eine kürzlich publizierte Untersuchung von Schelin vergleicht PLFT-TUMT (N=61) mit TURP (N=49) oder offener Adenomenukleation (N=10) bei insgesamt 120 BPS-Patienten mit einer relativ großvolumigen Prostata (mittleres Prostatavolumen 67-72 mL) nach akuter Harnverhaltung[22]. 78% der Patienten nach PLFT-TUMT waren nach 6 Monaten katheterfrei gegenüber 88% nach konventioneller Operation (TURP oder offene Adenomenukleation). Beide Methoden reduzierten die Symptomatik in den Bereich unter 7 Score Punkte (Behandlungsziel) (7.3 für PLFT-TUMT vs. 4.4 für konventionelle OP-Verfahren). Dieses war statistisch signifikant. Die geringe zusätzliche Reduktion des I-PSS unter Anwendung der Referenzmethoden von 2.9 Punkten entspricht dem Unterschied der Daten aus der Meta-Analyse. Eine klinische Relevanz dieses Unterschiedes besteht nicht, denn er wäre von den Behandelten kaum erfahrbar. Auch die BPS-spezifische QoL wurde sowohl durch PLFT-TUMT substanziell verbessert (IPSS QoL reduziert auf 1.4) wie auch durch die Referenzmethoden (IPSS QoL reduziert auf 0.8). Auch dies bestätigt die Daten aus der Meta-Analyse. Der mittlere Q_{max} wurde nach 6 Monaten ebenfalls, wie nach den Daten aus der Meta-Analyse zu erwarten, verbessert: Um 13.4 mL/s durch HE-TUMT und 18.0 mL/s durch die Referenzmethoden. Demgegenüber wurden schwere behandlungsbezogene AEs nur bei 2% nach PLFT-TUMT, aber bei 8% nach den Referenzmethoden beobachtet. Die häufigsten AEs nach PLFT-TUMT waren Harnwegsinfekte in 33% bzw. 22% nach den Referenzmethoden. Alle diese Parameter bestätigen die Daten der direkten Vergleichsstudie PLFT-TUMT/TURP mit einem Beobachtungszeitraum über 5 Jahre PLFT von Mattiasson[1].

3. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe wissenschaftlicher Literatur zur Begründung der Anmerkung.

Obwohl das IQWiG eine Meta-Analyse einer Reihe klinischer Studien durchgeführt hat, sind wir nicht der Ansicht, dass die daraus gezogenen Schlussfolgerungen der wirklichen Evidenz, abgebildet durch die verfügbaren Daten zur TUMT entsprechen.

Im Vorbericht des IQWiG wird festgestellt, dass die meisten Studien relevante Qualitätsmängel aufwiesen. Wir teilen diese Ansicht. Leider sind Qualitätsmängel gerade in Studien mit chirurgischen Inhalt eher die Regel als die Ausnahme. Eine ähnliche Analyse der Untersuchungen zur TURP z. B. im Vergleich zur offenen Operation würden ähnliche Qualitätsmängel aufweisen. Daher spricht die Urologie auch bei der TURP nicht mehr vom "Goldstandard", sondern eher von einer Referenzmethode. Andererseits wird, um eine Abschätzung über ein Therapieverfahren machen zu können, aus gutem Grund in Ermangelung einer "Highest Level Evidence" (I oder A) auf niedrigere Evidenzstufen zurückgegriffen[23, 24].

Ein schwerer methodischer Fehler liegt in der gemeinsamen Analyse eines nichtablativen Therapieverfahrens (Niedrigenergie-TUMT) mit einem ablativen Therapieverfahren (Hochenergie-TUMT, Gewebstemperaturen) im Vergleich zum Referenzverfahren: Niedrigenergie-TUMT erhebt nicht den Anspruch eines TUR-P-Äquivalents. Ein Vergleich dieses Verfahrens wäre allenfalls mit einer medikamentösen Therapie (z.B. Alpha-Blocker) sinnvoll. Niedrigenergie-TUMT ist als historische Entwicklungsvorstufe der Hochenergie-TUMT anzusehen und wird in diesem Zusammenhang nicht länger eingesetzt. Eine Nutzenbewertung macht also nur in dem Vergleich zwischen Hochenergie-TUMT und TURP Sinn. Die Datenvermischung in der Meta-Analyse des IQWiG verzerrt damit das Bild unzulässig zugunsten der Referenzmethode.

Ein schwerer methodischer Fehler liegt in der Interpretation der im Vorbericht zitierten Behandlungsunterschiede zwischen TUMT und dem Referenzverfahren. Es wurde nicht berücksichtigt, dass die in den Studien berichteten Unterschiede zugunsten der Referenzmethoden klinisch irrelevant sind. Um eine leichte Symptomenverbesserung gemessen mit dem IPSS überhaupt wahrnehmen zu können, muss sich ein medianer IPSS von 16 Punkten um wenigstens 3 Punkte verringern[14, 15]. Bei Patienten mit schweren Symptomen (I-PSS 20-35), muss eine Therapie den Score sogar um wenigstens 6.1 Punkte senken, bevor diese Verbesserung erfahrbar wird. Um eine moderate Verbesserung festzustellen, muss durch eine Therapie sogar eine Scoreverbesserung gegenüber dem Ausgangswert von 5.1 (mittelschwere Symptomatik) bzw. 8.7 Punkten (schwere Symptomatik) erfolgen. Es ist also unwahrscheinlich, ob die in den Vergleichsstudien aller im Vorbericht beurteilten Alternativverfahren, somit auch die mit HE-TUMT behandelten Patienten den Unterschied zugunsten der Referenzmethoden überhaupt bemerken könnten. Damit ist die Kernaussage des Vorberichts zu dieser Methode –wie zu allen anderen Alternativverfahren- falsch: Richtig ist, dass es sich hinsichtlich dieser Effizienzparameter um „klinisch“ äquivalente Therapieverfahren handelt.

Ein weiterer methodischer Fehler besteht in der vornehmlichen Analyse der Effektivitätsparameter IPSS, Lebensqualität, im Nachgang auch Uroflowmetrie, während die im Rahmen einer Operation für Patienten viel bedeutenderen Parameter wie Sicherheit der Anwendung, Invasivität und Morbidität nur geringe Beachtung finden: Geht man von einer subjektiv erfahrbaren Gleichwertigkeit der Verfahren (s.o.) hinsichtlich Symptomenverbesserung und QoL aus, muss der Hauptaugenmerk bei der Nutzenbewertung der einzelnen Verfahren auf einer möglichst schonenden, wenig invasiven Behandlung liegen. Methodisch ist zusätzlich zu kritisieren, dass keine genauere Wirtschaftlichkeitsanalyse erfolgte.

Nach Bewertung der selben publizierten Evidenz aus klinischen Studien wie in diesem Vorbericht kommen die Leitlinien der European Association of Urology (EAU[12]), der American Urological Association (AUA[25]) und der Deutschen Urologen[13] sowie die „International Consultations on Prostatic Diseases“ der WHO[26] zu dem Schluss, dass HE-TUMT eine Option zur Behandlung des benignen Prostatasyndroms (BPS) ist. Zusätzlich hat die Food and Drug Administration (FDA) der USA nach stringenten Prüfkriterien HE-TUMT als Behandlungsverfahren anerkannt. Es ist nur aufgrund der groben methodischen Mängel im Vorbericht nachvollziehbar, dass das IQWiG zu anderen Schlussfolgerungen in der Bewertung gelangt.

Zusammenfassung:

Nach der Analyse der Daten aus Vergleichsstudien zwischen HE-TUMT und TURP zeigen beide Verfahren bei BPS-Patienten eine vergleichbare Reduktion von Symptomen (LUTS), eine Verbesserung der Lebensqualität und eine Zunahme des maximalen Harnstrahls (Deobstruktion). Diese Effekte bleiben über einen Beobachtungszeitraum von 5 Jahren zumindest belegt für die PLFT-TUMT erhalten. Die Effekt auf die Symptomenreduktion ist signifikant besser unter der Referenzmethode TURP (zusätzliche I-PSS Reduktion um 1.3-4.7 Punkte), allerdings ist es unwahrscheinlich, dass dieser Unterschied von Patienten erfahrbar wäre. Daher ist eine klinische Relevanz für diesen Zusatzeffekt nicht gegeben. Zudem ist der zusätzliche geringe Vorteil in der Verbesserung der Lebensqualität nach der Referenzmethode gegenüber der HE-TUMT statistisch nicht signifikant und ebenfalls klinisch irrelevant. Wenn also diese Methoden für den Patienten keine erkennbaren Unterschiede im Behandlungsergebnis liefern, stellt sich die Frage, ob diese scheinbaren Nachteile der HE-TUMT gegenüber der Referenzmethode nicht hinsichtlich der Vorteile der HE-TUMT bei potentiellen und beobachteten Raten schwerer AEs wie Blutungen mit Transfusionsfolge (Infektionsgefahr etc.), Re-Hospitalisierung, TUR-Syndrom, Blasenhals- oder Harnröhrenstrikturen (Re-Hospitalisierung und Re-Operation erforderlich) nach der Referenzmethode überhaupt ins Gewicht fallen. Die Patientenpräferenz liegt hier jedenfalls eindeutig auf Seiten der HE-TUMT.

Ebenso muss von einem gesundheitsökonomischen Standpunkt aus davon ausgegangen werden, dass die HE-TUMT auch in Deutschland als Behandlung nicht mehr kosten wird, als die TURP. Nach den Daten aus anderen Gesundheitssystemen ist sogar eher von einem Kostenvorteil gegenüber der Referenzmethode auszugehen. HE-TUMT muss daher in der Gesamtbetrachtung als sichere, effektive, wirtschaftliche und damit sinnvolle Behandlungsalternative nicht nur gegenüber der Referenzmethode, sondern bei strikt ambulanter und narkosefreier Durchführung auch gegenüber anderen Alternativverfahren angesehen werden. In jedem Fall aber gilt sie als die Behandlungsoption der Wahl bei Hochrisikopatienten, die sonst dauerhaft Katheterträger bleiben müssten.

Somit sind die durch das IQWiG erarbeiteten Aussagen bezüglich dieser Behandlungsmethode bedingt durch die genannten methodischen Mängel zurückzuweisen. Dass diese methodischen Mängel auch alle anderen Therapieverfahren in gleicher Weise betreffen, ist der gesamte Vorbericht zurückzuziehen.

Stellungnahme zum Vorbericht IQWiG “Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

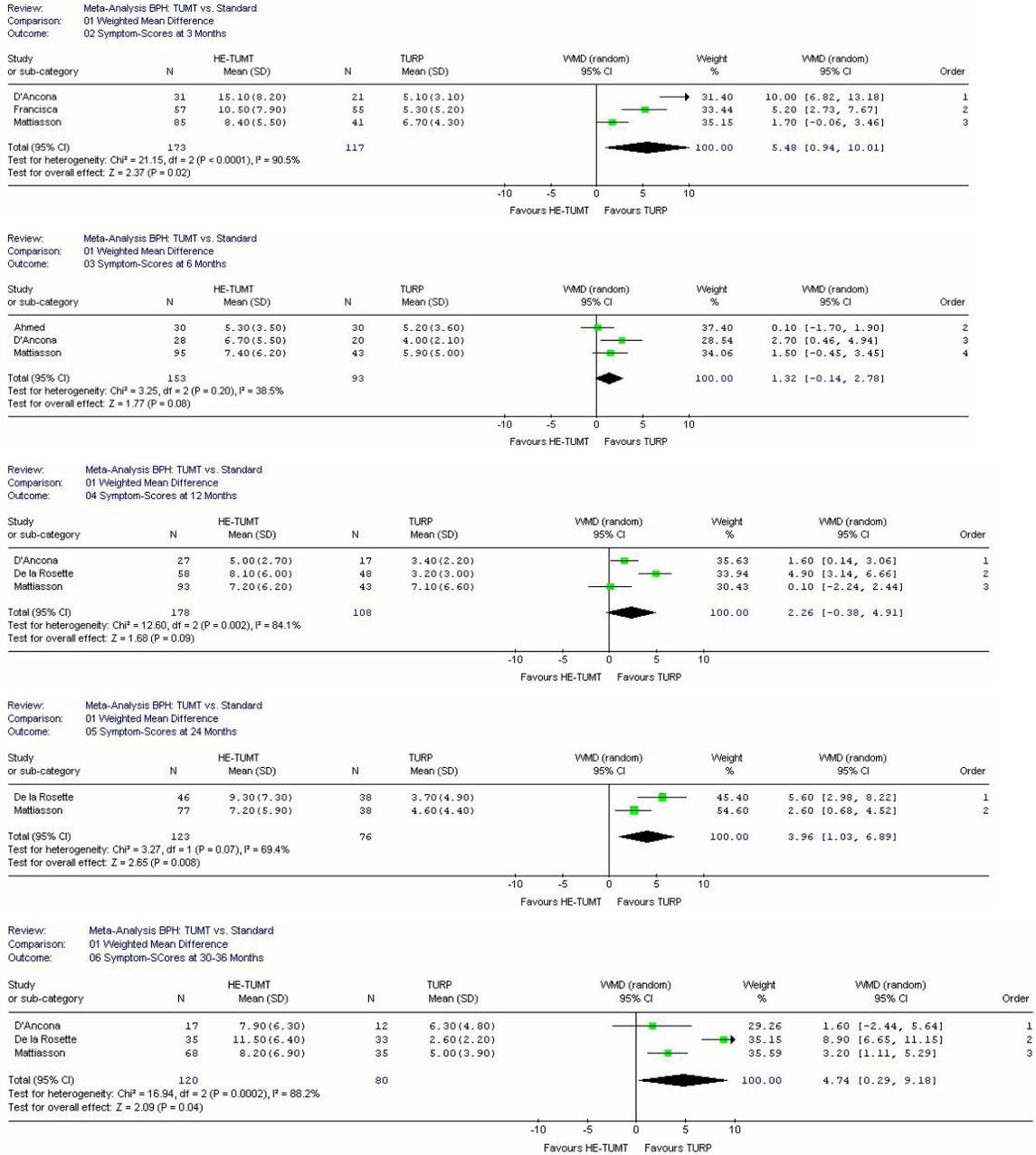


Abb. 1: Meta-Analyse für den Effekt von HE-TUMT und TURP auf den I-PSS

Stellungnahme zum Vorbericht IQWiG
 “Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

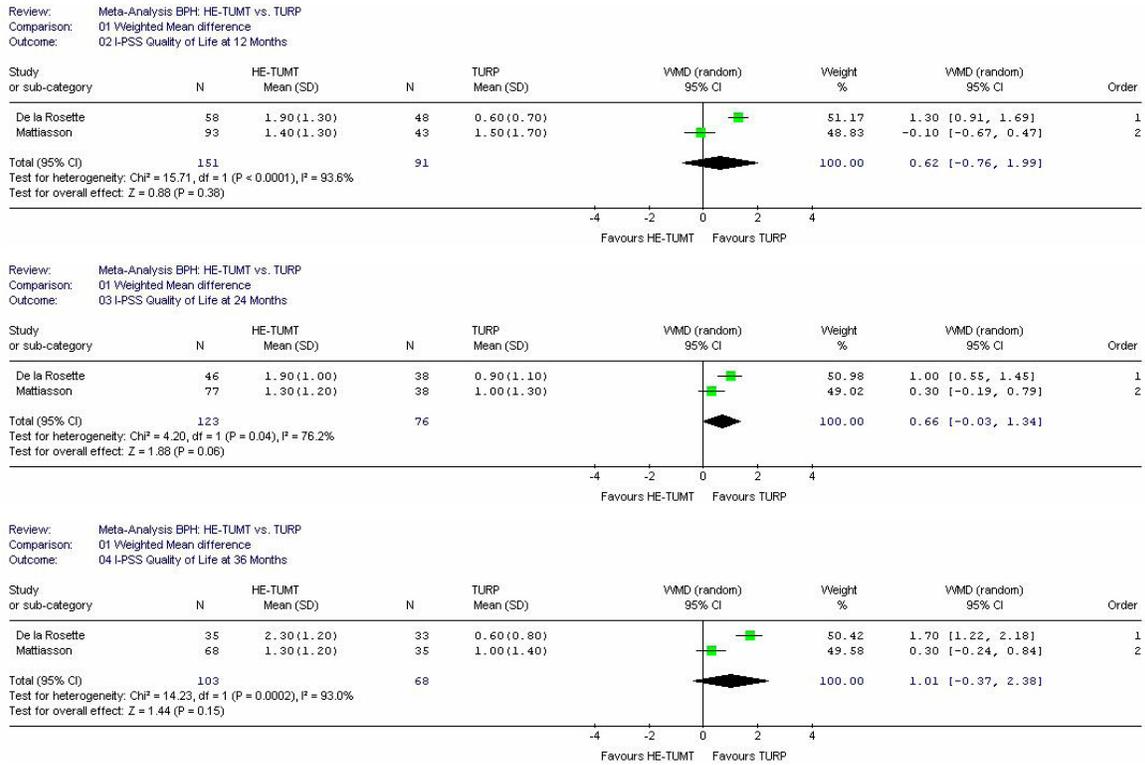


Abb. 2: Meta-Analyse für den Effekt von HE-TUMT und TURP auf I-PSS QoL.

Stellungnahme zum Vorbericht IQWiG

“Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

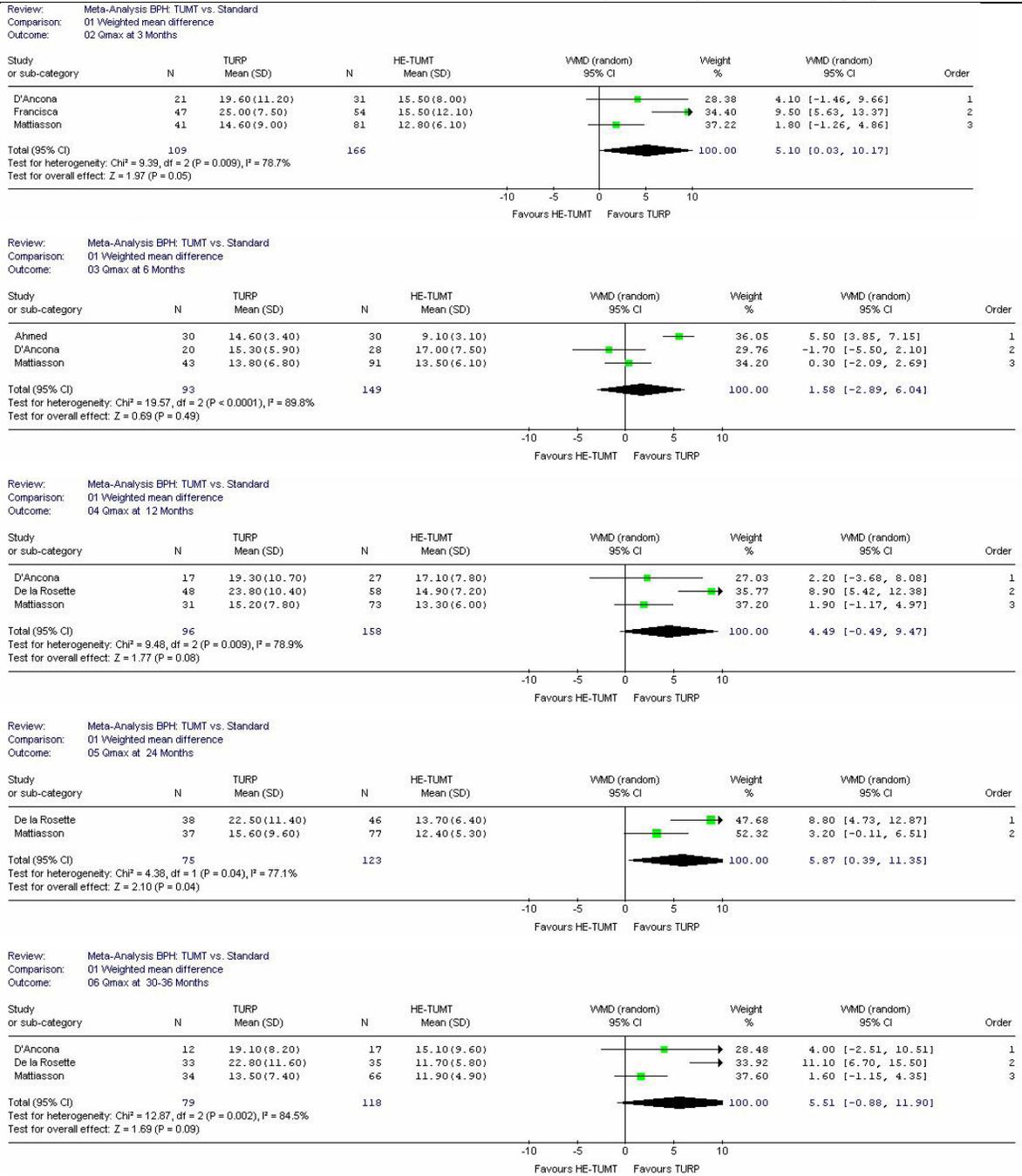


Abb. 3: Meta-Analyse für den Effekt von HE-TUMT und TURP auf Q_{max}

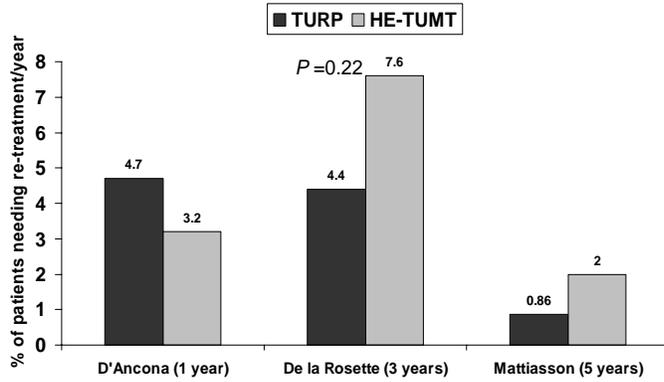


Abb. 4: Jährliche Re-Interventionsraten bei HE-TUMT und TURP [1, 3, 5]

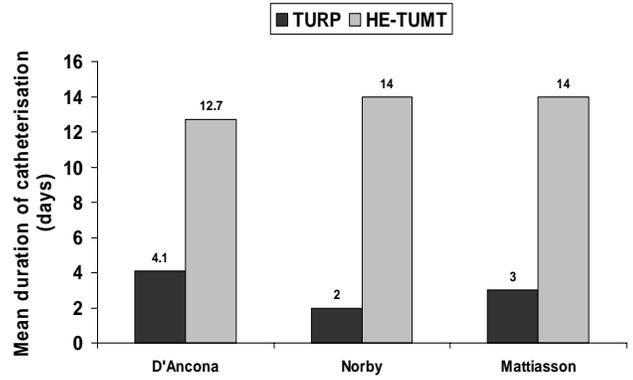


Abb. 5: Katheterisierungsdauer bei HE-TUMT und TURP [1, 4, 5].

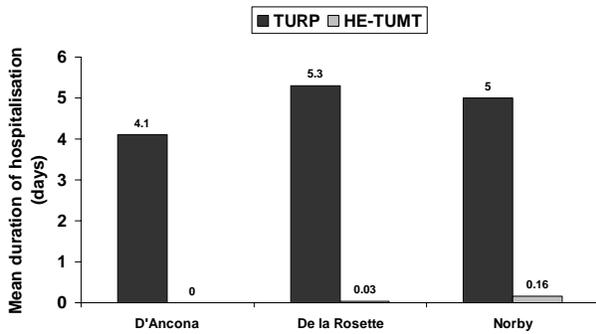


Abb. 6: Mittlere Hospitalisierungsdauer nach HE-TUMT und TURP [3-5]

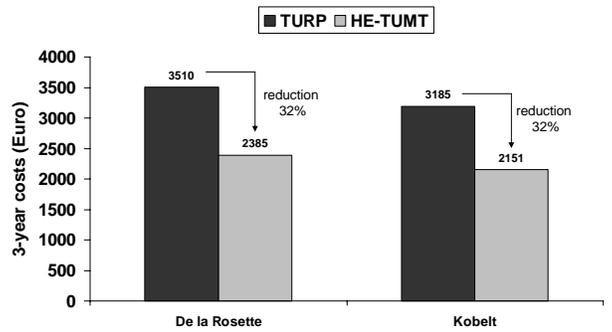


Abb. 7: Kostenvergleich von HE-TUMT und TURP über einen Zeitraum von 3 Jahren [2, 3]

Tab. 1: Der Effekt von HE-TUMT und TURP auf den I-PSS

Studie	Gerät	Score	HE-TUMT			TURP			P-Wert vs. TURP
			N	Mittelwert	SD	N	Mittelwert	SD	
Ahmed[6]	Prostatron 2.5	I-PSS	30			30			
○ Ausgangswert				18.5	4.2		18.4	4.8	
○ 6 Monate				5.3	3.5		5.2	3.6	
D'Ancona [7]	Prostatron 2.5	I-PSS							
○ Ausgangswert			31	18.3	6.3	21	16.7	5.6	
○ 3 Monate			31	15.1	8.2	21	5.1	3.1	
○ 6 Monate			28	6.7	5.5	20	4.0	2.1	
○ 12 Monate			27	5.0	2.7	17	3.4	2.2	
○ 30 Monate			17	7.9	6.3	12	6.3	4.8	
Francisca[8]	Prostatron 2.5								
○ Ausgangswert			68	20.1	6.5	68	20.8	6.2	
○ 3 Monate			57	10.5	7.9	55	5.3	5.2	< 0.01
○ 12 Monate			48	7.6	5.6	39	3.2	2.5	< 0.01
Floratos/de la Rosette[3, 9] ^a	Prostatron 2.5	I-PSS							
○ Ausgangswert			78	20	6.7	66	20	6.2	
○ 12 Monate			58	8.1	6.0	48	3.2	3.0	Sign
○ 24 Monate			46	9.3	7.3	38	3.7	4.9	Sign
○ 36 Monate			35	11.5	6.4	33	2.6	2.2	Sign
Nørby [4]	Prostatron 2.5								
○ Ausgangswert			37	20.5		22 ^b	21.3	6.6	
○ 6 Monate				9.2		22 ^b	6.8	5.7	
Mattiasson[1, 10, 11]	PLFT	I-PSS							
○ Ausgangswert			99	21.0	5.4	46	20.4	5.9	
○ 3 Monate			85	8.4	5.5	41	6.7	4.3	
○ 6 Monate			95	7.4	6.2	43	5.9	5.0	
○ 12 Monate			93	7.2	6.2	43	7.1	6.6	0.578
○ 24 Monate			77	7.2	5.9	38	4.6	4.4	0.014
○ 36 Monate			68	8.2	6.9	35	5.0	3.9	0.024
○ 48 Monate			56	7.1	5.4	30	6.4	6.6	0.289
○ 60 Monate			63	7.4	4.8	34	6.0	5.8	0.069

SD: Standard Deviation; ^a Hier wurden die Daten der Studie von de la Rosette verwandt, da diese auch die Basis für die Kostenanalyse bildeten^b TURP oder TUIP

Tab. 2: Der Effekt von HE-TUMT und TURP auf QoL

Studie	Score	HE-TUMT			TURP			P-Wert vs. TURP
		N	Mittelwert	SD	N	Mittelwert	SD	
Francisca[8]	General perception urinary difficulties*							
○ Ausgangswert		59	10.7	3.0	56	11.1	3.6	
○ 3 Monate		57	8.2	3.2	53	8.1	3.5	0.30
○ 12 Monate		32	7.4	2.7	43	5.6	1.0	< 0.01
Francisca[8]	Specific perception urinary difficulties							
○ Ausgangswert		59	11.3	4.4	56	12.4	4.8	
○ 3 Monate		57	8.3	3.6	50	7.1	3.0	0.02
○ 12 Monate		43	7.0	2.8	32	5.8	1.3	< 0.01
Francisca[8]	Sexual function							
○ Ausgangswert		56	14.2	6.0	56	14.9	4.7	
○ 3 Monate		53	14.5	5.6	51	14.4	4.3	0.15
○ 12 Monate		38	14.3	5.9	31	15.4	3.6	0.19
Francisca[8]	Daily activities							
○ Ausgangswert		58	12.4	4.0	56	13.8	3.9	
○ 3 Monate		56	10.4	3.1	53	9.8	3.1	0.01
○ 12 Monate		42	9.8	2.8	32	8.9	2.1	< 0.01
Francisca [8]	Psychological well-being							
○ Ausgangswert		57	26.4	4.4	56	27.3	5.3	
○ 3 Monate		56	27.9	4.0	52	27.9	4.6	0.63
○ 12 Monate		43	27.3	3.7	32	29.4	2.6	0.95
Francisca [8]	Social well-being							
○ Ausgangswert		55	13.4	3.8	54	12.5	3.4	
○ 3 Monate		53	13.6	4.4	50	12.0	3.9	0.69
○ 12 Monate		42	14.4	3.3	32	12.8	3.9	0.66
Francisca[8]	Experienced improvement							
○ Ausgangswert								
○ 3 Monate		70	12.2	5.9	69	11.5	6.7	0.58
○ 12 Monate		69	8.8	6.3	70	6.7	6.4	0.07
Floratos/ de la Rosette [3, 9] ^a	I-PSS QoL							
Floratos/de la Rosette[3, 9]								
○ Ausgangswert		78	4	0.9	66	4	1.1	

Stellungnahme zum Vorbericht IQWiG

“Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“; Version 1.0; 26-03-2007

Studie	Score	HE-TUMT			TURP			P-Wert vs. TURP
		N	Mittelwert	SD	N	Mittelwert	SD	
○ 12 Monate		58	1.9	1.3	48	0.6	0.7	Sign
○ 24 Monate		46	1.9	1.0	38	0.9	1.1	Sign
○ 36 Monate		35	2.3	1.2	33	0.6	0.8	Sign
Nørby[4]	I-PSS QoL ^b							
○ Ausgangswert		37	4 ^c	4-4 ^d	22 ^e	4 ^c	4-5 ^d	
○ 6 Monate			2 ^c		22 ^e	1 ^c	1-2 ^d	
Mattiasson [1, 10, 11]	I-PSS QoL							
○ Ausgangswert		99	4.3	1.0	46	4.2	1.1	
○ 3 Monate		84	1.5	1.4	41	1.1	1.6	
○ 6 Monate		93	1.3	1.4	42	1.0	1.5	
○ 12 Monate		93	1.4	1.3	43	1.5	1.7	0.972
○ 24 Monate		77	1.3	1.2	38	1.0	1.3	0.108
○ 36 Monate		68	1.3	1.2	35	1.0	1.4	0.162
○ 48 Monate		56	1.2	1.0	30	1.0	1.3	0.138
○ 60 Monate		63	1.1	0.9	34	1.1	1.2	0.841

SD: Standard Deviation; ^a Hier wurden die Daten der Studie von de la Rosette verwandt, da diese auch die Basis für die Kostenanalyse bildeten, *die engl. Originalitems wurden übernommen

^b in der Studie nicht angegeben, aber sehr wahrscheinlich I-PSS; ^c Median anstatt Mittelwert; ^d Spannweite anstatt SD; ^e TURP oder TUIP

Tab. 3: Der Effekt von HE-TUMT und TURP auf Q_{\max}

Studie	HE-TUMT			TURP			P-Wert vs. TURP
	N	Mittelwert	SD	N	Mittelwert	SD	
Ahmed[6]	30			30			
○ Ausgangswert		10.1	2.4		9.5	1.7	
○ 6 Monate		9.1	3.1		14.6	3.4	
D'Ancona [5, 7]							
○ Ausgangswert	31	9.3	3.9	21	9.3	3.4	
○ 3 Monate	31	15.5	8.0	21	19.6	11.2	
○ 6 Monate	28	17.0	7.5	20	15.3	5.9	
○ 12 Monate	27	17.1	7.8	17	19.3	10.7	
○ 30 Monate	17	15.1	9.6	12	19.1	8.2	
Francisca[8]							
○ Ausgangswert	68	9.6	3.0	68	7.9	2.8	
○ 3 Monate	54	15.5	12.1	47	25.0	7.5	< 0.01
○ 12 Monate	49	15.2	7.6	37	23.5	9.9	< 0.01
Floratos/de la Rosette [3, 9] ^a							
○ Ausgangswert	78	9.2	3.1	66	7.8	2.8	
○ 12 Monate	58	14.9	7.2	48	23.8	10.4	Sign
○ 24 Monate	46	13.7	6.4	38	22.5	11.4	Sign
○ 36 Monate	35	11.7	5.8	33	22.8	11.6	Sign
Nørby[4]							
○ Ausgangswert	37	9.1	4.2	22 ^b	9.6	3.2	
○ 6 Monate	37	13.5		22 ^b	20.6	12.8	
Mattiasson[1, 10, 11]							
○ Ausgangswert	79	7.6	2.7	35	7.9	2.7	
○ 3 Monate	81	12.8	6.1	41	14.6	9.0	
○ 6 Monate	91	13.5	6.1	43	13.8	6.8	
○ 12 Monate	73	13.3	6.0	31	15.2	7.8	0.565
○ 24 Monate	77	12.4	5.3	37	15.6	9.6	0.020
○ 36 Monate	66	11.9	4.9	34	13.5	7.4	0.584
○ 48 Monate	49	12.3	5.7	30	14.7	7.5	0.016
○ 60 Monate	61	11.4	4.9	32	13.6	7.8	0.217

SD: Standard Deviation; ^a Hier wurden die Daten der Studie von de la Rosette verwandt, da diese auch die Basis für die Kostenanalyse bildeten^b TURP oder TUIP

Tab. 4: Schwere AEs oder AEs in mindestens 5% der behandelten Patienten (HE-TUMT oder TURP)

	Ahmed [6]:		Nørby[4]:		D'Ancona[5]:		Wagrell[10]:		Mattiasson[1]:	
	6 Monate		6 Monate		Jahr 1		Jahr 1		Jahr 2-5	
	HE-TUMT	TURP	TUMT	TURP	HE-TUMT	TURP	HE-TUMT	TURP	HE-TUMT	TURP
	N=30	N=30	N=46	N=24	N=31	N=21	N=100	N=46	N=80	N=39
Serious ¹ AEs							2%	17%	4%	0%
○ Haematurie/ Bluttransfusion	0%	13%	2%	9%			1%	9%		
○ Harnwegsinfekt	3%	3%					0%	2%		
○ TUR-Syndrom			0%	5%			0%	2%		
○ Urosepsis							0%	2%		
○ Tamponade							0%	2%		
○ Harnverhalt							1%	0%		
Haematurie			0%	0%	0%	14%	13%	39%	6%	0%
Harnwegsinfekt	0%	7%	30%	14%	16%	4%	18%	20%	0%	3%
Urgency/ Speichersymptome (transiente)	100%		0%	0%	29%	19%	37%	13%	3%	13%
Inkontinenz			0%	5%			3%	13%	1%	5%
Harnverhalt	10%	7%	9%	5%			19%	13%	3%	0%
Impotenz	0%	21%	9%	14%			6%	11%	8%	15%
Retrograde Ejakulation	22%	63%	22%	50%						
PSA Anstieg			0%	0%					6%	0%
Urethrastriktur			0%	5%						
Blasenhalsenge	0%	3%								
Obstipation	100%	0%								

1: Definition der "International Committee on Harmonisation (ICH)": "any untoward medical event that results in death, is life-threatening, requires in-patient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability or incapacity, is cancer, or requires intervention to prevent permanent damage to body function or structure"

Literatur:

- [1] Mattiasson A WL, Schelin S, Nordling J, Richthoff J, Magnusson B, Schain M, Larson T, Boyle E, Duelund-Jacobsen J, Kroyer K, Ageheim H. . Five-year follow-up of Feedback Microwave Thermotherapy versus TURP for clinical BPH: a prospective randomized multicenter study. . *Urology*. 2007;69::91-7.
- [2] Kobelt G, Spangberg A, Mattiasson A. The cost of feedback microwave thermotherapy compared with transurethral resection of the prostate for treating benign prostatic hyperplasia. *BJU Int*. 2004 Mar;93(4):543-8.
- [3] De La Rosette JJ, Floratos DL, Severens JL, Kiemeney LA, Debruyne FM, Pilar Laguna M. Transurethral resection vs microwave thermotherapy of the prostate: a cost-consequences analysis. *BJU Int*. 2003 Nov;92(7):713-8.
- [4] Norby B, Nielsen HV, Frimodt-Moller PC. Transurethral interstitial laser coagulation of the prostate and transurethral microwave thermotherapy vs transurethral resection or incision of the prostate: results of a randomized, controlled study in patients with symptomatic benign prostatic hyperplasia. *BJU Int*. 2002 Dec;90(9):853-62.
- [5] D'Ancona FC, Francisca EA, Witjes WP, Welling L, Debruyne FM, de la Rosette JJ. High energy thermotherapy versus transurethral resection in the treatment of benign prostatic hyperplasia: results of a prospective randomized study with 1 year of followup. *J Urol*. 1997 Jul;158(1):120-5.
- [6] Ahmed M, Bell T, Lawrence WT, Ward JP, Watson GM. Transurethral microwave thermotherapy (Prostatron version 2.5) compared with transurethral resection of the prostate for the treatment of benign prostatic hyperplasia: a randomized, controlled, parallel study. *Br J Urol*. 1997 Feb;79(2):181-5.
- [7] D'Ancona FC, Francisca EA, Witjes WP, Welling L, Debruyne FM, De La Rosette JJ. Transurethral resection of the prostate vs high-energy thermotherapy of the prostate in patients with benign prostatic hyperplasia: long-term results. *Br J Urol*. 1998 Feb;81(2):259-64.
- [8] Francisca EA, d'Ancona FC, Hendriks JC, Kiemeney LA, Debruyne FM, de La Rosette JJ. A randomized study comparing high-energy TUMT to TURP: quality-of-life results. *Eur Urol*. 2000 Nov;38(5):569-75.
- [9] Floratos DL, Kiemeney LA, Rossi C, Kortmann BB, Debruyne FM, de La Rosette JJ. Long-term followup of randomized transurethral microwave thermotherapy versus transurethral prostatic resection study. *J Urol*. 2001 May;165(5):1533-8.
- [10] Wagrell L, Schelin S, Nordling J, Richthoff J, Magnusson B, Schain M, et al. Feedback microwave thermotherapy versus TURP for clinical BPH--a randomized controlled multicenter study. *Urology*. 2002 Aug;60(2):292-9.
- [11] Wagrell L, Schelin S, Nordling J, Richthoff J, Magnusson B, Schain M, et al. Three-year follow-up of feedback microwave thermotherapy versus TURP for clinical BPH: a prospective randomized multicenter study. *Urology*. 2004 Oct;64(4):698-702.
- [12] De la Rosette JJMCH AG, Madersbacher S, Perachino M, Thomas D, Desgrandchamps F, de Wildt M. . EAU guidelines on benign prostatic hyperplasia (BPH). . Updated March 2004. *Eur Urol*. 2001;40:256-63.
- [13] Berges R, Dreikorn K, Höfner K, Jonas U, Laval KU, Madersbacher S, et al. Leitlinien der Deutschen Urologen zur Therapie des Benigen Prostatasyndroms (BPS). *Urologe A*. 2003 May;42(5):722-38.
- [14] Barry MJ, Williford WO, Chang Y, Machi M, Jones KM, Walker-Corkery E, et al. Benign prostatic hyperplasia specific health status measures in clinical research: how much change in the American Urological Association symptom index and the benign prostatic hyperplasia impact index is perceptible to patients? *The Journal of urology*. 1995 Nov;154(5):1770-4.
- [15] Abramsson DALC, GRIFFITHS D.,VAN KERREBROECK P.,NISHIZAWA O, NITTI N., KEONG TATT F., TUBARO A., WEIN A. Lower Urinary Tract Symptom: Etiology, Patient Assessment and Predicting, Outcome from Therapy. n: Male Lower Urinary Tract Dysfunktion Evaluation and Management MCCONNELL,J ,ABRAMS, P, DENIS,L ,KHOURY, S ,ROEHRBORN C (eds):6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases. 2006;6:69-142.
- [16] Girman CJ, Panser LA, Chute CG, Oesterling JE, Barrett DM, Chen CC, et al. Natural history of prostatism: urinary flow rates in a community-based study. *J Urol*. 1993 Sep;150(3):887-92.

- [17] Dahlstrand C, Grundtman S, Pettersson S. High-energy transurethral microwave thermotherapy for large severely obstructing prostates and the use of biodegradable stents to avoid catheterization after treatment. *Br J Urol.* 1997 Jun;79(6):907-9.
- [18] Norby B, Nielsen HV, Frimodt-Moller PC. Cost-effectiveness of new treatments for benign prostatic hyperplasia: results of a randomized trial comparing the short-term cost-effectiveness of transurethral interstitial laser coagulation of the prostate, transurethral microwave thermotherapy and standard transurethral resection or incision of the prostate. *Scand J Urol Nephrol.* 2002;36(4):286-95.
- [19] Ackerman SJ, Rein AL, Blute M, Beusterien K, Sullivan EM, Tanio CP, et al. Cost effectiveness of microwave thermotherapy in patients with benign prostatic hyperplasia: part I-methods. *Urology.* 2000 Dec 20;56(6):972-80.
- [20] Blute M, Ackerman SJ, Rein AL, Beusterien K, Sullivan EM, Tanio CP, et al. Cost effectiveness of microwave thermotherapy in patients with benign prostatic hyperplasia: part II--results. *Urology.* 2000 Dec 20;56(6):981-7.
- [21] Bolmsjo M, Schelin S, Wagrell L, Larson T, de la Rosette JJ, Mattiasson A. Cell-kill modeling of microwave thermotherapy for treatment of benign prostatic hyperplasia. *J Endourol.* 2000 Oct;14(8):627-35.
- [22] Schelin S. Microwave thermotherapy in patients with benign prostatic hyperplasia and chronic urinary retention. *Eur Urol.* 2001 Apr;39(4):400-4.
- [23] Chapple CR. Clinical study of benign prostatic disease, current concepts and future prospects: randomized controlled trials versus real life practice. *Curr Opin Urol.* 2003 Jan;13(1):1-5.
- [24] Sackett DL SS, Richardson WS, Rosenberg W, Haynes RB. . *Evidence-Based Medicine.* 2000, . Churchill Livingstone. 2000:S1-11.
- [25] Committee. APG. AUA guideline on management of benign prostatic hyperplasia (2003). Chapter 1: diagnosis and treatment recommendations. .: *J Urol.* 2003;170:530-47.
- [26] DE LA ROSETTE J BS, BADLANI G., ELHILALI M., GRAVAS S., MUSCHTER R., NAITO S., NETTO N.R. New Minimally Invasive and Surgical Developments in the management of BPO. In: *Male Lower Urinary Tract Dysfunktion Evaluation and Management MCCONNELL,J ,ABRAMS, P, DENIS,L ,KHOURY, S ,ROEHRBORN C (eds):6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases.* 2006;6:195-234.

4.2.2 Lumenis GmbH

Autoren:

Hauke Albert Harms

Michael Martin

Adresse:

Lumenis (Germany) GmbH

Heinrich-Hertz-Str. 3

63303 Dreieich

Auftragsnummer:	--- N04-01 ---
--- Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie ---	

Stellungnahme.

Zu Holmium Laser Enukleation der Prostata (HoLEP)

Zuerst möchten wir dem Institut zu Ihrer gründlichen Analyse gratulieren. Zu den Aufgestellten Kriterien sind wir nicht in der Lage zu kommentieren. Wir möchten jedoch darauf hinweisen, dass mit denselben Daten und Veröffentlichungen die Welt Gesundheitsorganisation (WHO) und das britische „National Institute for Clinical Excellence“ (NICE) zu einem wesentlich positiveren Entschluss gekommen sind.

“HoLEP ≈ is equivalent to TURP/Adenomectomy”

“HoLEP has the same or even better results than the traditional surgery, with low morbidity, regardless of the prostate size and it is being promoted as the new gold standard for treatment of BPO.” (WHO)

“2.3.2 The Specialist Advisors considered holmium laser prostatectomy to be established practice and preferable in many cases to TURP, requiring a shorter stay in hospital.” (NICE)

“Holmium laser prostatectomy is a viable alternative to TURP and irrespective of any anatomical configuration.” (EAU §4.7.5.)

Wegen der klinischen und wirtschaftliche Vorteile gegenüber der TURP u.a.

- Kein TURP Syndrom
- Weniger Blutverlust
- Kürzerer Krankenhausaufenthalt
- Gewebe steht zur Verfügung für die Histologie
- Endoskopische Behandlung von Prostata mit 200 gm. Gewicht und mehr

würden wir gerne eine positivere Beurteilung dieses Verfahrens sehen.

4.2.3 Arbeitskreis BPH der Akademie der deutschen Urologen

Autoren:

Prof. Dr. Klaus Höfner
Prof. Dr. Kurt Dreikorn
Dr. Richard Berges
PD Dr. Oliver Reich
Prof. Dr. Rolf Muschter
Dr. Matthias Oelke
Dr. Wolfgang Rulf
Dr. Christian Tschuschke

Adresse:

Urologische Klinik, Evangelisches Krankenhaus
Virchowstr. 20
46047 Oberhausen

Stellungnahme zum Vorbericht
Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie
Auftragsnummer: --- N04-01 ---

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien

Siehe Literaturverzeichnis.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien

Fehlerhafte Interpretation der Ergebnisse

In der Diskussion des IQWiG-Berichtes finden sich zahlreiche Widersprüche bzw. Unklarheiten:
In *Tabelle 156: Zusammenfassende Darstellung der Ergebnisse der Nutzenbewertung* werden folgende Ergebnisse des Vergleiches von 11 nichtmedikamentösen gegenüber Standard geprüft und in die Nutzenbewertung eingegangenen Verfahren präsentiert:

Symptomatik: Für 5 Verfahren wird „kein Beleg für einen Unterschied“ gefunden (VLAP DB, HoLAP, HoLRP und TEAP) bzw. „liegt kein Beleg für einen Gruppenunterschied vor“ (HoLEP).

Für 6 Verfahren ist ein Vorteil des Standards „belegt“ (VLAP, TUMT) oder es „liegen Hinweise auf einen Vorteil des Standards vor“ (CLAP, Hybrid, ILK, TUNA).

Lebensqualität: Für 7 Verfahren ist entweder ein Vorteil gegenüber Standard „belegt“ (HoLRP), es wird „kein Beleg für einen Unterschied“ gefunden (TUNA, TEAP) oder es „liegt kein Beleg für einen Gruppenunterschied vor“ (CLAP, Hybrid, ILK, HoLEP). Nur für 2 (VLAP, TUMT) „liegen Hinweise auf einen Vorteil des Standards vor.“ 2 Verfahren wurden nicht einbezogen.

Krankenhausverweildauer: Für alle Verfahren ist entweder ein Vorteil gegenüber Standard „belegt“ (CLAP, VLAP, ILK, HoLEP), es „liegen Hinweise auf einen Vorteil gegenüber Standard vor“ (HoLRP), „Ein Vorteil ist prozedural bedingt immanent.“ (TUMT, TUNA) oder es wird „kein Beleg für einen Unterschied“ gefunden (VLAP DB, HoLAP, TEAP) bzw. „liegt kein Beleg für einen Gruppenunterschied vor“ (Hybrid).

Für 8 Verfahren ist entweder ein Vorteil gegenüber Standard „belegt“ (CLAP, HoLEP), es „liegen Hinweise auf einen Vorteil gegenüber Standard vor“ (HoLRP), wird „kein Beleg für einen Unterschied“ gefunden (VLAP DB, HoLAP, TEAP) oder es „liegt kein Beleg für einen Gruppenunterschied vor“ (VLAP, Hybrid). Nur für 1 Verfahren (TUMT) „liegen Hinweise auf einen Vorteil des Standards vor.“ 2 Verfahren wurden nicht einbezogen.

Fasst man die Ergebnisse unter dem Gesichtspunkt der klinischen Relevanz zusammen, so sind Vorteile oder vergleichbare Ergebnisse der beurteilten Methoden gegenüber Standard in der Beeinflussung der Symptomatik in 5/11, der Lebensqualität in 7/9, der Krankenhausverweildauer bei allen Verfahren und bei der Katheterisierungsdauer in 8/9 beurteilten Verfahren im IQWiG-Bericht dokumentiert. In *Tabelle 158* wird zusätzlich festgestellt, dass sich der IQWiG-Bericht ausser in einem Fall (Beurteilung VLAP vs. Standard in der Wirkung auf Symptome) in „Übereinstimmung“ oder nur „leichter Diskrepanz“ mit anderen systematischen Übersichten [8, 10, 11] befindet.

In der weiteren Diskussion finden sich jedoch dann inakzeptable Einschränkungen dieser Ergebnisse:

Auf S. 392 werden die o.g. Daten zur Nullaussage relativiert: „Formulierungen... wie „2 Verfahren hätten zu „ähnlichen Verbesserungen“ oder „ähnlichen Ergebnissen“ geführt, „wenn aus den zugrunde liegenden Studien kein Beleg für einen Vorteil zugunsten einer der Behandlungen abzuleiten war,... ist problematisch, da keine der im vorliegenden Bericht eingeschlossenen Studien erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudie geplant war und generell das Schließen von einem fehlenden Beleg für einen Unterschied auf eine Gleichwertigkeit unzulässig ist. Im vorliegenden Bericht wurde von „kein Beleg für einen Vorteil zugunsten einer der beiden Behandlungen“ gesprochen, woraus sich eine Vergleichbarkeit der Behandlungen nicht ableiten lässt“. In dieser für einen Kliniker kaum verständlichen Logik zeigt sich die eklatante Diskrepanz zwischen möglicherweise exakter biometrischer Interpretation

und klinischer Relevanz der Ergebnisse. Es könnte noch angehen, wenn die Aussage sich auf die rein biometrische Interpretation reduzieren würde. Im Gegenteil handelt es sich um eine Kernaussage des Berichtes, die wesentliche Begründung der Aussage des Fazits ist, dass „für keines der Verfahren letztendlich ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden.“ kann. Sowohl in der ICH Leitlinie E9 „Statistical Principles for Clinical Trials“ [12] als auch im IQWiG-Methodenpapier [5] findet sich eine ausführliche Darstellung der Wesensmerkmale einer Äquivalenzstudie oder einer Überlegenheitsstudie. Beide Publikationen legen dabei nicht fest, wann welcher der beiden Ansätze zu wählen ist. Bedenkt man den Umfang der Protokolle kontrollierter Studien ist schnell einsichtig, dass diese Informationsgüte bei der Erstellung von systematischen Übersichtsarbeiten nicht übernommen werden kann. Auch angesichts der marginalen Zusatzinformationen aus Sicht der klinischen Relevanz darf es nicht als Qualitätsmangel herangezogen werden, dass eine Planung der Studien als Äquivalenz- oder Nichtunterlegenheitsstudie in den herangezogenen randomisierten Studien nicht erkennbar ist. Das IQWiG sollte diesen Passus im Endbericht streichen.

Symptomatik und Lebensqualität:

„Fragen zur Lebensqualität kann der vorliegende Bericht aufgrund mangelnder Datenqualität bzw. unzureichender Messinstrumente in den Studien letztendlich nicht befriedigend beantworten. Das am häufigsten verwendete Instrument der in den vorliegenden Bericht eingeschlossenen Studien war die globale 1-Item-Frage des IPSS. Dem mehrdimensionalen Konstrukt der gesundheitsbezogenen Lebensqualität wird mit dieser Frage nicht ausreichend Rechnung getragen.“

In der Mehrzahl vom IQWiG identifizierten Studien wurde der IPSS verwendet. Auf Initiative des International Consensus Committee wurde Mitte der 90-er Jahre eine zusätzliche Frage zur Lebensqualität in den IPSS-Fragebogen aufgenommen. Der 1-Item-Score zur Erfassung der krankheitsspezifischen Lebensqualität war zum Zeitpunkt der Durchführung der Mehrzahl der Studien einzig validierter Standard und ist bis heute international und auch von der Deutschen Gesellschaft für Urologie als offizieller Bewertungsstandard für Miktionsbeschwerden und Lebensqualität bei BPS akzeptiert [6, 8, 9]. Eine Erfassung gesundheitsbezogener Lebensqualität mit generischen multidimensionalen Instrumenten mag zwar in einzelnen Fragestellungen hilfreich sein, für die Bewertung von Behandlungsergebnissen bei BPS haben sich diese Instrumente aber nicht durchgesetzt, da sie vielfach unpraktikabel sind und die Generalaussage nur wenig schärfen.

Im Übrigen ist der IPSS auch unter Berücksichtigung dieser Aspekte validiert worden und misst genau das, was er messen soll: Miktionsbeschwerden und den resultierenden Leidensdruck.

Darüber hinaus wurde bei der Bewertung aller Therapieverfahren völlig übersehen, dass die in den Studien berichteten Unterschiede zugunsten der Referenzmethoden klinisch irrelevant sind. Um eine leichte Symptomenverbesserung gemessen mit dem IPSS überhaupt wahrnehmen zu können, muss sich ein medianer IPSS von 16 Punkten um wenigstens 3 Punkte verringern [13]. Bei Patienten mit schweren Symptomen (IPSS 20-35), muss eine Therapie den Score sogar um wenigstens 6.1 Punkte senken, bevor diese Verbesserung spürbar ist. Um eine moderate Verbesserung festzustellen, muss durch eine Therapie sogar eine Scoreverbesserung gegenüber dem Ausgangswert von 5.1 (mittelschwere Symptomatik) bzw. 8.7 Punkten (schwere Symptomatik) erfolgen. Es ist also fragwürdig, ob die in den Vergleichstudien der Alternativverfahren behandelten Patienten den Unterschied zugunsten der Referenzmethoden überhaupt bemerken können. Damit sind die vom IQWiG als wichtigste Differenzierungsmerkmale eingestuften Parameter ungeeignet, diese Verfahren zu bewerten. Allein diese Fehleinschätzung aufgrund mangelnder Fachexpertise resultiert in einer grob falschen Kernaussage dieses Vorberichtes.

Krankenhausverweildauer:

Auf S. 389 wird festgestellt, dass „Insgesamt ... die Krankenhausverweildauer unter Standard länger als unter den im vorliegenden Bericht betrachteten Prüfinterventionen“ ist.“ *Zur Lebensqualität und Katheterisierungsdauer können keine einheitlichen Aussagen zugunsten der Standardtherapie oder der Prüfinterventionen abgeleitet werden.*“, was angesichts der Daten einer Fehlinterpretation entspricht.

Im Fazit (s. 395) wird festgestellt, dass „Obgleich einerseits für einen Teil der Verfahren prozedural bedingte und/oder durch Studien belegte Vorteile gegenüber einer Standardtherapie in Bezug auf die Krankenhausverweil- und teilweise auch auf die Katheterisierungsdauer geltend gemacht werden können,

kann andererseits unter der Prämisse, dass die Beeinflussung der Symptomatik und ggf. damit verbunden der Lebensqualität vordringliches Ziel der Behandlung eines BPS ist, für keines der Verfahren letztendlich ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden.“ Die Interpretation ist erneut völlig unverständlich und irreführend. Aus welchem Grund wird die Definition eines angeblichen Vorrangs eines Behandlungsziels dazu verwendet, andere deutlich ersichtliche Vorteile der untersuchten Methoden gegenüber Standard zu negieren. Diese Interpretation ist willkürlich und ohne jede Evidenz. Auf S. 40 der IQWiG-Methoden [5] stellt das Institut selbst fest, dass „Die Inanspruchnahme von Leistungen des Gesundheitssystems ... ebenfalls eine patientenrelevante Größe sein“ kann. Auf S. 392 heisst es weiter, dass „Der Vergleich der Prüfinterventionen mit der Standardtherapie dadurch erschwert“ war“, dass bei Belegen eines Vorteils hinsichtlich der Krankenhausverweildauer keine gleichgerichteten Belege bezüglich Symptomatik und Lebensqualität vorlagen. Für den Nachweis eines Zusatznutzens gegenüber Standard ist zumindest die Nichtunterlegenheit zu fordern. Es fanden sich jedoch in keiner Studie Überlegungen zu Äquivalenz- und Nichtunterlegenheitsbereichen.“ Einerseits wird hier ohne jede Evidenz behauptet, dass die Krankenhausverweildauer nur dann als Vorteil zu werten ist, wenn ein Zusammenhang mit der Veränderung von Symptomatik und Lebensqualität hergestellt werden kann, andererseits wiederholt sich die These von S.392.

Unerwünschte Ereignisse (UE):

Entsprechend S. 28 wurden „Für die Zielgröße „unerwünschte Therapienebenwirkungen und -komplikationen ... die in den Publikationen genannten unerwünschten Ereignisse (UEs) herangezogen.“ In Tabelle 157 erfolgt eine Zusammenfassung wie folgt:

„Hinweise auf einen Unterschied zugunsten der untersuchten Verfahren gegenüber Standard“ ergaben sich bei

- schwerer Blutung bei 5/5 berichteten Methoden (VLAP, ILK, HoLEP, TUMT, TUNA)
- EJD bei 5/5 berichteten Methoden (VLAP, Hybrid, ILK, TUMT, TUNA)
- ED bei 2/2 berichteten Methoden (ILK, TUNA)
- Strikturen bei 1/1 berichteter Methode (TUNA)

„Hinweise auf einen Unterschied zugunsten Standard gegenüber den untersuchten Verfahren“ ergaben sich bei

- Miktions-symptomatik bei 2/2 berichteten Methoden (HoLEP, TUMT)
- Retention bei 3/3 berichteten Methoden (VLAP, Hybrid, TUMT)
- Re-katheterisierung bei 2/3 berichteten Methoden (Hybrid, TUMT)
- HWI bei 2/2 berichteten Methoden (ILK, TUMT)
- Reinterventionen bei 4/5 berichteten Methoden (VLAP, Hybrid, ILK, TUNA)

Auf S. 389 des Berichtes wird bestätigt, dass „In der Gesamtbetrachtung der UE's für die Vergleiche mit Standard ... bezüglich sexueller Dysfunktionen und der behandlungsbedürftigen Blutungen ... Hinweise darauf vorliegen, dass die Prüfinterventionen der Standardtherapie überlegen sind“. Interessanterweise wird hier von Überlegenheit gesprochen, die doch auf S. 392 argumentativ abgelehnt wurde.

Auf S. 392 folgt fast zwangsläufig wiederum die entscheidende Relativierung mit dem Satz: „Aufgrund der nicht standardisierten und teils nur sehr rudimentären Darstellung der UE's in den Studien ist die Validität und Aussagekraft der Daten äußerst eingeschränkt.“

Laserverfahren:

Dem Text des Kapitels Laserverfahren der von der AWMF als S2 anerkannten Leitlinie zur Therapie des BPS [7] liegt eine der Literaturrecherche des IQWiG vergleichbare Recherche zugrunde.

Dabei wurden ausschließlich Studien von 1992 bis 2002 mit einer Mindestbeobachtungszeit von 3 Monaten mit mehr als 10 Patienten pro Behandlungsarm berücksichtigt. Mehrere und auch vom IQWiG selbst als relevant eingestufte Studien wurden in der Bewertung nicht berücksichtigt, da sie dem falschen Verfahren (Koagulation statt Vaporisation) zugeordnet wurden (z.B. Ref. 82, 83, 88, 91).

Daher muss an dieser Stelle bemerkt werden, dass das gesamte Kapitel über Laseranwendungen aufgrund grober Fehlzuordnungen neu erfasst werden muss. Die Aussagen des IQWiG zu diesem Thema sind als grob falsch einzustufen.

Da eine detaillierte Stellungnahme aufgrund der unverständlichen Beschränkung auf 10 Seiten an dieser Stelle nicht möglich ist und darüber hinaus eine inakzeptable Fristvorgabe von 4 Wochen für die Stellungnahme auf einen 500 Seiten umfassenden Vorbericht ohnehin eine wissenschaftlich optimale Neubearbeitung dieses Kapitels nicht erlaubt, muss auf eine Kurzbewertung durch Experten des Arbeitskreises BPH mit separater Kommentierung der zahlreichen Laserverfahren verwiesen werden, die wenigstens die grob falschen Zuordnungen benennt.

TUNA:

Seit 1995 wurden weltweit mehr als 100.000 Behandlungen mit der transurethralen Nadelablation (TUNA) durchgeführt, alleine 22.000 im Jahre 2006. Die Wirksamkeit der TUNA hinsichtlich der definierten patientenrelevanten Therapieziele ist durch Studien hoher Evidenz belegt. Der zusätzliche Nutzen im Vergleich zur Standardtherapie TURP leitet sich für eine selektionierte Patientenpopulation aus dem geringeren Nebenwirkungsprofil ab. Da eine detaillierte Stellungnahme aufgrund der unverständlichen Beschränkung auf 10 Seiten an dieser Stelle nicht möglich ist, muss auf eine Kurzbewertung durch Experten des Arbeitskreises BPH mit separater Kommentierung der TUNA verwiesen werden. Die Aussagen des IQWiG zu diesem Thema sind als grob falsch einzustufen.

TUMT:

Das IQWiG macht mangels urologischer Sachkenntnis den groben methodischen Fehler, ein nichtablatives Therapieverfahren (Niedrigenergie-TUMT) mit einem ablativen Therapieverfahren (Hochenergie-TUMT) in einer gemeinsamen Analyse mit dem postulierten Referenzverfahren zu vergleichen: Niedrigenergie-TUMT erhebt nicht den Anspruch eines TUR-P-Äquivalents. Ein Vergleich dieses Verfahrens wäre allenfalls mit einer medikamentösen Therapie (z.B. Alpha-Blocker) sinnvoll. Niedrigenergie-TUMT ist heute allenfalls als historische Entwicklungsvorstufe der Hochenergie-TUMT anzusehen, wobei letztere durch tatsächlichen Gewebeabtrag (wie bei TUR-P) über eine bloße Symptomenreduktion hinaus auch deobstruktiv wirksam ist. Ein Vergleich mit der TUR-P muss nach heutiger Einschätzung also allein auf die Hochenergie-TUMT beschränkt bleiben. Daher muss an dieser Stelle bemerkt werden, dass das gesamte Kapitel über TUMT aufgrund grober Fehlzusammenhänge neu erfasst werden muss. Die Aussagen des IQWiG zu diesem Thema sind als grob falsch einzustufen. Da eine detaillierte Stellungnahme aufgrund der unverständlichen Beschränkung auf 10 Seiten an dieser Stelle nicht möglich ist, muss auf eine Kurzbewertung durch Experten des Arbeitskreises BPH mit separater Kommentierung der TUMT verwiesen werden.

3. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe wissenschaftlicher Literatur zur Begründung der Anmerkung.

Verletzung wissenschaftlicher Grundsätze zur Beurteilung klinischer Evidenz:

Die Literatursammlung, welche dem Vorbericht zugrunde liegt, ist willkürlich und beruht auf der Auffassung des IQWiG, nur solche Studien seien für eine Evidenzbewertung zuzulassen, die den Anforderungen des IQWiG entsprechen. Diese Anforderungen sind rein willkürlich gesetzt und bewegen sich außerhalb der wissenschaftlichen Definition evidenzbasierter Medizin. Die Fokussierung auf diese Kriterien führt zu einer unzulässigen Verzerrung der tatsächlichen Evidenz zum Nachteil der Patienten.

David Sackett, ein Nestor und Experte der evidenzbasierten Medizin, definiert die EBM als die „Integration individueller klinischer Expertise mit der bestmöglichen Evidenz aus systematischer Forschung“. Hierbei wird häufig, so auch im Vorbericht des IQWiG, der erste Teil der Aussage von Sackett ausgelassen, wenn nicht sogar wesentlich unterdrückt, wobei Sackett diese Aussage sachgerecht und qualifiziert definiert: „Mit individueller klinischer Expertise meinen wir das Können und die Urteilskraft, die die Ärzte durch ihre Erfahrung und klinische Praxis erwerben“ [1]. Der Arzneimittelexperte Harald Schweim weist darauf hin, dass die Berichte und Äußerungen des IQWiG nur „eine“ von vielen wissenschaftlichen Fachmeinungen darstellen, die sich auch der Begutachtung der fachkompetenten Ärzte mit ihren praxisgeprägten Erfahrungen unterziehen müssen.“ (Ärztezeitung

2/2006). Daher kann nicht akzeptiert werden, dass bei der Analyse der Evidenz lediglich direkte Vergleichsstudien herangezogen, andere Daten aber nicht zur Bewertung zugelassen werden. Dies entspricht nicht der wissenschaftlichen Definition von EBM, da wertvolle Informationen ausgeschlossen werden [1]. Durch die Methodik des IQWiG wird das über Jahre gewonnene Fachwissen der Urologen ad absurdum geführt.

Es ist weder nachvollziehbar, aus welchem Grunde über 600 teilweise multizentrische Studien der Evidenzstufen II und III, insbesondere bei der Untersuchung von Teilaspekten (z.B. Wirkung auf die Obstruktion in urodynamischen Untersuchungen, z.B. Veränderung des Prostatavolumens) keine Berücksichtigung in der Bewertung finden, noch warum auch die vorhandene Literatur des Evidenzlevels I nahezu komplett aufgrund angeblicher Mängel im Studiendesign als nicht bewertbar eingestuft werden. In der Summe werden ca. 20 Jahre weltweiter klinischer Forschung auf das simple Ergebnis reduziert, es sei nicht möglich zu erkennen, ob die verschiedenen nicht medikamentösen Verfahren der Referenz-Therapie überlegen, unterlegen oder als gleichwertig zu betrachten seien.

Unannehmable Fristen

Das IQWiG hat den umfangreichsten Vorbericht in der Geschichte des Institutes im Internet am 26.03.2007 veröffentlicht. Das IQWiG verlangt durch seine Vorgaben, dass die Fachgesellschaft und andere auf den 500 Seiten umfassenden Bericht innerhalb von 4 Wochen eine Stellungnahme verfassen. Dies widerspricht eklatant der vom IQWiG selbst geübten Praxis der Verzögerung des Vorberichtes um 1,5 Jahre ohne Mitteilung von Gründen. Offensichtlich benötigte das IQWiG selbst unter Einsatz erheblicher Ressourcen beinahe 3 Jahre zur Bearbeitung, was die Komplexität des Themas unterstreicht. Das IQWiG unterstellt, dass über die Vorabveröffentlichung des Vorberichtes gegenüber der AWMF eine automatische Informationskaskade zu den Fachgesellschaften in Gang gesetzt wird. Eine Solche Informationskaskade besteht nicht. Eine Überprüfung der Informationsweitergabe erfolgte ebenso nicht. Selbst wenn eine offizielle Übergabe des Berichtes an die betreffende Fachgesellschaft erfolgt wäre, wäre die genannte Frist in Relation zum Umfang des Vorberichtes völlig unzureichend. Eine umfassende, wissenschaftlich fundierte Bearbeitung des Vorberichtes in der angegebenen Frist ist kaum möglich. Das strikte Beharren auf dem vorgegebenen Termin trotz Beschwerde der Deutschen Gesellschaft für Urologie und des Berufsverbandes der Deutschen Urologen unter Nennung nachvollziehbarer Gründe muss als unverantwortliches Desinteresse bzw. zielgerichtete Massnahme verstanden werden, um eine fundierte Stellungnahme der zuständigen Fachgesellschaft, die sich auch als Anwalt urologischer Patienten versteht, zu erschweren.

Selektion der Studien

Die IQWiG-Systematik enthält trotz ausführlich Beschreibung weder eine Erklärung, wodurch sich „relevante“ von „nicht relevanten“ Studien unterschieden, noch wie „Duplikate“ identifiziert und ausgeschlossen wurden und – besonders wesentlich – nach welchen Stichwörtern überhaupt gesucht wurde. Somit ist die Zahl primär ausgeschlossener Studien sehr hoch und übersteigt die Zahl der berücksichtigten Studien bei weitem. Dabei ist nicht nachvollziehbar, aufgrund welcher konkreten Gründe diese Auswahl erfolgte (vor den im Bericht beschriebenen ersten Schritt der Literatursuche). Die weiteren Auswahlsschritte werden im Bericht zwar vage beschrieben, auch, dass die Auswahl von 2 externen Gutachtern durchgeführt wurde. Entgegen der angeblichen Transparenz werden jedoch weder die Personen und die Qualifikation der Gutachter benannt, noch das System, nach dem diese ihre Bewertung vornahmen, noch die Gründe, die zum Ausschluss zahlreicher Studien als „nicht relevant“ führten. Es wird nicht einmal erwähnt, ob die Gutachter unabhängig voneinander bewerteten und was beispielsweise bei fehlendem Konsens geschah (Ein- oder Ausschluss der Studie?).

Die Subjektivität des Verfahrens wird auch im weiteren Bericht erkennbar, da zum einen doch nicht ausschliesslich RCTs, sondern gelegentlich auch CCTs Bewertung fanden, andererseits auch später einzelne Studien ausgeschlossen wurden und nicht zuletzt in der Interpretation der festgestellten Ergebnisse erneute Ausschlüsse zuvor eingeschlossener Studien erfolgten.

Im Detail wird unter „Kapitel 4.4.2 Gegenüberstellung der Ergebnisse der Einzelstudien“ ausgeführt, dass Studien ausgeschlossen wurden, „wenn sie auf weniger als 70 % der in der Studie eingeschlossenen Patienten basierten beziehungsweise wenn der Anteil der fehlenden Werte größer als 30 % war. Dieses Vorgehen wurde auch angewendet, wenn der Unterschied der fehlenden Werte zwischen den Gruppen größer als 15 % war. Die Festlegung auf 30 % ist zwar willkürlich, kann aber als vergleichsweise liberal angesehen werden.“ Diese Kriterien sind vor allem für die essentiell notwendige Beurteilung von Langzeitergebnissen relevant. Die Festlegung von Grenzwerten für die Ausfallquote gestaltet sich schwierig. Die Ausfallquote ist unter anderem abhängig von Anzahl, Lage und Spannweite der Erhebungszeitpunkte. Als Begründung wird [2] zitiert, „der Ausfallraten größer als 20 % als nicht mehr aussagekräftig“ betrachtet. Betrachtet man den Artikel genauer, so sind die auf Seite 784 angegebenen hohen Ansprüche zur Maximierung der Patientenzahl im Follow-up für die BPS-Population keinesfalls zu erfüllen, da nur ausreichend junge Patienten mit einem wahrscheinlichen F/U von 5 Jahren eingeschlossen werden dürften, eine für die BPS-Population nicht-repräsentative Gruppe. Was zunächst als hohes Qualitätsmerkmal der Studien erscheint, erweist sich im Ergebnis als unzulässige Einschränkung der Einschlusskriterien mit einem erheblichen Potential zum Selektionsbias.

Obwohl die Durchführung von RCT's zweifelsfrei essentiell notwendige Voraussetzung für die Einführung neuer Behandlungsoptionen ist, weist die Reduktion der Studienauswahl auf RCT's bei operativen / instrumentellen Verfahren zur Behandlung des BPS erhebliche Probleme auf, die bei [3] zusammengefasst werden. Das Konzept der Randomisierung bleibt insbesondere dann schwierig, wenn invasive (teilweise in Narkose durchgeführte) mit kaum invasiven Methoden (teilweise ohne Narkose bzw. stationär oder ambulant durchgeführte) verglichen werden.

Die Wünsche des Patienten nach einem weniger belastenden Verfahren müssen bei zutreffender Indikation zugunsten der Randomisierung ignoriert werden und sind Grund für schlechte Rekrutierung oder hohes Drop out. Meist liegen deshalb nur begrenzte Fallzahlen vor. Zusätzlich ist eine Verblindung schwierig und stößt an die Grenzen ethischer Überlegungen. In der Durchführung der Studien weist die unterschiedlich lange Lernkurve des Operateurs für die differenten Verfahren einen Bias unbekannter Größenordnung auf. Die Aussagekraft von RCT's für die klinische Praxis ist damit erheblich eingeschränkt und deshalb fehlerhaft. Das vom IQWiG getroffene Fazit zur Beurteilung aller nichtmedikamentöser lokaler Verfahren nur auf Grundlage von RCT's ist daher völlig inakzeptabel.

Erstellung des Vorberichts ohne Einbindung von Fachgesellschaft, Patientengruppen, Geräteherstellern

Die Methodik des IQWiG postuliert die Einbindung der Fachgesellschaft bei der Erstellung des Vorberichtes ebenso, wie die Kontaktaufnahme einzelner Autoren und ggf. der Herstellerfirmen.

Auf S. 107 wird unter „Allgemeinen Grundsätzen einer themenbezogenen Literaturrecherche“ in der „Auswahl der Datenquellen“ die „Kontaktaufnahme mit Fachexperten/Industrie/Patientenorganisationen etc.“ beschrieben. Im Bericht sind weder die Meinungen bekannter Selbsthilfegruppen noch die legitimerter Experten nationaler oder internationaler Fachgremien erkennbar eingeholt oder einbezogen worden. Der Kontakt zur Deutschen Gesellschaft für Urologie (DGU) reduzierte sich auf eine Anfrage im Jahr 2005 mit wenigen, im Kapitel 4.1.2. dargestellten Fragen zu nicht mehr durchgeführten Methoden. Die im Antwortbrief der DGU dezidiert angebotene weiterführende Zusammenarbeit wurde im Weiteren nicht genutzt. Die im „Projektplan“ genannten „externe Sachverständige... die an der Erstellung des Berichtsplans, an der Informationsbeschaffung und -bewertung sowie an der Erstellung des Vorberichts beteiligt waren sind der DGU unbekannt, von dieser Fachgesellschaft nicht legitimiert und offensichtlich nach eigenen nicht näher bezeichneten Kriterien des Instituts ausgewählt worden.

Die telefonische Kontaktaufnahme mit den entsprechenden Personen und Firmen ergab, dass hier nicht einzelne Ausschlüsse erfolgt sind, sondern ein systematischer Ausschluss aller Experten und Firmen. Niemand wurde im Gesamtzeitraum von 3 Jahren, die das IQWiG benötigte, um diesen Vorbericht zu erstellen, kontaktiert.

Dies resultiert in der falschen Abgrenzung akzeptierter von nicht mehr durchgeführten oder als obsolet bzw. experimentell geltenden Methoden, der fehlerhaften Einordnung physikalisch gleichwertiger Therapieverfahren wie Laser- und Elektrovaporisation, oder der fehlenden Unterscheidung ablativer von

nicht ablativen TUMT-Verfahren. In diesen grob falschen Zuordnungen wird in besonderem Masse die mangelhafte urologische Expertise sichtbar.

Falsche Definition von Behandlungszielen durch Fehlen urologischen Fachwissens:

Ziele der vorliegenden Untersuchung waren die

- vergleichende Nutzenbewertung verschiedener nichtmedikamentöser, lokaler Behandlungsverfahren des BPS mit den operativen Standardverfahren hinsichtlich patientenrelevanter Therapieziele,*
- vergleichende Nutzenbewertung verschiedener nichtmedikamentöser, lokaler Behandlungsverfahren des BPS untereinander hinsichtlich patientenrelevanter Therapieziele und die*
- vergleichende Nutzenbewertung verschiedener nichtmedikamentöser, lokaler Behandlungsverfahren des BPS mit anderweitigen Interventionen hinsichtlich patientenrelevanter Therapieziele.*

Obwohl das Erreichen patientenrelevanter Behandlungsergebnisse erklärtes Ziel des Berichtes ist, bleiben die Wünsche der Patienten nach geringerer Invasivität und Morbidität als sie mit den Referenzverfahren (TURP, TUIP, offenen Operation) potentiell verbunden sind [4], gänzlich unberücksichtigt, ebenso wie die Möglichkeiten der ambulanten Behandlung mit einigen Therapieverfahren. Dies ist ein grober Widerspruch zu den publizierten Methoden des IQWiG vom 19.12.2006 [5], wo zu patientenrelevanten Zielen auf den folgenden Seiten folgendes festgestellt wird:

S. 27: „Zur abschliessenden Darstellung der Ergebnisse gehören... eine Diskussion der klinischen Relevanz der Ergebnisse, die auf der Ermittlung patientenrelevanter Zielkriterien aufbauen sollte...“

S. 102: „die (grobe) Definition der Zielkriterien, insbesondere der patientenrelevanten Endpunkte... unter Einbeziehung externer Expertise und/oder Meinung, diese Definition verfeinert“ wird. „Hierbei werden regelhaft einzelne Betroffene, Patientenvertretungs- und/oder Verbraucherorganisationen für die themenbezogene Definition patientenrelevanter Endpunkte einbezogen.“

Die Interessen der Patienten bezüglich der Art der Behandlung, mit dem berechtigten Wunsch nach weniger invasiven oder ambulant durchführbaren Verfahren mit geringerer Morbidität sind in dem Vorbericht nicht ausreichend berücksichtigt worden, was nicht in Einklang mit dem § 139a Abs. 5 des SGB V (Neufassung GKV-WSG) steht. Gerade diese Gesichtspunkte haben zur Einführung dieser Behandlungsverfahren geführt, insbesondere bei Vorliegen spezieller Risikofaktoren und Begleiterkrankungen, die die Anwendung von Referenzverfahren ausschließen.

Die Gesamtaussage des Vorberichts ist somit irreführend. Bei der Verfeinerung der Definition von Behandlungszielen ist urologisches Fachwissen oder die Meinung von Patienten nicht eingeholt worden. Es findet sich ein Verstoß gegen die vom Institut selbst definierten Methoden.

Fehlende Sachkenntnis bei der Beschreibung der zugrunde liegenden Pathophysiologie

... „heute wird der Begriff „benignes Prostatasyndrom“ (BPS) verwendet, weil die Grösse der Prostata nur begrenzt Einfluss darauf hat...“ Es findet sich hier ein tiefes Unverständnis der Gründe für die Terminologie BPS, die sich aus der Pathophysiologie der Erkrankungen erklärt und in den Leitlinien der Deutschen Urologen unter dem Punkt Terminologie dargestellt ist [6].

„Diese Beschwerden können so unangenehm sein, dass Männer eine Behandlung wünschen...“

die wichtige imperative Indikation zu einer Behandlung, die in derartigen Fällen besteht, wird nicht erwähnt. Das überaus wichtige Problem der Blasenaustritts-Obstruktion (BOO) bleibt ebenfalls ohne Berücksichtigung.

„Das Spektrum reicht von anfänglichem Beobachten, ob sich die Beschwerden von selbst bessern, über Medikamente bis hin zu zahlreichen Operationsverfahren, bei denen die Prostata ganz oder teilweise entfernt wird.“ Die Prostata wird im Falle einer operativen Therapie des BPS niemals ganz entfernt. Der Tatsache, dass oft erst jahrelang medikamentös therapiert wird, und gerade die Therapieversager unter großem Leidensdruck nach einer schonenden Operation suchen, wird nicht erkannt. „Bislang ist aber unklar, welche Vor- und Nachteile diese Optionen haben.“

Die Vor- und Nachteile der Verfahren sind hinreichend bekannt und unter anderem in den aktuellen Leitlinien der Deutschen Urologen [7], den Leitlinien der American Urological Association [8] und den Leitlinien der European Association of Urology [9] dezidiert dargestellt.

Unkenntnis über Referenzverfahren und deren Alternative, Unkenntnis über physikalische Grundlagen

Bei der im Vorbericht als „Standard“ anerkannten Operationsverfahren fehlt die Erläuterung der sich zwar überlappenden, jedoch grundsätzlich unterschiedlichen Indikationsbereiche der Methoden. Stattdessen wird der Eindruck einer Gleichwertigkeit für alle Patienten mit BPS erweckt. Die methodische Beschreibung der TURP („... wird eine endoskopische Sonde über die Harnröhre in die Prostata eingestochen ...“) belegt die völlige Sachunkenntnis der Autoren.

Bei der Definition der Standardverfahren der operativen Therapie und in der Charakterisierung der Studien wird neben der transurethralen Vaporesektion der Prostata auch die transurethrale Elektrovaporisation der Prostata genannt (S. 27: *Die als Standardverfahren anerkannten Interventionen (TURP, TUIP, Adenomektomie, TVP, TUEVRP) wurden zu einer Gruppe zusammengefasst und innerhalb eines Kapitels dargestellt*). Eine schlüssige Begründung fehlt vollständig.

Die Gleichstellung der Vaporesektion mit der Resektion im Sinne einer Modifikation mag dabei noch angehen, da das wesentliche Element der Vaporesektion die Resektion darstellt. Völlig anders hingegen ist jedoch die Elektrovaporisation zu bewerten, bei der gerade im Hinblick auf die potentiell geringere Eingriffsmorbidität auf jegliche Resektion verzichtet und das Gewebe stattdessen durch Vaporisationseffekte entfernt wird. Bei der Bewertung nicht medikamentöser Behandlungsverfahren ist dieser Unterschied von höchster Wichtigkeit, denn es ergibt sich ein logisches Dilemma:

Nach der im Vorbericht geführten Argumentation stellt die Elektrovaporisation eine Modifikation der Elektroresektion dar. Damit stellt sich zwangsläufig die Frage, warum die Laservaporisation keine Modifikation der transurethralen Resektion der Prostata darstellt. Der zu erwartende Gewebeeffekt bei Anwendung der Elektrovaporisation bzw. der Laservaporisation unterscheidet sich allenfalls quantitativ, was bei operateurabhängigen Massnahmen durch Änderung der Applikationszeit bis zur Erzielung des gewünschten Ergebnisses angepasst wird. Somit muss auch die Laservaporisation eine Modifikation der Elektroresektion sein. Erst recht muss die Holmiumlaserresektion, die lt. Vorbericht in der Lage ist ist, „...*Prostatagewebe ähnlich wie bei einer TURP in kleinen Stücken herauszuschneiden...*“, eine Modifikation der transurethralen Elektroresektion sein. Auch hier liefern die publizierten experimentellen Daten zu den Schneideeigenschaften des Holmiumlasers und von Hochfrequenzstrom genügend evidenzgesicherte Informationen. Aus welchen Gründen im Vorbericht trotz gleichartiger Gewebeeffekte unterschiedlicher Energieträger (Laser bzw. Hochfrequenzstrom) diese nicht als Modifikation, grundverschiedene Gewebeeffekte (Resektion bzw. Vaporisation) derselben Energie (Hochfrequenzstrom) hingegen als Modifikation eines Verfahrens angesehen werden, bleibt unbelegt und völlig unklar und ist damit subjektiv und willkürlich. Die Beschränkung einer „Modifikation“ auf den Einsatz hochfrequenter elektrischer Energie entbehrt jeder wissenschaftlichen und biophysikalischen Grundlage. Für die trotz erkannter Ähnlichkeiten als eigenständiges Verfahren, nicht als Modifikation betrachteten Holmiumlaserresektion wird willkürlich postuliert, dass die klinische Wirkung der Geweberesektion mittels hochfrequentem Strom sich von der Resektion mit einem Laser unterscheidet. Hierfür gibt es weder praktische, noch theoretische Beweise. Unterschiede können lediglich in der technischen Ausführung der Resektion bzw. ihrer Quantität zu suchen sein und sind somit nicht methodisch bedingt, sondern anwenderabhängig.

Die Prinzipien der Kontakt-Laserablation sowie der visuellen Laserablation werden verfälscht dargestellt. Bei beiden Techniken wird tatsächlich keine Lasersonde in die Prostata eingeführt, vielmehr erfolgt die Applikation der Laserstrahlung jeweils von der Harnröhre aus, im ersten Fall mit dem Ziel der Gewebevaporisation, im zweiten Fall mit der Erzeugung einer grossvolumigen Koagulationsnekrose.

Im Kapitel „*Hochfokussierter Ultraschall (HIFU)*“ wird nicht darauf hingewiesen, dass es sich um ein obsoletes Verfahren handelt [7]. Auch die Methode der „*Wasserinduzierten Thermoerapie (WIT)*“ wird

in den Therapie-Leitlinien der Deutschen Urologen [7] nicht genannt und wird im Bericht nicht als klinisch irrelevantes experimentelles Verfahren erkannt und kenntlich gemacht.

Zusammenfassung

Der IQWiG-Bericht stellt im Fazit fest, dass „für keines der Verfahren ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden kann“.

Dies widerspricht dramatisch den geltenden Leitlinien nationaler und internationaler Fachgesellschaften, die die geringeren Nebenwirkungen als entscheidenden Vorteil der Verfahren gegenüber der Standardtherapie sehen und die der eigentliche Grund zur Entwicklung und weltweiten wissenschaftlichen Untersuchungen in den letzten 15 Jahren waren.

Obwohl das IQWiG selbst durchaus vergleichbare Effekte auf die Verbesserung der Lebensqualität und Vorteile der Methoden in Bezug auf die Krankenhausverweildauer und die Katheterisierungsdauer akzeptiert, werden wissenschaftlich valide essentielle Daten zu unerwünschten Ereignissen infolge einer dem Institut eigenen, weithin umstrittenen Interpretation evidenzbasierter Medizin in der Beurteilung gänzlich ignoriert. Obwohl das Erreichen patientenrelevanter Behandlungsergebnisse erklärtes Ziel des Berichtes ist, bleiben die Wünsche der Patienten nach geringerer Invasivität und Morbidität als sie mit den Standardverfahren (TURP, TUIP, offenen Operation) potentiell verbunden sind, gänzlich unberücksichtigt, ebenso wie die Möglichkeiten der ambulanten Behandlung mit einigen Therapieverfahren. Dies ist ein grober Widerspruch zu den publizierten Methoden des IQWiG vom 19.12.2006, wo festgestellt wird, dass „die (grobe) Definition der Zielkriterien, insbesondere der patientenrelevanten Endpunkte...unter Einbeziehung externer Expertise und/oder Meinung, ...verfeinert“ wird und „regelmäßig einzelne Betroffene, Patientenvertretungs- und/oder Verbraucherorganisationen für die...Definition patientenrelevanter Endpunkte einbezogen“ werden.

Im Bericht sind weder die Meinungen bekannter Selbsthilfegruppen noch die legitimer Experten nationaler oder internationaler Fachgremien erkennbar eingeholt oder einbezogen worden.

Insbesondere die unzureichende Abgrenzung akzeptierter von nicht mehr durchgeführten oder als obsolet bzw. experimentell geltenden Methoden, die fehlerhafte Einordnung physikalisch gleichwertiger Therapieverfahren wie Laser- und Elektrovaporisation oder die mangelnde Unterscheidung ablativer von nicht ablativen TUMT-Verfahren zeigen die mangelhafte urologische Expertise. Die Vorgaben des IQWiG zur Einbindung von Patientengruppen, Fachgesellschaften, Geräteherstellern etc. sind verletzt worden.

Der AK BPH der Akademie der Deutschen Urologen kommt zu der Einschätzung, dass der IQWiG-Bericht durch die unzureichende Berücksichtigung urologischen Sachverständes, die fehlende Berücksichtigung der Pathophysiologie des BPS, einer grob fehlerhaften Einschränkung patientenrelevanter Therapieziele und der damit im Ergebnis lückenhaften Berücksichtigung und Interpretation essentiell wichtiger wissenschaftlicher Daten schwere Mängel in der Methodik seiner Bewertung aufweist. Der Bericht ist deshalb in seinem Fazit zwangsläufig falsch und grob irreführend.

Dem IQWiG muss empfohlen werden, den Bericht zurückzuziehen und eine Neubewertung vorzunehmen.

Literatur

- [1] Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-Based Medicine. 2000, Churchill Livingstone. 2000:S1-11.
- [2] Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002;359:781-5.
- [3] Chapple CR. Clinical study of benign prostatic disease, current concepts and future prospects: randomized controlled trials versus real life practice. Curr Opin Urol. 2003 Jan;13(1):1-5.
- [4] Kaplan S, Naslund M. Public, patient, and professional attitudes towards the diagnosis of enlarged prostate: a landmark national US survey. 2006 [cited Int J Clin Pract 60: (10): 1157-1165, 2006 04.05.2007]; Available from: <http://www.blackwell-synergy.com/doi/pdf/10.1111/j.1742-1241.2006.01129.x>

Stellungnahme Arbeitskreis BPH der Akademie der Deutschen Urologen vom 09.05.2007

- [5] Bastian H, Bender R, Ernst AS, Kaiser T, Kirchner H, Kolominsky-Rabas P, et al. Methoden, Version 2.0 vom 19.12.2006. 2006 19.12.2006 [cited 04.05.2007]; Available from: http://www.iqwig.de/download/Methoden_IQWiG_V-2-0.pdf
- [6] Berges R, Dreikorn K, Höfner K, Jonas U, Laval KU, Madersbacher S, et al. Leitlinien der Deutschen Urologen zur Diagnostik des Benignen Prostatasyndroms (BPS). 2003 27.09.2005 [cited 04.05.2007]; Available from: <http://www.uni-duesseldorf.de/AWMF/II/043-034.htm>
- [7] Berges R, Dreikorn K, Höfner K, Jonas U, Laval KU, Madersbacher S, et al. Leitlinien der Deutschen Urologen zur Therapie des Benignen Prostatasyndroms. 2003 27.09.2005 [cited 04.05.2007]; Available from: <http://www.uni-duesseldorf.de/AWMF/II/043-035.htm>
- [8] Roehrborn CG, McConnell JD, Barry MJ, Benaim E, Bruskewitz RC, Blute ML, et al. Guideline on the Management of Benign Prostatic Hyperplasia (BPH). 2003 Updated 2006 [cited 04.05.2007]; Available from: <http://www.auanet.org/guidelines/bph.cfm>
- [9] de la Rosette J, Alivizatos G, Madersbacher S, Rioja Sanz C, Nordling J, Emberton M, et al. Guidelines on Benign Prostatic Hyperplasia. 2006 [cited 04.05.2007]; Available from: <http://www.uroweb.org/fileadmin/userupload/Guidelines/11%20BPH.pdf>
- [10] Hoffman RM, MacDonald R, Monga M, Wilt TJ. Transurethral microwave thermotherapy vs transurethral resection for treating benign prostatic hyperplasia: a systematic review. *BJU Int.* 2004 Nov;94(7):1031-6.
- [11] Ponholzer A, Marszalek M, Madersbacher S. Minimally invasive treatment of BPH: an update. *Eau Update Series.* 2004;2(1):24-33.
- [12] EMEA. ICH Topic E 9. Statistical Principles for Clinical Trials. Step 5. Note for Guidance on statistical principles for clinical trials. 1998.
- [13] Barry MJ, Williford WO, Chang Y, Machi M, Jones KM, Walker-Corkery E, et al. Benign prostatic hyperplasia specific health status measures in clinical research: how much change in the American Urological Association symptom index and the benign prostatic hyperplasia impact index is perceptible to patients? *The Journal of Urology.* 1995 Nov;154(5):1770-4.

4.2.4 Österreichische Gesellschaft für Urologie und Andrologie (ÖGU)

Autoren:

Doz. Dr. Stefan Madersbacher

Prof. Dr. G. Schatzl

Prim. Doz. Dr. Michael Rauchenwald

Adresse:

Abteilung für Urologie/Andrologie

Donauspital – SMZ Ost

Langobardenstrasse 122

1220 Wien

Österreich

ÖSTERREICHISCHE GESELLSCHAFT FÜR UROLOGIE UND ANDROLOGIE



Generalsekretär:
Univ.-Doz.Dr. Michael Rauchenwald
Abteilung für Urologie und Andrologie
Donauspital – SMZ Ost
A-1220 Wien, Langobardenstrasse 122
☎ (01) 28802 - 3700, Fax 3780
michael.rauchenwald@wienkav.at
ZVR-Zahl: 900565881

Prof. Dr. med. Peter T. Sawicki
Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
Dillenburger Str. 27
51105 Köln

EINGEGANGEN 21. Mai 2007

Wien, den 8.5.2007

Stellungnahme des Arbeitskreises Prostata der Österreichischen Gesellschaft für Urologie (ÖGU) sowie des Generalsekretärs der ÖGU zum Vorbericht „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie, IQWiG (26.3.2007)

Sehr geehrter Herr Prof. Sawicki,

Der Arbeitskreis Prostata der ÖGU hat die Veröffentlichung des o.g. Berichtes durch Ihr Institut am 03.04.2007 zur Kenntnis genommen und erachtet es für notwendig, folgende Stellungnahme zum o.a. Vorbericht abzugeben:

Seit über 20 Jahren existiert eine intensive Suche nach minimal invasiven Behandlungsverfahren für Patienten mit benignem Prostatasyndrom (BPS). Das weltweite Interesse an minimal invasiven Therapieverfahren bei BPS ist durch mehrere Umstände begründet.

1. **Demographie:** In den nächsten Jahren wird es zu einer deutlichen Zunahme der hochbetagten (80+) Patienten kommen. Für diese wichtige Patientengruppe ist eine schonende Behandlungsmethode mit minimalem

Blutungs- und TUR-Syndrom Risiko von großem Interesse. Die Tatsache, dass zum Beispiel die TUMT ohne Narkose und ambulant durchführbar ist, bedeutet einen großen Fortschritt für Patienten, die nicht narkosetauglich sind. Die Alternative für diese Patienten – folgt man den Empfehlungen des IQWiG - ist der lebenslange Dauerkatheter mit all seinen Komplikationsmöglichkeiten.

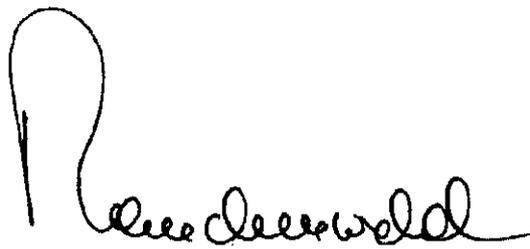
2. **Comorbidität:** Parallel zur demographischen Entwicklung sehen wir zunehmend multimorbide Patienten, denen eine konventionelle TURP zum Beispiel wegen einer Blutgerinnungsstörung oder oraler Antikoagulation nicht oder nur mit großem Risiko zugemutet werden kann. Für derartige Patienten ist heute der KTP-Laser weltweit die Methode der Wahl. Auch andere Verfahren (TUMT, TUNA, Holmiumlaser) können hier zur Anwendung kommen. Wenn diese Alternativen sicherer sind, warum sollten sie dann anderen Patienten vorenthalten werden. Wir sehen ein erhebliches ethisches Problem, wenn ein Patient nicht zwischen diesen Techniken wählen kann. Patienten möchten im Zeitalter von HIV nicht den Risiken einer Bluttransfusion ausgesetzt werden. In Deutschland wurden Transfusionsraten von 5-15% nach TURP berichtet. Auch hier liegt ein Grund für die Entwicklung dieser Verfahren.
3. **Patientenwunsch:** Viele Patienten insbesondere jüngere Patienten wünschen eine schonende Methode; so ist die retrograde Ejakulation, die überwiegend nach der TURP auftritt, für jüngere Patienten inakzeptabel. Auch die Vermeidung eines Spitalsaufenthaltes kommt dem Wunsch vieler Patienten entgegen.
4. **Technische Fortschritte:** Letztendlich ermöglicht der technische Fortschritt die Entwicklung minimal invasiver Therapieoptionen, wie der Mikrowellenthermotherapie, TUNA, dem KTP- oder Holmiumlaser

Auf Grund der oben angeführten Punkte werden heute einige der minimal invasiven Methoden, wie zum Beispiel die TUMT, die TUNA oder auch der Holmium oder KTP-Laser-TURP weltweit eingesetzt.

Der ***IQWiG-Vorbericht*** versucht die Wertigkeit der minimal invasiven Methoden anhand einer Metaanalyse zu bewerten. Dabei wurden zur Beurteilung primär komparative, randomisierte Studien (minimal invasive Therapie versus TURP) herangezogen. In weiten Teilen beschreibt der Bericht Methoden, die heute bereits wieder verlassen wurden und deshalb nicht mehr relevant sind. Darüber hinaus ist es methodisch problematisch, die große Anzahl von Phase II Studien nicht zu berücksichtigen, insbesondere da eine randomisierte Studie zwischen einer Methode, die eine Allgemeinnarkose benötigt (TURP) und einer, die in Lokalanästhesie durchgeführt werden kann (TUMT, TUNA) äußerst problematisch ist und deshalb nur ganz vereinzelt durchgeführt wurde. Problematisch erscheint auch die Tatsache, dass auf die Empfehlungen der diversen Leitlinien zur BPH-Therapie (Deutschen Urologen, Europäische und Amerikanische Urologengesellschaft) nicht eingegangen wird. Es ist auch unverständlich, weshalb der Bericht nicht **VOR** der Publikation den Experten (Urologen), die über Jahrzehnte BPH-Patienten behandeln und sich sowohl klinisch auch als grundlagenwissenschaftlich mit dem Thema befassen und BPH-Leitlinien verfassen, vorgelegt wurde.

Die Folge ist ein Bericht bzw. Schlussfolgerungen, die ganz wesentliche Aspekte in der Therapie des BPH völlig außer Acht lassen. Folgt man den Schlussfolgerungen des IQWiG, so besteht derzeit keine Indikation, diese Methoden einzusetzen. Sollte dies in Deutschland umgesetzt werden, würde dies einen bedeutender Rückschritt für manche Patientengruppen bedeuten: hochbetagte Patienten, Männer mit hohem Operationsrisiko oder solche mit Blutgerinnungsstörungen und Cochleaimplantaten bzw. implantierten Schrittmachern u.ä. könnten in Deutschland dann nicht mehr sicher behandelt werden. Diese Patienten haben die Wahl einer TURP mit sehr hohem Risiko (wer übernimmt hierfür die Verantwortung, das IQWiG?) oder einem lebenslangen Dauerkatheter. Auf den Patientenwunsch nach einer schonenden Behandlung kann in Deutschland nicht mehr eingegangen werden.

Zusammenfassend widersprechen die Empfehlung des IQWiG der gängigen klinischen Routine und allen wichtigen BPH-Leitlinien und bedeutet einen deutlichen Rückschritt in der Therapie dieses wichtigen Krankheitsbildes. Folgt man Ihrem Urteil, kann man sich als Europäer nur wünschen, in Österreich zu leben und nicht in Deutschland.



Prim. Univ. Doz. Dr. Stephan Madersbacher

Generalsekretär der Österreichischen Gesellschaft für Urologie und Andrologie



Univ. Doz. Dr. Stephan Madersbacher, ehem. Vorsitzender, Arbeitskreis

Prostata der ÖGU



Univ. Prof. Dr. Georg Schatzl, Vorsitzender, Arbeitskreis Prostata der ÖGU

4.2.5 Deutsche Krankenhausgesellschaft e. V.

Autoren:

Joachim Manz

Adresse:

Am Eichenhain 110
134645 Berlin

02.05.2007

Stellungnahme N04-01

„Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

1.) Im Vorbericht fehlende Originalstudien

2.) Anmerkungen zu Bewertung von Originalstudien im Vorbericht

Die Standardverfahren gemäß den Leitlinien sind: TUR-P, Adenomenukleation und Transurethrale Inzision. Die Indikationen für diese drei Verfahren werden anhand der individuellen Größe gestellt, so dass diese drei Standardverfahren nicht untereinander austauschbar sind, sondern durch die Patientenverhältnisse bestimmt werden.

Alle anderen Verfahren sind keine Standardverfahren. Es gibt jedoch Sonderindikationen für die Anwendung der Lasertherapie, vor allem dann, wenn Blutgerinnungsprobleme vorhanden sind. Dann besteht eine Indikation für die Laseranwendung, wobei ablativen Verfahren der Vorzug zu geben ist, d.h. die Holmiumresektion ist der Greenlightkoagulation vorzuziehen, da sie Gewebe entfernt und auch der histologischen Untersuchung auf etwaige unbekannte Karzinomanteile zuführen kann.

Letztlich sind alle anderen aufgeführten Verfahren experimentell, i.d.R. nur in Studien angewandt und gegenüber den Standardverfahren im Nachteil, da sie auf die Gewebegewinnung verzichten. Diese ist jedoch gerade bei einem Organ (Prostata), das für die häufigste Krebstodesursache des Mannes verantwortlich ist, essentiell, denn es ist bekannt, dass Tumore multifokal wachsen können, so dass die unbemerkte Beseitigung eines Herdes durch Koagulation keine Heilung bedeuten muss. Es ist deshalb unabdingbar, einen histologischen Befund zu haben, so dass die Nachsorge oder Weiterbehandlung optimiert erfolgen können.

Wait-and-see ist keine rationale Therapie, wenn nach Scheitern einer konservativen Therapie mit alpha-Blockern und/oder 5-alpha-Reduktasehemmern die Operationsindikation z.B. wegen unbeeinflussbaren Restharn besteht.

Das Abheben auf subjektive Beschwerden (z.B. unter Anwendung eines Scores) ist eine verkürzte Sicht. Objektive Parameter wie Restharnfreiheit, Normalisierung der Miktion, Reduktion einer Blasenwandhypertrophie, Beseitigung einer etwaigen Nierenstauung sind dagegen klinisch relevante Parameter. Ferner haben die Alternativverfahren bis dato im Langzeitverlauf über 10 und mehr Jahre nicht belegen können, dass sie den Standardverfahren zumindest gleichwertig sind. Im Gegenteil, die klinische Erfahrung lehrt, dass Rezidivadenome nach experimentellen Verfahren (einschließlich Greenlight-Laserung) nicht selten sind, auf jeden Fall häufiger als z.B. nach einer TUR-P.

Die bekanntesten Risiken einer TUR-P sind außerdem durch die bipolare Technik weitgehend beseitigt. Das Risiko des Entstehens einer Harninkontinenz durch eine operative Therapie ist dagegen allen Verfahren letztlich immanent. Die Erfahrung des Operateurs ist ein relevanterer Parameter für die Inkontinenzrate als das Verfahren an sich.

3.) Anmerkungen zur projektspezifischen Methodik

4.2.6 Kassenärztliche Bundesvereinigung

Autoren:

Dr. Paul Rheinberger

Dirk Horenkamp

Adresse:

Herbert-Lewin-Str. 2

10623 Berlin



Kassenärztliche
Bundesvereinigung

Körperschaft des öffentlichen Rechts

Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
- Stellungnahme zum Vorbericht Auftrag N 04/01 -
Prof. Dr. med. Peter T. Sawicki
Dillenburger Str. 27
51105 Köln

*Dezernat 1
Nutzenbewertung (HTA) ärztlicher Leistungen
Dirk Horenkamp, M.Sc.
Herbert-Lewin-Platz 2, Wegelystraße, 10623 Berlin
Postfach 12 02 64, 10592 Berlin*

*Tel.: 030 / 40 05 – 1112
Fax: 030 / 40 05 – 1190
E-Mail: DHorenkamp@kbv.de
www.kbv.de*

*Hk. / Ge. / AZ B 23
9. Mai 2007*

*nachrichtlich:
Dr. von Pritzbuer, G-BA-Siegburg*

Stellungnahme der KBV zum Vorbericht des IQWiG zu „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“ (Version 1.0, 26.03.2007, Auftrag N 04/01)

Sehr geehrter Herr Prof. Sawicki,

hiermit dürfen wir Ihnen unsere Stellungnahme zum o. g. Vorbericht übermitteln. Aus den Erfahrungen bisheriger Themengruppen und den hierzu vom IQWiG vorgelegten Endberichten möchten wir die Autoren des Vorberichtes zu „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“ bitten, die folgend aufgeführten Punkte bzw. Verbesserungsvorschläge im Endbericht zu berücksichtigen.

Zu den G-BA-Stellungnahmen:

1. Die vom G-BA eingeholten und an das IQWiG übermittelten Stellungnahmen sollten im Bericht berücksichtigt, d. h. zumindest in ihren wesentlichen Aussagen kurz zusammenfassend dargestellt werden.

Zur Notwendigkeit:

2. Die Prüfung und Darstellung der medizinischen Notwendigkeit der Standardtherapie und der im Bericht geprüften Vergleichstherapie fehlt gänzlich und sollte entsprechend § 20 Abs. 2 Verfahrensordnung dargestellt sein.

Mit freundlichen Grüßen

Dr. Rheinberger
- Dezernent -

D. Horenkamp
- Referent -

4.2.7 Deutsche Krankenhausgesellschaft e. V.

Autoren:

Dr. Nicole Schlottmann

Dr. Claudia Simon

Dr. Kai Schorn

Adresse:

Wegelystr. 3

10623 Berlin



DEUTSCHE
KRANKENHAUS
GESELLSCHAFT

Institut für Qualität und
Wirtschaftlichkeit im Gesundheitswesen
- Stellungnahme zum Vorbericht <N04-01>
Herrn Prof. Dr. med. Peter T. Sawicki
Dillenburger Str. 27
51105 Köln

Bundesverband
der Krankenhausträger
in der Bundesrepublik
Deutschland

Unser Zeichen

V/ Dr. Sim/ Dr. Scho/wut

Telefon

Durchwahl +49 (0) 30 398 01 -

1522

Telefax

+49 (0) 30 398 01 -

3510

Datum

04.05.2007

EINGEGANGEN 07. Mai 2007

Stellungnahme zum Vorbericht N04-01 „Nichtmedikamentöse lokale Verfahren zur Behandlung der Benignen Prostatahyperplasie“

Sehr geehrter Herr Professor Sawicki,
sehr geehrte Damen und Herren,

nachfolgend möchten wir zum Vorbericht N04-01 (Nichtmedikamentöse lokale Verfahren zur Behandlung der BPH) wie folgt Stellung nehmen:

Der Bericht stellt den Vergleich verschiedener, nichtmedikamentöser lokaler Behandlungsverfahren der BPH mit den operativen Standardverfahren und untereinander umfassend auf fast 500 Seiten dar. Erfreut stellen wir fest, dass das IQWiG in seiner Diskussion die Aussagen anderer aktueller systematischer Übersichten bzw. HTA-Berichte einbezieht und mit den Aussagen des IQWiG vergleicht. Daher haben wir zunächst nur einige allgemeine Anmerkungen zum Vorbericht vorzubringen, die vor allem die Lesbarkeit und den Umgang mit Stellungnehmenden betreffen.

Die vom IQWiG zunächst vorgegebene vierwöchige, jetzt fünfwöchige Stellungnahmefrist für die interessierte Fachöffentlichkeit ist zu kurz, um Stellungnehmenden eine fachlich fundierte Auseinandersetzung mit den Inhalten des Berichtes zu ermöglichen. Gerade mit Blick auf das zum 01.04.2007 in Kraft getretene GKV-WSG, in dem die Bedeutung der Stellungnahme sowie deren Berücksichtigung im Rahmen des Bewertungsverfahrens durch das IQWiG aufgewertet wurden, empfehlen wir hier zukünftig für Berichte ähnlichen Umfangs ein Überdenken der Stellungnahmefristen durch das IQWiG mit entsprechenden Verlängerungen.

Der Bericht, der 14 verschiedene therapeutische Verfahren überprüft, ist in seiner vorliegenden Gliederung äußerst komplex und daher schwierig zu lesen. So werden z. B. für jedes Verfahren für jeden damit vorgenommenen Vergleich das Studiendesign und die Studienpopulation sowie die Studien- und die Publikationsqualität beschrieben und im Anschluss die Ergebnisse zu den Therapiezielen bzw. eine Zusammenfassung der Ergebnisse des Vergleiches dargestellt. Dieses Vorgehen fordert ein ständiges

Vor- und Zurückblättern zwischen Ergebnissen und Studienqualität und erschwert damit die Lesbarkeit des Vorberichtes. Ferner bitten wir darum, dass das IQWiG seinen Berichten die von seinen Experten ausgefüllten Bewertungsbögen über die ausgewerteten Studien analog dem Vorgehen im Gemeinsamen Bundesausschuss anhängt, damit die Fachöffentlichkeit die einzelne Studienbewertung transparent und übersichtlich nachvollziehen kann.

Das IQWiG legt in seinem Berichtsplan vom 13.05.2005 als Zielgrößen für die Untersuchung die aus seiner Sicht patientenrelevanten Therapieziele Reduktion der irriativen und obstruktiven Symptome, Besserung bzw. Erhalt der krankheitsspezifischen Lebensqualität der Patienten, Reduktion unerwünschte Therapienebenwirkungen und Komplikationen, Reduktion der Dauer von Krankenhausaufenthalten und generell Vermeidung von Krankenhausaufenthalten, Reduktion der Dauer und Notwendigkeit der Katheterisierung sowie der Reduktion symptomatischer Harnwegsinfekte fest. Die von Ihnen genannten und bearbeiteten Therapieziele entsprechen im Wesentlichen denen anderer Übersichtsarbeiten, wobei diese gleich zu Beginn feststellen, dass die rein isolierte Betrachtung der einzelnen Therapieziele und die Gewichtung der Ergebnisse nach ihnen für die mit diesen Behandlungsverfahren zu erreichenden Patienten nur bedingt relevant sein können. So diskutiert z.B. der von Ihnen berücksichtigte HTA-Bericht der Agency for Healthcare Research and Quality von 2004, „für diese weniger invasiven Behandlungsmethoden mag es nicht nur eine einfache Frage der vergleichbaren Nutzenbewertung mit TURP sein, sondern eher die Frage, ob niedrigere Komplikations- und Nebenwirkungsraten einen angemessenen Kompromiss für vermutlich etwas geringere Symptomreduktion und einer vermutlich vorhandenen Notwendigkeit einer erneuten zukünftigen Behandlung wäre“. Dies sollte im vorliegenden Bericht diskutiert und beurteilt werden, da nur so eine Übertragung auf die Versorgungsrealität möglich erscheint. Im Übrigen geht aus dem Bericht nicht hervor, ob das Institut sich bei der Auswahl der patientenrelevanten Endpunkte an sein im Methodenpapier beschriebenes Verfahren gehalten hat. Dort ist festgelegt, dass „insbesondere einzelne Betroffene, Patientenvertretungs- und/oder Verbraucherorganisationen für die themenbezogene Definition patientenrelevanter Endpunkte einbezogen werden“.

Wie oben bereits ausgeführt sehen wir uns aufgrund der kurzen Fristsetzung nicht in der Lage, im Detail fachlich Stellung zu nehmen. Nichtsdestotrotz möchten wir im Folgenden weitere Aspekte exemplarisch zur Diskussion stellen:

1. Offensichtlich handelt es sich bei allen untersuchten Verfahren um Innovationen, die mit dem Ziel entwickelt worden sind, die Behandlung für den Patienten weniger invasiv mit geringerem Komplikationsrisiko bei ähnlicher Effektivität durchführen zu können. Vor diesem Hintergrund ist die Grundaussage des Berichtes, dass gerade zu den Unerwünschten Ereignissen (UE) keine belastbaren Aussagen zu machen sind, sehr unbefriedigend. Andere systematische Übersichtsarbeiten versuchen diesen Erkenntnismangel durch die zusammenfassende Darstellung von Fallserien zumindest für die UE zu beheben.
2. Fast durchgängig wird im Bericht das Fazit gezogen, dass aufgrund einer fehlenden Überlegenheit nicht der Schluss auf Gleichwertigkeit gezogen werden darf, da keine der Studien als Äquivalenzstudien angelegt war. Formal methodisch ist diese Feststellung zwar richtig, jedoch ist sie einerseits zu pauschal und andererseits für den Entscheidungsprozess im G-BA wenig hilfreich. Aufgrund der ausgewiesenen Expertise des Institutes sollte man erwarten, dass im

Rahmen der durchgeführten Metaanalysen die berechneten Konfidenzintervalle herangezogen und mögliche Irrelevanzbereiche diskutiert werden, um so zu einer differenzierteren Aussage bezüglich Äquivalenz zu kommen. Vorsichtige Schlussfolgerungen in der Weise, dass das Verfahren A mit großer Wahrscheinlichkeit dem Verfahren B bei Endpunkt X nicht unterlegen ist, wären für das Bewertungsverfahren im G-BA wesentlich besser operationalisierbar als die bisherige Sprachregelung des Institutes. Angesichts der implizierten Kritik an fehlenden Äquivalenzstudien wäre es ebenfalls hilfreich, wenn das Institut zumindest exemplarisch eigene Fallzahlberechnungen für solche Studien vorlegen könnte, um abschätzen zu können, ob solche Studien für die Zukunft in der klinischen Realität durchführbar wären. Vor diesem Hintergrund erscheint es besonders problematisch, wenn das IQWiG dann Formulierungen aus anderen HTA's wie „ähnliche Verbesserungen“ oder „ähnliche Ergebnisse“ gleichsetzt mit der Institutsformulierung „kein Beleg eines Vorteils zugunsten einer Behandlung“.

3. Ohne Anspruch auf Vollständigkeit fällt auf, dass das IQWiG in seiner Bewertung zumindest bei einzelnen Verfahren zu einer erheblich kritischeren Bewertung kommt als andere internationale systematische Reviews (z.B. positives Fazit zum HoLep-Verfahren durch AHRQ 2004, MSAC 2006, ASERNIP-S 2003; zu TUNA durch NICE 2003). Obwohl das Institut im Diskussionskapitel versucht, die internationalen Bewertungsverfahren zu würdigen, wird die Diskrepanz der Bewertung (Ausnahme VLAP vs. Standard) nicht ausreichend transparent.
4. Es ist zwar bekannt, dass Autoren dazu neigen können in ihren wissenschaftlichen Publikationen neue Verfahren eher in einem positivem Licht darzustellen, dennoch macht es nachdenklich, dass z.B. das Fazit der Autoren aller vom IQWiG bewerteten RCT's zum Vergleich HoLep vs. Standard im Abstract ein im Vergleich zum IQWiG eher positives Fazit ziehen. Mit dieser Anmerkung soll die Bewertung des IQWiG nicht diskreditiert werden, sondern es soll vielmehr deutlich gemacht werden, in welchem wissenschaftlichen Spannungsfeld sich das vorliegende Bewertungsverfahren möglicherweise befindet.
5. Zur Beschreibung der Größe der Prostata werden teilweise Grammangaben, teilweise Milliliterangaben getroffen (z. B. Seite 6 und Seite 42 Grammangaben, auf Seite 29 erfolgt die Vorgabe der Milliliterangaben). Diese kleinen Inkonsistenzen lassen sich sicherlich leicht beseitigen.
6. Auffallend ist zudem die teilweise wertende Wortwahl des IQWiG. So wird beispielsweise auf Seite 68 das Unterkapitel 5.2.3.1.5 (Unerwünschte Ereignisse des Vergleiches CLAP vs. Standardbehandlung) begonnen mit: „Nur 4 der 5 Studien berichteten über unerwünschte Ereignisse,...“. Da 4 von 5 Studien 80 % entspricht, erscheint uns das „nur“ wertend. In wissenschaftlichen Texten sollte das Ergebnis sachlich und neutral dargestellt werden, um dem Leser keinen falschnegativen Eindruck über die Studienlage zu vermitteln.

Wir hoffen, dass wir Ihnen mit unseren Anmerkungen einige Anregungen für die Überarbeitung des Vorberichtes und die Erstellung weiterer Berichte geben konnten, weisen im Übrigen jedoch darauf hin, dass es sich bei den vorstehenden Ausführungen nicht um eine abschließende Stellungnahme handelt, da eine umfassende formale und inhaltliche Prüfung des Berichtes in der Kürze der vom Institut zur Verfügung gestellten Zeit nicht möglich war. Wir behalten uns daher vor, in das weitere Verfahren ggf. weitere Aspekte einzubringen.

Mit freundlichen Grüßen
Der Hauptgeschäftsführer
In Vertretung



Dr. N. Schlottmann
Geschäftsführerin Dezernat Medizin

Im Auftrag



Dr. med. C. Simon
Referentin Dezernat Medizin



Dr. med. K. Schorn
Referent Dezernat Medizin

Anlagen

4.2.8 Medtronic GmbH

Autoren:

Christian Weinrank

Wolfgang Frisch

Stefan Zenk

Adresse:

Emanue-Leutze-Str. 20

40547 Düsseldorf



Medtronic

**Stellungnahme zum
IQWiG-Vorbericht N04-01
Nichtmedikamentöse lokale Verfahren
zur Behandlung der BPH ¹⁾**

Mai 6, 2007

¹⁾ Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.
Nichtmedikamentöse lokale Verfahren zu Behandlung der benignen Prostatahyperplasie.
Vorbericht N04-01. Köln: IQWiG; 2007

Grundsätzliches

In dem 500-seitigen IQWiG-Vorbericht N04-01, Version 1.0 vom 26.03.2007 werden auf Basis einer aufwendigen Literaturrecherche zumeist randomisierte, kontrollierte Studien von insgesamt vierzehn Therapieinterventionen zur Behandlung der benignen Behandlung der BPH (BPS) im Vergleich zur „Standardbehandlung“ TURP, aber auch im Vergleich untereinander ausgewertet. Im IQWiG-Berichtsplan N04-01, Version 1.0 vom 13. Mai 2005 wird dabei im Vorfeld auf behandlungsbedürftige Patienten mit benigner Prostatahyperplasie (ICD-10-GM-2007 N40) abgestellt, ohne weitere Einschränkungen in Hinblick auf den Schweregrad der Erkrankung zu treffen. Sowohl die im Berichtsplan aufgeführte Methode der Informationsbeschaffung als auch die Ein- und Ausschlusskriterien für die Einbeziehung identifizierter Studien entsprechen den in der Gesundheitstechnologiebewertung methodisch anerkannten und in der Forschungspraxis etablierten Verfahren.

Die Schlussfolgerung des Institutes ist, „... dass sich für keines der 14 nichtmedikamentösen lokalen Verfahren zur Behandlung des BPH eine Aussage treffen lässt, ob sie hinsichtlich der Linderung der Symptome und Verbesserung der Lebensqualität besser als oder gleichwertig mit einer Standardtherapie sind.“

Dieses Resümee reduziert die insgesamt konstruktiven Ansätze der IQWiG-Wissenschaftler um ein nicht unerhebliches Ausmaß. Es stellt sich die Frage, warum das Institut die vielen Vergleiche (CLAP vs. VLAP, VLAP vs. WW, ILK vs. TUMT, ILK vs. TUNA, HoLRP vs. VLAP, HoLEP vs. HoBNI, TUMT vs. Sham, Tumt vs. WW, TUMT vs. medikamentöse Behandlung, TEAP vs. ILK, TEAP vs. TUNA, TEAP vs. ILK,) angestellt hat, wenn diese Gegenüberstellungen in der zusammenfassenden Bewertung mit keinem Satz Beachtung finden.

Darüber hinaus wird die knappe Zusammenfassung nicht der begrüßenswerten Vielfalt der in Abschnitt 4.1.3 des Vorberichts benannten Zielgrößen (Irritative und obstruktive Symptome, gesundheitsbezogene Lebensqualität der Patienten, unerwünschte Therapienebenwirkungen und – komplikationen, Dauer der Krankenhausaufenthalte, Dauer und Notwendigkeit der Katheterisierung, symptomatische Harnwegsinfekte) gerecht.

In Anbetracht der Vielfalt der verglichenen Verfahren und der uneinheitlichen Anwendung von Selektions- und Auswertungskriterien wird nachfolgend ausschließlich auf die für das Verfahren der Transurethralen Nadelablation (TUNA) relevanten Aspekte Bezug genommen.

Identifizierung relevanter Studien

Die im IQWiG-Vorbericht N04-01 referenzierten Literaturdatenbanken sind für eine Bewertung des TUNA-Verfahrens gängig und ausreichend. Positiv hervorzuheben ist der Ansatz von drei parallel durchgeführten Suchstrategien (Verwendung der Suchstrategie des AHRQ-HTA, die eigene Suchstrategie und eine separate Suchstrategie für Sekundärpublikationen) sowie die erheblich erweiterte Suchstrategie der Nachrecherche.

Unklar bleibt, ob innerhalb der 2144 Studien, die als „sicher nicht relevant“ erachtet wurden, Studien unzulässig ausgeschlossen wurden.

Viele Sekundärpublikationen haben zusätzliche Literaturquellen verwenden können, die von der IQWiG-Arbeitsgruppe nicht identifiziert worden sind, oder nach Identifikation ausgeschlossen wurden. Als Beleg dient der HTA-Bericht der französischen Gesundheitsbehörde Haute Autorité de Santé (HAS).¹ In dem Bericht werden Schlussfolgerungen zusammengetragen, die nationale Gesundheitsbehörden im Rahmen der TUNA-Gesundheitstechnologiebewertung getroffen haben (siehe Tabelle 1).

Jeder dieser Arbeiten ergänzt die im IQWiG-Vorbericht enthaltenen „gescreenten systematischen Übersichten“ (Anlage C). Dabei lassen sich gerade diese Publikationen für eine versorgungspolitische Diskussion nutzen. Es fällt auf, dass die Schlussfolgerungen dieser Technologiebewertungen differenziert sind, ohne einen Nutzen des TUNA-Verfahrens, zum Beispiel unter dem Vorwand mangelnder Evidenz, grundsätzlich in Frage zu stellen. Hier ist das IQWiG gehalten, die Konditionen, unter denen das TUNA-Verfahren in Anwendung gebracht werden kann, sorgfältiger auszuarbeiten.

Es wird daher dringend darum gebeten, bei der Erstellung des IQWiG- Endberichts die Erfahrungen der HAS-Projektgruppe in der Bewertung des TUNA-Verfahrens² einzuholen und damit auch die vom IQWiG angestrebte Annäherung an dieses Institut³ zu konkretisieren.

Da die Ergebnisse der HAS-Arbeitsgruppe in einem nachfolgenden Arbeitsgruppentreffen mit zwölf Medizinexperten (8 Urologen, 2 Allgemeinmediziner, 2 Radiologen)⁴ extern validiert wurden, lässt sich über diese Plattform zusätzliche Expertise einholen.

¹ http://www.has-sante.fr/portail/upload/docs/application/pdf/rapport_tuna-tmtu.pdf
[Link führt zu Bericht in französischer Sprache. Der Bericht ist auch in Englisch erhältlich]

² Dr. Jean-Dominique Doublet, Dr. Denis Jean David, Christine Devaud, Renee Cardoso, Felix Muller, Dr. Sun Hae Lee-Robin

³ <http://www.iqwig.de/nice-has-und-iqwig-wollen-kooperieren.385.html>

⁴ Dr. Cavrois, Lille; Prof. Descotes, Grenoble; Prof. Desgrandchamps, Paris; Dr. Dugit-Gros, Cergy Saint-Christophe; Dr. Fantoni, Lille; Dr. Glemain, Nantes; Dr. Houn, Trets; Dr. Lapray, Lyon; Dr. Le Doze, Biarritz; Prof. Lemaitre, Lille; Prof. Ravery, Paris; Prof Saint, Amiens

Tabelle 1. Schlussfolgerungen der Empfehlungen und Evaluationsberichte über die Effektivität von TUNA

Behörde, Jahr	Schlussfolgerung
MSAC, 2002 ⁵	TUNA appears to have a short-term therapeutic benefit with statistically significant improvements of the subjective symptoms and objective signs. The improvement is initially comparable to that resulting from TURP*, but after 6 to 12 months, the objective measurements such as the peak urine flow rate and symptom scores are significantly better after TURP. The retreatment rate is between 5 and 30% at 2 years.
SBU, 2002 ⁶	TUNA has favourable short-term effects on bladder obstruction. A longer follow-up shows that 20% of patients require additional treatment at 2 years.
ANAES, 2003 ⁷	Radio frequency treatment is an experimental procedure which is being perfected and evaluated.
AUA, 2003 ⁸	TUNA is a therapeutic option. It is an effective treatment which results in a partial improvement of symptoms.
NICE, 2003 ⁹	The available studies showed that TUNA was an effective and safe technique in the short term. The long-term efficacy was not demonstrated. The retreatment rate was between 4% and 30% at two years.
'Clinical Evidence', 2004 ¹⁰	TUNA is less effective than TURP but results in less retrograde ejaculation and bleeding. It may be performed on an outpatient basis. The anaesthetic choices are variable. The long-term effects were not correctly evaluated.
CEDIT, 2004 ¹¹	CEDIT considers that the benefit of TUNA therapy is demonstrated in relation to TURP (less effective but fewer incapacitating adverse events) but not in relation to the medical treatment (notably to combination therapy)
EAU, 2004 ¹²	TUNA is a simple and effective technique and may be performed under local anaesthesia in a large number of cases. It leads to symptom improvement in about 50% to 60% of cases and the peak urine flow rate increases from 50% to 70%. The clinical efficacy was only demonstrated in one sole randomised trial, and the evidence of efficacy in the long term is limited.

*: TURP = transurethral resection of the prostate

⁵ Medical Services Advisory Committee. Transurethral needle ablation (TUNA) for the treatment of benign prostatic hyperplasia. Canberra: MSAC; 2002.

⁶ Swedish Council on Technology Assessment In Health Care. TUNA Transurethral needle ablation for BPH. Stockholm: SBU; 2002.

⁷ Agence nationale d'accréditation et d'évaluation en santé. Prise en charge diagnostique et thérapeutique de l'hypertrophie bénigne de la prostate. Paris: ANAES; 2003.

⁸ American Urological Association, Roehrborn CG, McConnell JD, Barry MJ, Benaim E, Bruskewitz RC, *et al.* Guideline on the management of benign prostatic hyperplasia (BPH). Linthicum (MD): AUA; 2003.

⁹ National Institute for Clinical Excellence. Transurethral radiofrequency needle ablation of the prostate. London: NICE; 2003.

¹⁰ Webber R. Benign prostatic hyperplasia. *Clin Evid* 2004;(12):1231-50.

¹¹ Comité d'évaluation et de diffusion des innovations technologiques, Verdeaux S, Tiah D, Baffert S, Charpentier E, Wizen S, *et al.* Dispositif médical TUNA (Traitement par radiofréquence de l'hypertrophie bénigne de la prostate). Paris: AP-HP; CEDIT; 2004.

¹² European Association of Urology, de la Rosette F, Madersbacher S, Alivizatos G, Rioja Sanz C, Emberton M, *et al.* Guidelines on benign prostatic hyperplasia. Arnhem: EAU; 2004.

Berücksichtigung relevanter Studien

Nach eigenen Angaben hat die Projektgruppe des IQWiG Daten von Nicht-RCTs erst dann extrahiert und berücksichtigt, wenn wenig RCTs im Bereich der jeweiligen Therapie vorlagen. Im Falle des Verfahrens TUNA wurden die identifizierten und eingeschlossenen RCTs aufgrund grober Mängel im Studiendesign verworfen. Es ist daher zu postulieren, dass spätestens jetzt die identifizierten CCT-Studien^{13 14 15} und weitere in den Übersichtsarbeiten der Tabelle 1 enthaltenen Publikationen Eingang in die Bewertung finden müssen. Das vom IQWiG durchgeführte iterative Selektionsverfahren ist ungeeignet, da dieses Vorgehen CCT-Studien ausschließt bevor eine Bewertung der randomisierten Studien erfolgt.

Beurteilung der Studienqualität

Im Falle der TUNA-Therapie ist nicht transparent dargestellt, warum die Studienqualität der vier Studien grobe Mängel aufweist bzw. mit welchen Bewertungsmaßstäben hier Auswahlentscheidungen und Aussagen getroffen wurden. Eine strukturierte, gleichförmige Bewertung der Originalpublikationen findet in vielen Bereichen nicht statt. Es ist beispielsweise wenig zielführend, wenn unter dem Aspekt Concealment bei der Studie von Hindley (2001) aufgeführt wird, dass der Hinweis auf die Verwendung blickdichter Umschläge fehlt (siehe Tabelle 136). Auch der Aspekt der fehlenden Angaben zur Verblindung des nachbehandelnden Personals (welches denn?) berücksichtigt nicht die in der Praxis übliche Berichterstattung. Setzt man derartig strenge Maßstäbe für jede Studie in gleicher Weise als Entscheidungskriterium (mit welchem Einfluss auch immer) an, blieben nahezu alle Publikationen unberücksichtigt.

Ebenfalls nicht nachvollzogen werden kann die kategorische Nichtberücksichtigung der Langzeitergebnisse (> 18 Monate) der Studie von Hill 2004. Eine signifikante Reduktion der IPSS-Werte zu Studienbeginn von 23,9 auf kurzfristig 10,1 (2-4 Monate) und eine Stabilisierung dieses Wertes im Langzeitverlauf auf 11,0 (5-8 Monate) und 11,1 (9-17 Monate) lässt einen stabilen Wert auch für die Zeit nach 18 Monaten vermuten. Eine Ausfallrate von >30 % ist bei einer Einmalbehandlung, eines Durchschnittsalters von 66 Jahren mit assoziierter Mobilität, einer raschen Statusverbesserung nach Baseline, geringen Nebenwirkungen, einer IPSS Recall-Period von 1 Monat sowie einem ungewöhnlich langen FollowUp von fünf Jahren erklärbar. Positiv fällt hier die Eigenberechnung des p-Wertes durch die IQWiG-Projektgruppe aufgrund Unplausibilität des publizierten p-Wertes auf. Hier zeigt sich, dass sorgfältig nachgearbeitet werden kann und so auch LOCF-Berechnungen der Originaldaten Berücksichtigung finden könnten. Indes werden ITT, Fallzahlplanung und DropOut-Kriterien im Bericht nur vage erörtert.

¹³ Arai Y, Aoki Y, Okubo K, et al. Impact of interventional therapy for benign prostatic hyperplasia on quality of life and sexual function: a prospective study. J Urol. 2000 Oct;164(4):1206-11.

¹⁴ Schatzl G, Madersbacher S, Djavan B et al. Two-year results of transurethral resection of the prostate versus four 'less invasive' treatment options. Eur Urol. 2000 Jun;37(6):695-701.

¹⁵ Minardi D, Benedetto Galosi A, Yehia M, Cristalli A, Hanitzsch H, Polito M, et al. Transurethral resection versus minimally invasive treatments of benign prostatic hyperplasia: results of treatments. Our experience. Arch Ital Urol Androl 2004;76(1):1-8.

Datenextraktion und Berichterstattung

Der Bericht besticht durch eine sehr ausführliche Darstellung der Studienergebnisse. Der Darstellung der Diagnostik- und Therapieverfahren der BPS sowie der Methoden zur Erstellung des Berichtes sind zahlreiche Aspekte für eine differenzierte Bewertung jeder einzelnen Technologie zu entnehmen.

So lassen sich besondere Untersuchungsschwerpunkte z.B. in Hinblick auf Prostatagröße, Symptomatik, Eindringtiefe, Interventionshäufigkeit, Höhe der Wärmeapplikation, erforderliche Zusatzbehandlungen (Anästhesie, Blasenkatheter, Beseitigung der Abflusshindernisse), Kombination von Therapieverfahren und Komplikationsprofile, identifizieren.

Eine Vielzahl von Lebensqualitätsmessinstrumenten werden aufgeführt und verständlich erklärt.

Diese dargestellte Vielfalt der zu berücksichtigenden Aspekte offenbart dabei gleichzeitig, dass der Untersuchungskorridor zwar sorgfältig beschrieben wird, dann aber auch praktikabel hätte eingegrenzt werden müssen.

Fazit

Das unter Abschnitt 5.9.3.1.6 „Zusammenfassung TUNA vs. Standardbehandlung“ dargestellte Resümee der IQWiG Arbeitsgruppe ist in dieser Form nicht akzeptabel. Die Arbeitsgruppe des IQWiG sollte sich die Konsequenzen ihrer Schlussfolgerungen deutlich machen: Es wird resümiert: „Lebensqualität der Patienten: Ein Vorteil zugunsten einer der beiden Behandlungen ist durch diese Studien nicht belegt“. Die logische Schlussfolgerung dieser Aussage ist es, am Standardverfahren festzuhalten – unabhängig von patientenrelevanten Erkenntnissen.

Gleichwohl gilt: Für die Therapieform TUNA sind bedeutende Statusverbesserungen nachgewiesen worden. Die Akzeptanz des TUNA-Verfahrens ist durch positive Kostenerstattungsentscheidungen – wenn auch zum Teil unter Einschränkungen - in den USA (FDA), in Frankreich (HAS), in England (NICE) sowie Australien (MSAC) und in Hinblick auf die Kodierung in den USA (CPT), in Frankreich (CCAM) und in Australien (MBS) belegt.

Ein wesentlicher Aspekt bei der Beurteilung der TUNA-Therapie ist die geringe Komplikationsrate. In den im IQWiG-Vorbericht berücksichtigten vier Studien zur TUNA wird in der Tabelle 139 die Anzahl der unerwünschten Ereignisse angegeben (Tabelle 2).

Tabelle 2. TUNA vs. Standardbehandlung – Unerwünschte Ereignisse

Komplikationen	TUNA [N]	Standardtherapie [N]
Ejakulationsdysfunktion	5	76
Erektile Dysfunktion	1	24
Inkontinenz	2	12
Strikatur	1	13
Bluttransfusion/ schwere Blutung	0	22

Neben der Komplikationsrate spielt die Schwere der möglichen Komplikationen eine bedeutende Rolle für die Therapieentscheidung des Patienten. Insbesondere können Komplikationen, wie Inkontinenz, erektile und ejakuläre Dysfunktion sowie schwere Blutungen Patienten davon abhalten, sich einer Standardtherapie zu unterziehen.

Eine Meta-Analyse von Boyle et al.¹⁶ von 14 Studien zur TUNA-Therapie mit mehr als 1200 Patienten zeigt konsistente klinische Ergebnisse, insbesondere die klinisch relevanten Symptom-Verbesserungen (12 IPSS point improvement). Diese wiederum stehen in enger Korrelation zur Verbesserung der Lebensqualität.

Zlotta et al.¹⁷ berichtet in einer Fallstudie mit 188 Patienten über eine Re-Interventionsrate von 16.5% innerhalb von 5 Jahren. Zusammen mit einer zusätzlichen medikamentösen Therapie ergibt sich eine Non-Responderquote von 23.3% innerhalb von 5 Jahren. In der RCT Studie von Hill¹⁸ werden 14% als Non-Responderquote für einen Nachbeobachtungszeitraum von 5 Jahren beschrieben.

In den Leitlinien der Deutschen Urologen¹⁹ wird die Wahrscheinlichkeit für eine Re-Intervention nach einer Standardtherapie (TURP) mit 12-15% innerhalb von 8 Jahren angegeben. Bezieht man Re-Interventionen aufgrund von schweren Blutungen und anderen Komplikationen mit ein, dürfte die Re-Interventionsrate nach TURP in ähnlicher Größenordnung liegen. Insofern sind die in Tabelle 1 erfassten Teilaussagen zu Re-Interventionen zu relativieren (MSAC, NICE).

Minimalinvasive Verfahren wie die TUNA-Therapie sind geeignete, alternative Verfahren. TUNA dient als ergänzende Therapie für

- BPH Patienten, die nicht auf Medikamente hinreichend reagieren
- BPH Patienten, die nicht lebenslang Medikamente einnehmen möchten
- BPH Patienten, bei denen eine Operation kontraindiziert ist
- BPH Patienten, die einen TURP-Eingriff aufgrund der zu erwartenden Komplikationen ablehnen.

Dabei ist das TUNA-Verfahren im Vergleich zur Standardtherapie mit weniger unerwünschten Ereignissen und geringeren Kosten verbunden. Der Aspekt der Kosten-Effektivität, auf den im

¹⁶ Boyle P, Robertson C, Vaughan ED, Fitzpatrick JM. A meta-analysis of trials of transurethral needle ablation for treating symptomatic benign prostatic hyperplasia. *Br J Urol* 2004;94(1):83-8.

¹⁷ Zlotta AR, Giannakopoulos X, Maehlum O, Ostrem T, Schulman CC. Long-term evaluation of transurethral needle ablation of the prostate (TUNA) for treatment of symptomatic benign prostatic hyperplasia: clinical outcome up to five years from three centers. *Eur Urol* 2003;44(1):89-93.

¹⁸ Hill B, Belville W, Bruskewitz R, Issa M, Perez-Marrero R, Roehrborn C, et al. Transurethral needle ablation versus transurethral resection of the prostate for the treatment of symptomatic benign prostatic hyperplasia: 5-year results of a prospective, randomized, multicenter clinical trial. *J Urol* 2004;171(6 Pt 1):2336-40.

¹⁹ Leitlinien der Deutschen Urologen: Therapie des Benigen Prostatasyndroms (BPS).AWMF-Leitlinie Nr. 043/035. Februar 2003

Methodenbericht (Version 2) des IQWiG gegenüber der Vorgängerversion wesentlich stärker Wert gelegt wird, wird durch die oben zitierte Schlussfolgerung der IQWiG-Arbeitsgruppe konterkariert.

Es wird daher dringend an die Wissenschaftler des IQWiG appelliert, die Vorteilhaftigkeit der Anwendung einer TUNA-Technologie differenzierter zu erarbeiten.

Beispielsweise wird ein sexuell aktiver Mann mit symptomatischen BPH mittleren Alters von dem geringeren Risikoprofil der TUNA besonders profitieren können. Dies impliziert soziale Aspekte genauso wie den weiteren Ressourcenverbrauch dieses Patienten.

Seit 1995 sind weltweit mehr als 100.000 Patienten mit TUNA versorgt worden. Allein in 2006 sind es über 22.000. Die nachgewiesenen geringeren Komplikationen sollten schon für sich bestätigen, warum diese Methode bei den behandelnden Urologen eine hohe Akzeptanz hat.

4.3 Stellungnahmen von Einzelpersonen

4.3.1 Prof. Dr. Klaus Höfner

Adresse:

Urologische Klinik, Evangelisches Krankenhaus
Virchowstr. 20
46047 Oberhausen

**Stellungnahme zum Vorbericht
Nichtmedikamentöse lokale Verfahren zur
Behandlung der benignen Prostatahyperplasie**

Auftragsnummer: --- N04-01 ---

zum Verfahren

Transurethrale Nadel-Ablation (TUNA)

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien

siehe Literaturverzeichnis.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien

Fehlerhafte Interpretation der Ergebnisse

Die Wirksamkeit und geringe Morbidität der Transurethralen Nadel-Ablation in der Behandlung von Patienten mit BPS ist in zahlreichen kontrollierten und randomisierten Studien und deren Publikation in peer reviewed Zeitschriften dokumentiert worden und als innovatives Verfahren von internationalen Expertengremien und nationalen und internationalen Leitlinien anerkannt. Die TUNA wurde aufgrund dieser Daten und den Erfahrungen aus breiter klinischer Anwendung im Jahr 1996 von der US-amerikanischen FDA zugelassen. Die TUNA bietet darüber hinaus für bestimmte Patientengruppen eine Überlegenheit hinsichtlich der nach Patientenpräferenzen gewichteten Therapieziele. Darin begründet sich die medizinische Notwendigkeit und Zweckmäßigkeit der Methode. Diese Ergebnisse befinden sich auch im Falle der TUNA im deutlichen Widerspruch zur Gesamtaussage des IQWiG-Vorberichtes.

Die Methoden des IQWiG zur Studienauswahl sind zwar klar definiert, im Detail jedoch wegen der unzulässigen Reduktion auf RCT's sowohl aus wissenschaftlicher Perspektive als auch aus Sicht der klinischen Relevanz deutlich eingeschränkt und verfälschen

deshalb die medizinische Evidenz. Sogar Studien des Evidenzlevels I werden vom Institut als grob mangelhaft eingestuft.

Zur generellen Kritik des IQWiG an der mangelnden Studienqualität auch im Falle der TUNA sei speziell zur Bruskewitz-Studie [2] angemerkt, dass das Design dieser multizentrischen RCT-Studie maßgeblich durch die FDA-Behörde im Rahmen der Zulassung der TUNA in den USA vorgegeben wurde und deren Ergebnisse wesentlich zur Akzeptanz der TUNA durch die FDA beigetragen hat.

Die Aussagen der vom IQWiG berücksichtigten Studien zur TUNA lassen sich wie folgt zusammenfassen:

Einfluss auf die Lebensqualität:

Eine RCT-Studie stellt fest, dass die TUNA und die transurethrale Resektion hinsichtlich der Verbesserung der Lebensqualität bei einem Nachbeobachtungszeitraum von 12 Monaten gleich effektiv sind [2]. Die beobachteten Verbesserungen der Lebensqualität bleiben auch im 5-Jahresverlauf dauerhaft [3]. Eine zweite RCT kann bei einem Nachbeobachtungszeitraum von 18 Monaten keinen statistisch signifikanten Unterschied hinsichtlich der Verbesserung der Lebensqualität zwischen der TURP- und der TUNA-Gruppe ermitteln [4].

Auch der Vorbericht des IQWiG akzeptiert die Aussagen der Studien und bemerkt, dass hinsichtlich der Lebensqualität „*keine statistisch signifikanten Gruppenunterschiede*“ (IQWiG S. 344) festzustellen sind.

Die in diesem Zusammenhang vom IQWiG geäußerte Kritik am Messinstrument der Lebensqualität und der dadurch begründeten Relativierung der Ergebnisse wurde schon in der Stellungnahme des Arbeitskreises BPH zurückgewiesen.

Irritative und obstruktive Symptomatik:

Hinsichtlich des IPSS-Scores, der die irritative und obstruktive Symptomatik erfasst, kann die RCT-Studie von Cimentepe [4] keine statistisch signifikanten Unterschiede zwischen TURP und TUNA-Kohorte beobachten. In dieser Studie wird bestätigt, dass das TUNA-Verfahren hinsichtlich der Verbesserung der Symptomatik der BPH so effektiv ist wie die TURP.

Bei [2] zeigen TUNA und TURP bei einem Nachbeobachtungszeitraum von einem Monat eine ähnliche und

statistisch signifikante Abnahme des Symptomenscores, nach 12 Monaten wurde ein statistisch signifikanter Unterschied zugunsten der TURP beobachtet: der IPSS nahm in der TURP-Gruppe um 15 Punkte ab, während in der TUNA-Gruppe eine Abnahme um 13,6 Punkte zu verzeichnen war. Dieser Unterschied bei [2] mag zwar als statistisch signifikant einzustufen sein, ist aber klinisch wegen einer höheren Schwelle der Wahrnehmung von Unterschieden in der Symptomatik nicht relevant, wie in der Stellungnahme des Arbeitskreises BPH ausgeführt.

Ähnliche klinische Erfahrungen werden auch in einer Meta-Analyse von Boyle et al mit mehr als 1200 Patienten beschrieben: die Verbesserungen des internationalen Symptomenscores (IPSS) durch die TUNA wird als vergleichbar mit der TURP bezeichnet [5].

Zur Aussage des IQWiG: *„In allen 4 Studien zeigte sich, abgesehen von Cimentepe 2003 nach 24 Monaten, zu allen Zeitpunkten eine numerisch geringer ausgeprägte Symptomatik nach der Standardbehandlung im Vergleich zu den Ergebnissen nach TUNA (IQWiG S.342).“* Auch hier wird die fehlende klinische Relevanz des rein biometrisch erfassbaren Unterschiedes der Gruppen durch das IQWiG nicht kommentiert.

Komplikationen

Wesentlich für die Patientenentscheidung zu einer bestimmten Therapie ist die Zielsetzung hinsichtlich der in Betracht zu ziehenden Komplikationsraten. Hier ergibt sich ein klarer Vorteil zugunsten der TUNA im Vergleich zur Standardtherapie. Während eine erektile Dysfunktion für die TUNA nur in einem Fall vom IQWiG auf den Seiten 348-349 beschrieben wird, werden für die TURP 24 Fälle genannt. Ähnliches gilt für die Ejakulationsdysfunktion sowie Strikturraten. Während an gleicher Stelle vom IQWiG für das Standardverfahren 22 Fälle mit schweren Blutungen bzw. Bluttransfusionen beschrieben werden, konnten bei der TUNA keine derartigen Komplikationen beobachtet werden.

Eine analoge Beobachtung findet sich bei der *Inkontinenz*: 2 Fälle bei TUNA, 9 bei der Standardtherapie (IQWiG S. 348-349). In der RCT-Studie von Hill [3] werden der TUNA signifikant geringere unerwünschte Ereignisse (adverse events) zugeschrieben als der TURP. Explizit benannt werden in diesem Zusammenhang die Häufigkeit von erektiler Dysfunktion, Inkontinenz und Striktur. In der Studie von Bruskewitz [2] wird die Inkontinenzrate der TURP mit 3.6%, die der transurethralen Nadelablation mit 0.0% angeben.

Die Feststellung des IQWiG, dass „*sich kaum Unterschiede zwischen den Gruppen (TUNA und Standardtherapie) hinsichtlich der Inkontinenz fanden*“ (IQWiG S. 347), mag allenfalls noch als mathematische Interpretation durchgehen, ist weder aus klinischer Sicht, noch aus Sicht des Patienten, der einem Aufklärungsgespräch unterzogen werden soll, in irgendeiner Art und Weise nachvollziehbar. Neben der ED ist die potenzielle Gefahr der Inkontinenz trotz geringer Wahrscheinlichkeit des Auftretens einer der Hauptgründe für Patienten, eine TURP abzulehnen. Besonders wesentlich ist in diesem Zusammenhang die fehlende Wertung der Schwere von Komplikationen.

Am Beispiel der Inkontinenz zeigt sich besonders, wie irreführend die Berücksichtigung rein biometrischer Daten zur Beurteilung eines interventionellen Verfahrens ist und zwangsläufig außer Zweifel stehende klinische Ergebnisse falsch oder in falschem Zusammenhang präsentiert werden.

Die Kritik an der Darstellung der unerwünschten Ereignisse kann ebenfalls nicht nachvollzogen werden. Es wird kritisiert, dass „*häufig nicht differenziert werden konnte, ob Patienten mehrfach genannt wurden, da die Studien lediglich die Häufigkeit des Auftretens spezieller UEs darstellten (IQWiG S. 392)*.“ Es ist unverständlich, wieso allein dieser angebliche Mangel ausreichen soll, um ein Fehlen der Aussagekraft zur Einstufung der klinischen Wertigkeit der TUNA zu konstatieren. Gleiches gilt übrigens auch für andere hier nicht diskutierte Verfahren. Die Morbidität als wichtiges patientenrelevantes Therapieziel wird auch im Fazit des IQWiG-Vorberichtes deutlich bagatellisiert. Die Feststellung, „*Daten zu unerwünschten Ereignissen konnten nur eingeschränkt interpretiert werden, da deren Berichterstattung generell unbefriedigend war (IQWiG S. 395)*“, ist sachlich falsch und vom IQWiG zu korrigieren.

Die höheren Re-Interventionsraten der TUNA im Vergleich zur Standardtherapie werden im IQWiG Vorbericht dem günstigeren Nebenwirkungsprofil der TUNA kommentarlos gegenübergestellt: „*In der Gesamtschau liegen Hinweise auf eine Häufung von Ejakulations- und erektilen Dysfunktionen, Strikturen und behandlungsbedürftigen Blutungen unter Standard vor. Demgegenüber ergaben sich, bezogen auf Reinterventionen, Hinweise auf einen Gruppenunterschied zugunsten Standard (IQWiG S. 348)*.“ Es ist im Allgemeinen davon auszugehen, dass potentielle Re-Interventionen und in Betracht zu ziehende (schwere) Komplikationen einen individuell sehr unterschiedlichen Einfluss auf die Patientenpräferenz haben werden. Auch hier findet sich erneut ein Beleg für die mangelnde Berücksichtigung der

Patientenpräferenzen sowie der klinischen Belange durch das IQWiG.

Die Kritik an den höheren Re-Interventionsraten der TUNA muss darüber hinaus relativiert werden: Eine vom IQWiG nicht berücksichtigte Studie von Zlotta et al. [6] mit 188 Patienten berichtet über die Notwendigkeit einer zusätzlichen Behandlung in 23.3% der Fälle, davon wurden 6,4% medikamentös behandelt. Die Ausfallrate wird in dieser Studie mit einem Nachbeobachtungszeitraum von 5 Jahren mit 12/188 (6.4%) angegeben.

Es ist eine bekannte Tatsache, dass Re-Interventionsraten zur Beurteilung der TUNA und aller anderen nichtmedikamentöser lokaler Verfahren vor allem im Langzeit-Followup von hoher Bedeutung sind. Allerdings müssen diese Raten der Re-Intervention mit denen des Standardverfahrens TURP verglichen werden.

In den Leitlinien der Deutschen Urologen wird die Wahrscheinlichkeit für eine Re-Operation nach TURP mit 12-15% innerhalb von 8 Jahren angegeben [7], zusätzliche medikamentöse Therapien sind hier nicht eingerechnet. Re-Interventionen wegen starker Blutungen (siehe Abschnitt Komplikationen) infolge TURP müssen zusätzlich berücksichtigt werden, was den Unterschied weiter relativiert.

Dauer von Krankenhausaufenthalten

Hier liegt ebenfalls ein Vorteil zugunsten der TUNA vor [S. 347, IQWiG]. Cimentepe stellt die kürzere Hospitalisierungszeit zusammen mit der schnellen Wiederaufnahme normaler Aktivitäten als weiteren Vorteil der TUNA dar [4].

3. Anmerkungen zur projektspezifischen Methodik unter jeweiliger Angabe wissenschaftlicher Literatur zur Begründung der Anmerkung.

Die Auswahl der Literatur beschränkte sich auf 4 RCT's, denen das IQWiG „grobe Mängel“ bescheinigte. Die auf S. 341 geäußerte Kritik dass bei [1] eine „Beschreibung“ fehlt, „ob die verwendeten Umschläge blickdicht waren“, illustriert an einem einfachen Beispiel die mehr als grenzwertige methodische Spitzfindigkeit des IQWiG.

Fazit:

Der IQWiG-Vorbericht beschreibt die Ziele der Nutzenbewertung wie folgt: *„Aus Sicht der Patienten sollte eine Behandlung vor allem die Beschwerden beim Wasserlassen lindern und die Lebensqualität verbessern, dabei möglichst selten unerwünschte Ereignisse mit sich bringen.“ Für Operationsverfahren sei zudem die Frage relevant, ob ein stationärer Krankenhausaufenthalt nötig ist und wie lange er dauert [IQWiG S. iv].*

Die hier dargelegten Argumente belegen, dass mit der TUNA für ein selektioniertes Patientengut diese patientenrelevanten Therapieziele erreicht werden.

So bewertet eine nicht vom IQWiG berücksichtigte Meta-analyse [5] die TUNA als attraktive Behandlungsoption für folgende drei Patientengruppen:

1. BPH-Patienten, die eine lebenslange medikamentöse Therapie ablehnen
2. BPH-Patienten, die für die Standardtherapie nicht geeignet sind
3. BPH-Patienten, die eine Standardtherapie aufgrund der bekannten Risiken (z.B. Inkontinenz, schwere Blutungen, erektile Dysfunktion) ablehnen

Durch eine Erweiterung des Behandlungsspektrums der BPH kann der medizinische Fortschritt im Sinne der Patienten berücksichtigt werden.

Literatur

- [1] Hindley RG, Mostafid AH, Brierly RD, Harrison NW, Thomas PJ, Fletcher MS. The 2-year symptomatic and urodynamic results of a prospective randomized trial of interstitial radiofrequency therapy vs transurethral resection of the prostate. *BJU Int.* 2001 Aug;88(3):217-20.
- [2] Bruskewitz R, Issa MM, Roehrborn CG, Naslund MJ, Perez-Marrero R, Shumaker BP, et al. A prospective, randomized 1-year clinical trial comparing transurethral needle ablation to transurethral resection of the prostate for the treatment of symptomatic benign prostatic hyperplasia. *J Urol.* 1998 May;159(5):1588-93; discussion 93-4.
- [3] Hill B, Belville W, Bruskewitz R, Issa M, Perez-Marrero R, Roehrborn C, et al. Transurethral needle ablation versus transurethral resection of the prostate for the treatment of symptomatic benign prostatic hyperplasia: 5-year results of a prospective, randomized, multicenter clinical trial. *J Urol.* 2004 Jun;171(6 Pt 1):2336-40.
- [4] Cimentepe E, Unsal A, Saglam R. Randomized clinical trial comparing transurethral needle ablation with transurethral resection of the prostate for the treatment of benign prostatic hyperplasia: results at 18 months. *J Endourol.* 2003 Mar;17(2):103-7.
- [5] Boyle P, Robertson C, Vaughan ED, Fitzpatrick JM. A meta-analysis of trials of transurethral needle ablation for treating symptomatic benign prostatic hyperplasia. *BJU Int.* 2004 Jul;94(1):83-8.
- [6] Zlotta AR, Giannakopoulos X, Maehlum O, Ostrem T, Schulman CC. Long-term evaluation of transurethral needle ablation of the prostate (TUNA) for treatment of symptomatic benign prostatic hyperplasia: clinical outcome up to five years from three centers. *Eur Urol.* 2003 Jul;44(1):89-93.
- [7] Leitlinien der Deutschen Urologen zur Therapie des benignen Prostatasyndroms; *Urologe [A]* (2003):42; 722-738

4.3.2 Prof. Dr. Rolf Muschter

Adresse:

Diakoniekrankenhaus
Else-Averdieck-Str. 17
27356 Rotenburg / W.

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

Dieser Kommentar zum Report „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“ des Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) bezieht sich allein auf die Kapitel 5.2, 5.3, 5.4, 5.5, 5.6 und 5.7 zu den verschiedenen Laserverfahren.

1. Benennung von zusätzlichen, im Vorbericht nicht genannten, relevanten Studien

Dies ist in Anbetracht der Vielzahl der nicht berücksichtigten Studien leider nicht möglich. Bei Stichwortsuche „benign prostatic hyperplasia and laser“ ausschließlich in der Datenbank „Unabridged Medline 1966-2007 (März)“ werden 1937 Treffer genannt [1]. Auch bei weiterer Einschränkung der Suchbegriffe durch „laser coagulation“, „laser vaporization“, „interstitial laser“, „greenlight laser“, oder „holmium laser“, jeweils mit der Verknüpfung „and BPH“ kommt zu jeweils ca. 2000 Treffern. Seitens der Autoren des Vorberichtes des IQWiG sind – auch bei Zugrundelegung einer großen Zahl mit den Suchbegriffen nicht korrekt erfasster Studien – mehrere hundert relevante Veröffentlichungen nicht berücksichtigt worden.

Allein ein Themenheft des Journal of Endourology (17/8, 2003) enthält zu den verschiedenen Laserverfahren der Non-Kontakt- und Kontakt-Koagulation 74 [2], der Non-Kontakt- und Kontakt-Vaporisation 57 [3] und der interstitiellen Koagulation 40 Literaturzitate [4]. Die 2006 erschienene Monographie „Male Lower Urinary Tract Dysfunction“ (Editors: McConnell J et al.) der „6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases“ enthält für das Kapitel „New Minimally Invasive and Surgical Developments in the Management of BPO“ (de la Rosette J et al.) 217 Zitate [5]. Dabei wurden ausschließlich Studien von 1992 bis 2002 mit einer Mindestbeobachtungszeit von 3 Monaten mit mehr als 10 Patienten pro Behandlungsarm berücksichtigt.

Dem Text des Kapitels Laserverfahren der von der AWMF als S2 anerkannten Leitlinie zur Therapie des BPS liegt eine der Literaturrecherche des IQWiG vergleichbare Recherche zugrunde [6].

Mehrere und auch vom IQWiG selbst als relevant eingestufte Studien wurden in der Bewertung durch die Autoren des Vorberichtes nicht berücksichtigt, da sie dem falschen Verfahren (Koagulation statt Vaporisation) zugeordnet wurden (z.B. Ref. 82, 83, 88, 91). Daher muss an dieser Stelle bemerkt werden, dass das gesamte Kapitel über Laseranwendungen aufgrund grober Fehlzuordnungen neu erfasst werden muss. Die Aussagen des IQWiG zu diesem Thema sind als unrichtig einzustufen.

2. Bewertung und Interpretation der im Vorbericht in die Nutzenbewertung eingeschlossenen Studien

Auswahl und Zuordnung der Studien

Abgesehen von den oben genannten primären Ausschlüssen einer erheblichen Zahl relevanter Publikationen durch eine erheblich voreingeschränkte Suchstrategie – die benutzten Stichwörter werden im Vorbericht nicht erwähnt, so dass hier mangelnde Transparenz besteht – ist auch die Literatúrauswahl, welche dem Vorbericht zugrunde liegt, willkürlich und beruht auf der - dem Charakter der evidenzbasierten Medizin widersprechenden - Auffassung der Autoren des Vorberichtes, nur Studien für die Bewertung zuzulassen, die Evidenzlevel I erreichen bzw. Studien niedrigerer Evidenzlevel auszuschließen, da jeweils eine ausreichende Zahl des hohen Evidenzlevels vorlag. Bei der Bewertung wurde die Qualität der eingeschlossenen Studien regelmäßig kritisiert, die mögliche Konsequenz, niedrigere Evidenzlevel zuzulassen, wurde außer Acht gelassen. Die Betrachtung ausschließlich des Evidenzlevels I ist willkürlich gesetzt und bewegt sich völlig außerhalb der wissenschaftlichen Definition evidenzbasierter Medizin. Die Fokussierung auf dieses Kriterium führt zu einer unzulässigen Verzerrung der tatsächlichen Evidenz. David Sackett, Nestor und Experte der evidenzbasierten Medizin (EBM), definiert diese als die „Integration individueller klinischer Expertise mit der bestmöglichen Evidenz aus systematischer Forschung“ [7]. Mit der Beschränkung auf Evidenzlevel I wird im Vorbericht des IQWiG der erste Teil der Aussage von Sackett unterdrückt, die dieser sachgerecht und qualifiziert definiert („Mit individueller klinischer Expertise meinen wir das Können und die Urteilskraft, die die Ärzte durch ihre Erfahrung und klinische Praxis erwerben.“ Sackett DL, 2000) [7]. Die Beschränkung auf Evidenzlevel I entspricht daher nicht der wissenschaftlichen Definition von EBM, da wertvolle Informationen

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

ausgeschlossen werden (Sackett DL 2000) [7]. Die Autoren des Vorberichtes verletzen somit bei der Literatursuche die wissenschaftlichen Grundsätze der Evidenzbasierten Medizin in grober Weise. Durch die Methodik des IQWiG wird das über Jahre gewonnene empirische Wissen ausgeschlossen. Dabei ist keine nachvollziehbare Begründung gegeben worden, warum über 500 uni- und multizentrische Studien der Evidenzlevel II und III, insbesondere bei der Untersuchung von relevanten Teilaspekten (z.B. Wirkung auf die Obstruktion in urodynamischen Untersuchungen, z.B. Veränderung des Prostatavolumens) in der Bewertung keine Berücksichtigung finden.

Mehrere und auch vom IQWiG selbst als relevant eingestufte Studien wurden in der Bewertung deshalb letztlich nicht berücksichtigt, da sie dem falschen Verfahren (Koagulation statt Vaporisation) zugeordnet wurden (z.B. 82, 83, 88, 91).

Die Systematik, die das IQWiG für die Literatursuche benutzt hat ist zwar ausführlich beschrieben, enthält aber weder eine Erklärung, wodurch sich „relevante“ von „nicht relevanten“ Studien unterschieden, noch, wie „Duplikate“ identifiziert und ausgeschlossen wurden und – besonders wesentlich – nach welchen Stichwörtern überhaupt gesucht wurde. Somit ist die Zahl primär ausgeschlossener Studien sehr hoch und übersteigt die Zahl der berücksichtigten Studien bei weitem. Dabei ist mangels Transparenz nicht nachvollziehbar, aufgrund welcher konkreter Gründe diese Auswahl erfolgte (vor dem im Bericht beschriebenen ersten Schritt der Literatursuche). Die weiteren Auswahlverfahren werden im Bericht zwar vage beschrieben, auch, dass die Auswahl von 2 externen Gutachtern durchgeführt wurde, es werden jedoch weder die Personen und die Qualifikation der Gutachter benannt, noch das System, nach dem diese ihre Bewertung vornahmen, noch die Gründe, die zum Ausschluß zahlreicher Studien als „nicht relevant“ führten. Es wird nicht einmal erwähnt, ob die Gutachter unabhängig voneinander bewerteten und was beispielsweise bei fehlendem Konsens geschah (Ein- oder Ausschluß der betreffenden Studie?). Die Subjektivität des Verfahrens wird auch im weiteren Bericht erkennbar, da zum einen doch nicht ausschließlich RCTs, sondern gelegentlich auch CCTs Bewertung fanden, andererseits auch später einzelne Studien ausgeschlossen wurden und nicht zuletzt in der Interpretation der festgestellten Ergebnisse erneute Ausschlüsse zuvor eingeschlossener Studien erfolgten, da deren Qualität angeblich unzureichend sei.

Insgesamt erfolgt somit unzulässigerweise nach bereits nicht transparenter Studiensuche eine nachträgliche Bewertung der Studien mit sekundärem Ausschluß von Studien bzw. einzelner Studienergebnisse während des Bewertungsverfahrens. Dies ist willkürlich und unwissenschaftlich.

Fehlende Einbindung von Patientenvertretern, Fachgesellschaften, Experten und Geräteherstellern

Die Methodik des IQWiG postuliert die Einbindung der Fachgesellschaft bei der Erstellung des Vorberichtes ebenso, wie die Kontaktaufnahme einzelner Autoren und ggf. der Herstellerfirmen. Im Vorbericht sind allerdings weder die Meinungen bekannter Selbsthilfegruppen noch die legitimerter Experten nationaler oder internationaler Fachgremien erkennbar eingeholt oder einbezogen worden. Die im „Projektplan“ genannten „externe Sachverständige..., die an der Erstellung des Berichtsplans, an der Informationsbeschaffung und -bewertung sowie an der Erstellung des Vorberichts beteiligt waren“ sind von den nationalen und/oder europäischen urologischen Fachgesellschaften nicht legitimiert und nach intransparenten und nicht bezeichneten Kriterien des IQWiG ausgewählt worden. Die telefonische Kontaktaufnahme mit Patienten- und Selbsthilfegruppen, den national und international bekannten Experten sowie den Herstellerfirmen medizinischer Lasersysteme ergab, dass hier nicht einzelne Ausschlüsse erfolgt sind, sondern ein systematischer Ausschluss aller Patientengruppen, Experten und Firmen [8].

Nicht vorurteilsfreie Darstellung der Standardtherapie

Erstaunlicherweise wird die transurethrale Resektion der Prostata (TURP) im Vorbericht mit dem unwissenschaftlichen Terminus „Goldstandard“ bezeichnet. Dieser in den Wörterbüchern der deutschen Sprache nicht vorkommende Begriff [9] spricht für die Voreingenommenheit der Autoren des Vorberichtes, die damit offensichtlich versuchen, dem Begriff des „Standards“, der nach den Definitionen der Wörterbücher der deutschen Sprache lediglich den Begriffen „Maßstab, Richtschnur, Norm, Qualitäts-

oder Leistungsniveau“ entspricht [9], und somit als zutreffend verstanden werden könnte, durch die Hinzufügung der Vorsilbe „Gold“ eine höhere Wertigkeit zu geben bzw. eine Steigerung des Begriffes „Standard“ zu implizieren.

Unwissenschaftlich definierter Therapieziele

Die Autoren des Vorberichtes definieren „*patientenrelevante Therapieziele*“, ohne dass deren Evidenz belegt wird. Unzweifelhaft sollte eine Behandlung „*aus Sicht des Patienten ... vor allem die Beschwerden beim Wasserlassen lindern und die Lebensqualität verbessern,*“ aber auch „*dabei möglichst selten unerwünschte Ereignisse mit sich bringen ...*“. Die wesentlichen Behandlungsziele der Vermeidung einer Progression der Erkrankung mit dem Auftreten von Komplikationen sowie die Vermeidung eines operativen Eingriffes werden nicht erwähnt. Auch fehlt jede Evidenz für die These, dass „*für Operationsverfahren ... zu dem die Frage relevant [ist], ob ein stationärer Krankenhausaufenthalt nötig ist, wie lange er dauert und ob und wie lange ein Patient einen Blasenkatheter tragen muß ...*“, auch wenn der klinisch tätige Arzt dieser Betrachtung vorbehaltlos zustimmen dürfte. Die in den Studien hierzu festgestellten Vorteile mehrerer Verfahren werden allerdings unter Hinweis auf die angeblich schlechte Studienqualität konsequent ignoriert.

Unrichtige Beschreibung und Kategorisierung von Verfahren und resultierenden Gewebewirkungen

Eine erhebliche Verletzung wissenschaftlicher Grundsätze besteht in der unrichtigen Beschreibung von Verfahren und Techniken, einer teilweise falschen Zuordnung von Studien in die falsche Kategorie (z.B. wurden mehrere Studien zur Kontaktlaservaporisation nicht an dieser Stelle, sondern gemeinsam mit Verfahren zur Laserkoagulation bewertet, sind also einerseits beim einer unzutreffenden Kategorie in die Bewertung eingeflossen und fehlen andererseits in der Bewertung der richtigen Kategorie). Die bei der Anwendung eines Verfahrens resultierenden Gewebefeffekte werden bei der Kategorisierung der Verfahren völlig außer Acht gelassen, obwohl dies in allen internationalen Konsensuskonferenzen und bei der Erstellung von Therapieempfehlungen bzw. Leitlinien längst Standard ist [5,6,10,11]. Die Autoren des Vorberichtes verharren auf der antiquierten Auffassung, dass nicht der mit einem Verfahren erzielte qualitative und quantitative Gewebefeffekt, sondern die eingesetzte Energieform die klinische Wirkung bestimmt und fasst daher unterschiedliche Verfahren fälschlich in einer Kategorie zusammen. Als Beispiel sei die Definition der Standardtherapie genannt, als gemeinsame Kategorie der Inzision (=keine Gewebeablation), Resektion (=mittlere bis vollständige Adenomentfernung) und Vaporisation (=mittlere bis vollständige Gewebeablation) mit dem gemeinsamen Merkmal der transurethralen Anwendung von Hochfrequenzstrom; mit der weiteren Zuordnung der Enukleation (=vollständige Adenomentfernung) als offenes Operationsverfahren [5,6,10,11].

Leider wird auch der Betrachtung klinisch nicht mehr durchgeführter (z.B. da die Systeme aufgrund der eingestellten Marktteilnahme der Herstellerfirmen nicht mehr verfügbar sind) oder als obsolet bzw. experimentell geltender Methoden breiter Raum gewidmet [5,6,10,11]. Den Autoren des Vorberichtes ist dabei entgangen, dass einige der untersuchten Verfahren bei gutem Grundkonzept aufgrund klinischer Nachteile bzw. inakzeptabler Kosten nur temporär eingesetzt wurden, bevor das Konzept weiterentwickelt wurde. Insbesondere die Lektüre des Kapitels des Vorberichtes zum Hybrid Laser lässt erkennen, dass die Autoren keine Vorstellung von den operativen Techniken und den erzielten Gewebefeffekten haben, indem hier die verschiedensten (von den jeweiligen Studien-Autoren als „Hybrid“ bezeichneten) Verfahren, deren einzige Gemeinsamkeit eine Kombination mehrerer Energiequellen oder Applikationstechniken darstellt, zu einem „Verfahren“ zusammengefasst werden. Dieses unwissenschaftliche Vorgehen setzen die Autoren des Vorberichtes konsequent fort, indem sie physikalisch gleichwertige Therapieverfahren, wie z.B. Laser- und Elektrovaporisation oder transurethrale Resektion und Laserresektion unterschiedlich klassifizieren, hingegen – wie oben bereits beispielhaft erwähnt - die transurethrale Resektion und die offene Adenomenukulation, die ein völlig unterschiedliches Indikationsspektrum besitzen und völlig unterschiedliche Verfahren darstellen, zum „Standard“ zusammenfassen, und sich nicht scheuen, die von ihrer physikalischen, aber erst recht biologischen Wirkung völlig verschiedenen Verfahren der

transurethralen Inzision der Prostata und die transurethrale Elektrovaporisation der Prostata diesem Standard zu untergliedern! In diesen grob falschen Zuordnungen wird die ungenügende urologische Expertise sichtbar. Für die Tatsache, dass bei der Definition der Standardverfahren der operativen Therapie neben der transurethralen Vaporesektion der Prostata auch die transurethrale Elektrovaporisation der Prostata genannt wird, fehlt eine schlüssige Begründung vollständig. Die Gleichstellung der Vaporesektion mit der Resektion im Sinne einer Modifikation mag dabei noch angehen, da das wesentliche Element der Vaporesektion die Resektion darstellt. Völlig anders hingegen ist jedoch die Elektrovaporisation zu bewerten, bei der gerade im Hinblick auf die potentiell geringere Eingriffsmorbidität auf jegliche Resektion verzichtet und das Gewebe stattdessen durch Vaporisationseffekte entfernt wird [12]. In der medizintechnischen Historie der jüngeren Vergangenheit ist die Elektrovaporisation aus dem Konzept der Laservaporisation abgeleitet worden, dabei wurde versucht, dass „minimal-invasive“ gegenüber der TURP potentiell komplikationsärmere Verfahren der Vaporisation technisch einfacher und vor allem erheblich preiswerter im Vergleich zur Laseranwendung zu erreichen.

Bei der Bewertung nicht medikamentöser instrumenteller Behandlungsverfahren, insbesondere der Laserverfahren ist der Unterschied zwischen Vaporisation und Resektion/Enukleation von höchster Wichtigkeit, denn es ergibt sich ein logisches Dilemma. Nach der im Vorbericht geführten Argumentation stellt die Elektrovaporisation eine Modifikation der Elektroresektion dar. Damit stellt sich zwangsläufig die Frage, warum die Laservaporisation keine Modifikation der transurethralen Resektion der Prostata, sondern ein eigenes Verfahren darstellt. Der zu erwartende Gewebeeffekt bei Anwendung der Elektrovaporisation bzw. der Laservaporisation unterscheidet sich allenfalls quantitativ, was bei operateurabhängigen Maßnahmen durch Änderung der Applikationszeit bis zur Erzielung des gewünschten Ergebnisses angepasst wird. Somit muss auch die Laservaporisation eine Modifikation der Elektroresektion sein. Erst recht muss die Holmiumlaserresektion, die lt. Vorbericht „...*Prostatagewebe ähnlich wie bei einer TURP in kleinen Stücken herauszuschneiden...*“ in der Lage ist, eine Modifikation der transurethralen Elektroresektion sein. Auch hier liefern die publizierten experimentellen Daten zu den Schneideeigenschaften des Holmiumlasers und von Hochfrequenzstrom genügend evidenzgesicherte Informationen [3]. Aus welchen Gründen im Vorbericht trotz gleichartiger Gewebeeffekte unterschiedlicher Energieträger (Laser bzw. Hochfrequenzstrom) diese nicht als Modifikation, grundverschiedene Gewebeeffekte (Resektion bzw. Vaporisation) derselben Energie (Hochfrequenzstrom) hingegen als Modifikation eines Verfahrens angesehen werden, bleibt unbelegt und völlig unklar und ist damit subjektiv und willkürlich. Die Beschränkung einer „Modifikation“ auf den Einsatz hochfrequenter elektrischer Energie entbehrt jeder wissenschaftlichen und biophysikalischen Grundlage. Für die trotz erkannter Ähnlichkeiten als eigenständiges Verfahren, nicht als Modifikation betrachteten Holmiumlaserresektion wird willkürlich postuliert, dass die klinische Wirkung der Geweberesektion mittels hochfrequentem Strom sich von der Resektion mit einem Laser unterscheidet. Hierfür gibt es weder praktische, noch theoretische Beweise. Unterschiede können lediglich in der technischen Ausführung der Resektion bzw. ihrer Quantität zu suchen sein und sind somit nicht methodisch bedingt, sondern anwenderabhängig.

Fehlerhafte Interpretation der Ergebnisse der Kapitel 5.2.-5.7. des Vorberichtes zur Bewertung der Laserverfahren

ad Kapitel 5.2. des Vorberichtes: Das Prinzip der Kontakt-Laserablation wird im Vorbericht völlig verfälscht dargestellt. Bei dieser Technik wird tatsächlich keine Lasersonde in die Prostata eingeführt, vielmehr erfolgt die Applikation der Laserstrahlung von der Harnröhre aus mit dem Ziel der Gewebevaporisation [3,6].

Zur Bewertung der Kontaktlaservaporisation, die in der biophysikalischen Wechselwirkung, also der Wirkung auf das Gewebe qualitativ exakt der als Modifikation des Standardverfahrens begriffenen Elektrovaporisation entspricht, wurden lediglich insgesamt 5 RCTs, die in 13 Publikationen beschrieben wurden, in die Nutzenbewertung eingeschlossen. Nach Auffassung des IQWiG hatten alle 5 randomisierten Studien relevante Qualitätsmängel. Somit wurde „von einer quantitativen

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

Zusammenfassung der Ergebnisse im Verlauf wegen der geringen Studienanzahl und der heterogenen Datenlage abgesehen.“ Doch wird gefunden: „Ein Vorteil der CLAP hinsichtlich der Lebensqualität ist nicht belegt.“ ...“Insgesamt ergibt sich, bezogen auf die mittlere Krankenhausverweildauer, ein Vorteil zugunsten CLAP.“ ...„Insgesamt ergibt sich für transurethrale Katheter ein Vorteil zugunsten CLAP, da die mittlere Katheterisierungsdauer um 0,6 bis 1 Tag kürzer war.“ Den Beweis der Überlegenheit von „CLAP“ oder Standard kann das IQWiG nicht konstatieren, jedoch auch nicht die Gleichwertigkeit, da die Studien hierfür nicht konzipiert waren. Mit dieser Aussage werden patientenrelevante Therapieziele, wie ein kürzerer Krankenhausaufenthalt und eine kürzere Katheterverweilzeit, als Vorteil für „CLAP“ ohne konkrete Begründung ausgeschlossen, letztlich mit dem willkürlichen Argument, dass sich dies in der Lebensqualität nicht ausdrücke, da hier kein Vorteil für eines der beiden Verfahren, aber auch keine Gleichwertigkeit abzuleiten sei. Entsprechendes gilt für die Bewertung der Obstruktion. „Bemerkenswert ist, dass sich bei den methodisch etwas besseren Studien ... jeweils kein Effekt zeigte.“ „Ein Vorteil zugunsten einer der Behandlungen kann aus diesen Ergebnissen nicht abgeleitet werden, Gleichwertigkeit jedoch auch nicht...“

Die Willkür der Literaturoauswahl der Autoren des Vorberichtes beeindruckt auch in diesem Kapitel in spezieller Weise, z.B.: „Ming 2002 und Horninger 1997: Diese Studien wurden nicht in die Nutzenbewertung aufgenommen, da eine ausreichende Anzahl RCTs für diesen Vergleich identifiziert werden konnte.“ (!) Zunächst als relevant eingestufte Studien wurden nachträglich willkürlich ausgeschlossen!

ad Kapitel 5.3. und 5.4. des Vorberichtes: Auch die Prinzipien der visuellen Laserablation sowie des „Hybrid“ Lasers werden falsch dargestellt. Bei beiden Techniken wird tatsächlich keine Lasersonde in die Prostata eingeführt, vielmehr erfolgt die Applikation der Laserstrahlung jeweils von der Harnröhre aus, im ersten Fall mit dem Ziel der Erzeugung einer großvolumigen Koagulationsnekrose [2,6], im zweiten Fall wird in der Regel eine koagulierende Verfahren mit einem vaporisierenden Verfahren oder einer Resektion kombiniert.

In den Kapiteln zu VLAP und zum Hybrid-Laser werden Techniken untersucht, die derzeit aus verschiedenen Gründen keine Relevanz mehr besitzen, insbesondere, da sie aus heutiger Sicht lediglich die Vorstufe weiterer Entwicklungen darstellten bzw. zu anderen Verfahren weiterentwickelt wurden. Auffällig ist in diesen Kapiteln jedoch, dass die Zuordnung einzelner Studien eklatante Fehler aufweist. So wurden beispielweise mehrere Studien zur „CLaSP“ (= Contact Laser Prostatectomy“ = Kontaktlaservaporisation“ im Kapitel zur „VLAP“ betrachtet und bewertet, diese Studien fehlen konsequenterweise im richtigen Kapitel (s.o.).

Auch in diesem Kapitel werden publizierte Daten willkürlich als „nicht plausibel“ zurückgewiesen: „...Des Weiteren war der Unterschied der Ausfallraten zwischen den Gruppen VLAP DB und Standard bereits nach 3 Monaten größer als 15 %, so dass die entsprechenden Ergebnisse in dem vorliegenden Bericht nicht dargestellt wurden...“ „...Auf die Darstellung ... wurde im vorliegenden Bericht verzichtet, da ... die SDs zum anderen auffällig niedrig im Vergleich zu den anderen Studien waren...“ Die hohen Ausfallraten der meisten Studien, insbesondere der Langzeituntersuchungen, wird wiederholt kritisiert. Die Autoren des Vorberichtes lassen bei dieser zunächst berechtigt erscheinenden Kritik jedoch außer Acht, dass die BPH bzw. das BPS grundsätzlich in weit überwiegender Zahl Patienten mit hohem Lebensalter betrifft und zudem das Studiendesign für die Untersuchung nichtoperativer alternativer Verfahren häufig speziell auf die Therapie multimorbider Patienten ausgerichtet war. Ausfallraten müssen somit zwangsläufig hoch sein.

Unter dem Oberbegriff „Hybrid-Laser“ werden im Vorbericht heterogene Studien zusammengeführt, deren Gemeinsamkeit darin besteht, dass die Autoren der jeweiligen Studien ein Laserverfahren mit einem zweiten Laserverfahren oder einem anderen Verfahren kombiniert angewendet haben. Die Heterogenität blieb auch den Autoren des Vorberichtes nicht verborgen („...2 Studien (Carter 1999, Tuhkanen 2001) nutzten eine Kombination von KTP und Nd:YAG-Laser mit unterschiedlichen Laserfasern. In der Studie von Abdel-Khalek 2003 wurde ein Nd:YAG-Laser mit einer saphirbesetzten Laserfaser eingesetzt. In der Studie von Carter 1999 erhielten die Patienten der zu vergleichenden Gruppe unterschiedliche Begleitmedikationen, was die Vergleichbarkeit einschränkte...“). Dies hinderte die Autoren jedoch nicht an einer zusammenfassenden Bewertung. Keines der Hybrid-Verfahren hat eine nennenswerte

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

Verbreitung gefunden, so dass auch eine separate Bewertung der verschiedenen Kombinationen sinnlos ist. Keinesfalls macht es aber Sinn, derartige Verfahren zusammenzufassen. Dieses Kapitel ist somit ein sehr gutes Beispiel für die fehlende Sachkenntnis der Autoren des Vorberichtes.

ad Kapitel 5.5. des Vorberichtes: Auch für die interstitielle Laserkoagulation finden die Autoren des Vorberichtes in einzelnen Studien Vorteile: *„...zeigte 3 Monate nach der Behandlung einen numerischen Unterschied zugunsten der ILK, die Werte der beiden Gruppen glichen sich aber im weiteren Zeitverlauf aneinander an... Die Ergebnisse zeigten, dass die Sexualfunktion der Patienten nach einer Standardbehandlung numerisch schlechter war als nach einer ILK.“* Dann allerdings wird *„Von einer quantitativen Zusammenfassung der Ergebnisse ... wegen der geringen Studienzahl und der heterogenen Datenlage abgesehen.“*

Folgende Gruppenunterschiede werden gefunden: *„Die mittlere bzw. mediane Krankenhausverweildauer war in allen Studien nach der ILK kürzer als nach der Standardbehandlung. Aus den Ergebnissen ergibt sich der Beleg eines Vorteils der ILK bezüglich der Krankenhausverweildauer mit einem Unterschied von im Mittel 0,5 bis 5,3 Tagen....Schwerwiegende Komplikationen wie Bluttransfusionen bzw. schwere Blutungen wurden lediglich ... jeweils ausschließlich unter der Standardbehandlung beobachtet...Insgesamt liegen Hinweise auf eine Häufung von Ejakulations- und erektilen Dysfunktionen sowie behandlungsbedürftiger Blutungen unter Standard vor. Demgegenüber ergaben sich, bezogen auf Reinterventionen und Harnwegsinfektionen, Hinweise auf einen Gruppenunterschied zugunsten Standard.“*

Auch bei diesem Verfahren verwundert die Literaturlauswahl. Ohne weitere Begründung nicht erwähnt wurde die RCT von Liedberg F et al (Scand J Urol Nephrol 37, 2003, 494-497), Begründet mit der ausreichenden Anzahl von RCTs wurden alle zitierten CCTs aus der Bewertung ausgeschlossen, weitere CCTs mit zusammen ca. 1000 Patienten wurden nicht erwähnt, u.a. Conn RL et al, Daehlin L und Hedlund H, Knoll T et al, Terada N et al, Muschter R et al, de la Rosette JJMCH et al, Floratos L et al, Greenberger M und Steiner MS, Henkel TO et al sowie Krautschick AW et al. [13,14,15,16,17,18,19,20,21,22,23]. Trotz der von den Autoren des Vorberichtes zunächst als ausreichend befundenen Anzahl von RCTs und den sich daraus ergebenden oben zitierten Gruppenunterschieden wurde abschließend keine eindeutige Aussage zum möglichen Nutzen der ILK getroffen.

ad Kapitel 5.6. des Vorberichtes: In diesem Kapitel begründen die Autoren des Vorberichtes zumindest, warum *„...Keine der für die Untersuchung der PVP identifizierten Studien – alle 3 Studien verglichen die PVP mit der TURP – ...“* in die Nutzenbewertung aufgenommen wurde. Der wesentliche Grund ist, dass die Studien noch nicht vollständig abgeschlossen werden konnten, da es sich bei der PVP um ein sehr junges Verfahren handelt. Hierzu erfolgt kein Kommentar. Auch, warum in diesem Fall keine der vorhandenen CCTs [24,25] für zumindest eine vorläufige Nutzenbewertung herangezogen wurde, bleibt erneut intransparent und unklar.

ad Kapitel 5.7. des Vorberichtes: Auch bei der Bewertung der Holmiumlaserverfahren erfolgt im Bericht eine umfassende Diskussion von obsoleten Verfahren, die als Entwicklungsstufe zur HoLEP nur eine historische Bedeutung besitzen. Im Bericht wird u.a. kritisiert, dass in den Studien der Hersteller des Lasers nicht erwähnt wird. Dies ist zwar aus heutiger Sicht zu beanstanden, damals gab es jedoch lediglich einen Hersteller medizinischer Holmiumlasersysteme.

Bei der im Vorbericht als „Standard“ anerkannten Operationsverfahren fehlt jede Erläuterung der sich zwar überlappenden, jedoch grundsätzlich unterschiedlichen Indikationsbereiche der Verfahren. Stattdessen wird der Eindruck einer Gleichwertigkeit für alle Patienten mit BPH erweckt. Die methodische Beschreibung der TURP (*„... wird eine endoskopische Sonde über die Harnröhre in die Prostata eingestochen ...“*) belegt wieder die völlige Sachkenntnis der Autoren. Ebenfalls nur mit Wohlwollen als korrekt zu bezeichnen ist die Darstellung der technischen Prinzipien der verschiedenen Laserverfahren. Für die Anwendung des Holmium-Lasers im Form der HoLEP wird zumindest teilweise korrekt vermerkt, dass *„...Prostatagewebe ähnlich wie bei einer TURP in kleinen Stücken herauszuschneiden...“* Absicht des Verfahrens ist. Die Autoren des Vorberichtes erwähnen jedoch an keiner Stelle, dass ein möglicher Vorteil der HoLEP darin könnte (und nach Ansicht aller nach EBM bewertenden internationalen Expertengruppen, Leitlinien und Konsensuskonferenzen liegt)

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

[5,6,10,11], dass der Indikationsbereich der HoLEP nicht nur dem der TURP, sondern auch dem der offenen Operation entspricht, und konsequent in einigen RCTs ein transurethrales Verfahren gegen ein offen-chirurgisches getestet wurde.

In dieser Bewertung konstatieren die Autoren des Vorberichtes: *„Für den Vergleich der HoLEP vs. Standardbehandlung gingen 7 RCTs ...in die Nutzenbewertung ein...Alle Studien zeigten relevante Qualitätsmängel, die als „grob“ eingestuft wurden. In 3 Studien wurden als Standardbehandlung die Adenomektomie eingesetzt, in den anderen Studien die TURP und 1-mal zusätzlich die TUVRP.“* ...*„Nur 1 Studie (Kuntz 2004b) zeigte einen statistisch signifikanten Gruppenunterschied nach 6 und 12 Monaten zugunsten der HoLEP. Die anderen Studien wiesen im Zeitverlauf lediglich numerische Unterschiede ohne einheitliche Richtung zugunsten einer der Behandlungen auf...“* Dann wird der Leser mit folgender Feststellung (zur Zusammenfassung der Studienergebnisse der transurethralen Laserenukleation, der transurethralen Elektroresektion und der offenen Adenomektomie) erstaunt: *„...Die quantitative Zusammenfassung mittels Meta-Analysen verdeutlichte das heterogene Bild... Gründe für die große Heterogenität ließen sich nicht erkennen. Lediglich nach 24 Monaten waren die Ergebnisse homogen, es lag zu diesem Zeitpunkt kein statistisch signifikanter Unterschied zwischen den Gruppen vor. Die vorliegenden Ergebnisse lassen keine Aussage bezüglich eines Vorteils zugunsten einer der beiden Behandlungen zu, aber auch nicht bezüglich einer Gleichwertigkeit.“* Dass die Operateursabhängigkeit der Ergebnisse – auch der Standardverfahren – allein Grund genug für die Heterogenität der Studienergebnisse sein könnte, kommt den Autoren des Vorberichtes nicht in den Sinn *„...Gründe für die große Heterogenität ließen sich nicht erkennen...“* Der tatsächliche Grund für heterogene Ergebnisse dürfte jedoch in den völlig unterschiedlichen Patientenkollektiven der einzelnen Studien zu suchen sein, bei denen die HoLEP einerseits mit der TURP, andererseits bei entsprechend anderer Indikation – sehr großvolumigen Prostatae – mit der offenen Enukleation verglichen wird. Natürlich wird dann seitens der Autoren des Vorberichtes trotz des umfangreichen Studienmaterials erneut keine Aussage zum möglichen Nutzen der HoLEP gegeben, obwohl für dieses Verfahren sogar RCTs im Vergleich zu einer offenen Operation vorliegen! Zumindest zur Krankenhausverweildauer (*„...Nach einer HoLEP war die Verweildauer, verglichen mit der TURP, 1 Tag, verglichen mit der Adenomektomie etwa 3 bis 7 Tage kürzer...“*), konnten sich die Autoren zu einer Aussage verleiten lassen, auch ein zweiter Unterschied (*„...Nach einer HoLEP war die Katheterisierungsdauer, verglichen mit der TURP, etwa 1 Tag, verglichen mit der Adenomektomie zwischen 3 und 7 Tagen kürzer...“*) wird wahrgenommen, wengleich nicht als Vorteil der HoLEP bewertet, da dies nicht der Definition der Zielkriterien der Autoren des Vorberichtes entsprach, im übrigen ist die Bewertung unisono. *„...Die Resultate zur Symptomatik und Lebensqualität waren heterogen. Aus dem fehlenden Beleg eines Vorteils bezüglich dieser Kriterien lässt sich keine Gleichwertigkeit der Behandlungen ableiten, da keine der Studien erkennbar als Äquivalenz- oder Nichtunterlegenheitsstudie konzipiert war...“*

Diese Analyse ist nicht nachvollziehbar. Kritisiert wird die Heterogenität der Ergebnisse, für die sich seitens der Autoren des Vorberichtes keine Gründe erkennen lassen. Bereits bei Standard werden so unterschiedliche Verfahren wie eine transurethrale Elektroresektion und die offene-chirurgische Enukleation gemeinsam bewertet, diese erheblichen Differenzen z.B. der Behandlungsmorbidität und der Lebensqualität innerhalb des Standards [5] müssen sich natürlich auch bei der Zusammenfassung der Ergebnisse finden. Allein die Schlussfolgerung der fehlenden Gleichwertigkeit zeigt die Absurdität des allein statistischen und nicht an den Zielvorstellungen der Patienten orientierten Bewertungsverfahrens. Kaum vorstellbar für einen klinisch tätigen Arzt kann die Schlußfolgerung sein, dass trotz fehlender Aussage bezüglich der Gleichwertigkeit zweier Verfahren aus Patientensicht nicht dasjenige allein deshalb überlegen sein muß, weil ein offener Eingriff nicht erforderlich ist, sondern ein transurethraler Eingriff ausreicht; abgesehen von damit verbundenen offensichtlichen Unterschieden wie der Katheterverweilzeit und der Krankenhausverweildauer. Letzteres fanden auch die Autoren des Vorberichtes: *„Bezüglich der Krankenhausverweildauer liegt ein Vorteil für eine kürzere Verweildauer für HoLEP gegenüber den Standardbehandlungen vor.“* Die anschließende Interpretation relativiert diesen Vorteil wieder (s.o.) und misst ihm keine Bedeutung zu.

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

Aus Sicht der Autoren des Vorberichtes durch Daten belegen ließ sich weiterhin: „...*Im Hinblick auf die irritative Miktions symptomatik liegt ein Hinweis auf ein geringeres Auftreten unter HoLEP vor. Darüber hinaus zeigte sich ein Hinweis bezüglich behandlungsbedürftiger Blutungen zugunsten HoLEP in den Studien, die als Standard die Adenomektomie einsetzen...*“ Dennoch konstatiert der Vorbericht: „...*Ein Vorteil zugunsten einer der Behandlungen bezüglich der Lebensqualität sowie der Sexualfunktion ist durch diese Daten nicht belegt...Aus dem fehlenden Beleg eines Vorteils bezüglich dieser Kriterien lässt sich keine Gleichwertigkeit der Behandlungen ableiten...*“

Kommentar zum Vergleich der Verfahren untereinander

Mit akribischem Fleiß und genauer Systematik werden die vergleichenden Studien der Verfahren untereinander beschrieben. Bei der seitens der Autoren des Vorberichtes unisono festgestellten schlechten Datenqualität und der eingeschränkten bzw. unmöglichen Vergleiche der Verfahren gegenüber dem Standard werden Vergleiche der Verfahren untereinander zur reinen Farce und dienen lediglich dem Beleg des Einhaltens einer vorgegebenen Systematik.

Ad Kapitel 6 (Diskussion)

In der Diskussion des Vorberichtes des IQWiG finden sich weitere zahlreiche Widersprüche bzw. Unklarheiten bezüglich der Bewertung der Laserverfahren:

In „*Tabelle 156: Zusammenfassende Darstellung der Ergebnisse der Nutzenbewertung*“ werden folgende Ergebnisse des Vergleiches von 11 nichtmedikamentösen gegenüber Standard geprüft und in die Nutzenbewertung eingegangenen Verfahren präsentiert:

Symptomatik: Für 5 Verfahren wird „*kein Beleg für einen Unterschied*“ gefunden (VLAP DB, HoLAP, HoLRP und TEAP) bzw. „*liegt kein Beleg für einen Gruppenunterschied vor*“ (HoLEP).

Für 6 Verfahren ist ein Vorteil des Standards „*belegt*“ (VLAP, TUMT) oder es „*liegen Hinweise auf einen Vorteil des Standards vor*“ (CLAP, Hybrid, ILK, TUNA).

Lebensqualität: Für 7 Verfahren ist entweder ein Vorteil gegenüber Standard „*belegt*“ (HoLRP), es wird „*kein Beleg für einen Unterschied*“ gefunden (TUNA, TEAP) oder es „*liegt kein Beleg für einen Gruppenunterschied vor*“ (CLAP, Hybrid, ILK, HoLEP). Nur für 2 (VLAP, TUMT) „*liegen Hinweise auf einen Vorteil des Standards vor.*“ 2 Verfahren wurden nicht einbezogen.

Krankenhausverweildauer: Für alle Verfahren ist entweder ein Vorteil gegenüber Standard „*belegt*“ (CLAP, VLAP, ILK, HoLEP), es „*liegen Hinweise auf einen Vorteil gegenüber Standard vor*“ (HoLRP), „*Ein Vorteil ist prozedural bedingt immanent.*“ (TUMT, TUNA) oder es wird „*kein Beleg für einen Unterschied*“ gefunden (VLAP DB, HoLAP, TEAP) bzw. „*liegt kein Beleg für einen Gruppenunterschied vor*“ (Hybrid).

Katheterisierungsdauer: Für 8 Verfahren ist entweder ein Vorteil gegenüber Standard „*belegt*“ (CLAP, HoLEP), es „*liegen Hinweise auf einen Vorteil gegenüber Standard vor*“ (HoLRP), es wird „*kein Beleg für einen Unterschied*“ gefunden (VLAP DB, HoLAP, TEAP) oder es „*liegt kein Beleg für einen Gruppenunterschied vor*“ (VLAP, Hybrid). Nur für 1 Verfahren (TUMT) „*liegen Hinweise auf einen Vorteil des Standards vor.*“ 2 Verfahren wurden nicht einbezogen.

Fasst man die Ergebnisse für die Laserverfahren unter dem Gesichtspunkt der klinischen Relevanz zusammen, so sind Vorteile oder vergleichbare Ergebnisse der beurteilten Methoden gegenüber Standard in der Beeinflussung der Symptomatik in 5/11, der Lebensqualität in 7/9, der Krankenhausverweildauer bei allen Verfahren und bei der Katheterisierungsdauer in 8/9 beurteilten Verfahren im Vorbericht des IQWiG dokumentiert. In Tabelle 158 wird zusätzlich festgestellt, dass sich der Vorbericht des IQWiG außer in einem Fall (Beurteilung VLAP vs. Standard in der Wirkung auf Symptome) in „*Übereinstimmung*“ oder nur „*leichter Diskrepanz*“ mit anderen systematischen Übersichten befindet. In der weiteren Diskussion finden sich jedoch dann nicht weiter begründete Einschränkungen dieser Ergebnisse: „*In den systematischen Übersichtsarbeiten fanden sich häufig Formulierungen, die diskussionswürdig sind. So wurde beispielsweise beschrieben, 2 Verfahren hätten zu „ähnlichen Verbesserungen“ oder „ähnlichen Ergebnissen“ geführt, wenn aus den zugrunde liegenden Studien kein Beleg für einen Vorteil zugunsten einer der Behandlungen abzuleiten war. Diese Formulierung ist deshalb problematisch, da keine der im vorliegenden Bericht eingeschlossenen Studien erkennbar als Äquivalenz- bzw. Nichtunterlegenheitsstudie geplant war und generell das Schließen*

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

von einem fehlenden Beleg für einen Unterschied auf eine Gleichwertigkeit unzulässig ist. Im vorliegenden Bericht wurde von „kein Beleg für einen Vorteil zugunsten einer der beiden Behandlungen“ gesprochen, woraus sich eine Vergleichbarkeit der Behandlungen nicht ableiten lässt. Die folgende Diskussion geht davon aus, dass Formulierungen wie „ähnliche Verbesserungen“ oder „ähnliche Ergebnisse“ in anderen systematischen Übersichtsarbeiten inhaltlich dasselbe Ergebnis beschreiben wie die in diesem Bericht verwendete Formulierung „kein Beleg eines Vorteils zugunsten einer der Behandlungen“ In dieser für einen Kliniker und vor allem für Patienten kaum verständlichen Logik zeigt sich die eklatante Diskrepanz zwischen möglicherweise exakter biometrischer Interpretation und klinischer Relevanz der Ergebnisse. Es könnte noch angehen, wenn die Aussage sich auf die rein biometrische Interpretation reduzieren würde. Im Gegenteil handelt es sich um eine Kernaussage des Berichtes, die wesentliche Begründung der Aussage des Fazits ist, dass „für keines der Verfahren letztendlich ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden.“ kann.

Zusammenfassung

Der IQWiG-Bericht stellt im Fazit fest, dass „für keines der Verfahren ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden kann“.

Dies widerspricht für mehrere Laserverfahren den geltenden Leitlinien nationaler und internationaler Fachgesellschaften, die die geringeren Nebenwirkungen als entscheidenden Vorteil der Verfahren gegenüber der Standardtherapie sehen und die der eigentliche Grund zur Entwicklung und weltweiten wissenschaftlichen Untersuchungen in den letzten 15 Jahren waren [5,6,10,11].

Obwohl das IQWiG selbst durchaus vergleichbare Effekte auf die Verbesserung der Lebensqualität und Vorteile der Methoden in Bezug auf die Krankenhausverweildauer und die Katheterisierungsdauer akzeptiert, werden wissenschaftlich valide essentielle Daten zu unerwünschten Ereignissen in der Beurteilung gänzlich ignoriert. Obwohl das Erreichen patientenrelevanter Behandlungsergebnisse erklärtes Ziel des Berichtes ist, bleiben die Wünsche der Patienten nach geringerer Invasivität und Morbidität als sie mit den Standardverfahren (TURP, TUIP, offenen Operation, jeweils mit verschiedenem Indikationsbereich) potentiell verbunden sind, gänzlich unberücksichtigt, ebenso wie die Möglichkeiten der ambulanten Behandlung mit einigen Therapieverfahren. Dies ist ein grober Widerspruch zu den vom IQWiG selbst publizierten Methoden.

Ich komme als Mitglied des AK BPH der Akademie der Deutschen Urologen, als Mitglied der 4th und 5th International Consultation on Benign Prostatic Hyperplasia, der 6th International Consultation on New Developments in Prostate Cancer and Prostate Diseases und als Board Member der European Society of Uro-Technology, Official Section of the European Society of Urology zu der Einschätzung, dass der Vorbericht des IQWiG durch die unzureichende Berücksichtigung urologischen Sachverstandes, die fehlende Berücksichtigung der Pathophysiologie des BPS, einer grob fehlerhaften Einschränkung patientenrelevanter Therapieziele und der damit im Ergebnis lückenhaften Berücksichtigung und Interpretation essentiell wichtiger wissenschaftlicher Daten schwere Mängel in der Methodik seiner Bewertung aufweist. Der Bericht ist deshalb in seinem Fazit zwangsläufig falsch und grob irreführend.

Dem IQWiG muss empfohlen werden, den Bericht zurückzuziehen und eine Neubewertung vorzunehmen.

Literatur

- [1] Literaturrecherche vom 08.05.2007: Datenbank: Unabridged Medline 1966-2007 (März), Suchprogramm: Knowledge Host Browser Gateway, ARIES KG, Langenfeld; Browser: Microsoft Internet Explorer [2] Muschter R: Free-Beam and Contact Laser Coagulation. J Endourol 17 (2003) 579-585 [3] Tan AHH, Gilling PJ: Free-Beam and Contact Laser Soft-Tissue Ablation in Urology. J Endourol 17 (2003) 587-593 [4] Laguna MP, Alivizatos G, de la Rosette JJMCH: Interstitial Laser Coagulation Treatment of Benign Prostatic Hyperplasia: Is It to Be Recommended? J Endourol 17 (2003) 595-600 [5] de la Rosette JJMHC, Baba S, Badlani G, Elhilali M, Gravas S, Muschter R, Naito S, Netto NR: New Minimally Invasive and Surgical Developments in the Management of BPO. Chapter 6, pp.195-233. In: McConnell J, Abrams P, Denis L, Houry S, Roehrborn C (Editors): Male Lower Urinary Tract Dysfunction, Edition 2006, Health Publications - 2006, Distributor EDITIONS 21,

Stellungnahme zum Vorbericht IQWiG: „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“, Auftragsnummer --- N04-01 ---

Paris, 2006 (ISBN 0-9546956-6-6) [6] Berges R, Dreikorn K, Höfner K, Jonas U, Laval KU, Madersbacher S, Michel MC, Muschter R, Oelke M, Pientka L, Tschuschke C, Tunn U: Leitlinien der Deutschen Urologen zur Therapie des Benigen Prostatasyndroms. 2003 27.09.2005 [cited 04.05.2007]; Available from: <http://www.uni-duesseldorf.de/AWMF/II/043-035.htm> [7] Sackett DL, SS, Richardson WS, Rosenberg W, Haynes RB. Evidence-Based Medicine 2000. Churchill Livingstone. 2000, pp.1-11 [8] pers. Telefonprotokolle, 10.04.2007, 11.04.2007, 12.04.2007 [9] Duden Die deutsche Rechtschreibung Band 1, 22. Auflage, Dudenverlag, Mannheim-Leipzig-Wien-Zürich, 2000 [10] Roehrborn CG, McConnell JD, Barry MJ, Benaim E, Bruskewitz RC, Blute ML, et al: Guideline on the Management of Benign Prostatic Hyperplasia (BPH). 2003 Updated 2006 [cited 04.05.2007]; Available from: <http://www.auanet.org/guidelines/bph.cfm> [11] de la Rosette J, Alivizatos G, Madersbacher S, Rioja Sanz C, Nordling J, Emberton M, et al: Guidelines on Benign Prostatic Hyperplasia. 2006 [cited 04.05.2007]; Available from: <http://www.uroweb.org/fileadmin/userupload/Guidelines/11%20BPH.pdf> [12] Namiki K, Shiozawa H, Arai Y, Miki M: Basic Study on Transurethral Electrosurgical Vaporization of the Prostate. In: Koshihara K, Miki M, Terachi T, Uchida T (Editors): Recent Advances in Endourology 2: Treatment of Benign Prostatic Hyperplasia, Springer, Tokyo-Berlin-Heidelberg-New York, 2000, pp.147-156 [13].Liedberg F, Adell L, Hagberg G, Palmqvist IB. Interstitial laser coagulation versus transurethral prostate resection of the prostate for benign prostatic enlargement--a prospective randomized study. Scand J Urol Nephrol 37: 494-497, 2003 [14] Conn RL, Kursh ED, Williams JC, Steiner M, Issa M: One Year Outcomes of Interstitial Laser Coagulation of the Prostate in 100 Men: A Multicenter Experience. J Urol 161, Suppl: 390, 1999 [15] Daehlin L, Hedlund H: Interstitial Laser Coagulation in Patients with Lower Urinary Tract Symptoms from Benign Prostatic Obstruction: Treatment under Sedoanalgesia with Pressure Flow Evaluation. BJU Int 84: 628-636, 1999 [16] Knoll T, Michel MS, Trojan L, Bross S, Alken P, Kohrmann KU. Langzeitergebnisse der interstitiellen Laserkoagulation zur Behandlung der benignen Prostatahyperplasie. Akt Urol 34: 48-51, 2003 [17] Terada N, Arai Y, Okubo K, Ichioka K, Matsui Y, Yoshimura K, Terai A. Interstitial laser coagulation for management of benign prostatic hyperplasia: long-term follow-up. Int J Urol 11: 978-982, 2004 [18] Muschter R, Hofstetter A. Die interstitielle Laserkoagulation mit dem Diodenlaser der Wellenlänge 830 nm zur Therapie der benignen Prostatahyperplasie – High- versus Low-Volume-Koagulation. Lasermedizin 15: 72-80, 2000 [19] de la Rosette JJMCH, Muschter R, Lopez MA, Gillatt D: Interstitial Laser Coagulation in the Treatment of Benign Prostatic Hyperplasia Using a Diode-Laser System with Temperature Feedback. Br J Urol 80: 433-438, 1997 [20] Floratos DL, Sonke GS, Francisca EA, Kiemeny LA, Debryne FM, de la Rosette JJ: Long-term follow-up of laser treatment for lower urinary tract symptoms suggestive of bladder outlet obstruction. Urology 56: 604-609, 2000 [21] Greenberger M, Steiner MS: The University of Tennessee Experience with the Indigo 830e Laser Device for the Minimally Invasive Treatment of Benign Prostatic Hyperplasia: Interim Analysis. World J Urol 16: 386-391, 1998 [22] Henkel TO, Niedergethmann M, Alken P: Laserinduced Interstitial Thermotherapy (LITT): In vitro and in vivo Studies. J Endourol 9, Suppl.1: S56, 1995 [23] Krautschick AW, Koehrmann KU, Henkel TO, Michel MS, Alken P: Interstitial Laser Coagulation in Benign Prostatic Hyperplasia: A Critical Evaluation after 2 Years of Follow-Up. Urol Int 62: 76-80, 1999 [24] Fu WJ; Hong BF; Wang XX; Yang Y; Cai W; Gao JP; Chen YF; Zhang CE: Evaluation of greenlight photoselective vaporization of the prostate for the treatment of high-risk patients with benign prostatic hyperplasia. Asian J Androl 2006 May;8(3): 367-371 [25] Sulser T; Reich O; Wyler S; Ruszat R; Casella R; Hofstetter A; Bachmann A: Photoselective KTP laser vaporization of the prostate: first experiences with 65 procedures. J Endourol 2004 Dec; 18(10):976-81

4.3.3 Prof. Dr. J. J. M. C. H. de la Rosette

Adresse:

Academisch Medisch Centrum, Universiteit van Amsterdam
Meibergdreef 9
1105 A7 Amsterdam
Niederlande



Betr.: Stellungnahme zum Vorbericht über „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“ des Institutes für Qualität und Wirtschaftlichkeit im Gesundheitswesen vom 26-03-2007

Als Vertreter Kommissionen für das Krankheitsbild der benignen Prostatahyperplasie (BPH) der Europäischen Urologenorganisation (European Association of Urology, EAU) und Weltgesundheitsorganisation (WHO) habe ich den Vorbericht des deutschen Institutes für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) zur Kenntnis genommen. Der mehr als 500seitige Vorbericht versucht, die nichtmedikamentösen lokalen Verfahren zur Therapie der BPH hinsichtlich Qualität und Wirtschaftlichkeit gegenüber den etablierten Behandlungsverfahren (TURP, TUIP, Adenomenukleation, etc.) zu bewerten.

Die Schlußfolgerungen des Institutes hinsichtlich dieser neuen Therapieverfahren stehen im Gegensatz zu international anerkannten Beurteilungen [1-5] und auch im Gegensatz zu den Empfehlungen der deutschen Urologen [6]. Zum einen werden umfangreiche Analysen von Therapieverfahren gemacht, die zwar in der Vergangenheit durchgeführt, aber heute wieder vollständig verlassen wurden (z.B. VLAP, HIFU, WIT). Zum anderen werden etablierte minimal-invasive Verfahren zur Behandlung der BPH (z.B. Hochenergie TUMT, HoLEP, TUNA) als qualitativ minderwertig eingestuft, da der Behandlungserfolg, gemessen am symptomatischen Ergebnis, geringer ist als die Standardverfahren.

Zur Bewertung wurden größtenteils Studien verwendet, die randomisiert Standardverfahren gegen neue lokalen Verfahren vergleichen und somit der Evidenzstufe I angehören (definiert nach dem Bewertungssystem des *US Department of Health and Human Services*). Bei diesen Untersuchungen ist zu berücksichtigen, daß die Standardoperationen zur Grundausbildung jedes Urologen gehören, jedoch Operateure mit den neuen lokalen Therapieverfahren noch keine ausreichenden Erfahrungen haben. Diese Unerfahrenheit und Lernkurve der neuen Operationstechniken sind zwar schwer meßbar, dürften aber zu den „schlechteren“ Ergebnissen der neuen lokalen Therapieverfahren beitragen. Andererseits können Zentren mit langjähriger Erfahrung mit einer neuen lokalen Behandlungstechnik (z.B. Holiumlaser-Enukleation der Prostata) hervorragende, den Standardverfahren oft überlegende Ergebnisse präsentieren [7, 8]. Diese jedem operativ tätigen Urologen bekannte Tatsache ist im Bericht jedoch nicht eingegangen.

Der Bericht des IQWiG berücksichtigt nur ungenügend die Wünsche und die Lebensqualität der Patienten. Viele der neuen lokalen Therapieverfahren haben den Vorteil, daß sie die

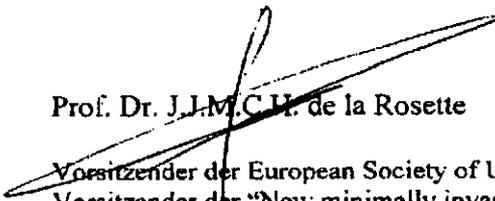
Symptomatik bei geringer Morbidität reduzieren. Der berechtigte Wunsch der Patienten hinsichtlich geringen Nebenwirkungen bei der Behandlung, kurzer Krankenhausverweildauer und Erhalt der sexuellen Funktion wird somit außer Acht gelassen. Auch wird bei den Schlußfolgerungen nur ungenügend berücksichtigt, daß manche Männer aufgrund Komorbiditäten nicht für die Standardtherapien geeignet sind. Würden die Empfehlungen des IQWiG in Deutschland umgesetzt werden, könnten manche Männer nur noch durch dauerhafte oder intermittierende Katheterisierung der Blase behandelt werden. Diese Maßnahme dürfte jedoch nicht im Interesse des IQWiG und der Lebensqualität der Patienten sein.

Es ist davon auszugehen, daß betroffene Männer die Entscheidung gegen die neuen lokalen Therapieverfahren nicht akzeptieren werden. Somit würden diese Männer Rat und Hilfe im benachbarten Ausland suchen, wo die lokalen Therapieverfahren zur Behandlung der BPH akzeptiert sind und auch durchgeführt werden. Die Reise ins Ausland dürfte nur für Betroffene eine Option sein, die wohlhabend sind und somit selber die Therapie bezahlen könnten. Die Entscheidung gegen die lokalen Therapieverfahren würde eine 2-Klassenmedizin schaffen, die auch nicht im Sinne des IQWiG und des deutschen Gesundheitssystems sein dürfte.

Im Bericht des IQWiG ist die Wirtschaftlichkeit der neuen lokalen Therapieverfahren nur unzureichend beschrieben worden. Geht man davon aus, daß die neuen lokalen Therapieverfahren ambulant oder kurzstationär und mit lokaler Anästhesie durchführbar sind, nicht aber einen teuren stationären Aufenthalt und eine Narkose wie die Standardverfahren zur Folge haben, ist der Kostenaufwand der neuen lokalen Therapien trotz höherer Re-Interventionsrate geringer. Nicht berücksichtigt sind außerdem die Folgekosten, die aufgrund der Komplikationen der Standardverfahren entstehen können (z.B. Bluttransfusionen, Behandlung transfusionbedingter Infektionskrankheiten, Behandlung einer erektilen Dysfunktion oder Harninkontinenz). Diese Faktoren wurden jedoch von den US amerikanischen Kollegen mitberücksichtigt, weshalb dort für die ambulante oder kurzstationäre Durchführung der neuen lokalen Therapieverfahren gegenüber den Standardoperationen finanziell überproportional hoch honoriert wird (TUMT).

Der Bericht des IQWiG ist in vielen Teilbereichen und in der Gesamtaussage nicht nachvollziehbar. Eine Überarbeitung des Berichtes mit fachkompetenter urologischer Unterstützung und mit Vertretern von Selbsthilfegruppen wird seitens der BPH-Gruppe der EAU empfohlen.

Mit freundlichem Gruß,



Prof. Dr. J.J.M. C.H. de la Rosette

Vorsitzender der European Society of Urotechnology (ESUT)
Vorsitzender der "New minimally invasive and surgical developments of BPH" der WHO
Ehemaliger Vorsitzender der BPH Leitlinien der European Association of Urology (EAU)
Direktor der Urologischen Klinik, Academisch Medisch Centrum der Universität Amsterdam

Referenzen

1. de la Rosette J, Alivizatos G, Madersbacher S, et al. (2001) EAU guidelines on benign prostatic hyperplasia (BPH). *Eur Urol* 40: 256-63
2. Madersbacher S, Alivizatos G, Nordling J, et al. (2004) EAU 2004 guidelines on assessment, therapy and follow-up of men with lower urinary tract symptoms suggestive of benign prostatic obstruction (BPH guidelines). *Eur Urol* 46: 547-545
3. AUA practice guidelines committee (2003) AUA guideline on management of benign prostatic hyperplasia (2003). Chapter 1: Diagnosis and treatment recommendations. *J Urol* 170: 530-547
4. de la Rosette J, Baba S, Badlani G, et al. (2006) New minimally invasive and surgical developments in the management of BPO. In: McConnell J, Abrams P, Denis L, Khoury S, Roehrborn C (Hrsg.) Male lower urinary tract dysfunction. 6th International Consultation on new developments in prostate cancer and prostate diseases. Editions 21, Paris, ISBN 0-9546956-6-6
5. Nickel JC, Herschorn S, Corcos J, et al. (2005) Canadian guidelines for the management of benign prostatic hyperplasia. *Can J Urol* 12: 2677-2683
6. Berges R, Dreikom K, Höfner K, et al. (2003) Guidelines of German urologists on therapy of benign prostatic syndrome. *Urologe A* 42: 722-738
7. Kuntz RM, Ahyai S, Lehrich K, Fayad A (2004) Transurethral holmium laser enucleation of the prostate versus transurethral electrocautery resection of the prostate: a randomized prospective trial in 200 patients. *J Urol* 172: 1012-1016
8. Kuntz RM, Lehrich K, Ahyai S (2004) Transurethral holmium laser enucleation of the prostate compared with transvesical open prostatectomy: 18-month follow-up of a randomized trial. *J Endourol* 18: 189-191

4.3.4 Dr. Christoph von Zastrow

Adresse:

Hildesheimer Str. 202
30519 Hannover

Von: Christoph.von-Zastrow@mdkn.de
Gesendet: Freitag, 20. April 2007 11:48
An: info@iqwig.de
Betreff: Vorbericht Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie

Sehr geehrte Damen und Herren,

in den letzten Jahren habe ich wiederholt Stellungnahmen zu Methoden abgegeben, die auch in diesem Bericht beurteilt werden, unter anderem im Auftrag der bundesweiten Expertengruppe zur Methoden- und Produktbewertung (der SEG 7) und des MDS.

Aufgrund der dabei gewonnenen Erfahrungen ergeben sich zu der Darstellung folgende Anregungen:

- Bei den operativen Standardverfahren TUR-P und Adenomektomie wird das entfernte Gewebe histopathologisch aufgearbeitet. In bis zu 20% der Fälle wird ein inzidentes Prostatakarzinom als Zufallsbefund diagnostiziert (1), dies kann dann zu weiteren Therapieoptionen, abhängig von den Umständen des Einzelfalls wie der Lebenserwartung, führen. Bei den meisten bewerteten Verfahren werden diese Tumoren nicht erkannt. Da das Prostatakarzinom meist von den Randbereichen ausgeht, in denen die Hitze der Laserverfahren deutlich geringer ist, wachsen diese Tumore weiter und werden dann erst in einem späteren und eventuell fortgeschrittenen Stadium entdeckt. Es ist deshalb wichtig, das bei Studien angegeben wird, in welcher Anzahl in dem Behandlungsarm mit konventioneller Operation ein inzidentes Karzinom entdeckt wurde. Es ist dann davon auszugehen, dass in der anderen Gruppe ohne postoperative histopathologische Untersuchung eine ähnliche Zahl nicht erkannt wurde. Da ca. 25% der Prostatakarzinome ein PSA unter dem sogenannten Grenzwert von 4ng/ml haben und durch eine Stanzbiopsie ca 70% der Karzinome erkannt werden, löst die Empfehlung, vor Verfahren ohne Histologie durch eine Stanzbiopsie ein Karzinom auszuschließen, nicht das Problem.
- Die verschiedenen Verfahren unterscheiden sich auch in der Rate an Folgeoperationen in den Jahren nach dem Ersteingriff. Ob nach einer Operation in den kommenden drei Jahren 5% oder 50% erneut operiert werden müssen, ist unter dem Gesichtspunkt der Wirtschaftlichkeit ein wichtiger Teilaspekt, Stichwort "Therapiekaskade". Teilweise sind Angabe hierzu in der Literatur nicht verwertbar, wenn eine Prozentzahl genannt wird ohne Angabe eines Zeitraumes. So sind aus der Begutachtung Verläufe bekannt, in denen erst eine HE-TUMT durchgeführt wird und innerhalb eines Jahres eine TUR-Prostata folgt.
- Bei den Verfahren ohne Gewebeabtrag ergeben sich besondere Probleme bei der Qualitätssicherung. So war es bei der BQS für die Qualitätssicherung der TUR-P ein Kriterium, welche Gewebemenge entfernt wurde. Einrichtungen, bei denen grundsätzlich immer nur wenige Gramm Prostatagewebe entfernt wurden, standen im Verdacht, die Leistung unvollständig erbracht zu haben. Gibt es keine postoperative Histologie, dann ist nicht beurteilbar, welche Menge des Organs zum Beispiel mittels PVP abgetragen wurde. Durch die Entfernung von minimalen Gewebemengen kann man die Zahl der Operationen pro Tag steigern und es treten kaum Nebenwirkungen auf, man muß die Maßnahme nur öfter wiederholen.
- Viele Autoren verwenden die Harnstrahlmessung vor und nach einer Operation als Parameter zur Erfolgsmessung. Nun ist das Ergebnis einer solchen Messung übungsabhängig, vergleichbar der Lungenfunktionsmessung, bei der der Patient mehrfach in den Apparat atmet, bevor das Ergebnis verwertbar ist. Bei einer Harnstrahlmessung kommt es zu gleichbleibenden Ergebnissen erst nach der 7. Messung (2). Wenn man also bei einem Patienten einmal vor der Operation und anschließend alle 3 Monat erneut bei Kontrollen mißt, ist diese "Besserung" auch ein Übungseffekt und nur teilweise Folge der Operation.
- Der sogenannte Symptomscore IPSS ist nur eingeschränkt verwertbar. Beispiel: es handelt sich überwiegend um Patienten im höheren Lebensalter, das Symptom "Nykturie" kann in diesem Alter unterschiedlichste Ursachen haben. Neben kardialen Ursachen ist zu berücksichtigen, dass die zirkadiane Schwankung der Urinausscheidung, aufgrund derer ein Mensch nachts deutlich weniger Urin pro Stunde ausscheidet als tags, auch bei Herzgesunden im Alter verschwindet (3). Aus der

Pflegebegutachtung ist darüber hinaus bekannt, dass man bei der Frage nach den nächtlichen Toilettengängen von Patienten und und deren Partner- /innen unterschiedliche Angabe bekommt, viele erinnern sich nicht an diese nächtlichen Ereignisse. Der Symptomscore ist auch bei Frauen anwendbar, z. B. bei Interstitieller Cystitis. Patienten mit Überlaufblase können sehr geringe Punktwert aufweisen. Der Score gibt den Leidensdruck eines Patienten wieder ohne Angabe ob dieser urologisch bedingt ist. Die Operationsindikation ist nicht vom Punktwert ableitbar. Gefragt wird im IPSS auch nach dem "Restharngefühl". Restharn ist mittels Ultraschall meßbar. Was besagt bei dem IPSS ein Restharngefühl bei einem Patienten, bei dem Restharn definitiv ausgeschlossen wurde? Der Symptomscore liefert Zahlen und der Glaube ist weit verbreitet, überall wo Zahlen stehen handelt es sich um Wissenschaft. Erforderlich ist eine kritische Bewertung der Meßinstrumente.

Für Rückfragen stehe ich gern zur Verfügung.

Mit freundlichen Grüßen

Dr. C. v. Zastrow
Urologe
Allgemeinarzt
Sozialmedizin
Qualitätsmanagement

Literatur:

1. Jocham, D., Miller, K.: Praxis der Urologie; Thieme Verlag, Stuttgart New York 2003, Band II, Seite 184
2. Jepsen, J. V. et al.: Variability in urinary flow rate and prostate volume. An investigation using the placebo arm of a drug trial, The Journal of Urology, Vol 160,1699.-1694, November 1998.
3. Blanker, M., et,al.: Normal values and determinants of circadian urine production in older man: a population based study; The Journal of Urology, Vol 168, 1453-1457, October 2002.

MDKN - Geschäftsbereich Consulting
Christoph von Zastrow
E-Mail: Christoph.von-Zastrow@mdkn.de
Tel.: +49 (0)511 8785 2213
Web: <http://www.mdkn.de>

5 Dokumentation der wissenschaftlichen Erörterung – Teilnehmerliste, Tagesordnung und Protokoll

Teilnehmerliste

wissenschaftliche Erörterung zum Berichtsplan N04-01 :

„Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

**am 03.07.2007 im IQWiG
11:00–15:45 Uhr**

Moderation: Dr. Thomas Kaiser

Teilnehmerliste:

Name	Organisationen, Unternehmen
Dr. Richard Berges	Arbeitskreis BPH der Akademie der Deutschen Urologen
Katharina Biester	IQWiG
Barbara Buchberger	Universität Duisburg-Essen
Rebecca Deppisch	Universität Duisburg-Essen
Prof. Klaus Höfner	Arbeitskreis BPH der Akademie der Deutschen Urologen
Dr. Thomas Kaiser	IQWiG
PD Dr. Michael Kulig	IQWiG
PD Dr. Stefan Lange	IQWiG
Prof. Rolf Muschter	Arbeitskreis BPH der Akademie der Deutschen Urologen
Dr. Roman Schiffner	Kassenärztliche Bundesvereinigung

Name	Organisationen, Unternehmen
Dr. Petra Schnell-Inderst	Universität Duisburg-Essen
Dr. Kai Schorn	Deutsche Krankenhausgesellschaft
Dr. Guido Skipka	IQWiG
Thilo Rörtgen	Sitzungsdokumentarischer Dienst, Landtag NRW
Dr. Volker Rohde	Bad Schwartau (HELIOS Klinik)
Christian Weinrank	Medtronic GmbH
Dr. Christoph von Zastrow	Hannover
Stefan Zenk	Medtronic GmbH

Tagesordnung

wissenschaftliche Erörterung zum Berichtsplan N04-01 :

**„Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostata-
hyperplasie“**

**am 03.07.2007 im IQWiG
11:00–15:45 Uhr**

Moderation: Dr. Thomas Kaiser

Begrüßung

Einleitung

- TOP 1** Vorstellung des Auftrags und der Vorgehensweise des IQWiG
- TOP 2** Zuordnungen von Studien zu Therapietechnologien
- CLasP-Studie (CLAP / VLAP)
 - Hybrid-Verfahren
 - Hochenergie (HE)-TUMT und Niedrigenergie (NE)-TUMT
- TOP 3** Modifikationen der TURP
- Was wird als Modifikation der TURP betrachtet?
- TOP 4** Studiendesign / Studienauswahl
- Einbezug von nicht randomisierten, kontrollierten Studien in die Nutzenbewertung
- TOP 5** Gleichwertigkeit / irrelevante Unterschiede / klinische Relevanz
- Definition irrelevanter Gruppenunterschiede
- TOP 6** Verschiedenes

Protokoll

wissenschaftliche Erörterung zum Berichtsplan N04-01 :

„Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“

am 03.07.2007 im IQWiG

11:00–15:45 Uhr

Moderator Dr. Thomas Kaiser: Meine sehr verehrten Damen und Herren! Willkommen im Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen zur Erörterung des Berichtsplans des Auftrags N04-01 „Nichtmedikamentöse lokale Verfahren zur Behandlung der benignen Prostatahyperplasie“. Zu dem Punkt „Erörterung des Berichtsplans“ werde ich gleich noch etwas sagen.

Zuvor ein paar Bemerkungen zu formalen Aspekten: Die Erörterung wird aufgezeichnet. Von der Erörterung wird ein Wortprotokoll erstellt. Die Zustimmung Ihrerseits zur Aufzeichnung und Veröffentlichung eines solchen Wortprotokolls ist Voraussetzung für die Teilnahme. Unabhängig davon, dass Sie dem schon zugestimmt haben, möchte ich jetzt noch einmal kurz nachfragen, ob das soweit von Ihrer Seite in Ordnung ist. Sollte das nicht der Fall sein, können Sie leider an der Erörterung nicht teilnehmen. - Ich sehe keinen Widerspruch, sodass wir mit der Erörterung beginnen können.

Dass ein Wortprotokoll erstellt wird, bedeutet auch, dass wir eine gewisse Rededisziplin brauchen. Je weniger Sie durcheinander reden, umso deutlicher und klarer kann hinterher das Wortprotokoll erstellt werden. Ich bitte Sie also darum, dann zu sprechen, wenn Sie aufgerufen werden. Außerdem bitte ich Sie darum, vor Ihrem jeweiligen Redebeitrag Ihren Namen zu nennen, damit später eindeutig zu identifizieren ist, wer was gesagt hat, und nicht Redebeiträge anderen Personen zugeordnet werden.

Des Weiteren möchte ich bemerken, dass wir für ca. 13 Uhr eine Pause eingeplant haben, wofür wir einen Imbiss vorbereitet haben. Möglicherweise ist dann die Sitzung schon beendet. Unabhängig davon, ob die Sitzung unterbrochen oder beendet ist, können Sie am Imbiss teilnehmen.

Sie haben die Tagesordnung vor sich liegen. Ich habe bereits gesagt, dass ich zu dem Punkt „Erörterung des Berichtsplans“ noch etwas sagen möchte. Wir haben auf den Formblättern, die Sie unterschrieben haben, und dem Leitfaden, den wir zu dem Stellungnahmeverfahren veröffentlicht haben und der für Sie zur Information zum Stellungnahmeverfahren diente, erwähnt, dass es sich um ein kombiniertes Stellungnahmeverfahren zum Berichtsplan und Vorbericht handelt, so wie der Berichtsplan, also die Methodik im Vorbericht abgebildet ist -

der Vorbericht enthält ja nicht nur die Ergebnisse, sondern auch die Methodik -, also entsprechend dem vorab formulierten Berichtsplan plus gegebenenfalls Änderungen, die gesondert gekennzeichnet sind.

Wie Sie sicherlich wissen, gab es in den letzten Monaten aufgrund der gesetzlichen Änderung durch das GKV-WSG eine Diskussion. In diesen gesetzlichen Änderungen ist zunächst nur für Arzneimittelbewertungen festgeschrieben worden, dass eine Beteiligung zu allen relevanten Verfahrensschritten zu erfolgen hat, was so vorher gesetzlich nicht festgeschrieben worden war. Das bedeutet, dass es zukünftig eine gesonderte Stellungnahmemöglichkeit zu Berichtsplänen und zu Vorberichten geben soll. In unseren Methoden, Version 2.0 vom 19. Dezember letzten Jahres haben wir dem schon vorgegriffen und hatten schon damals ein getrenntes Stellungnahmeverfahren beschrieben, allerdings nur für diejenigen Projekte, die nach dem 19. Dezember begonnen wurden, was ja bei diesem Projekt nicht der Fall war. Unabhängig davon haben wir uns dazu entschieden, dass wir auch für die bereits laufenden Projekte, auch zu Nichtarzneimittelbewertungen – dabei handelt es sich ja hier –, ein gesondertes Stellungnahmeverfahren vornehmen. Dieser Entschluss ist erst vor wenigen Tagen gefallen, so dass das in den damaligen Unterlagen so noch nicht vermerkt war, sondern damals noch als kombiniertes Stellungnahmeverfahren. Also auch für dieses Projekt machen wir eine gesonderte Stellungnahme zu Berichtsplan und Vorbericht. Sie haben ja umfangreich zur Methodik, wie im Berichtsplan formuliert, Stellung genommen, sodass wir die Stellungnahmen, die jetzt eingegangen sind, als Stellungnahmen zur Methodik, zum Berichtsplan betrachten, entsprechend das auch heute erörtern. Es wird sich daran eine Veröffentlichung der Stellungnahmen, die Sie abgegeben haben, eine Würdigung von unserer Seite, eine Veröffentlichung des Protokolls dieser Erörterung und ein überarbeiteter Vorbericht anschließen. Inwieweit dieser Vorbericht überarbeitungsbedürftig ist, hängt natürlich davon ab, was in den Stellungnahmen an berechtigten Punkten zur Methodik gekommen ist und was möglicherweise heute noch besprochen wird. Das zum groben Ablauf, wie es mit diesem Projekt weitergeht. Haben Sie dazu Fragen?

Dr. Kai Schorn: Habe ich das richtig verstanden? Wir haben eine Stellungnahme zum Vorbericht abgegeben. Sie werden jetzt aber gewertet als Stellungnahme zum Berichtsplan. Ist das richtig?

Moderator Dr. Thomas Kaiser: Nicht ganz. Sie haben eine Stellungnahme zum Vorbericht und Berichtsplan abgegeben. So haben Sie auch das Formblatt unterschrieben und so stand es auch im Leitfaden. Die Aspekte, die Sie zum Berichtsplan genannt haben, werden jetzt diskutiert und erörtert.

Stefan Zenk: Ich habe eine Frage, die sich auch auf die Tagesordnung bezieht im Hinblick auf den Berichtsplan. Ich vermisse die TUNA-Therapie. Hat das ein Ziel, dass die nicht draufsteht, im Hinblick auf die Überarbeitung des Vorberichts?

PD Dr. Stefan Lange: Wir haben nicht alle Verfahren gesondert aufgelistet. Es ist ja nicht nur TUNA, TUMT, sondern zum Beispiel auch CLAP, VLAP. Da, wo jetzt aus unserer Sicht methodisch besondere Probleme bestanden und angesprochen worden sind, wie unter TOP 2 formuliert, wie Verfahren zu Technologien zuzuordnen sind – da ergaben sich bei TUNA eigentlich keine Probleme. Von daher sind die einzelnen Verfahren hier nicht aufgeführt. Wenn Sie aber noch einen Punkt ergänzen oder besprechen mögen, dann können wir das gerne unter TOP 6 tun.

Moderator Dr. Thomas Kaiser: Noch eine grundsätzliche Bemerkung: Wir erörtern hier nicht den Berichtsplan – das ist ein Missverständnis –, sondern Ihre Stellungnahmen, die zum Berichtsplan eingegangen sind. Das heißt, insbesondere erörtern wir die Unklarheiten, die sich für die Projektgruppe aus den Stellungnahmen ergeben haben. Wenn sich also für die Projektgruppe keine Unklarheiten ergeben haben, Ihre Stellungnahmen klar waren, dann gibt es dazu auch keine Nachfragen, dann müssen wir das auch nicht erörtern. Es geht auch nicht darum, hier einen Konsens zu finden, sondern es geht darum, die Stunden, die wir hier sitzen, zu nutzen, Unklarheiten aus den Stellungnahmen zu beseitigen, die sich durch die schriftlichen Formulierungen ergeben haben. Nichtsdestotrotz ist es, wie Herr Lange schon richtig erwähnt hat, möglich, dass unter Punkt „Verschiedenes“ weitere Punkte angesprochen werden, die so dezidiert in der Tagesordnung nicht aufgeführt sind.

Prof. Klaus Höfner: Ich habe noch eine Nachfrage. Unsere Stellungnahmen heißen, wenn ich es richtig sehe, „Stellungnahme zum Vorbericht“. So haben wir im Prinzip unsere Stellungnahmen überschrieben. Es ist mir jetzt nicht ganz transparent, warum Sie zwischen Berichtsplan und Vorbericht unterscheiden und wo Sie die Trennung vornehmen. Wie differenzieren Sie Vorbericht und Berichtsplan?

Moderator Dr. Thomas Kaiser: Das ist relativ einfach zu differenzieren, nämlich durch die Abschnitte. Der Abschnitt Kapitel 4 schildert die Berichtsmethodik. Der Abschnitt „Kapitel 3“ schildert die Zielsetzung des Berichtes. So ist das auch in dem jeweiligen Berichtsplan zu allen Projekten. Sie waren dazu aufgerufen, sowohl zum Berichtsplan als auch zum Vorbericht Stellung zu nehmen. Auch wenn Sie Ihre Stellungnahme mit „Stellungnahme zum Vorbericht“ überschrieben haben, haben Sie dennoch zu Aspekten des Berichtsplans Stellung genommen. Das lässt sich aus den Stellungnahmen eindeutig identifizieren. Genau diese Punkte erörtern wir heute.

Dr. Kai Schorn: Ich möchte trotzdem noch einmal feststellen, dass auch unsere Stellungnahme von der Deutschen Krankenhausgesellschaft „Stellungnahme zum Vorbericht“ heißt und dass es in Ihrem Anschreiben sehr missverständlich war, dass man auch gleichzeitig zum Berichtsplan Stellung nehmen konnte. Das war sozusagen nicht klar formuliert. Ich möchte mich eigentlich dagegen verwehren, unsere Stellungnahme dahingehend zu deuten, dass wir zu dem Berichtsplan Stellung genommen hätten. Zum Berichtsplan hätten wir noch anders Stellung nehmen können als zum Vorbericht. Die Stellungnahme heißt „Stellungnahme zum

Vorbericht“ und ist auch nur als Stellungnahme zum Vorbericht von unserer Seite zu werten. Sie können das hinterher nicht nach Gusto umdeuten.

Moderator Dr. Thomas Kaiser: Dazu möchte ich Folgendes bemerken: Auch wenn Sie Ihre Stellungnahme mit „Stellungnahme zum Vorbericht“ überschrieben haben, sind Sie ausdrücklich zur Stellungnahme zum Vorbericht und Berichtsplan aufgerufen worden. Ob Sie zum Berichtsplan Stellung nehmen, steht Ihnen natürlich frei. Diese Gelegenheit hatten Sie. Wenn Sie sie nicht wahrgenommen haben, dann ist das so, dann nehmen wir das zur Kenntnis. Sie können natürlich gerne auch nicht an der Erörterung teilnehmen, wenn Sie sagen, wir haben sowieso nicht zum Berichtsplan Stellung genommen und es ist nicht relevant, dass ich heute an der Erörterung teilnehme. Aber grundsätzlich hatten Sie die Gelegenheit. Es war Ihnen freigestellt, das wahrzunehmen. Entsprechend werden wir das heute erörtern.

Dr. Kai Schorn: Habe ich Sie richtig verstanden, dass ich an der Erörterung nicht teilnehmen soll, wenn ich nicht zum Berichtsplan Stellung genommen habe?

Moderator Dr. Thomas Kaiser: Herr Schorn, es macht keinen Sinn, wenn Sie dazwischenreden. Bitte vorab immer den Namen sagen, wie am Anfang gesagt. Ich habe nicht gesagt, Sie sollen nicht Stellung nehmen. Ich habe Ihnen freigestellt, an der Erörterung teilzunehmen oder nicht. Das können Sie sich überlegen.

PD Dr. Stefan Lange: Zur Glättung der Wogen: Herr Schorn, das ist eigentlich gar nicht böse gemeint von uns, sondern im Gegenteil. Wir haben hier schon im relativ vorauseilenden Gehorsam im letzten Jahr den Wunsch des Gesetzgebers vorhergesehen und für zukünftige Verfahren die Möglichkeit eröffnet, das getrennte Stellungnahme- und gegebenenfalls Anhörungsverfahren durchzuführen. Wir haben uns jetzt nach Rücksprache mit anderen Gremien überlegt, dass es vielleicht sinnvoll wäre, für die Verfahren, die wir bedauerlicherweise nicht mehr bis Mitte bzw. bis Ende des ersten Quartals dieses Jahres abschließen konnten, zu sagen, dann machen wir das jetzt hier ganz genauso. Das heißt, das ist eher ein Angebot als eine Restriktion oder eine Missachtung. Wir werden also noch ein zusätzliches Stellungnahmeverfahren eröffnen. Wir wollen das jetzt nur trennen, wie das auch im GKV-WSG für die Arzneimittel vorgesehen ist, einerseits Stellung nehmen und gegebenenfalls Anhörung zur Methodik und dann ein zweites Mal zu den Ergebnissen. Das werden wir jetzt hier genauso machen. Das heißt, es wird einen überarbeiteten Vorbericht geben und dann gibt es noch einmal eine Stellungnahmemöglichkeit. Das ist also nicht böse gemeint, Herr Schorn.

Dr. Kai Schorn: Das ist ja auch sehr löblich, es kommt nur ein bisschen missverständlich herüber. Wenn Sie das so darstellen, dass Sie aus den Stellungnahmen zum Vorbericht die Aspekte diskutieren wollen, die die Berichtsmethodik betreffen, dann ist das in Ordnung. Aber wenn Sie hinterher deklarieren, wir hätten explizit damit zum Berichtsplan Stellung genommen, dann ist das nicht in Ordnung. Wenn wir sagen, o. k., wir treffen uns hier heute, eigentlich haben wir alle, die hier am Tisch sitzen, mit den Überschriften klar geäußert, wir nehmen Stellung zum Vorbericht und nicht mehr zum Berichtsplan, und wir diskutieren jetzt

aufgrund Ihres Anliegens die methodischen Aspekte, weil Sie das sozusagen anbieten, und können noch einmal Stellung nehmen zu dem überarbeiteten Vorbericht, dann ist das löblich. Wenn man das so klar darstellt, dann ist das auch in Ordnung. Aber ich möchte nicht, dass unsere Stellungnahme umgedeutet wird als auch Stellungnahme zum Berichtsplan. Das war sie definitiv nicht.

Moderator Dr. Thomas Kaiser: Abschließende Äußerung, dann ist dieser Punkt verlassen: Ich erinnere noch einmal daran, was das Formblatt beinhaltet hat, welche Überschrift, was in dem Leitfaden stand. Sie haben das Formblatt unterschrieben. Sie haben damit unterschrieben, dass Sie den Leitfaden zur Kenntnis genommen haben. Dort ist explizit zur Stellungnahme auch zum Berichtsplan aufgerufen worden. Wenn Sie das nicht wahrgenommen haben, dann ist das Ihre Entscheidung. Aber nichtsdestotrotz werden heute die Stellungnahmen zum Berichtsplan diskutiert. Sie hatten Gelegenheit, dazu Stellung zu nehmen.

Dr. Roman Schiffner: Es wird also dann noch einmal eine weitere Erörterung zum Vorbericht geben?

Moderator Dr. Thomas Kaiser: Richtig. Es wird einen überarbeiteten Vorbericht geben. Dieser überarbeitete Vorbericht wird wieder veröffentlicht. Dazu wird es wieder eine Gelegenheit zur Stellungnahme geben. Sollten sich aus diesen Stellungnahmen wieder Unklarheiten ergeben – die Erörterung ist ja optional –, dann wird wieder eine Erörterung durchgeführt.

Wir kommen nun zur inhaltlichen Arbeit und zur Besprechung der methodischen Punkte, also zu den Aspekten, wie im Berichtsplan genannt, und zu den Dingen, zu denen Sie Stellung genommen haben und aus denen sich Unklarheiten für die Projektgruppe ergeben haben.

Ich rufe auf:

Tagesordnungspunkt 1:

Vorstellung des Auftrags und der Vorgehensweise des IQWiG

Katharina Biester: Gegenstand des Auftrags war die Darstellung des aktuellen Wissensstandes zu den nichtmedikamentösen lokalen Verfahren zur Behandlung der BPH. Ziel der Untersuchung ist eine vergleichende Nutzenbewertung der nichtmedikamentösen lokalen Verfahren gegenüber operativen Verfahren, der nichtmedikamentösen lokalen Verfahren untereinander und der nichtmedikamentösen lokalen Verfahren gegenüber anderen Interventionen, zum Beispiel Medikamente oder Sham-Interventionen.

Moderator Dr. Thomas Kaiser: Sie sehen ja die Tagesordnung vor sich liegen.

Ich rufe auf:

Tagesordnungspunkt 2:

Zuordnungen von Studien zu Therapietechnologien

Hier sind 3 Unterpunkte genannt. Wir beginnen mit der

CLasP-Studie (CLAP/VLAP).

Katharina Biester: Wir waren der Ansicht, dass die VLAP in dieser Studie eingesetzt wurde, weil unter anderem „non contact visual laser ablation“ in einer der Publikationen beispielsweise aus einem Abstract hervorging. In den Stellungnahmen wurde aber angemerkt, dass diese Studie der CLAP zuzuordnen sei. Wir würden gerne nachfragen, ob wir einen Hinweis übersehen haben oder wie diese Studie zuzuordnen ist.

Dr. Richard Berges: Unser Problem ist im Moment, dass unser Laserexperte fehlt, Herr Muschter, der darüber detailliert aussagen kann. Er ist draußen. Ich weiß nicht, warum er nicht auf der Liste stand. Da muss irgendein Fehler passiert sein. Er ist eingeladen worden.

Katharina Biester: Er hat aber nicht zugesagt.

Moderator Dr. Thomas Kaiser: Bis auf Herrn von Zastrow sind wir vollständig, was die Eingeladenen und Zugesagten angeht. Es gab ja einen Brief an jeden. Diesen Brief haben Sie bekommen und Ihre Teilnahme entsprechend zugesagt. Eine Zusage von dem von Ihnen erwähnten Kollegen gab es nicht. Insofern ist die Teilnahme nicht möglich. Es ist eine nichtöffentliche Veranstaltung. Die Bedingungen sind vorab so geschildert worden.

Dr. Richard Berges: Die Zusage von Herrn Muschter habe ich persönlich hier abgegeben. Ich persönlich bin hier gewesen und habe zusammen mit seinen nachgeforderten Unterlagen die Zusage an die Sekretärin abgegeben. Dort waren drin seine Unterschriften, seine Zusage zur Teilnahme und seine nachgeforderten Literaturstellen.

Moderator Dr. Thomas Kaiser: Jetzt haben wir ein Problem.

Dr. Richard Berges: Und Herr Muschter steht vor der Tür und hat sich den Tag freigegeben.

Katharina Biester: Die Unterlagen, die nachgeforderten Literaturstellen und das Formblatt habe ich erhalten. Zu dem Zeitpunkt hatten wir die Einladung noch gar nicht geschrieben. Wir haben die Einladungen erst rausgeschickt, nachdem wir die fehlenden Unterlagen erhalten haben. Deswegen bin ich mir nicht sicher, ob die Zusage dabei sein kann.

Dr. Richard Berges: Ich bin jetzt auch nicht sicher. Ich bin nur sicher, dass ich die Sachen hier vorbeigebracht habe. Ich bin davon ausgegangen, dass seine Zusage mit drin war.

Moderator Dr. Thomas Kaiser: Ich denke, es ist unproblematisch, auf eine entsprechende Anfrage zu antworten, in der steht: Wir bitten um Rückmeldung innerhalb von so und so viel Tagen. Wir müssen diese Veranstaltung ja auch organisieren. Es kann also in den Unterlagen, wenn die Einladung noch gar nicht herausgegangen ist und damit auch der Termin noch nicht feststand, keine Bestätigung zur Teilnahme enthalten sein. Ich denke – es bleibt uns leider nichts anderes übrig, denn sonst werden wir diesen Punkt immer wieder und an anderen Stellen haben; so sind die Kautelen der Erörterung; diese können wir nicht immer wieder ändern – , dass wir jetzt mit den Punkten fortfahren sollten.

Prof. Klaus Höfner: Natürlich haben Sie Ihre Regularien. Auf der anderen Seite sollten Sie zur Kenntnis nehmen, Herr Muschter ist im Arbeitskreis, er ist sozusagen unser Experte, der jetzt draußen vor der Tür steht. Es ist also die Frage, ob man in diesem Falle die Regularien zur Kenntnis nimmt und dieses auch im Protokoll kommentiert. Aber es ist vielleicht überlegenswert, ob man nicht der Sache wegen Herrn Muschter hereinbittet, damit er seine Expertise hier darstellen kann. Das möchte ich zu bedenken geben.

Moderator Dr. Thomas Kaiser: Ich denke, es ist die beste Lösung, dass die Punkte einzeln noch einmal notiert werden und dann bei Nachfragen eine entsprechende schriftliche Anfrage an Herrn Muschter gestellt wird.

Dr. Richard Berges: Es ist halt für die Sache wichtig. Herr Muschter ist der Einzige, der wirklich im Detail die Hunderte von verschiedenen Therapiesystemen, Lasersystemen erklären kann. Es ist auch für einen Urologen extrem schwierig, die einzelnen Methoden auseinanderzuhalten. Wenn Sie da auf second best, also auf uns, zurückschalten, dann bekommen Sie möglicherweise unscharfe Antworten.

Moderator Dr. Thomas Kaiser: Zu den Punkten 2 und 3 sollten wir eine Ausnahmeregelung machen, weil es für diese Punkte erforderlich ist.

PD Dr. Stefan Lange: Ich möchte kurz ergänzen, damit Sie nicht den Eindruck haben, das sei blöder und dummer Formalismus. Wir haben wegen solcher Dinge, Teilnahme von Personen, die eigentlich nicht teilnahmeberechtigt sind, unangenehme Auseinandersetzungen, um es vorsichtig auszudrücken. Es geht also bis zu Gerichtsverfahren. Das ist der Grund. Das hängt damit zusammen, diese Personen müssen sich mit unseren Regularien einverstanden erklären. Wenn wir das vorher nicht wissen, dann sind einfach diese Sachen nicht vernünftig prüfbar. Wir wollen das nicht so, aber wir sind in mehreren Verfahren letztlich dazu gezwungen worden.

Prof. Klaus Höfner: Es geht nicht darum, das zu kritisieren. Wir kennen ja auch Regularien. Aber verstehen Sie auch uns. Es geht uns um die Sache. Das ist keine Kritik an Ihren Regularien.

Ich habe noch eine Frage zum Verfahren. Sie haben ja eine strikte Vorgabe der Tagesordnung gemacht. Es gibt 2 Punkte, die man sicherlich einbauen kann, die wir für sehr wesentlich halten, und zwar ist das ein methodisches Problem. Es geht um die Frage, was in diesem Zusammenhang patientenrelevante Ziele sind und was Standard ist. Das sind zwei wesentliche Punkte, über die man getrennt reden sollte, wobei sich das durch das ganze Papier zieht. Ich denke, diese 2 Fragen betreffen alle weiteren Aussagen. Deswegen ist das für uns ein zentraler Punkt.

Moderator Dr. Thomas Kaiser: Wie zu Beginn erwähnt, ist es durchaus unter „Verschiedenes“ möglich, noch Punkte anzusprechen. Noch einmal zum Ziel der Erörterung: Es geht nicht darum, den kompletten Berichtsplan zu erörtern, sondern darum, Unklarheiten zu beseitigen, die sich aus Ihren Stellungnahmen ergeben haben. Wenn sich aus Ihren Stellungnahmen bezüglich dieser Aspekte keine Unklarheiten ergeben haben, dann muss man das nicht besprechen. Sie haben ja Stellung genommen. Entscheidend sind die schriftlichen Stellungnahmen. Diese Unklarheiten aus den schriftlichen Stellungnahmen werden erörtert. Nichtsdestotrotz können wir diese Punkte unter „Verschiedenes“ ansprechen. Bevor wir noch einmal die Tagesordnung diskutieren, bitte ich darum, alles in den Punkt „Verschiedenes“ zu verschieben. Wir werden die Tagesordnung durchgehen, und unter „Verschiedenes“ kommen wir zu dem Punkt noch einmal zurück. Dann kann sich dazu jeder noch einmal melden.

Ich möchte nun unseren neuen Gast, Herrn Muschter, begrüßen. Das, was gerade eben mit Ihnen passiert ist, ist kein böser Wille von uns, sondern es liegt daran, dass Sie auf die Einladung nicht reagiert haben, dass Sie keine Zusage zur Teilnahme gegeben haben, aus welchem Grund auch immer, und dass wir bestimmte Regularien haben. Wir haben auf Bitten Ihrer Kollegen allerdings eine Ausnahme gemacht, weil Ihre Kollegen gesagt haben, dass sie Ihre Expertise für die beiden Tagesordnungspunkte 2 und 3 benötigen. Wir sind natürlich sehr erfreut, dass wir von Ihrer Expertise jetzt profitieren und Sie zu den entsprechenden Rückfragen Stellung nehmen können. Ich weise Sie darauf hin, dass es eine Aufzeichnung gibt. Die Aufzeichnung läuft bereits. Es wird von dieser Erörterung auch ein Wortprotokoll geben. Die Zustimmung zu Aufnahme und Wortprotokoll ist Voraussetzung für die Teilnahme. Sie haben zwar schon im Formblatt unterschrieben, dass Sie diese Teilnahmebestimmungen anerkennen. Nichtsdestotrotz möchte ich Ihnen Gelegenheit geben, wenn Sie inzwischen Ihre Meinung geändert haben, Ihre Zustimmung zu widerrufen. Aber ich sehe, das ist nicht der Fall. Dann können wir jetzt tatsächlich in die Sache einsteigen.

Frau Biester, bitte geben Sie noch einmal den Einstieg.

Katharina Biester: Wir hatten bezüglich der CLasP-Studie die Frage, dass wir der Ansicht waren, dass in der Studie die VLAP eingesetzt wird. In den Stellungnahmen wurde jedoch angesprochen, dass die Studie der CLAP zuzuordnen sei. Wir fragen nun nach, ob wir in der Studie einen Hinweis übersehen haben, weil den Hinweis, den wir unter anderem gefunden

haben, war, dass „noncontact visual laser prostate ablation“ beschrieben wurde. Darüber wollen wir mit Ihnen reden.

Prof. Dr. Rolf Muschter: Das Problem bei diesen frühen Studien ist, dass die Nomenklatur sehr unscharf verwendet wurde. Das heißt, die Autoren und die Anwender der Lasertechnologie haben etwas gemacht, das sehr schlecht definiert war, und haben die damals vorhandene Nomenklatur darauf angewendet. Das Problem ist, dass man das auch im Nachhinein methodisch nicht sauber trennen kann. Das ist das Problem von vielen dieser frühen Studien, dass da „Laserprostatektomie“ oder „visuelle Laserablation“ oder „Vaporisation“ darübersteht und das tatsächlich eingesetzte Verfahren dieser Nomenklatur nicht genau folgt. Das ist ein ganz großes Problem. Das ist aber letztlich inhaltlich, wenn man diese Art von Techniken zusammenfasst, unerheblich, weil sie von der Wirkung her äquivalent sind. Letztlich ist der Name, dem der Autor seinem Verfahren gegeben hat, sekundär. Immer dann, wenn es sich um transurethrale Laseranwendung mit Einsatz eines Neodym:Yag-Lasers – etwas anderes gab es in den 90er Jahren nicht – handelt, ist das relativ wirkungsäquivalent.

PD Dr. Stefan Lange: Ich habe eine Nachfrage. Wir haben ja die Aufteilung Kontakt, Nonkontakt bzw. CLAP, VLAP. Sie würden also sagen, das braucht man nicht zu trennen?

Prof. Dr. Rolf Muschter: Doch. Kontakt und Nonkontakt sind unterschiedliche Techniken. Solange „Kontakt“ darüber steht, ist es Kontakt. Bei Nonkontakt ist bei den frühen Studien manchmal diese Faser im Kontakt eingesetzt worden. Nonkontakt ist also nicht immer Nonkontakt. So ist das zu verstehen gewesen, was ich gerade gesagt habe. Bei Nonkontakt halten Sie eine Laserfaser in die Harnröhre und bestrahlen mit einer rechtwinklig abstrahlenden Faser die Oberfläche. Wenn Sie im Nonkontaktverfahren nach dem damals üblicherweise verwendeten 4-Quadranten-Schema vorgehen, dann bewegen Sie die Faser nicht, sondern bestrahlen einen einzigen Punkt. Je nach Abstand der Faser zur Oberfläche, je nach Laserleistung und Absorption dieser Laserleistung im Gewebe kommt es entweder zu einer langsamen, tiefen Koagulation des Gewebes oder zu einer frühen Karbonisation des Gewebes mit konsekutiv sämtlicher zusätzlicher Energieabsorption an der Oberfläche, sodass es einem Vaporisationseffekt gleichkommt. Aus dieser Beobachtung haben die Autoren geschlossen, dann kann man auch gleich näher herangehen und diesen Vaporisationseffekt nutzen, haben aber nicht die Technik anders genannt, weil sie in den 4 Quadranten in der gleichen Prozedur manchmal vaporisierten und manchmal koagulierten. Postuliert wurde als Ergebnis immer eine tiefe Koagulationsnekrose. Im Unwissen, dass das mit einer frühen Karbonisation nicht zu erzielen ist, sind auch diese heterogenen klinischen Ergebnisse zu erklären. Das ist einfach ein methodisches Problem, weil die damaligen Anwender gar nicht wussten, was der Laser da ganz genau tut. Die Variablen, die man im Labor untersuchen kann, kann man in die Klinik gar nicht richtig umsetzen. Das ist das Problem dabei.

PD Dr. Stefan Lange: Ich muss noch einmal nachfragen, denn das ist wichtig. Ich mache es einfach: Kontakt ist immer Kontakt und Nonkontakt kann auch Kontakt sein?

Prof. Dr. Rolf Muschter: Ja.

PD Dr. Stefan Lange: Ich meine das nicht ironisch, sondern sage das nur, um das zu verstehen. Wir wollen ja versuchen, Ordnungskriterien zu haben, also Ordnung hineinzubringen. Dann wäre ein sinnvolles Ordnungskriterium, zu sagen, wenn in einer Arbeit beschrieben ist, dass sie mit einem Kontaktlaser gearbeitet haben, dann würde man das dem Verfahren zu Kontaktlaser zuordnen, und wenn das nicht so klar beschrieben ist, dann ist es quasi der Rest, also Kontakt versus Rest.

Prof. Dr. Rolf Muschter: Richtig.

PD Dr. Stefan Lange: Wir haben uns das ja nicht selber ausgedacht. Manchmal schien es klar, aber andere haben in ähnlichen Arbeiten, systematischen Übersichtsarbeiten, zum Beispiel ein relativ aktuelles Werk von den Amerikanern, von der AHRQ, die gleiche Einteilung wie wir vorgenommen. Würden Sie sagen, das entspricht nicht dem heutigen Wissensstand?

Prof. Dr. Rolf Muschter: Kontakt war eine definierte Art von Kontaktspitze. Das ist ein definiertes Verfahren; dazu gibt es überhaupt keine Frage. Die Verfahren, die als VLAP, LAP, TULAP, TUEP usw. benannt worden sind, haben deswegen unterschiedliche Namen, weil die Applikationsspitzen für den Laser - -⁵ Man hat einen Lichtleiter, der die Laserenergie vom Lasergerät ins Körperinnere überträgt. Am Ende dieser Laserfaser ist eine mit dem Produkt, mit der Faser fest verbundene Applikationsspitze, die die Laserstrahlung rechtwinklig, im 45-Grad-Winkel oder in irgendeiner Form umleitet, die jetzt je nach Hersteller und Umleitungsprinzip ein unterschiedliches Aussehen und unterschiedliche Eigenschaften hat. Das ganz simple Ursprungsmodell, das zu der Bezeichnung VLAP geführt hat, ist eine Metallspitze aus Gold mit einem Spiegel, sodass die Laserenergie um etwa 90 Grad umgelenkt wurde.

Kurz darauf hat eine andere Firma ein Refraktionsumleitungsprinzip eingesetzt. Das heißt, das ist ein Glaskolben, in dem ein Glasprisma die Laserstrahlung umleitet. Dadurch, dass man einen Glaskolben hat, kann man das bewegen. Wenn man dann in Kontakt kommt, dann verhält sich das wie eine Kontaktspitze, weil eine Kontaktspitze von vornherein ein Glaskolben ist. Die Anwender dieses Verfahrens, obwohl das auch noch VLAP hieß, weil das Produkt, das sie eingesetzt haben, auch für VLAP benutzt wurde, haben dann im Grunde genommen so eine Mischtechnik eingeführt. Aber Sie können es heute nicht anders klassifizieren, weil alle diese Arbeiten letztlich diesem gewissen Zufallsprinzip unterliegen, weil auch der, der Nonkontakt arbeiten wollte, unter Umständen mit der Spitze in Kontakt gekommen ist. Das ist einfach eine Frage der eingesetzten Technologie. Kontakt ist eine ganz klar definierte Technologie und das andere muss man als Nonkontakt führen, aber letztlich zusammenfassen.

Dr. Richard Berges: Vielleicht sollte man das auf eine höhere Ebene heben bezüglich der verschiedenen Techniken, Gewebe zu entfernen. Ich glaube, das ist vom Verständnis her ganz

⁵ Abbrechen des Satzes

wichtig. Grundsätzlich gibt es die Möglichkeit, primär und sekundär ablativ Verfahren einzusetzen. Der Unterschied liegt auf der Hand. Einmal wird während der Operation das Gewebe sofort entfernt und einmal wird es durch eine Nekrosebildung zerstört. Letztendlich ist es technisch völlig egal, wie man diese beiden Verfahrensweisen anwendet. Eigentlich sind es alles nur Abwandlungen der TURP, was die primär ablativen Verfahren betrifft, und es ist egal, welche Technik man nimmt. Oder es sind Abwandlungen von Hitzeverfahren, die sekundär ablativ wirken. Auch da ist es eigentlich egal, welche Hitzequelle man benutzt, weil das Gewebe immer auf die gleiche Weise zerstört wird. Deswegen sind die technischen Nuancen, die dort erfunden wurden, gerade bei den Lasern, die ja so mannigfaltig sind, wahrscheinlich Spielereien, die Sie klinisch nirgends wiederfinden werden. Wichtig ist, ob es sich um ein primär ablativ Verfahren oder um ein sekundär ablativ Verfahren handelt, weil das auch im Outcome und im Verlauf Unterschiede bietet. Alles andere ist wahrscheinlich klinisch nicht relevant.

Prof. Klaus Höfner: Eine Ergänzung: Wir sprechen auch von nicht ablativen Verfahren gegenüber (...) ⁶ symptomatische Verbesserungen erreichen, indem durch Hitze Nerven zerstört werden. Da funktioniert eine symptomatische Verbesserung, die man ja haben will, einfach durch eine Nervenzerstörung, auch durch Hitze. Wenn man es genau nimmt, geschieht das sowohl über nicht ablativ Verfahren gegenüber ablativen Verfahren, und die sind sofort ablativ oder verzögert ablativ, primär oder sekundär ablativ, wie man das auch immer bezeichnet. Die Unterscheidung der primären und sekundären Ablation hat erhebliche Konsequenzen für den Patienten, da die sekundäre Ablation ja zunächst einmal über eine Entzündung oder Schwellung der Drüse läuft. Das heißt, da wird eine Hitze auf die Drüse gesetzt, was sekundär eine Nekrose erzeugt und zu einer Schrumpfung führt – das nennen wir sekundäre Ablation –, aber zwischendurch entsteht eine Zunahme der Beschwerden des Patienten. Das heißt, er hat vermehrt Beschwerden. Erst nachdem er über den Berg an Beschwerden gestiegen ist, geht er in die letztlich definitive symptomatische Besserung. Das ist unterschiedlich ausgeprägt, aber wichtig ist – das ist auch eine Frage der Aufklärung des Patienten –, dass er weiß, er wird nicht sofort einen Benefit haben, sondern der Benefit wird sekundär eintreten. Das geht in alle Verfahren hinein. Das heißt, wir haben im Moment diese Erklärung bei den Laserverfahren. Das betrifft auch alle Hitzemethoden, wie Herr Berges das schon gesagt hat. Das heißt, es macht eigentlich mehr Sinn, diese Einteilung vorzunehmen, aber diese Einteilung findet sich zurzeit in keinem Papier. Das heißt, wir sind eigentlich die Einzigen, die diese Terminologie verwenden. Ich sage das aber zum Verständnis für Sie. Es ist wegen der Einordnung ganz wichtig, weil das extreme Konsequenzen für die Einordnung dieser Verfahren hat.

Prof. Dr. Rolf Muschter: Das Problem ist, dass diese Techniken einer Evolution unterlagen und dass viel später in Konsensuskonferenzen eher auf die Gemeinsamkeiten eingegangen worden ist als auf die Unterschiede. Man hat anfangs nach eingesetzter Energie eine klare Trennung vollzogen und hat gesagt: Laser und elektrische Energie ist etwas ganz anderes.

⁶ In der Aufzeichnung nicht zu verstehen.

Dann ist man später allmählich zu der Erkenntnis gelangt, dass es nicht entscheidend ist, ob man mit einem Laser, mit elektrischer Energie oder mit irgendeiner anderen Wärmequelle arbeitet, sondern dass entscheidend ist, was tatsächlich am Gewebe erzeugt wird und dass Letzteres mit wenigen Ausnahmen anwenderabhängig und nicht systemabhängig ist. Das heißt, ob ich fünfmal oder nur dreimal einen Punkt bestrahle, führt einfach zu einem anderen Ergebnis, ohne dass das Verfahren einen anderen Namen bekommt, weil ich die gleiche Energiequelle einsetze. Dieses Wissen ist erst spät in die Literatur eingegangen, aber in den alten Studien natürlich nicht wiederzufinden, weil man das Wissen zu dem Zeitpunkt, als die Studien gemacht worden sind, noch nicht hatte. Retrospektiv ist also diese Einteilung, die wir genannt haben, jetzt eingesetzt worden, und retrospektiv kann man heute die Studien so klassifizieren. Das wird in den entsprechenden Leitlinienkommissionen, in verschiedenen Ländern auch so gemacht, also die Einteilung in thermische sekundär ablativ Verfahren und in das, was primär ablativ wirkt, wo man also während des Eingriffs tatsächlich Gewebe entfernt. Dies ist natürlich ganz besonders anwenderabhängig, weil ich ja entscheide, wann ich aufhöre. Das ist auch bei dem Standard der Transurethralen Resektion der Fall. Da gibt es natürlich einen Lehrbuchstandard, wie man das machen soll, aber das Verfahren heißt nicht anders, wenn man es sozusagen nur halb macht, was klinisch durchaus vorkommt, weil man manchmal durch äußere Umstände gezwungen ist, vorher aufzuhören, oder von vornherein sagt, ich mache, weil der Patient nur noch eine begrenzte Lebenserwartung hat, nur eine unvollständige Behandlung, um keine Risiken zu haben. Das ist also eine Entscheidung des jeweiligen Operateurs, ohne dass sich das im Namen des Verfahrens ausdrückt.

Dr. Volker Rohde: Wir als Berater haben uns ja auch Gedanken gemacht und uns zugegebenmaßen bei der Zuordnung sehr schwergetan. Ich komme darauf zurück, was Sie gesagt haben. Nicht abweichend von dem, was wir verfasst haben, finde ich jetzt eine andere Kategorisierung bei der Zuordnung von VLAP zu CLAP, denn die physikalischen Prinzipien, dass bei Nd:Yag-Lasern natürlich neben der dominierenden Nekrose-Korrelation auch eine geringartige Vaporisation vorgeht, haben wir ja besprochen gehabt. Ich denke, wir kommen da an Grenzen bei einigen Verfahren, wo wir uns jetzt darüber streiten, was in der praktischen Anwendung deutlich ins zweite Glied zurückgerutscht ist. Das heißt, wir diskutieren hier um eine Technologie, die kaum noch Anwendung findet, und haben jetzt eine historische Bewertung von alten Studien, die wir jetzt in irgendeiner Form lösen müssen. Ich sehe aus der Diskussion nicht, wie wir es hätten anders machen sollen.

PD Dr. Stefan Lange: Genau das ist die Frage. Wir haben das ja aufgegriffen, weil es ein relativ prominenter Kritikpunkt war, dass hier eine fälschliche Zuordnung vorgenommen worden sei, insbesondere dieser CLasP-Studie. So richtig verstanden habe ich es nicht. Es gibt Kontaktlaserverfahren und dann im Prinzip den Rest, der gewissermaßen ein Sammelsurium ist. Aber da würde doch die CLasP-Studie eher zu dem Rest und nicht zu einem Kontaktlaserverfahren gehören, weil es ja dort eben nicht explizit als ein Kontaktlaserverfahren bezeichnet worden ist. Das war quasi der Ausgangspunkt unseres Problems. Aber im Prinzip können wir die Diskussion zu dem Punkt jetzt beenden.

Prof. Dr. Rolf Muschter: Dieser Kritikpunkt besteht insofern zu Recht und ist wichtig, als Sie die Transurethrale Elektrovaporisation zum Standard mit hinzugenommen haben. Dieses ist aber im Grunde genommen klinisch das Gleiche wie eine Kontaktlaservaporisation. Hier ist das Abgrenzungsproblem. Alles, was nicht explizit mit Kontaktspitzen gemacht worden ist, darf man nicht guten Gewissens als Kontaktvaporisation bezeichnen. Deswegen grenze ich das jetzt so ab, indem ich sage, das ist der Rest. Aber die CLasP-Studie hat ja schon im Namen dieses Kontakt. Das ist ganz klar eine Kontaktstudie. Ich weiß nicht, ob Sie das inhaltlich nicht darin wiedergefunden haben.

PD Dr. Stefan Lange: Das Akronym „CLasP“ steht nicht für Kontaktlaser, sondern für „conservative management, laser, transurethral resection of the prostate“. Im Abstract zu all diesen Studien steht explizit, dass es sich um ein Nonkontaktverfahren handelt. Das ist unser Problem dabei. Aber wir haben das ja jetzt ausführlich besprochen. Den Punkt – da komme ich auf Ihren Einwand von vorhin zurück – TUEVRP, also die Vaporisationsverfahren, die wir ja noch zum Teil als Standard geführt haben, wollten wir unter dem nächsten Tagesordnungspunkt besprechen.

Moderator Dr. Thomas Kaiser: Damit ist TOP 2 Unterpunkt 1 geklärt.

Wir kommen damit zum Unterpunkt 2.

Hybrid-Verfahren

Katharina Biester: Dieser Unterpunkt richtet sich ebenfalls an Herrn Muschter. Da kam in den Stellungnahmen der Hinweis, dass eine Vermischung verschiedener Kombinationen von Laserverfahren medizinisch nicht sinnvoll sei. Da wollten wir genauer nachfragen, warum und wie man das vielleicht sonst lösen kann.

Prof. Dr. Rolf Muschter: Der Begriff „Hybrid“ stammt von Graham Watson. Der ist etwa im Jahr 1990 publiziert worden. Das ist keine randomisierte Studie. Die haben Sie also bestimmt nicht berücksichtigt. Er hat damals verschiedene Laser miteinander kombiniert. Er hat einen Grünlichtlaser und einen Nd:Yag-Laser zusammen eingesetzt. Es haben dann später Versuche stattgefunden, die verschiedenen Wirkungen des Lasers auf das Gewebe miteinander zu kombinieren. Manche Autoren haben das dann „Hybrid“ genannt. Dann gab es Versuche, den Nachteil der Laserverfahren – das muss man dahinter verstehen - -⁷ Der Nachteil der Laserverfahren, die nicht primär ablativ waren, war, dass sie zunächst eine Schwellung hervorgerufen haben und die Patienten zunächst eine noch höhere Symptomatik hatten als vor dem Verfahren, weil sie sekundär ablativ waren. Diese Übergangszeit musste mit einem Katheter überbrückt werden. Das hat man insbesondere in den angloamerikanischen Ländern als großes Problem angesehen und hat versucht, diese Katheternotwendigkeit zu vermeiden. Aus diesem Grund hat man versucht, es irgendwie zu gewährleisten, dass man doch sofort ablativ

⁷ Abbrechen des Satzes

wirkt. Deswegen hat man gesagt, wir machen das minimalere, schonendere, unblutigere Koagulieren in die Tiefe mit einem Laser und nehmen dann die Resektionsschlinge und nehmen gerade so viel Gewebe weg, dass der Patient keinen Katheter mehr braucht, also eine Kombination aus Laser und TUR. Das wurde dann auch „Hybrid“ genannt. Sie haben hier also Hybrid als Kombination verschiedener Laser, Sie haben Hybrid als Kombination von Laser mit Vaporisation und Koagulation und Hybrid als Kombination von Lasereinsatz und Resektion. Wenn Sie den Begriff „Hybrid“ als Sammelbegriff benutzen, dann werfen Sie ganz viele, vollkommen unterschiedliche Techniken in einen Topf, die vor allem im Hinblick auf ihre Wirkungen und Nebenwirkungen nicht vergleichbar sind, weil die einen eben nicht sofort abtragen, die anderen das schon tun. Da das ohnehin keine standardisierten Techniken sind, die im Einsatz sind – niemand würde heute sagen, ich benutze standardmäßig ein Hybrid-Verfahren –, ist das im Grunde genommen eine historische Problematik.

PD Dr. Stefan Lange: Das kann ich nachvollziehen. Kurz zur Erläuterung: Die zu bewertenden Verfahren, so wie sie in dem Bericht aufgelistet sind, haben wir uns nicht selbst ausgedacht, sondern das war ein Auftrag vom Gemeinsamen Bundesausschuss, dem wiederum ein Antrag von einer der Bänke aus dem Gemeinsamen Bundesausschuss zugrunde lag, sodass solche Begrifflichkeiten möglicherweise auch historisch bedingt auftauchen.

Wir haben in unserem Bericht 4 Studien dem Hybrid-Verfahren zugeordnet, wobei man jetzt auch wieder sagen könnte: Hybrid ist quasi keine geschützte Bezeichnung wie vielleicht TUMT oder TUNA, sondern eine Kombination verschiedenster Verfahren, also ein Sammelurium. Wäre es denn sinnvoll, gegebenenfalls eine Unterteilung vorzunehmen, zum Beispiel in Verfahren, die Sie auch angesprochen haben, Nd:YAG-Laser und Greenlightlaser – das wäre ja ein klares Vorgehen, das teilweise passiert ist –, oder vielleicht – ich glaube, das haben wir aber nicht betrachten können – die Kombination einer richtigen Abtragung und einem Laserverfahren oder eine Kombination von einem Kontakt- und einem Nichtkontaktverfahren? Würde das, wenn man das trennen würde, Sinn machen, oder sollte man gar nicht die Kategorie „Hybrid“ einführen? Was macht man dann? Dann hat man 4 Einzelstudien, wo man „Sonstiges“ sagt. Aber das ist auch eine Kategorie.

Prof. Dr. Rolf Muschter: Im Grunde genommen sind es Einzelfälle. Ob man die vier zu einer Gruppe zusammenfasst oder einzeln betrachtet, es sind im Grunde genommen unbedeutende Verfahren, über die irgendwann ein Autor eine Veröffentlichung gemacht hat. Dann ist das verlassen worden. Es hat niemand im großen Stil nachgeahmt. So gesehen macht eine Bewertung im Sinne eines Vergleichs zu einem Standard wenig Sinn. Letztlich würde ich so etwas schlichtweg eliminieren.

Moderator Dr. Thomas Kaiser: Wenn es zum zweiten Unterpunkt keine weiteren Fragen gibt, dann kommen wir zum dritten Unterpunkt.

Hochenergie (HE)-TUMT und Niedrigenergie (NE)-TUMT

Katharina Biester: Da wurde angemerkt, dass wir nicht in Hochenergie- und Niedrigenergie-TUMT eingeteilt haben. Wir haben das nicht gemacht, weil das ein Spezifikum der Prostatrongeräte ist mit der Software 2.0, 2.5 und für die anderen Geräte, die wir in den Studien gefunden haben, das nicht eindeutig operationalisiert ist. Das würden wir gerne mit Ihnen besprechen.

Dr. Richard Berges: Es ist ein großes Problem der gesamten TUMT-Literatur, dass diejenigen, die als Anwender solche Studien publizieren oder auch Meta-Analysen publiziert haben, hier keine saubere Trennung vorgenommen haben. Grundsätzlich gilt aber das Gleiche, was wir vorher gesagt haben. Die TUMT zählt zu den sekundär ablativen Verfahren, und zwar nur die Hochenergie-TUMT, weil nur die Hochenergie-TUMT so viel Energie in die Prostata bringt, dass tatsächlich eine Gewebsnekrose entsteht. Alles, was unterhalb dieser Energieschwelle abläuft, ist ein nichtablatives Verfahren, was generell in einen solchen Vergleich zur TURP, wo wir Gewebe entfernen, nicht zugehört. Das heißt, eigentlich – das haben wir in diesem Bericht gemacht – muss man diese Studien aussortieren, die mit Niedrigenergie-TUMT arbeiten. Das sind tatsächlich alle Prostatronstudien mit der Version 2.0. Zum Verständnis: Das Gerät kam zunächst – auch das ist ja eine historische Entwicklung – mit der Softwareversion 2.0 auf den Markt. Die Software erlaubt nur eine so hohe Temperatur wie bei 2.0. Das geht gerade einmal bis 40, 45 Grad. Das heißt, hier werden Sie keine Gewebsnekrose in der Prostata erzeugen. Dann hat man das erkannt. Man hat gesehen, dass hier allenfalls eine symptomatische Therapie vorliegen kann, aber keine Gewebsablation erfolgen kann.

Man hat dann das Gerät mit einer neuen Softwareversion aufgerüstet, die erlaubt, mehr Energie abzugeben. Dadurch wurde es zur Hochenergie-TUMT. Die anderen Geräte, zum Beispiel Targis, die sonst noch auf den Markt kamen, haben das von vornherein berücksichtigt. Bei denen ist es also immer Hochenergie-TUMT gewesen, allerdings hat man – auch dort historisch bedingt – nicht immer die volle Energie gefahren. Das ist geräteabhängig. Bei manchen Geräten können Sie die Energieeinträge regeln, sodass Sie möglicherweise trotz eines Versuchs der Hochenergie-TUMT nur eine Niedrigenergie-TUMT erzielt haben, wenn Sie nicht genügend Energie eingebracht haben. Deswegen gibt es eine dritte Weiterentwicklung bei den Hochenergie-TUMT-Geräten, und das ist die PLFT-TUMT. Die erlaubt als einziges Gerät, tatsächlich im Gewebe ankommende Temperatur zu messen. Das heißt, man ist nicht mehr blind bezüglich des Energieeintrags, sondern das kann man jetzt während der Behandlung genau monitoren und damit die Mikroenergieleistung selbst oder teilweise auch die Behandlungsdauer adaptieren. Das ist ein wesentlicher Unterschied. Das heißt, Sie haben sozusagen bei dem PLFT-TUMT immer eine Hochenergie-TUMT, also auch immer ein sekundär ablatives Verfahren. Bei den Targis- und Prostatronversionen haben Sie meistens eine Hochenergie-TUMT, solange Sie bei Prostatron 2.5 geblieben sind. Die anderen Studien zählen nicht – das ist Historie –, also die Version 2.0; die müsste herausgerechnet werden. Das ist meistens deswegen, weil der Energieeintrag ganz wesentlich vom Wärmeabfluss abhängig ist. Das ist

genauso wie beim Auto die Wasserkühlung. Der Blutfluss in der Prostata ist so erheblich unterschiedlich von Patient zu Patient, dass, wenn Sie da nicht über eine Messung der Temperatur adaptieren, Sie in vielen Fällen nur eine Niedrigenergie-TUMT hinbekommen und damit das Gesamtergebnis natürlich verschlechtern. So ist das in der Technik zu sehen.

PD Dr. Stefan Lange: O. k. Damit haben Sie uns schon ein bisschen weitergeholfen. Es ist ja ein Problem, dass es auch in der Literatur nicht vernünftig beschrieben ist. Auch andere Autoren, die solche systematischen Arbeiten gemacht haben, wie wir auch, also zum Beispiel die Amerikaner, wurden dadurch veranlasst, zu sagen, verzichten wir auf die Einteilung Niedrigenergie und Hochenergie. Die haben ein anderes Ordnungssystem genommen, nämlich Kühlung und Nichtkühlung. Da bleibt aber nur das PLFT-Gerät übrig. Das ist vielleicht auch nicht der Brüller, wenn man eines gegen den Rest hat. Deswegen haben wir auch darauf verzichtet.

Ich habe noch eine Nachfrage. Sie sagen, die Einteilung müsste sich eigentlich auf die tatsächlich im Gewebe erreichte Temperatur beziehen. Ich weiß nicht, ob es da einen Cut-off gibt, wo man sagt, unter 40 oder 45 Grad ist Niedrigenergie und ab 50 Grad – dann bleibt allerdings ein Graubereich übrig, was auch blöd ist – ist Hochenergie. Das wäre eine Frage.

Eine zweite Frage: Bei dem Prostatron ist das ja auch für uns relativ simpel gewesen, 2.0 und 2.5, wenn es denn so steht. Es gibt ältere Arbeiten von Mitte der 90er-Jahre, aber – davon können wir ausgehen – das werden Niedrigenergie-TUMT-Geräte sein. Jetzt sagen Sie, Targis oder Urowave oder das schwedische System ECP werden meistens Hochenergie sein. Das hängt ein bisschen davon ab, wie die das in der Studie selber gemacht haben.

Dr. Richard Berges: Den Urowave kenne ich selber. Der Urowave ist in sich in den ersten Auslieferungen immer eine Niedrigenergie-TUMT gewesen. Bei den anderen Geräten bin ich mir nicht sicher. Die habe ich nie selbst angewendet. Ich weiß das aber vom Urowave. Der Urowave von Dornier ist dann auch mal aufgerüstet worden in seiner Software. Letztendlich bleibt es immer das Gleiche. Es ist unsicher aufgrund der fehlenden Informationen über den tatsächlich ins Gewebe eingebrachten Hitzebetrag, was dort herausgekommen ist. Das sieht man auch an den Studiendaten. Die sind in ihren Bandbreiten, im Outcome, sehr weit gestreut. Das liegt daran, dass jeder Patient anders ist. Diese Streuung kann man eingrenzen.

Sie fragten noch nach der Temperatur, ab der wirklich eine Ablation erfolgt. Dazu gibt es sehr gute Untersuchungen, die wir auch zitiert haben, wo gezeigt wird, ab wie viel Grad Temperatur in der Prostata es wirklich zu einer verlässlichen Nekrose kommt. Das sind etwa 55 Grad, die Sie dort erzeugen müssen, und das konstant.

PD Dr. Stefan Lange: Ich habe noch eine Nachfrage. Nun haben wir das Problem, dass dieser entscheidende Punkt bedauerlicherweise in den Publikationen bei der Methodenbeschreibung nicht angegeben wird, was ja eigentlich eine sinnvolle Maßnahme wäre. Jetzt wäre es schön, wenn Sie uns helfen könnten, ein wirklich eindeutiges Ordnungskriterium zu erstellen, was wir tatsächlich verwenden können. Sie haben den Urowave genannt. Wir hätten – ein

bisschen extrapoliert, weil da eine relativ hohe urethrale Temperatur erreicht wird, was ja eigentlich nicht so schön ist, aber wenn man das überträgt, dass dann auch entsprechend intraprostatisch höhere Temperaturen erzielt werden – es eher dem Hochenergie-Verfahren zugeordnet. Wir haben das TMx-2000-Gerät. Da wird sogar explizit 50 bis 55 Grad angegeben. Das ist also an der Grenze. Was machen wir damit?

(Dr. Richard Berges: Niedrigenergie!)

– Niedrigenergie. – Sie würden sagen, Hochenergie ist eigentlich nur das PLFT-Gerät und Prostatron 2.5?

Dr. Richard Berges: Und Targis. Das sind die 3 gängigen Hochenergie-TUMT-Apparate, die diese Temperaturen erreichen können, wobei nur bei dem PLFT-Gerät die Erreichbarkeit auch überprüft werden konnte.

Rebecca Deppisch: Ich habe eine Frage zu dem Targis. Sie hatten gesagt, dass das auch eine gewisse historische Entwicklung hat. Jetzt haben wir das Problem, dass wir zum Teil auch ältere Studien, beispielsweise Larson et al. von 1998, mit verwenden, die auch ein Targis-Gerät einsetzen. Wäre das eher dann noch Niedrigenergie? Wann gab es die Wandlung?

Dr. Richard Berges: Wann genau es eine Wandlung gab, kann ich nicht sagen. Vielleicht kann das Herr Muschter noch aus dem Gedächtnis kramen. Das Targis-Gerät ist primär als Hochenergie-TUMT angelegt worden. Das heißt, man wollte Gewebetemperaturen von über 55 Grad erzielen. Man erzielt das auch bei einem Teil der Patienten. Wie groß der Anteil der Patienten ist, kann man mit Sicherheit nicht definieren, weil ein Monitoring der Temperatur in der Prostata fehlt. Aber es ist primär als Hochenergie-TUMT ausgelegt, das Prostatron 2.5 ebenfalls.

Prof. Klaus Höfner: Vielleicht muss man ein bisschen mehr sagen zum Verständnis der Technologie. Es gibt nur ein einziges Gerät, das PLFT-Gerät, das die Temperatur in der Prostata exakt misst. Das heißt, wenn Sie den präzisesten Standard ansetzen, dann müssen Sie sagen, bei allen anderen Geräten ist es eigentlich unklar, wie viel Temperatur in die Prostata appliziert wird. Es ist aber in der Entwicklung so gewesen, man hat die Temperatur zunehmend erhöhen wollen, um mehr Energie in die Prostata hineinzubekommen. Folglich haben wir uns zunehmend von der Niedrigenergie-TUMT auf die Hochenergie-TUMT bewegt.

Die Geräte funktionieren so: Wenn man nicht die Temperatur in der Prostata misst, dann kann man sie natürlich auch nicht monitoren. Das heißt, alle diese Geräte arbeiten mit einem theoretischen Modell der Prostatatemperschwelle. Das funktioniert so: Wir haben eine Temperaturmessung in der Harnröhre und im Mastdarm. Dazwischen liegt die Prostata. Nur diese beiden Messfühler werden gemessen. Alle diese Geräte sind gekühlt, eine Wasserkühlung im Harnröhrenkatheter. Durch die Kühlung wird die Temperatur zunächst heruntergeregelt, die eigentlich sehr hoch sein sollte. Es entsteht dann eine Temperaturspitze in der Prostata selbst.

Die Geräte arbeiten mit einem Alarm. Das heißt, wenn die am Gerät eingestellte Temperatur überschritten wird, sowohl in der Harnröhre als auch im Mastdarm, dann schalten die Geräte ab. Das ist also ein autoregulatives System. Das heißt, wir liegen in der Harnröhre bei 43,5 Grad und im Rektum bei 42,5 Grad. Dann wird abgeschaltet. Das ist bei allen High-Energie-TUMT-Geräten konstant. Welche exakte Temperatur dieses Abschalten in der Prostata erzeugt, hängt von Gewebeeigenschaften, Durchblutung, Größe der Prostata etc. ab. Trotzdem ist die Konstanz, wo wir gesagt haben, das ist Hochenergie-TUMT, im Prinzip die Temperaturschwellen der Geräte, die an diesen Stellen gemessen worden sind. Etwas anderes konnte bei diesen Geräten gar nicht definiert werden.

Bei den anderen Geräten, zum Beispiel bei der Version 2.0, war es so, dass wir in der Harnröhre mit 42,5 Grad gemessen haben. Da ist also eine ganz andere Temperaturschwelle gewesen. Die hat man dann erhöht, weil man gesagt hat, wir können die Temperatur erhöhen, ohne dem Patienten Schaden zuzufügen. Die Temperaturschwellen, das Abschalten der Geräte, ist deswegen eingeführt worden, damit keine Hitzeschäden an anderen Stellen entstehen. Am Anfang hat es schwere Schäden am Mastdarm gegeben, weil dort die Temperatur zu hoch war. Deswegen hat man Temperaturschwellen eingeführt, ein autoregulatives System erzeugt.

Noch einmal: Die Hochenergie-Definition liegt an diesen Temperaturschwellen. Die optimale technische Lösung ist, die Temperatur in der Prostata zu messen. Das ist die ideale Lösung.

PD Dr. Stefan Lange: Es gibt ein Gerät – das haben unsere Recherchen ergeben –, das praktisch ohne Kühlung arbeitet, nämlich das TMx-2000, während alle anderen, wie Sie völlig richtig gesagt haben, mit Kühlung arbeiten. Ich kann das sehr gut nachvollziehen. Letztendlich müsste man schauen, ob sich in der historischen Entwicklung Veränderungen ergeben, was Effekte angeht. Das ist ja vielleicht ein ganz sinnvolles Vorgehen. Man müsste jetzt annehmen, dass die neuen Studien mit den neueren Verfahren stärkere Effekte erzielen, was wir natürlich zu den eingesetzten Vergleichsverfahren bewerten können, als ältere Studien mit älteren Verfahren.

Dr. Richard Berges: Man wird vielleicht nicht stärkere Effekte, sondern in einer Patientenpopulation von 100 Behandlungen kongruentere Effekte erzielen, weil, sobald die Schwelle von 55 Grad erreicht wird, Gewebe zerstört wird. Jetzt ist es praktisch abhängig vom Temperaturabtransport. Blutfluss in der Prostata ist hier die variable Größe, die erheblich ist. Vom Hitzeabtransport ist dann der Effekt abhängig. Das heißt, Sie werden, wenn Sie 100 Hochenergie-TUMT-Behandlungen machen, ein paar Patienten haben, die sehr gut werden. Dann haben Sie ein Mittelfeld und dann haben Sie ein paar, bei denen das überhaupt nicht funktioniert, und zwar genau die, wo Sie die erwünschte Temperatur eben nicht erreicht haben. So wird sich das fortsetzen. Sie haben also letztlich eine geringe Streuung.

PD Dr. Stefan Lange: Das war genau meine Frage. Sie würden also sagen, der Effekt an sich – absolut gesehen – verändert sich nicht wesentlich, also im Mittel Reduktionen, wenn

wir IPSS nehmen, um 10, 12 Scorepunkte. Aber die Streuung wird geringer. Das heißt, der Effekt bezogen auf die Streuung wird größer, weil die Variabilität nicht mehr so weit ist.

Dr. Richard Berges: Für die PLFT-TUMT. Leider im Moment nur da.

Prof. Klaus Höfner: Eine kurze Ergänzung: Man muss dabei immer feststellen, dass es nicht hundertprozentig gesichert ist, dass wir bei einem bestimmten Temperaturlevel, zum Beispiel 55 bis 60 Grad, unterschiedliche subjektive Ergebnisse bekommen im Sinne der subjektiven Patientenverbesserung, der Symptomatik. Das ist nicht belegt. Wir wissen nur, dass die Hochenergie-TUMT besser ist als die Niedrigenergie-TUMT. Wir wissen aber nicht, wie viel Energie wir brauchen, damit ein optimales subjektives Ergebnis entsteht. Diese genaue Schärfe, die Sie verlangen, Herr Lange, werden Sie in der Studie nicht finden. Wir wissen nur global, eine hohe Energie mit einem Optimum bei der PLFT, wenn man das als Optimum bezeichnet, ist sicher nicht das Optimum in Bezug auf die Verbesserung der Symptomatik, wobei das noch lange nicht heißt, dass auch ein anderes Hochenergie-TUMT-Gerät das nicht auch im Einzelfall kann.

Moderator Dr. Thomas Kaiser: Aber das ändert ja nichts daran, dass bei den Geräten, wo es zunächst einmal durch die technischen Voraussetzungen möglich und wahrscheinlich ist, dass ein solcher Erfolg erreicht wird, im Mittel das subjektive Ergebnis als besser zu erwarten ist, wenn es denn tatsächlich durch die Behandlung dazu kommen sollte, als bei den Niedrigenergiegeräten.

Prof. Klaus Höfner: Exakt. Das ist eine klare Botschaft.

Dr. Richard Berges: Es ist ebenfalls zum Beispiel für die TURP völlig ungeklärt, wie viel Gewebe wir entfernen müssen, um einen entsprechenden symptomatischen Effekt zu erzielen. Dort gibt es ganz unterschiedliche Ansichten. Die amerikanische Welt, der ganze angloamerikanische Raum reseziert nur sehr spärlich – das sehen Sie an den Resektionsgewichten in den publizierten Studien –, während in Deutschland sehr viele Jahre sehr ausgiebig reseziert wurde, also sehr viel höhere Resektionsgewichte. Es gibt aber keine Untersuchung, die wirklich sagt, das ist notwendig. Das gilt nicht nur für die TUMT, sondern auch für die TUNA und die Laserverfahren. Wir wissen nicht, wie viel wir entfernen müssen, um eine optimale Symptomreduktion zu bekommen.

PD Dr. Stefan Lange: Ich entnehme Ihrer Äußerung, da fehlen uns Untersuchungen, um das genau festzustellen, also auch für die TURP. Man könnte ja so einen Vergleich mal machen, also auf der einen Seite wie man es üblicherweise in Deutschland macht, nämlich viel Gewebe entfernen, und auf der anderen Seite, wie es die Amerikaner machen, nämlich weniger entfernen. Das wäre doch toll.

Dr. Richard Berges: Das wäre sicherlich eine Untersuchung wert, wobei es Ansätze solcher Untersuchungen gibt. Die Untersuchungen, die dazu gemacht wurden, haben keine Unter-

schiede gezeigt, was die Symptomverbesserung betrifft. Das ist auch der Grund, warum wir davon abgekommen sind, in Deutschland immer eine vollständige Resektion zu verlangen. Sie brauchen diesen irren Gewebeabtrag, den wir in den 70er- und 80er-Jahren noch gemacht haben, nicht. Das ist klinisch irrelevant. Sie müssen halt so viel Gewebe entfernen, dass Sie nachher einen Effekt erzielt haben. Dann ist es wahrscheinlich so, dass sich erst in der Langzeitfolge – wahrscheinlich ist das weit über 5 Jahre hinaus, eher 10 Jahre – Unterschiede ergeben, denn gutartiges Gewebe kann nachwachsen, und wenn Sie wenig reseziieren, dann haben Sie vielleicht in 10 Jahren das Gleiche noch einmal zu tun.

Prof. Dr. Rolf Muschter: Eine derartige Studie ist vor Jahren konzipiert und von der zuständigen Ethikkommission in Deutschland nicht genehmigt worden. Man kann natürlich hilfsweise fragen, wo denn immer vollständig entfernt wird. Das ist bei der offenen Enukleation. Man kann das randomisiert mit der Transurethralen Resektion vergleichen. Hierzu gibt die Literatur Studien her. Solche randomisierten Vergleiche gibt es. Die sind vom Ergebnis her mit geringen Unterschieden zugunsten der offenen Enukleation verlaufen. Aber das ist eine hilfsweise Abweichung, weil man natürlich hier Äpfel mit Birnen dahingehend vergleicht, dass man offen randomisiert das Gewicht begrenzt. Man geht also aus der eigentlichen Indikation heraus. Denn in dem randomisierten Vergleich kann ich nicht 150-Gramm-Drüsen nehmen, die ich transurethral nicht reseziieren kann. Das heißt, diese Studien haben alle den Nachteil, dass die beiden Verfahren eigentlich unterschiedliche Indikationsbereiche haben.

Dr. Volker Rohde: Urologisch ist hinzuzufügen, es ist nicht nur das Volumen, sondern auch die Lokalisation. Da gibt es die berühmte Lokalisation des parakolikulären Anteils, der abgetragen werden muss. Da gibt es sicherlich auch ein paar technische Besonderheiten der Resektionstechnik. Das ist eben nicht nur die Abtragung des Volumens des Gewebes. Aber es würde den Rahmen sprengen, das jetzt im Detail zu diskutieren.

PD Dr. Stefan Lange: Vielleicht kommen wir darauf unter TOP 3 zu sprechen, denn ich glaube, dass es ein wichtiger Punkt ist. Natürlich hat es auch bei der TURP eine Entwicklung über die letzten 10 bis 20 Jahre gegeben, die man natürlich bei entsprechenden Vergleichen mit berücksichtigen muss, was zum Beispiel die Invasivität des Verfahrens angeht.

Moderator Dr. Thomas Kaiser: Wir gehen nun zu dem Punkt 3 über. Ich weiß nicht, ob es zu TOP 2 Unterpunkt 3 – Hochenergie, Niedrigenergie – noch Nachfragen gibt, was die Einteilung angeht. Wir haben festgestellt, dass es Geräte gibt, bei denen man feststellen kann, ob man das erreicht hat, was man erreichen wollte. Es gibt Geräte, die haben die theoretische Möglichkeit, das zu erreichen, ohne wirkliche Kontrollmöglichkeit. Es gibt aber auch Geräte, die von ihrer technischen Voraussetzung her nicht geeignet sind, tatsächlich ablativ zu wirken. Ist das so eine grobe Dreiteilung?

(Zuruf: Ja, richtig!)

Dann gehen wir jetzt über zum

Tagesordnungspunkt 3:

Modifikationen der TURP

– Was wird als Modifikation der TURP betrachtet?

Prof. Klaus Höfner: Wir kommen jetzt langsam in die Richtung, wo es in die pathophysiologischen Grundlagen geht. Ich möchte das nicht vertiefen, denn darüber könnten wir den ganzen Tag reden. Es gibt aber ein Faktum, das wir sicherlich bedenken sollten, auch im Umgang mit diesen Studien. Wir haben ein objektives Kriterium der Prostatahyperplasie, und zwar ist das die Obstruktion, was eigentlich die Prostata an obstruktiver Potenz bildet, also der eigentliche Krankheitswert unabhängig von dem, was der Patient selbst spürt. Dafür wissen Sie ja genau, dass die Obstruktion, wenn der Ductus ejaculatorius⁸ eingeengt ist, dort eine Widerstandserhöhung entsteht. Das korreliert aber leider in keiner Weise mit der subjektiven Symptomatik. Es gibt keinen Zusammenhang. Jetzt haben wir das Problem, dass wir ganz klar dann über Ablation reden müssen, wenn ein Patient eine Obstruktion hat. Dann ist das Faktum Ablation eigentlich zu nennen. Anders ausgedrückt: Wenn keine Obstruktion da ist – etwa 50 % der BPH-Populationen haben keine Obstruktion –, muss man eigentlich nicht abladieren, sondern kann über andere Effekte, beispielsweise Hitze, Nerven, einen solchen Effekt erreichen. Das heißt, das, was Sie angesprochen haben, Herr Kaiser, zeigt das ganze Dilemma im Verständnis. Das heißt, wenn wir ein Kriterium nehmen und sagen, wir wollen jetzt nur nach der Symptomatik gucken, können wir das mit der Ablation, mit einer der Ablation immanenten Morbidität erreichen und wir können es ohne Ablation erreichen bei einem nicht obstruktiven Patienten. Ich sage das nur zum Verständnis. Da gibt es sehr wenige Hinweise in den Studien. Das sind alles nicht randomisierte Studien, wo wir nachgucken, was diese Verfahren eigentlich an der Obstruktion tun. Das Maximum ist die Ablation, die Enukleation. Wir wissen, das ist das Maximum einer Ablation, auch mit der höchsten Morbidität. Deswegen haben diese Verfahren aufgrund einer maximalen Ablation auch die größte deobstruierende Wirkung. Wenn ich ein nicht ablatives Verfahren nehme, zum Beispiel Medikamente, die nichts abladieren, die subjektiv durchaus funktionieren, und dann eine entsprechende Indikation - -⁹ Dieses Problem können Sie nur sehr vage beantworten. Aber es bedeutet, wenn ich nur nach der subjektiven Symptomatik des Patienten gucke und dabei die Morbidität des Verfahrens ausklammere, dann habe ich immer ein Problem, weil ich sagen muss: Ich kann das mit dem einen wie mit dem anderen Verfahren gleichermaßen erreichen. Das wollte ich in die Diskussion bringen.

Jetzt komme ich zur Modifikation der TURP. Die Modifikation der TURP - da kommen wir im Prinzip in die Frage der Ablation - -⁹ Die Ablation abladiert. Deswegen können nur alle Verfahren, die auch Ablation leisten, eine Modifikation der TURP sein, wenn überhaupt. Ich

⁸ In der Aufzeichnung nicht deutlich zu verstehen.

⁹ Abbrechen des Satzes

weiß nicht, ob ich mich einigermaßen vernünftig ausgedrückt habe. Das Problem ist, Sie sagen, wir können das Gleiche erreichen. Sie reden immer über Symptome. Wir reden dann über die Symptomatik und Lebensqualität. Das ist der Punkt, den wir diskutieren. Da können wir halt mit sehr vielen verschiedenen Verfahren das Gleiche erreichen. Der einzige Unterschied ist, welches Maß an Morbidität wir einsetzen, um dieses zu erreichen.

Moderator Dr. Thomas Kaiser: Das ist natürlich die Kehrseite der Medaille.

Prof. Klaus Höfner: Das ist das, was wir diskutieren. Wir kommen ja hoffentlich darauf zurück, dass wir dann über die medizinischen Nebenwirkungen reden. Ich habe Ihre Bemerkung aufgenommen, Herr Kaiser, dass Sie gesagt haben, wir können es mit sehr viel verschiedenen Dingen tun, wenig Hitze, viel Hitze, gar keine Hitze.

Moderator Dr. Thomas Kaiser: Das bezog sich nur auf die Einteilung dieser Hoch- und Niedrigenergie. Es hatte nicht das gesamte Gebiet eingeteilt.

Was sind denn die konkreten Punkte und Fragen der Projektgruppe?

Katharina Biester: Das sind eigentlich die Punkte, die bereits gerade angesprochen wurden: Was wird als Modifikation der TURP gesehen und wo ist die Trennlinie? Wir hatten ja in TUVRP und TVP eingeteilt, also nicht nur wir. In den Stellungnahmen kam dann der Hinweis, dass auch die Laservaporisation in der logischen Folge als Modifikation gesehen werden müsse.

Dr. Richard Berges: Das ist das Dilemma, was wir vorhin mit der größeren Einteilung versuchten, darzustellen. Wenn Sie primär abladieren, also ein primär ablatives Verfahren einsetzen, dann ist es, außer dass Sie eine offene Operation machen, die technisch anders durchgeführt ist, eigentlich egal, welches Verfahren Sie einsetzen. Es ist egal, ob Sie mit Strom oder Laser eine Vaporisation erzeugen, eine Bandschlinge oder eine Elektroschlinge nehmen, die ganzen Modifikationen nehmen, die die ureigene TURP in sich erfahren hat, koagulieren, das Schneiden und was es sonst noch alles gibt. Letztendlich tun Sie immer das Gleiche: Sie entfernen transurethral Gewebe. Wenn man den Zulassungstext sozusagen genau interpretiert, dann muss man eigentlich sagen, dass alle diese Verfahren nie einer Zulassung oder einer Bewertung bedurften, weil sie letztlich nur Modifikationen der TURP sind. Man schneidet halt Gewebe, statt mit einer Schlinge, mit einer Bandschlinge, oder man verdampft es. Man macht immer das Gleiche, es wird immer transurethral Gewebe entfernt. Letztendlich bedarf diese ganze Sitzung keiner Erörterung, weil es eigentlich alles Modifikationen der gleichen Technik sind, die keiner Zulassung bedürfen. Nun gibt es natürlich Verfahrensunterschiede, nämlich die verwendeten Techniken, aber das Ende ist immer das Gleiche.

PD Dr. Stefan Lange: Das konnte ich der Diskussion bis jetzt noch nicht so ganz entnehmen. Wir haben ja gerade festgestellt, zum Beispiel die TUMT, ältere Verfahren sind ja ablativ, Transurethrale Nadelablation weiß ich nicht so genau. Der eigentliche Punkt ist ja folgender:

Wir sind in dem Bericht davon ausgegangen, wie auch andere, es gibt so etwas wie ein Standardverfahren. Wir haben das „Goldstandard“ genannt. Man kann es „Referenzstandard“ nennen. Das ist eigentlich die TURP. Zu dieser TURP hat es jetzt Weiterentwicklungen gegeben, bzw. nicht zu TURP, sondern zu der möglichst schonenderen Methode, Prostatagewebe oder die Ursache des Krankheitsbildes zu entfernen. Daraus ergibt sich für uns die Notwendigkeit, die neueren Verfahren gegenüber dem, was schon da ist, in ihrem Stellenwert zu beurteilen. Wir hätten natürlich sagen können, wir betrachten nur Arbeiten, die die neueren Verfahren gegenüber der TURP vergleichen. Das wäre eine Möglichkeit. Nach Sichtung der Literatur ergab sich für uns der Eindruck, dass eben auch von der Schwere des Eingriffs her die Elektrovaporisationsverfahren im Prinzip einen ähnlichen Stellenwert haben wie die TURP, das heißt, eigentlich eine Weiterentwicklung des Standards sind, aber eben auf einem anderen Prinzip. Das wäre der Punkt, den wir versuchen müssten, herauszuarbeiten. Das mit der Ablation und Nichtablation hilft mir noch nicht so wahnsinnig viel bei der Problematik.

Prof. Klaus Höfner: Den Fehler, den wir kritisiert haben – es geht ja um die Stellungnahme –, ist folgender: Sie haben in die Transurethrale Resektion die Vaporesektion hineingenommen. Das ist ein Verfahren, das auch mit einer Schlinge arbeitet, aber beim Schneiden gleichzeitig vaporisiert. Das kann man, denke ich, tolerieren. Sie sagen, wir entfernen Gewebe mit einer elektrischen Schlinge, die nur einen anderen Strom und eine andere Stromqualität hat, aber es wird Gewebe entfernt. Das heißt, Sie haben hinterher Gewebe zur Verfügung, das sie zum Pathologen schicken können. Das ist also ein sehr ähnliches Verfahren. Die Elektrovaporisation ist eine Vaporisation, die sich nur von der Energiequelle von der Laservaporisation oder dem Kontaktlaser unterscheidet. Auf jeden Fall wird Gewebe verdampft und nicht reseziert. Der Effekt der Ablation ist der gleiche. Das heißt, Sie erhalten eine Ablation, indem Sie Gewebe resezieren oder enukleieren mit der E nukleation. Auf jeden Fall wird immer Gewebe entfernt. Man hat es dann zur Verfügung. Das ist die eine Sache. Das sind also die E nukleation, die transurethrale Vaporesektion und die transurethrale Resektion. Aber Sie haben von den reinen vaporisierenden Techniken nur die Elektrovaporisation zum Standard gezählt. Deswegen haben wir gesagt, wenn Sie das tun, dann müssen Sie alle anderen vaporisierenden Verfahren wie Laser auch dort hineinpacken, denn die machen nichts anderes, nur die Energiequelle ist unterschiedlich. Ob ich das mit elektrischem Strom mache oder Laserenergie dazunehme, der Vaporisationseffekt ist der Gleiche. Das heißt, wir haben keine Gewebeentfernung, sondern eine Verdampfung von Gewebe. Auch die Vaporesektion ist ein bisschen etwas Ähnliches, weil auch der Schnitt, das heißt das Durchziehen der Schlinge durch die Prostata, vaporisiert. Das heißt, man sagt – ich weiß nicht genau –, ein Drittel, ein Viertel des Gewebes wird verdampft, aber trotzdem wird reseziert. Man kann also sagen, das geht in die Richtung der Resektion. Es ist nicht genau das Gleiche, aber darauf kann man sich einigen. Alles andere ist halt komplett different.

Prof. Dr. Rolf Muschter: Ich wollte kurz den Hintergrund darlegen, der zu dieser Entwicklung geführt hat. Wir haben den Standard Transurethrale Resektion seit den 30er-, 40er- Jahren in Amerika, seit den 60er-Jahren in Deutschland übernommen. Natürlich ist das eine

permanente Entwicklung gewesen, die dazu führt, dass man eigentlich das Wort „Standard“ kaum benutzen kann. Wenn Sie ein ganz einfaches Lehrbuch nehmen, dann finden Sie 4 oder 5 verschiedene Techniken, wie man das machen kann. Dieses Verfahren war mit einer gewissen Grundmorbidity verbunden, die sich auch in den gesamten Entwicklungen bis zu den 90er-Jahren nicht wesentlich bessern konnte. Wir hatten eine Untergruppe von Patienten, und zwar gar nicht wenige, die eine fortgeschrittene Symptomatik hatten bis hin zur Harnverhaltung, schwerst multimorbide waren und wo die Anästhesisten gesagt haben, diesen relativ morbiden Eingriff, Transurethrale Resektion, lassen wir nicht zu. Diese Patienten wurden mit einer Dauerharnableitung versorgt und vegetierten dann in ihren Pflegeheimen mit einem Dauerkatheter vor sich hin, erzeugen viele Kosten, müssen immer wieder gewechselt werden. Es ist einfach eine ganz schlechte Lebensqualität für die Patienten.

Das war der Ausgangspunkt, sodass man sich gefragt hat, wie man den gleichen Effekt mit einer geringeren Morbidity erreichen kann. Das wurde mit diesen verschiedenen Technologien versucht. Nun kommt natürlich die unglückliche Entwicklung. Wenn ich eine Hochtechnologie einführe, ob das Mikrowellen oder Laser sind, dann kann ich das nur tun, indem ich Geld investiere, und zwar relativ viel. Das muss ich ja irgendwo wieder herausbekommen. Ich argumentiere jetzt einmal, obwohl ich keine Firma repräsentiere, den Standpunkt der Firma. Dann muss ich auch etwas dafür tun, mein Verfahren möglichst von anderen, schon existierenden Verfahren abzugrenzen und zu zeigen, was bei meinem Verfahren besser ist. So erklärt sich letztlich die Vielfalt der unterschiedlichen Techniken und der unterschiedlichen Hersteller, die aber im Grunde genommen, wenn man das Ganze weniger analytisch, sondern mehr synthetisch betrachtet, sehr viel mehr Gemeinsamkeiten als Unterschiede aufweisen. Diesem trägt natürlich die wissenschaftliche Literatur nur bedingt Rechnung, weil jeder Anwender, der jetzt eine Maschine in seine Klinik bekam und diese ausprobieren konnte – ich nenne das so, wie das wirklich zum Teil gelaufen ist – und dann eine Anwendungsstudie publiziert, dieses Verfahren natürlich als einzigartig beschreibt. Er hat auch selber das Gefühl, es ist ein einzigartiges Verfahren. Das ist die Erklärung für diese Vielfalt. Dass wir sagen, wie gerade Herr Berges, im Grunde genommen ist das alles eine Modifikation der TURP, ist das Ergebnis, das klinisch tatsächlich erzielt wird. Ihnen bleibt tatsächlich nur die Wahl: Entweder Sie definieren die Transurethrale Resektion, indem Sie sagen, da wird elektrischer Strom oder Hochfrequenzenergie eingesetzt, um Gewebe zu exzidieren, herauszuschneiden, zu reseziieren, zu enukleieren – wie man das nennt, ist nachher unerheblich; vielleicht ist sogar das erheblich, aber lassen wir es erst einmal dabei –, oder ob Sie es vaporisieren. Wenn Sie dort einen Unterschied sehen, dann dürfen Sie die Elektrovaporisation nicht als Modifikation der transurethralen Resektion verstehen. Wenn Sie aber sagen, das Endergebnis ist für mich wichtig, nämlich dass Gewebe entfernt worden ist, und zwar unmittelbar und sofort, dann muss ich zwangsläufig auch andere Energieformen zulassen. Denn dass sich eine Laservaporisation von einer Elektrovaporisation unterscheidet, ist in keiner experimentellen Studie gezeigt worden. Im Gegenteil: Es gibt experimentelle Studien, die die Einheitlichkeit dieser Effekte zeigen. Die sind natürlich alle nicht randomisiert. Das sind experimentelle, tierexperimentelle

Studien, Laboruntersuchungen. Das ist also nicht Level 1. Das ist die Grundproblematik, die wir haben.

Jetzt gehe ich einmal ans andere Ende und sage, es gibt ja ein Verfahren namens Lasere-
nukleation. Hier wird ein anderes Werkzeug und eine andere Art von Technik genutzt, die
Prostata zu entfernen, als bei einer Transurethralen Resektion. Deswegen muss man das ir-
gendwie als eigenständiges Verfahren begreifen. Aber vom Ergebnis her entspricht das, wie
der Name sagt, letztlich der offenen Enukleation. In einer wissenschaftlichen Bewertung ist es
jetzt sehr schwierig, diese Techniken als unterschiedlich anzusehen bei gleichem biophysika-
lischen Effekt, indem man eine Studie nimmt, wo durch die Patientenselektion, durch die
technische Ausführung große Unterschiede oder weniger große Unterschiede herauskommen.
Sie haben selber in manchen Studien die Heterogenität der Ergebnisse kritisiert. Das ist völlig
richtig. Das ist eben so heterogen, weil hier solche Dinge zusammengefasst werden, wo die
Autoren versucht haben, Spezialgruppen für sich darzustellen. Das ist das Problem bei dieser
ganzen Technologie. Entweder man geht auf das Werkzeug und sagt, Elektrizität ist etwas
anderes als Laser, oder man guckt nach dem tatsächlichen Effekt. Dann muss aber Resektion
auch Resektion sein, egal ob ich das mit diesem oder jenem mache.

Moderator Dr. Thomas Kaiser: Eine kurze Zwischenbemerkung: Das könnte man ja auch
so interpretieren, dass man sagt – hier werde ich sicherlich Gegenäußerungen bekommen –:
Wozu brauche ich dann all diese ganzen neuen Sachen und technischen Neuerungen, wenn
das sowieso alles dasselbe ist und dasselbe Ziel erreicht.

Prof. Klaus Höfner: Darauf kommen wir noch zu sprechen.

Moderator Dr. Thomas Kaiser: Irgendeinen Vorteil müssen sie ja haben, und diesen muss
man auch evaluieren und darstellen können. Das ist ja die Frage dieses Berichts: Welchen
Vorteil haben sie, und wie bewertet man es? Gegen welche andere Alternative stellen Sie es
in welchen Studien dar?

Dr. Richard Berges: Man kann das Ganze auch von einer anderen Seite aufzäumen. Es ist
für die einzelnen Therapieverfahren sehr schwierig, über die wir hier sprechen, bestimmte
Patientengruppen zu identifizieren, die davon besonders profitieren würden. Der einzige Un-
terschied – damit kommen wir zu der Frage, wo der eigentliche Wert dieser Verfahren liegt –
liegt in der unterschiedlichen potenziellen Morbidität der einzelnen Verfahren. Das klinische
Outcome ist für die Patienten von der Erfahrbarkeit her für alle Verfahren gleich, solange Sie
Gewebe abladieren. Die klinischen Verläufe dazu sind unterschiedlich, wenn man von primä-
rer oder sekundärer Ablation ausgeht, und das Outcome ist am Ende relativ gleich. Unter-
schiede sind ganz eindeutig zu finden in der Morbidität und der Invasivität der einzelnen
Verfahren. Aus diesem Grund sind sie auch entwickelt worden.

Das große Problem ist, dass die meisten dieser Verfahren anwenderabhängig sind. Das heißt,
Sie haben eine Lernkurve, die teilweise schmerzlich verläuft, nämlich für den Patienten, der

beim Anfänger landet. Sie leiden immer anwenderabhängig auch in der Folge, auch wenn man diese Techniken erlernt hat. Es gibt eben gute und schlechte Operateure – das ist in der Chirurgie generell so – und es gibt viele dazwischen. Von daher hat man aus diesem Grunde versucht, diese Anwenderabhängigkeit zurückzuschrauben. Das ist in Teilen gelungen, in Teilen auch nicht. Da muss man eigentlich die Unterschiede suchen. Man muss sie nicht in der Symptomverbesserung suchen. Das ist für den Patienten nicht erfahrbar. Ob er eine TUR, eine Laservaporisation, eine Laserenukleation oder eine Elektrovaporisation bekommt, das ist für den Patienten im Outcome nicht erfahrbar. Erfahrbar ist aber, ob er weniger Blutkonserven in der Wahrscheinlichkeit bekommen muss, ob er kürzer im Krankenhaus bleiben muss, ob er inkontinent, impotent etc. wird, und das alles in Unterschieden, wie schnell er wieder zu Hause ist, wie lange der Heilungsverlauf ist, über den wir überhaupt nichts wissen. Auch die große Untersuchung in Deutschland, die eine Vorreiterrolle hat, nämlich die BQS-Studie mit 34 000 TUR-behandelten Patienten, hört an der Kliniktür auf. Sie wissen danach nichts mehr. Die mittlere Verweildauer war, glaube ich, 6 Tage. Was sich in der Praxis vollzieht, ist ein ganz anderer Schuh, auch bei der TURP. Da muss man suchen. An der Stelle muss man eigentlich schauen, ob diese Verfahren eine Berechtigung haben. Und deswegen gibt es die.

Stefan Zenk: Ich möchte Stellung nehmen zu Ihrer Aussage, Dr. Kaiser. Die Frage kommt ja auf, wo die Notwendigkeit betrachtet wird. Im publizierten Berichtsplan steht ja, es gibt eine Nutzen- und eine Notwendigkeitsbewertung. An welcher Stelle sprechen wir über die Notwendigkeitsbewertung?

PD Dr. Stefan Lange: Das ergibt sich letztlich aus dem Nutzen. Das kann ja nur dann beantwortet werden, wenn überhaupt ein Nutzen festgestellt wird bei Fragen der Notwendigkeit. Die Bewertung der medizinischen Notwendigkeit – hierauf möchte ich nur kurz eingehen, da sie nicht Gegenstand der Stellungnahmen und der Tagesordnung ist – ist problematisch, weil sie auch noch Werte mit einschließt. Man kann gewisse Definitionen aus dem Sozialgesetzbuch ableiten, aber das ist ein bisher nicht ganz scharf definierter Begriff, dem wir uns nichtsdestotrotz in dem Bericht zu nähern versuchen. Da wird ein Aspekt, der uns helfen wird, auch das Ergebnis der Erörterung hier bzw. auch gegebenenfalls die Stellungnahme zum modifizierten Vorbericht sein, genau diesen Notwendigkeitsaspekt herauszuarbeiten.

Aber ich würde jetzt gerne auf die Historie und auf das zurückkommen, was Herr Berges gesagt hat. Letztlich ist es, wie es Herr Kaiser bereits gesagt hat, Gegenstand des Berichts gewesen. Das hatten wir vor Augen. Natürlich ist es ein Ziel der neueren Verfahren gewesen, die Morbidität zu reduzieren. Jetzt sagen Sie so nonchalant, man sieht auch im Outcome keinen Unterschied. Das wäre ja genau das, was zu prüfen ist. Das heißt, wenn man einer neuen Methode eine alte Methode gegenüberstellt, sie einführen will in die Versorgung, einen bestimmten Vorteil postuliert, dann gehört erstens dazu, dass man diesen Vorteil auch nachweist, und zweitens, dass bestimmte Nachteile wie auch immer ausgeschlossen werden. Hier konkret: Eigentlich bräuchte man Untersuchungen adäquater Art, die wissenschaftlich belegen, dass die Morbidität reduziert wird, so es denn nicht sofort einsichtig ist, zum Beispiel weil es kein

TUR-Syndrom geben kann bei der TUMT, aber was sexuelle Funktionsstörungen und Mikti-
onsproblematik angeht, die obstruktive Problematik usw., also alles das, was wir als uner-
wünschte Ereignisse oder Komplikationen nennen. Das wäre das eine, was man bräuchte,
nämlich einen Beleg für einen Vorteil. Auf der anderen Seite muss aber auch nachgewiesen
sein, dass das, was den Patienten eigentlich interessiert und weshalb er auch zum Urologen
geht, weil er Beschwerden hat, durch das neuere Verfahren günstig beeinflusst wird. Man
muss aufpassen, dass man nicht schon das voraussetzt und sich damit der Überprüfung ent-
zieht, denn dann wäre es natürlich relativ einfach.

Ich gebe zu bedenken, was wir vorhin auch festgestellt haben, es hat ja nicht nur bei den loka-
len minimalinvasiven Verfahren eine Entwicklung gegeben – das haben wir bei der TUMT
bis hin zu Temperatursteuerungsgeräten innerhalb der Prostata festgestellt –, sondern auch bei
der TURP. Das heißt, hier sind auch schonendere Techniken entwickelt worden.

Letzter Punkt: Wir können bei der Bewertung von nichtmedikamentösen und medikamentö-
sen Verfahren die Heterogenität der Anwendung nicht aus der Bewertung ausschließen. Dass
es Anwender gibt, die das mehr oder weniger tun, muss in die Bewertung mit hinein. Das
wird sich gegebenenfalls in Studien auch abbilden, wenn sie einigermaßen groß sind. Es ist
genauso, dass wir bei einem medikamentösen Verfahren auch mit abbilden müssen, dass die
Patienten das Medikament nicht so, wie vorgeschrieben, morgens um 8 und abends um 8 ein-
nehmen, sondern die einen nehmen nur die Hälfte und dann erst um 10 und die anderen neh-
men direkt 3 Pillen. Das gehört zu einem Verfahren dazu. Genauso gehört zu einem
chirurgischen Verfahren natürlich die Variabilität der Qualität der Anwendung des Verfahrens
dazu. Das darf man nicht ganz als Exklusivität des chirurgischen Verfahrens ansehen, dass
man die deshalb vielleicht nicht bewerten könnte, weil man erst mal so viel lernen muss.

Dr. Richard Berges: Ich gehe mit Ihnen vollkommen d'accord, dass man eigentlich all diese
Fragen, die in den Raum gestellt sind, prospektiv, randomisiert, wissenschaftlich untersuchen
könnte. Einige Sachen sind aber so evident, dass es eigentlich schon fast schildbürgermäßig
wäre, zu sagen, das wollen wir prospektiv untersuchen. Zum Beispiel: Sie machen eine PLFT-
TUMT und brauchen dazu keine Narkose. Der Patient geht nach einer halben Stunde Behand-
lung nach Hause und hat nach 3 Monaten den gleichen symptomatischen Qmax oder de-
obstruktiven Outcome wie nach einer TURP, wo er 5 Tage oder 10 Tage im Krankenhaus
gelegen hat. Letztlich kann man alle solche Sachen schön zeigen, aber die Frage ist, wieweit
man das muss. Oder: Wer als Operateur eine Resektion durchführt, eine ganz normale Stan-
dard-TURP, der weiß sehr genau, wie viel das bluten kann, wie sehr man hinterhersein muss,
dort zu koagulieren. Daran scheitern ja auch viele. Wenn man dann eine Laservaporisation
oder eine Elektrovaporisation durchführt und die Spülflüssigkeit die ganze Zeit klar bleibt,
dann brauchen Sie nicht zu zeigen, das ist ein Unterschied. Das ist so evident, dass das jeder
Operateur sofort sieht. Der Patient geht dann auch ganz anders aus solch einer Operation her-
aus. Natürlich können Sie sagen, das müssen wir alles in großen Untersuchungen prüfen.
Dann haben Sie aber etwas vor, was nachher in seinem Nutzen relativ gering ist, weil Sie das

sofort sehen. Das ist die Diskrepanz, die da herrscht. Anders ist das bei einer Medikamentenstudie. Da sehe ich einen ganz anderen Zusammenhang.

PD Dr. Stefan Lange: Ich möchte kurz nachhaken. Das haben wir ja auch beschrieben, vielleicht nicht in dem Detail mit dem PLFT. Bestimmte Sachen kann man ambulant durchführen und auch ohne Narkose, und wir haben die Krankenhausaufenthaltsdauer bei der TURP. Das braucht man nicht zu prüfen. Dafür brauchen wir keine Studie. Das Ergebnis steht fest. Das haben wir aber auch in anderen Bereichen. Es ist eben das typische Problem der Vergleichbarkeit von weniger invasiven zu invasiven Verfahren, also Hernienchirurgie, Cholezystektomie, Appentektomie. Oft ist es ja so, dass man sich dann durch das weniger invasive Verfahren andere Probleme einkauft, übrigens auch Komplikationsarten. Vielleicht ist das hier nicht der Fall, aber könnte ja sein. Das heißt, darauf muss man gucken. Das ist ein Aspekt. Man muss also sauber trennen zwischen dem, was Sie gesagt haben, was evident, was augenfällig ist, was sich eigentlich einer Überprüfung entzieht, weil das Ergebnis feststeht, und dem, was zwar sehr wahrscheinlich scheint, aber doch nicht so sicher, dass man es dennoch prüfen könnte und sollte. Das wären schwerwiegendere Komplikationen, auch Blutungskomplikationen, die vielleicht doch bei dem einen oder anderen Verfahren gesehen werden. Dann braucht man auch keine großen Studien, um das nachzuweisen, wenn der Effekt zu groß ist. Aber dann hat man es einmal in 2 oder 3 Studien sauber dokumentiert.

Dann wäre noch folgende Frage: Wenn das so ist, wie Sie sagen, dass das so augenfällig ist, dann frage ich mich, warum wir in Deutschland 30 000 TURP haben? Das ist ja eigentlich skandalös.

(Prof. Klaus Höfner: 60 000!)

– 60 000. – Wie kann man die Menschen mit diesem irrsinnigen Verfahren mit dieser wahnsinnigen Komplikationsmöglichkeit behandeln – ich sage es etwas überspitzt –, wenn doch eigentlich diese herrlichen Verfahren zur Verfügung stehen, die für die Patienten ohne größere Probleme das gleiche Ergebnis erzielen lassen? Da spüre ich einen gewissen Widerspruch.

Prof. Klaus Höfner: Wir sind jetzt an einem ganz entscheidenden Punkt der Diskussion, den wir als sehr wesentlich erachten. Wir haben jetzt das Problem, das ich so aus der Agenda nicht erkennen kann. Deswegen müssen wir uns darüber einigen, ob wir jetzt darüber reden, denn zur Relation von Nutzen und Risiko gibt es sehr viel zu sagen.

Um auf Ihre Frage zurückzukommen, Herr Kaiser, was eigentlich der Vorteil ist – das ist ja eigentlich das, worüber wir diskutieren –: Wenn wir darüber jetzt diskutieren, dann können wir das jetzt tun. Dann können wir alle Punkte abhaken, die aus meiner Sicht angesprochen werden müssen. Man kann jetzt aus meiner Sicht nicht so richtig erkennen, wo wir an diesem Punkt offiziell gelandet sind. Deswegen spreche ich das noch einmal an.

Moderator Dr. Thomas Kaiser: Völlig richtig. Ich meine, dass wir mit dem TOP 3 eigentlich durch sind. Ich weiß nicht, ob es dazu noch konkrete Äußerungen gibt. Nachfragen gibt es, glaube ich, nicht. Das Grundverständnis der verschiedenen Positionen oder nicht verschiedenen Positionen ist, denke ich, klar geworden. Wir sind jetzt eigentlich bei der Frage, was den Nutzen ausmacht. Welche Parameter machen Nutzen aus, vielleicht auch Studiendesign/Studienauswahl? Welche Studien benötigt man eigentlich, um einen postulierten Vorteil tatsächlich nachzuweisen? Denn das ist die Aufgabe des Berichts, also nicht festzustellen, ob etwas postuliert ist, sondern zu zeigen, welchen Nachweis es für den postulierten Vorteil gibt.

PD Dr. Stefan Lange: Ich schlage vor, dass wir den TOP 4 um diese Fragen erweitern, also Nutzenkriterien und sich daraus ergebenden Implikationen, was zum Beispiel auch Studiendesigns angeht. Es könnte ja sein, wir brauchen gar keine Studie – das ist dann das einfachste Design –, weil man eben nicht prüfen muss. Wenn es zwangsläufig ist, dass jemand ein paar Tage im Krankenhaus ist, und es auf der anderen Seite zwangsläufig ist, dass das Verfahren innerhalb von 3 Stunden erledigt ist und der Mensch wieder nach Hause gehen kann, dann würde ich auch sagen, dass man dann keine Studie braucht. Das ist das einfachste Studiendesign. Aber genau das könnte man unter Punkt 4 behandeln.

Moderator Dr. Thomas Kaiser: Das kann man unter dem vorhandenen Punkt 4 subsumieren.

Ich schlage vor, dass wir nun in die Mittagspause eintreten.

(Unterbrechung von 13 Uhr bis 13:30 Uhr)

Moderator Dr. Thomas Kaiser: Herr Muschter, ich meine, dass wir auch am Nachmittag die Ausnahmeregelung fortsetzen sollten. Ich hatte den Eindruck, dass die Diskussion von Ihrer Anwesenheit sehr profitiert hat und dass auch die Projektgruppe einen wesentlichen Input bekommen hat, sodass Sie herzlich eingeladen sind, auch am Nachmittag an der Diskussion teilzunehmen.

Wir kommen nun zu

Tagesordnungspunkt 4:

Studiendesign/Studienauswahl

- Einbezug von nichtrandomisierten kontrollierten Studien in die Nutzenbewertung

Tagesordnungspunkt 5:

Gleichwertigkeit/irrelevante Unterschiede/klinische Relevanz

- Definition irrelevanter Gruppenunterschiede

Katharina Biester: Wir hatten ja in den Methoden festgelegt, dass wir RCTs und nicht randomisierte, kontrollierte Studien (CCTs) mit paralleler Kontrollgruppe einbeziehen und die CCTs dann einbezogen wurden, wenn nicht ausreichend RCTs oder gar keine RCTs vorlagen, und die CCTs als Voraussetzung eine Berücksichtigung von prognostischen Faktoren vorgenommen haben mussten in den CCTs. Da wurde angemerkt, dass man bei qualitativ mangelhaften RCTs dann auf CCTs hätte zugreifen sollen. Das ist der eine Punkt, den wir unter diesen TOP ansprechen möchten.

Der andere Punkt, der sich auch auf andere Studientypen als RCTs bezieht, sind die unerwünschten Ereignisse. Auch hier wurde angemerkt, dass man auf andere Studientypen als RCTs zugreifen müsse, um das zu bewerten.

Das würden wir jetzt gerne mit Ihnen besprechen, vielleicht zunächst den ersten Punkt.

Stefan Zenk: Es geht um den Punkt, der uns beschäftigt hat, warum so viele Studien als „grob“ mangelhaft bezeichnet wurden. Wir haben daraus abgeleitet, wenn die schon alle so schlecht sind, dann hätte man auf die größere Zahl von nicht randomisierten Studien zugreifen sollen. Eine klare Definition unsererseits kann ich jetzt nicht liefern. Wir haben uns gefragt: Es ist so viel Evidenz da, die RCTs als Evidenzklasse 1 wird als „grob mangelhaft“ eingeordnet und damit ja ein Stück weit verworfen. Warum geht man dann nicht auf das nächste Level?

PD Dr. Stefan Lange: Einerseits eine Erläuterung und andererseits eine Nachfrage:

Die Bezeichnung „grob mangelhaft“ ist etwas unglücklich, die wir in Zukunft in unseren Berichten nicht mehr verwenden wollen. Dahintersteckt, dass bestimmte Charakteristika der Studiendesignaspekte, Ausfallraten usw., zu einer höheren Ergebnisunsicherheit führen. Ein Prinzip einer solchen Bewertung ist eben, dass man versuchen muss, diesen Aspekt der Ergebnissicherheit, Ergebnisunsicherheit zu den postulierten Effekten in Bezug zu setzen, um

trotz einer vielleicht hohen Ergebnisunsicherheit zu einer Aussage kommen zu können oder auch nicht, jedenfalls Vorschläge zu machen, wie die Studien eine höhere Ergebnissicherheit liefern könnten. Das als Erläuterung.

Bei randomisierten Studien geht man davon aus – das ist nicht von uns erfunden –, dass die Ergebnissicherheit am höchsten ist, weil das die bisher einzig bekannte Methode ist, die es relativ einfach erlaubt, vergleichbare Gruppen zu schaffen. Natürlich kann man fragen: Kann es nicht sein, wenn die RCTs so durchgeführt sind, dass sie trotzdem eine höhere Ergebnisunsicherheit haben, dass es vom Evidenzgrad her niedrigere Studien gibt, die aber zu einer besseren Ergebnissicherheit führen, weil sie eben so gut durchgeführt sind? Das ist eine völlig berechtigte Frage. Hier ist unsere Frage, ob Sie uns gegebenenfalls entsprechende Studien nennen können. Wir stellen uns, wie wir es auch im Berichtsplan definiert haben, eine vergleichende Studie vor – das ist, glaube ich, das Wesentliche –. Zumindest was bestimmte Aspekte angeht, kann man ja den Stellenwert einer Therapie letztendlich nur im Vergleich zu der Alternative sehen, und das eben möglichst zeitlich parallel, damit solche Zeiteffekte nicht eine Rolle spielen, wie wir es vorhin diskutiert haben. Wir wären sehr dankbar, wenn Sie uns Hinweise auf entsprechende nicht randomisierte Studien geben könnten.

Stefan Zenk: Für den Fall TUNA, den ich vertrete, möchte ich Boyle und Zlotta nennen. Das sind Arbeiten, die da durchgeführt wurden, die herangezogen werden könnten. Darüber hinaus würde ich auf die Arbeiten, die die französische HAS geleistet hat, verweisen, die eine ganze Vielzahl von Studien in ihre Bewertung mit einbezogen hat.

Moderator Dr. Thomas Kaiser: Wir müssen zwischen den Ergebnissen und der Methodik trennen. Generell zunächst: Welche Studien sind aus grundsätzlichen theoretischen Überlegungen für diesen Bericht geeignet, und welche Studien kommen aufgrund dieser theoretischen Überlegungen tatsächlich später in die Bewertung?

Katharina Biester: Zu den 3 Arbeiten, die Sie gerade genannt haben:

Boyle ist eine Meta-Analyse. Das ist also keine Studie, kein CCT, die wir zur Nutzenbewertung herangezogen hätten. Ich meine auch, dass dort keine systematische Literaturrecherche hinterlag.

Der Report von HAS ist auch ein HTA, also auch keine Studie, die zur Nutzenbewertung herangezogen wird.

Die Arbeit von Zlotta ist keine vergleichende Studie in dem Sinne, dass zwar 2 Gruppen behandelt werden, aber beide mit TUNA, mit unterschiedlichen Kathetern. Das ist in unserer Methodik ein Ausschlusskriterium, weil es ein Vergleich innerhalb eines Verfahrens ist.

Moderator Dr. Thomas Kaiser: Das ist nicht Gegenstand des Auftrags.

Ich denke, wir sollten jetzt nicht auf einzelne Studien eingehen, denn das ist die Folge dieser Diskussion, welche Studien hineinkommen. Gibt es zu dem grundsätzlichen Punkt, Einbezug von Studien, wie im Berichtsplan geschildert, erstens aus den Stellungnahmen noch Nachfragen und zweitens von den Stellungnehmenden noch Anmerkungen?

Prof. Klaus Höfner: Sie haben ja im Prinzip eine sehr transparente Studienauswahl. Die kann man ja dem Papier entnehmen; das ist ja auch in Ihren Methoden klar dargestellt. Das heißt, im Wesentlichen sind RCTs und nur dann, wenn es nicht passt, CCTs zu verwenden. Was eigentlich fehlt, ist ein Katalog der Fragestellungen, der uns sagt, welche Dinge interessieren uns am Ende und wann kommen wir, um dieses zu untersuchen, aus dieser strikten Trennung heraus und akzeptieren einen niedrigeren Studienlevel. Beispiel: Es gibt die Gruppe der Patienten mit Harnverhalt, die also nicht Wasser lassen können, die eine ganz spezielle Gruppe sind und die nachweislich nach TUNA und TUMT - -¹⁰ Da gibt es Studien, ohne die jetzt detailliert nennen zu wollen. Ich erkläre nur das Prinzip, weil wir ja über Methoden reden. Hier können wir sagen, diese Patienten profitieren – das ist nachgewiesen – prospektiv, nicht einmal retrospektiv, zum Teil prospektiv, zum Teil retrospektiv analysiert, insofern, als sie hinterher Wasser lassen können. Das ist eine Gruppe von Patienten, die in Pflegeheimen sind, die einen Katheter haben müssen. Insofern besitzen diese Ergebnisse eine extreme soziale Relevanz. Wenn einen das Problem interessiert, dann muss man auf ein Level zurückgehen, wo ich sage: Das ist das, was ich wissen will. Wenn Sie das in Ihren Methoden transparent machen, einen Anforderungskatalog haben, unabhängig von dem prinzipiellen Ansatz – das ist ja ein transparenter Ansatz –, dass Sie sagen, ich gehe nur von dem Studienlevel aus, habe aber einen Katalog von Fragen, die ich mir selbst stelle und danach wähle ich die Studien aus und gucke danach, was der Outcome ist in dieser Richtung, dann wäre es sehr viel günstiger und für den, der solch einen Bericht liest, auch sehr viel transparenter, wenn Sie sagen: Wir haben das und das getan und nach diesen Kriterien die Studien ausgewählt, unabhängig von unserem hohen Standard der wissenschaftlichen Evidenz, was die Studien machen. Das würde ich vorschlagen. Dieser Katalog müsste abgestimmt werden. Das muss man auch einmal mit Experten durchgehen, was da wichtig wäre, es sei denn, Sie wissen es selbst.

PD Dr. Stefan Lange: Es ist ja unter anderem Sinn und Zweck unserer heutigen Runde, Hinweise zu erlangen. Ich glaube, dass es ein vernünftiges Vorgehen ist, zu versuchen, bestimmte Fragestellungen oder Endpunkte zu identifizieren und gegebenenfalls daraus abzuleiten, welcher Studientyp uns die größte Ergebnissicherheit liefert und durchführbar oder nicht durchführbar ist. Ich habe das jetzt mit der Retention so verstanden, als ob hier vielleicht keine Vergleiche möglich seien. Das könnte man jetzt zum Beispiel heranziehen. Kann man bei Patienten, wie Sie sie geschildert haben, für die Beschreibung des Stellenwerts, zum Beispiel der TUMT oder der TUNA - -¹⁰ Was ist die Alternative für diese Patienten, und kann man zu dieser Alternative nicht trotzdem eine Studie machen, oder gibt es gar keine Alternative? Das könnten wir ja versuchen, an solchen Beispielen durchzugehen.

¹⁰ Abbrechen des Satzes

Moderator Dr. Thomas Kaiser: Eine kurze Zwischenbemerkung: In dem Methodenpapier 2.0 sind bestimmte Situationen geschildert worden, wo nicht randomisierte Studien zu Fragestellungen von Interventionen trotzdem hinreichend ergebnissicher sind. Das Ziel unserer Arbeit und auch ein sinnvolles Vorgehen ist nicht, zu sagen, wenn man nichts hat, dann guckt man sich halt irgendetwas anderes an, sondern das Ziel ist ja, zu einer hinreichend ergebnissicheren Aussage zu kommen. Da macht es keinen Sinn, sich Sachen anzuschauen, wo man von vornherein weiß, die sind nicht hinreichend ergebnissicher. Es macht aber Sinn, sich zu überlegen, welches sind denn Fragestellungen, wo ich von einem derart dramatischen Effekt ausgehe, dass mir bestimmte Studienformen hinreichende Ergebnissicherheiten liefern, auch wenn sie nicht randomisiert sind. Wir hatten in der Mittagspause ein Beispiel von einer Erkrankung, die zwangsläufig zum Versterben führt, also deterministische Erkrankungen. Da sind einige Fallserien, die zeigen, dass eine Therapie bei 90 % der Patienten zum Überleben führt, ausreichend. Das muss man dann nicht randomisieren. Sie haben gesagt, man müsste einen Katalog feststellen. Es macht keinen Sinn, einen solchen Katalog aus dem Hörensagen festzustellen, sondern man braucht dafür gewisse wissenschaftliche Grundlagen, also Anhaltspunkte, entweder aus dem Verfahren selbst, aus der Erkrankung oder aus solchen Studien, die nicht randomisiert sind, die einen solchen dramatischen Effekt zeigen, dass man sagt, hier gibt es einen solchen Anhaltspunkt, dass es Sinn macht, in die nicht randomisierten Studien zu schauen, weil hier dramatische Effekte zu erwarten sind und die systematische Betrachtung solcher Studien mir dann im Gesamtergebnis eine hinreichende Ergebnissicherheit liefert. In dem neuen Methodenpapier wird genau dieser Punkt konkreter geschildert sein. So müsste man an diesen Katalog herangehen.

Stefan Zenk: Zunächst möchte ich Frau Biester antworten. Zlotta und den HTA von Frankreich habe ich genannt. Dahinter verbirgt sich ja eine ganze Reihe von Studien. Wie Prof. Höfner sagte, müsste man anhand der Fragestellung eine einzelne Studie herausgreifen, um Ihre Frage zu beantworten.

Ich möchte auch Herrn Dr. Kaiser antworten. Die Frage nach einem Katalog können wir sicher hier nicht beantworten, denn das kann nicht Gegenstand dieser Sitzung sein. Es wäre wünschenswert, wenn es ein weiteres Meeting gäbe, wo man an dieser Fragestellung arbeitet. Ich halte es für sinnvoll, genau in der Art, wie Sie es ausgeführt haben.

Moderator Dr. Thomas Kaiser: Wir gehen davon aus, dass in den Stellungnahmen alle relevanten Studien genannt wurden, die zu einer Veränderung des Ergebnisses führen. Das heißt, diese Studien müssen benannt worden sein. Sie haben ja die Gelegenheit zur Stellungnahme gehabt. Vielleicht können Sie ja in der Diskussion noch darauf hinweisen, welches die Studien sind, die den Anhaltspunkt für einen solchen dramatischen Effekt geben.

Stefan Zenk: Wir haben eine retrospektive und eine prospektive Betrachtung. Ich habe Sie mit dem Punkt „Katalog“ in der prospektiven Richtung denkend verstanden.

Moderator Dr. Thomas Kaiser: Aber natürlich basierend auf dem vorhandenen Material, das Ihnen ja schon vor 2 Monaten bekannt war.

PD Dr. Stefan Lange: Wir sollten das jetzt konkret angehen. Nehmen wir einmal das Szenario, das Herr Prof. Höfner geschildert hat: ältere Menschen in einem Pflegeheim mit einem akuten Harnverhalt. So habe ich Sie verstanden.

(Prof. Klaus Höfner: Katheterversorgte Patienten!)

– Katheterversorgte Patienten. – Man könnte ein minimalinvasives Verfahren anwenden, man könnte den Katheter belassen oder trotzdem eine TURP machen. Das wären vermutlich die Alternativen oder nicht, weil man sagt, die sind so multimorbide, da verbietet sich beispielsweise die TURP, sodass das als Alternative nicht infrage käme. Dann wäre aber noch die Alternative, den Katheter zu belassen, und könnte dann sagen, ich müsste doch jetzt eigentlich diese Frage untersuchen, müsste überlegen, was die entscheidenden Zielkriterien sind. Die Möglichkeit der Katheterentfernung ist klar. Da würden vermutlich gewisse Beobachtungen an 20, 30 Patienten ausreichen. Dann ist der Katheter weg und die können wieder Wasser lassen. Wenn das die einzige Fragestellung ist, dann, glaube ich, wird das keine Frage sein. Da kann man sich nur fragen, ob es noch andere Möglichkeiten gibt, das zu erzielen, zum Beispiel Blasentraining. Oder man belässt alles so, wie es ist, und überlegt sich die Outcomes, die man sich angucken will. Das ist aber aus meiner Sicht schon eine sehr spezielle Fragestellung. Dafür sind ja die Verfahren letztlich nicht entwickelt worden. Sie werden ja vermutlich in 90 bis 95 % der Fälle für den Patienten entwickelt, der noch auf seinen beiden Beinen laufen kann und seine Beschwerden beim Wasserlassen hat, was nicht mehr gut medikamentös in den Griff zu bekommen ist oder wenn er keine Medikamente einnehmen will. Man muss also aufpassen, dass man nicht zu sehr in irgendwelche Nischenindikationen hineingeht, weil man dann irgendwann in Einzelfallbetrachtungen kommt.

Dr. Richard Berges: Wir gelangen nun in der Diskussion in die Indikationsstellung. Ich glaube, das ist ein ganz wesentlicher Aspekt und auch eines der großen Probleme, die wir mit dem gesamten Erkrankungsbild haben. Sie haben, wenn Sie die operative Technik generell betrachten, nur sehr wenige Patienten, die in die Kategorie „absolute Operationsindikation“ fallen. Das sind vielleicht 10 % der BPS-Patienten. Das sind die Harnwegsinfekte, Patienten mit Blasenstein, mit rezidivierenden Harnverhaltungen und mit rezidivierenden Blutungen. Diese Patienten müssen nach allen unseren Erfahrungen operiert werden. Alle anderen Patienten – das sind 90 % der BPS-Patienten – können operativ oder medikamentös therapiert werden, wobei medikamentös meistens nur über einen gewissen Zeitraum gut geht. Ich glaube, es gibt kein Krankheitsbild, das bezüglich der einsetzbaren Medikamente besser untersucht wurde als die BPS. Wir sehen, dass wir hier nur einen ganz marginalen Therapieeffekt erzielen, was die Symptomreduktion, die Flowverbesserung, die Deobstruktion betrifft, egal welche Medikamente wir einsetzen. Das heißt, letztendlich ist die Diskussion zu führen: Wann wird denn eigentlich der Beginn für die Indikation einer Operation wichtig? Diesen genauen Zeit-

punkt können Sie nicht definieren. Es gibt keine Untersuchung, die sagt, du musst operiert werden und du nicht. Das sind alles Kann-Bestimmungen oder auch, wenn Sie so wollen, Patientenpräferenzen. Der eine sagt, ich möchte keine Tabletten mehr nehmen oder er möchte überhaupt keine nehmen, wenn er sie noch nie genommen hat. Davon gibt es in Deutschland wenige. Die meisten haben eine medikamentöse Therapie hinter sich und sagen, das hilft nicht mehr. Unter dem Szenario wird dann die Indikation gestellt. Aber es ist letztendlich eine weiche Indikation. Es ist in vielen Fällen keine Muss-Indikation. Wenn Sie in die Leitlinien weltweit gucken, dann stellen Sie fest, dass mittlerweile auch nicht mehr die Symptomstärke der ausschlaggebende Punkt für den Beginn einer Therapie ist, sondern der dadurch abgeleitete Leidensdruck des Patienten. Da liegt letztendlich die Krux der ganzen Geschichte, denn die ist höchst variabel. Sie oder wir müssen dann entscheiden, welches Verfahren man anwendet. Da sind wir dann wieder an der Stelle und sagen: Natürlich ist es ein solches, wenn man schon eine Kann-Bestimmung hier einführt, das möglichst schonend ist. Das ist das, was der Patient wünscht. Dann sitzen wir wieder da. Sie sehen das Problem? Aus diesem Problem heraus muss man sich die Studien angucken.

Moderator Dr. Thomas Kaiser: Um einem Missverständnis vorzubeugen: Wir als Institut entscheiden hier nicht, welches Verfahren eingesetzt werden soll, sondern wir beschreiben die Evidenzlage. Die Entscheidung liegt woanders.

Dr. Richard Berges: Wir als Arzt entscheiden ja. Das „wir“ war rein auf die klinische Situation bezogen.

Prof. Klaus Höfner: Ich möchte noch etwas zur Situation der Studien sagen. Das Problem, Herr Kaiser, das Sie ansprechen, ist, dass wir natürlich die Möglichkeit hatten, das zu liefern. Aber man muss bitte Verständnis haben. Wir haben ja die Fristen kritisiert. Ich will das auch nicht ausweiten. Das ist natürlich eine unglaubliche Arbeit, in einem Gremium innerhalb von 4 Wochen Ihnen in einem 500-Seiten-Bericht die Studien zu liefern, auch in einer gewissen Präzision und Systematik, um Ihnen zu helfen.

Trotzdem bin ich der Meinung, wenn Sie Evidenz wollen und diese Evidenz nicht auf die logische und nachvollziehbare Studienausswahl beschränken, das heißt RCTs und CCTs oder CCTs nur dann, wenn RCTs nicht, dann werden Sie ein gewisses Bild der Evidenz bekommen, wie wir das verstehen, der klinischen Relevanz der Evidenz. Das ist ja auch in Ihrem Methodenpapier lange erklärt, was das eigentlich ist. Dass das sehr unscharf ist, was wir diskutieren, ist klar. Aber ich sage noch einmal: Wenn man die Evidenz so versteht, dass man sagt, wir wollen beispielsweise auch eine Nischenindikation untersuchen, um bei diesem Beispiel zu bleiben, dann muss man diese Studien gezielt nehmen und gucken, was dran ist, was wir daraus entnehmen können, ohne jetzt auf den wirklich statistischen Level zu gucken und zu sagen, was für uns akzeptabel ist. Das ist eine ganz andere Fragestellung. Diesen Katalog, wie gehe ich von meiner strikten Studienausswahl weg oder wo gehe ich sogar hin und sage, auch eine retrospektive Studie kann eine Aussage in der und der Weise haben – da sind wir

keine Experten. Da müssen wir die Statistiker fragen, welcher Level akzeptabel ist, wo ich den Evidenzlevel ansetze, welche Evidenz für uns klinisch noch interpretierbar ist oder in einer gewissen Aussage für eine bestimmte Nischenindikation verwertbar. Das sind alles Fragen, die wir nicht in zehn Minuten und auch nicht innerhalb der vier Wochen beantworten können. Deswegen sage ich: Wir können dieses Problem nicht in dieser Sitzung lösen. Das ist ein sehr komplexes Problem. Da muss man zumindest auf einen Level des Verständnisses kommen. Wir haben in unserem Bericht geschrieben, nämlich zur Auswahl der Studien auf Seite 6 oben, das heißt die Ausfallraten von mehr als 70 %, wo man Selektionsbias diskutiert, wo man bewusst nicht auf diesen Level kommt. Man hat eben gewisse Probleme, wenn man eine Auswahl in dieser Weise vornimmt und dann bestimmte Festlegungen setzt. Das ist sicherlich transparent und in keiner Weise anzugreifen oder wissenschaftlich kritisch zu bewerten. Aber man hat damit Probleme, weil das Bild, das ich am Ende habe, aus unserer Sicht sehr eingeschränkt ist und eine verzerrte Situation darstellt. Es mag sein, dass das aus Ihrer Sicht anders ist. Für uns ist eine Verzerrung da, die für uns und letztlich auch für den Patienten nicht akzeptabel ist. Deswegen muss man einen größeren Zeitraum haben, um das sehr präzise und spezifisch festzulegen, dass Sie damit arbeiten können.

PD Dr. Stefan Lange: Ich glaube, wir müssen eines feststellen: Es ist ein Problem, das wir nicht nur beim BPS, sondern auch in vielen anderen Bereichen haben. Sie haben das jetzt „Nischenindikation“ genannt, also der Versuch des Herunterbrechens fast auf Individuen herunter, wenn man es einmal extrem formuliert. Das können wir natürlich nicht leisten. Das kann man im Rahmen eines solchen Berichtes nicht leisten; das hat meines Wissens bis jetzt auch keine Agentur der Welt geschafft.

(Moderator Dr. Thomas Kaiser: Das ist auch nicht Gegenstand des Auftrags!)

– Das ist auch nicht Gegenstand des Auftrags. – Natürlich ist das, was wir betrachten wollen, die mehr oder weniger breite Anwendung, die natürlich mehr oder weniger scharfe Indikationen hat, hier möglicherweise weniger scharfe Indikationen, was sich aber wiederum in entsprechenden Studien widerspiegelt.

Ich gehe jetzt mal auf die Historie zurück: Als wir diesen Auftrag bekommen haben, habe ich mich relativ gefreut, weil es ja eine nicht ganz so alltägliche Situation in operativen Fächern ist, dass man überhaupt randomisierte Studien zur Verfügung hat, und das sind ja hier eine ganze Reihe. Das heißt, das Prinzip an sich ist erkannt, nämlich dass es wichtig ist, um zu aussagefähigen Ergebnissen zu kommen. Von daher müssen wir versuchen, davon wegzukommen, zu sagen, wir müssen versuchen, einzelne Nischen zu beschreiben. Das können und wollen wir mit dem Bericht nicht und war auch nicht Gegenstand des Auftrags. Wir wollen die tatsächlich breite Anwendung in der Versorgung. Hier würde ich auch nicht auf einzelne Studien eingehen. Für mich war nur wichtig, ob sie so weit genannt sind. Wir gucken dann noch einmal, wie wir die einschätzen können. Dazu, der Mensch, der seine Beschwerden mit Medikamenten nicht mehr richtig in den Griff bekommt und sich langsam mit dem Gedanken

trägt, etwas operativ machen zu lassen, kann es jetzt verschiedene Fragestellungen geben. Einerseits war das die Frage der Morbidität durch das Verfahren selbst oder auch seine Symptomverbesserung. Gibt es Situationen, wo uns andere Studien als randomisierte Studien eine höhere Ergebnissicherheit liefern können?

Moderator Dr. Thomas Kaiser: Ich möchte auch noch eine Bemerkung machen. Sie haben eben gesagt, man hätte ein verzerrtes Bild bei der engen Betrachtung. Mir ist da direkt die Idee des Fensters gekommen, und dann gibt es noch etwas darum herum. Wir haben eigentlich das Ziel, durch Glas, durch das man durchgucken kann, zu schauen, um zu wissen, was dahinterliegt. Das kann dazu führen, dass man nur durch die Mitte des Fensters schaut und durch die äußeren Bestandteile des Fensters nicht, die dreckig oder verwischt sind, weil man nicht weiß, ob dahinter ein Baum oder ein Auto ist. Das heißt, genau dafür sind diese Kriterien da, das Glas so transparent wie möglich zu halten. Sollte sich jetzt herausstellen, dass an den Außenkanten irgendwo doch noch eine Stelle ist, wo man transparent genug durchschauen kann – genau das ist jetzt der Diskussionspunkt –, dann macht es Sinn, sich durch diesen Punkt die Welt draußen anzuschauen. Es macht aber keinen Sinn, sich stundenlang vor dem Fenster zu bewegen und zu gucken, was ich eventuell hinter den nicht transparenten Sachen sehe. Genau das ist die Frage: Gibt es transparente Flecken, wo man hinreichend sicher das Ergebnis dahinter betrachten kann? Da müssen wir rangehen und nicht uns das Fenster grundsätzlich komplett angucken, auch wenn es dreckig ist.

Dr. Richard Berges: Da habe ich eine Verständnisfrage: Mich würde das in die Nischenindikation führen, zum Beispiel Patient, Harnverhaltung, multimorbide.

Moderator Dr. Thomas Kaiser: Es kann aber auch zu einem dramatischen Effekt bezüglich bestimmter Komplikationen oder ähnlichen Dingen führen – es muss nicht zwingend eine Nischenindikation sein –, also zu einem Vorteil für einen bestimmten relevanten Endpunkt, der so dramatisch ist, zum Beispiel versterben, was jetzt hier nicht der Fall ist, aber es könnte ja sein, dass es hinreichend sicher ist, ihn auch in nicht randomisierten Studien zu beobachten. Das muss nicht zwingend die Nischenindikation sein.

Prof. Klaus Höfner: Es ist unbenommen, Herr Kaiser, was Sie sagen. Wir bemühen uns ja, zu verstehen, dass man eine Evidenz haben muss, um eine klinische Aussage zu treffen. Das ist ja auch unser Anliegen. Hier gibt es keinen Dissens. Das, was Sie sagen, ist ganz klar zu unterschreiben. Wir sehen aber schon in bestimmten Bereichen – wir kommen ja gleich zur Morbidität – das Problem, dass wir mit diesen Dingen, wenn ich den Bericht richtig gelesen habe, zu einer verzerrten Darstellung kommen. Das ist der Punkt, über den wir sprechen. Deswegen meinen wir, dass die Instrumente so strikt, wie sie gehandhabt sind, auch so transparent – das ist ja hervorragend, dass das alles so transparent ist –, automatisch zu einer verzerrten Darstellung führen in der Interpretation, die ich an bestimmten Stellen lese. Das ist die Lesart. Wir haben keinen Dissens, dass wir sagen, wir brauchen Instrumente, die genau trans-

parent und validiert sein müssen. Damit müssen wir klarerweise auch leben. Insofern gibt es keinen Dissens. Das wollte ich noch einmal darstellen.

Moderator Dr. Thomas Kaiser: Ihr Punkt ist der, zu sagen, ihr guckt zwar richtig durch, aber ihr beschreibt hinten einen Baum, obwohl dort ein Auto steht?

Prof. Klaus Höfner: Ja, vielleicht trifft es das.

Moderator Dr. Thomas Kaiser: Nur zum Verständnis, weil es ist ja eigentlich Diskussion der Ergebnisse des Vorberichts.

Prof. Dr. Rolf Muschter: Ich weiß nicht, ob Sie das getan haben, was man als Bewerter, wenn man ein Buchkapitel zur TUR schreiben muss, tut. Man sucht sich die Studien zur TUR und wird in erster Linie in den Armen der randomisierten Studien gegen minimalinvasive Verfahren fündig. Wenn Sie jetzt diesen ganzen Studienpool, diese Arme, separat zur TUR, nehmen und die in eine Art von – statistisch nicht wörtlich zu nehmen – Meta-Analyse packen, dann stellen Sie fest, dass Sie unglaubliche Heterogenitäten in diesen Studienarmen haben. Sie finden am Schluss in Studien TUR-Arme, wo das minimalinvasive Verfahren unterlegen war, wo aber der TUR-Arm in einer anderen Studie dem minimalinvasiven Verfahren unterlegen gewesen wäre, wäre das die Randomisierung gewesen. Ich will damit sagen: Die Qualität der Behandler findet sich auch in diesen randomisierten Studien wieder. Nun kommt das Problem: Randomisierte Studien sind klinisch immer schwierig umzusetzen, erst recht, wenn man dann schon Erfahrungswerte hat. Wir haben ja vorhin in der nicht protokollierten Pause gesagt: Wenn ein Patient seinen Nachbarn im Bett sieht und sieht, wie wunderbar es dem geht, dann sagt er: Randomisieren will ich nicht mehr. Ich will dieses Verfahren haben. Da geht es mir ja blendend. – Das ist das Problem, das uns an der Randomisierung oft gestört hat, dass wir Patienten aufgrund ihrer Präferenz verschiedenen Gruppen zuordnen mussten. Jetzt komme ich dazu, ich habe eine randomisierte Studie, fahre zeitlich parallel eine nicht randomisierte Kohortenstudie in der gleichen Institution, gleicher Operateur, alles gleich, und komme im Grunde genommen zu den gleichen oder zu vergleichbaren Ergebnissen – das ist natürlich statistisch nicht zulässig – das ist nun einmal nicht randomisiert. Jetzt kann ich aber doch aus dieser Studie enorme Erkenntnisse ableiten, die über das Randomisierte hinausgehen. Das Randomisierte untersucht doch, wie gut das im Verhältnis zum Standard ist oder wie sich die beiden Gruppen untereinander verhalten. Die nicht randomisierte Studie will doch darüber hinaus nur zeigen, wie das in einem größeren Kollektiv ist, ob das Verfahren machbar und sicher ist. Das sind ja im Grunde genommen die Fragestellungen. Die stehen dann in der Publikation auch vorne drin. Das heißt dann meistens: „Feasibility and safety of...“ usw. Ist dieses Material denn nicht zu verwenden? Ich verstehe das nicht ganz. Sie haben Recht: Es ist eine Studie, die statistisch deutlich weniger wert ist als die randomisierte Studie ist. Aber sie bringt doch Erkenntnisse. Diese Erkenntnisse müssen doch irgendwo berücksichtigt werden, wenn man Verfahren bewertet. Ich kann das doch nicht ausblenden, erst recht bei guten Erfahrungen in einer Institution, die vielleicht nicht belegbar gleich dem Standard sind, aber die

doch immerhin so gut sind, dass ich sage, dieses kann ich meinem Patienten auch weiterhin in Studienform anbieten und das untersuchen, dass man dann Subgruppen bildet und sagt, bei dieser Gruppe gucken wir nach einem ganz bestimmten Parameter, nämlich ob das Verfahren die Obstruktion, Harnverhalte, besonders große Prostatae behandeln kann. Diese Liste können Sie endlos fortsetzen. Dann bekomme ich doch auch wieder Antworten. Sind diese Antworten denn alle so, dass man sie nicht berücksichtigen darf, nur weil sie nicht einer randomisierten Studie entstammen? Das verstehe ich einfach nicht.

Moderator Dr. Thomas Kaiser: Das ist nicht der Grund. Die Frage ist, wie sicher diese Antworten sind.

Dr. Guido Skipka: Ich bin ein bisschen erstaunt. Das Ziel der Randomisierung ist ja nicht, irgendetwas Neues mit irgendetwas Altem zu vergleichen. Das Ziel einer randomisierten Studie ist, zu klären, welche von 2 wei möglichen Ansätzen, Strategien besser ist. Sie haben das Beispiel von Patienten genannt, die den Nachbarn sehen und sagen, dass sie das auch wollen. Wenn er überhaupt nicht bereit ist, irgendetwas anderes zu nehmen, dann brauche ich auch keine randomisierte Studie. Für den stellt sich überhaupt keine Frage. Aber so wollen wir sicherlich nicht vorgehen. Ich denke schon, man hat ein Standardverfahren oder eine Referenz und jetzt hat man etwas Neues entwickelt. Dann stellt sich doch die Frage, ob das Neue dem Alten überlegen oder gleichwertig ist. Die Frage ist doch nur interessant und weiterzuverfolgen, wenn ich Patienten habe, wo ich diese beiden Strategien wahlweise anwenden kann. Wenn das nicht der Fall ist – da gebe ich Ihnen vollkommen Recht –, dann brauche ich keine randomisierte Studie.

Sie fragten, ob man aus Studien niedrigerer Evidenz keine Informationen holen kann. Natürlich können wir das. Dadurch bekommt man ja erst die Informationen, um neue Verfahren zu entwickeln. Aber noch einmal, wenn wir vor dem Punkt stehen, dass wir etwas Altes und etwas Neues haben, das dem Patienten angeboten werden kann. An dieser Stelle ist eine Randomisierung möglich, sonst habe ich keine Fragestellung.

PD Dr. Stefan Lange: Um das auf den Punkt zu bringen: Sie haben das ja sehr schön beschrieben. Wir müssen aufpassen, denn wir vermischen wieder ein bisschen Ergebnisse mit Methodik. Aber ich glaube, das ist in Ordnung, denn die Ergebnisse haben wir nicht neu erfunden, diese lagen ja schon vorher vor.

Die Unterschiede in den Komplikationsraten für ein und dasselbe Verfahren, TUR, zwischen verschiedenen Studien, aber auch im Effekt im Vergleich zu einem nicht invasiven Verfahren oder weniger invasiven Verfahren, das ist erst einmal so da. Das ist Fakt. Ich habe Sie so verstanden, dass Sie sagen, deswegen können wir zum diesbezüglichen Stellenwert zum Beispiel der TUR gar nicht so viel sagen, weil das so heterogen ist. Wir können eigentlich viel mehr sagen aus der Begleitforschung, was vielleicht auch den alltäglichen Gebrauch des Verfahrens widerspiegelt und wo vielleicht weniger Heterogenität da ist. Genau das ist für uns der Punkt: Wo bekommen wir mehr Ergebnissicherheit her, aus dem fairen Vergleich – das ist das einzi-

ge Ziel der Randomisierung, dass man einen fairen Vergleich machen kann, dass die Ausgangsbedingungen für die Gruppen, die man vergleichen will, gleich sind – oder aus einer Beobachtungsstudie, wo wir in aller Regel eben nicht wissen, wie die Patienten ausgewählt worden sind, was das für eine Selektion ist, ob man da unbewusst oder auch bewusst besonders günstige Patienten mit dem Verfahren behandelt hat, aus welchen Gründen auch immer, woran es liegt, dass ich ganz andere Komplikationsraten sehe, oder, wenn ich die TUR nehme, ob in der Praxis nicht so fein wie in der RCT ausgewählt wird, sondern etwas gröber behandelt wird, sodass ich mehr Komplikationen, größere Schwierigkeiten sehe, die ich aber dann, wenn ich die gleichen Patienten in meiner randomisierten Studie hätte, auch bei den minimalinvasiven Verfahren sehe? Das ist genau die Frage: Wo können mir solche Studien außerhalb eines randomisierten fairen Vergleichs eine größere Ergebnissicherheit bieten?

Prof. Dr. Rolf Muschter: Das hat Herr Berges schon gesagt. Manche Effekte sind evident. Es ist einfach klar, dass ich ein Verfahren ambulant einsetzen kann, was mit einem anderen nicht geht.

PD Dr. Stefan Lange: Sie vermischen etwas. Das war jetzt nicht die Frage.

Prof. Dr. Rolf Muschter: Das ist ja die Zusatzinformation. Selbstverständlich kann man auf die randomisierte Studie nicht verzichten. Da haben Sie völlig Recht. Die sind ja auch für die Verfahren, die hier untersucht worden sind, durchgeführt. Es geht nur darum, dass die anderen Studien, die nicht randomisierten Studien, die zusätzliche Informationen liefern, aus dieser Untersuchung ausgeblendet sind. Das ist das, worauf ich im Moment hinausmöchte. Die Frage, die Sie eigentlich haben, wie ich es verstanden habe, Nutzenbewertung, steht nicht Nutzenbewertung gegenüber Standard als Überschrift in diesem Bericht. Genau das ist aber das, was gemacht wird. Es wird ein Standard gesetzt und untersucht, ob es einen zusätzlichen Nutzen gegenüber dem Standard gibt. Aber Nutzen ist doch etwas Abstraktes. Ich gebe ein Beispiel: Sie lassen 2 Autos randomisiert 1 km fahren. Die eine Gruppe bekommt einen 8-Zylinder-Mercedes und die andere Gruppe bekommt einen VW Polo. Dann wird die Gruppe mit dem 8-Zylinder-Mercedes aller Voraussicht nach schnellere Zeiten fahren. Kommt dann daraus, dass der VW Polo ein schlechtes Auto ist, mit dem man nicht fahren kann? Es kommt doch nur heraus, dass das ein langsames Auto ist, aber nicht, dass das ein Auto ist, das nichts taugt. Das ist das Problem, das ich mit diesem absurden Beispiel klarlegen möchte. Die Randomisierung sagt ja nur, das eine ist besser als das andere, aber nicht, das andere ist schlecht oder nicht verwertbar. Es zeigt uns ja sogar die Größenordnung, in der es unterlegen ist als ganz wertvolles Ding. Wir können natürlich dann in den anderen Studien untersuchen, die nicht randomisiert sind, ob dieser Nutzen nicht doch noch groß genug ist, um es Patienten anbieten zu können, um damit die ausreichende Zufriedenheit zu erzielen bei möglicherweise anderen Vorteilen, die ich in der randomisierten Studie vielleicht nur teilweise gesehen, aber zum Teil auch schon gesehen habe, nämlich weniger Blutung.

PD Dr. Stefan Lange: Ich gebe Ihnen völlig Recht. Wir betrachten hier Aspekte des Zusatznutzens. Sonst würden wir nicht gegen Standards vergleichen, weil wir davon ausgehen, es gibt ein Standardverfahren, das einen Nutzen hat. Die Frage ist, ob uns die neuen Verfahren einen zusätzlichen Nutzen versprechen, den wir nicht schon durch den Standard haben, was eben weniger Schaden bedeuten kann, bei aber gleich günstigen Symptomreduktionen. Das wäre dann ein Zusatznutzen. Vielleicht haben wir uns da noch nicht präzise genug ausgedrückt.

Den Aspekt des Zusatznutzens sollte man aber auch wieder vergleichend machen, und zwar möglichst fair vergleichend. Wir sollten trotzdem versuchen, es zu trennen von Aspekten, wo Sie sagen, die sind selbstevident. Vielleicht gibt es auch welche, die nicht so selbstevident sind. Das ist sehr wichtig, dass wir das nicht vermischen. Man müsste die klar benennen. Sie haben gesagt, Kategorisierung, ob es sekundär oder primär ablativ ist, oder Krankenhausaufenthalte.

Dr. Richard Berges: Vieles ist jetzt angesprochen worden. Es sind 2 Seiten einer Medaille. Es ist natürlich klar, dass man zunächst randomisierte Studien betrachten muss, um den Zusatznutzen, wo auch immer der liegen soll, gegenüber einem etablierten Verfahren, sage ich mal ganz vorsichtig - ¹¹ Es ist ja eines, das alle anwenden. Da muss man zunächst einmal gucken, ob man das mit dem Verfahren erreichen kann und wo der Zusatznutzen ist.

Der zweite Aspekt über den Zusatznutzen hinaus ist, was der eigentliche Nutzen ist. Da muss man in die Masse, in die Breite gehen. Ich glaube, da ist das große Problem bei chirurgischen Studien, die Patienten zu rekrutieren in der Größenordnung, wo man sie dann noch randomisiert haben will, dass es wahrscheinlich wenig Zusatzinformationen gibt. Wir sehen in allen minimalinvasiven Verfahren: Das, was in den randomisierten Studien an Effekten erzielt wurde, wurde nachher auch in den Massenstudien wiedergefunden. Da – hier gebe ich Ihnen Recht – kann man dann nicht mehr von Zusatznutzen sprechen, sondern nur noch von Nutzen.

PD Dr. Stefan Lange: Bei TUMT haben wir Studien, die nach unserer reinen Lehre im Vergleich zu Sham-Interventionen sogar den Nutzen nachgewiesen haben. Das ist ja eine Art Placebo.

(Dr. Richard Berges: Das war Niedrigenergie-TUMT!)

– Genau. Deswegen ist es so schade, dass Sie das mit der Niedrigenergie-TUMT so blöd finden, weil wir da aus meiner Sicht sehr wertvolle Erkenntnisse herausziehen können. Man sollte, auch wenn Sie es ethisch verwerflich finden, sich noch einmal überlegen, ob das Sham-Prinzip nicht auch in zukünftigen Studien möglich wäre. Denn wenn man sich zum Beispiel Effekte ansieht, und zwar TUMT versus Sham und TUMT versus Watchful Waiting, so sieht man, dass die Sham-Prozedur schon einen deutlichen Effekt hat gegenüber dem Watchful

¹¹ Abbrechen des Satzes

Waiting. Das ist uns allen klar. Da spielen noch ganz andere Effekte, wie Zuwendung, also Kontexteffekte, eine Rolle, und zwar keine geringe Rolle. Der Nutzen ist überhaupt nicht infrage gestellt. Es geht einzig und allein um den Zusatznutzen. Noch einmal die Frage: Gibt es für den Aspekt Zusatznutzen, der auch weniger Schaden bedeuten kann – Voraussetzung dafür ist, dass sie von der Symptomatik her zu ähnlichen Ergebnissen führen; das müssen wir dann definieren –, Situationen, wo wir sagen, da können uns andere Studien außer RCTs entweder eine größere Ergebnissicherheit liefern oder es ist ganz klar, dass wir darauf verzichten können, weil die Effekte so sind, dass ich draußen aus dem Fenster sehe, dass der Zug da langfährt? Ich fände es gut, wenn wir das heute noch hinbekommen könnten, zu sagen, bei den sekundär ablativen Verfahren müssen die Patienten eine gewisse zeitlang einen Katheter tragen, was ein Nachteil gegenüber Verfahren ist, bei denen sie das nicht müssen, jedenfalls für diejenigen, die das stört. Die meisten Männer stört das. Für einen gewissen Anteil von Männern ist das jedenfalls eine unangenehme Vorstellung. Dann ist das für sie ein Zusatznutzen, wenn sie das nicht müssen. Das muss man natürlich dann nicht in Studien prüfen, weil das durch das Verfahren so vorgegeben ist.

Moderator Dr. Thomas Kaiser: Ich wüsste aber trotzdem gerne, ob diese Verfahren den Nutzen machen.

Prof. Dr. Rolf Muschter: Um beim Beispiel Auto zu bleiben: Da hat das andere Auto wahrscheinlich deutlich weniger Benzin verbraucht. Das wäre ja jetzt sozusagen gegeneinander abzuwägen. In Ihrem Bericht steht ja, wo sozusagen der andere Aspekt vorteilhaft ist. Ich empfinde das nur als ausgeblendet. Die Studien haben kürzere Verweildauer und kürzere Katheterverweilzeit gezeigt. Ich empfinde das nicht als neutral bewertet. Das klingt immer so: Die Studie ist eigentlich nichts wert – bisschen übertrieben ausgedrückt und polemisiert –, weswegen wir die Effekte nicht wirklich bewerten können. Das ist das, was für mich als Leser dieses Vorberichts als Eindruck hinterlassen bleibt, dass ich einfach das Gefühl habe: Sie sehen ganz klar, das eine Auto fährt schneller. Sie sehen aber nicht, dass weniger Benzin verbraucht wird, oder Sie sehen das zwar, aber es wird nicht zur Kenntnis genommen, weil Sie sagen, das schnelle Fahren ist das Wichtigste.

Moderator Dr. Thomas Kaiser: Wenn das so wäre. Wenn man aber beides vorab definiert, ich möchte gucken, wie schnell es fährt, wie viel Benzin es verbraucht, wie komfortabel es ist, also sozusagen alles das, was für mich bei der Bewertung eines Autos relevant ist, dann würde das mit reinkommen.

Prof. Dr. Rolf Muschter: Aber das haben doch die Studienautoren getan.

Moderator Dr. Thomas Kaiser: Das ist aber der andere Punkt, und das ist der Punkt der Ergebnisbewertung. Es ist wertvoll, dass Sie jetzt noch einmal diesen Hinweis geben, weil das natürlich auch für die Bearbeitung des Vorberichts wichtig ist. Das hat aber mit der grundsätzlichen Methodik erst einmal nichts zu tun, sondern das hat mit der Interpretation der ausgewählten Studien sowie der Darstellung und Abwägung der Ergebnisse gegeneinander zu tun.

Dr. Roman Schiffner: Ich bin Herrn Lange dankbar, weil wir ja mitunter Schwierigkeiten mit der Definition der Begriffe haben. Insofern wäre mein Petition, dass Sie die Begriffe „Zusatznutzen“, „Nutzen“ – das haben wir ja auch im G-BA versucht zu definieren – im Fazit definieren, so wie Sie es jetzt gesagt haben, Nutzen ja, aber Zusatznutzen vielleicht nicht, damit es klar definiert und nachvollziehbar für die nachfolgenden Gremien ist. Das ist ganz wichtig. Das war mir bei der Lektüre des Vorberichts so nicht klar. Das hat definitiv Auswirkungen auf die Interpretation in den Themengruppen.

Prof. Klaus Höfner: Ich versuche das jetzt zu konkretisieren, um zu sagen, wo wir spezielle Probleme sehen. Vielleicht diskutieren wir das einmal an der Substanz.

Das erste Problem, das man als Laie nicht versteht, ist die Notwendigkeit der Definition der patientenrelevanten Ziele, das ja in Ihrem Methodikpapier eine sehr große Rolle einnimmt. Wir haben das erkannt: Für Sie ist patientenrelevantes Ziel die Symptomatik und die Lebensqualität.

(Moderator Dr. Thomas Kaiser: Nicht nur!)

Was wir nicht so sehr erkennen, ist, dass Sie die UEs, Katheterisierungs- und Krankenhausverweildauer als patientenrelevantes Ziel identifizieren. Ich möchte einen Satz von Seite 395 im Fazit zitieren:

„Obgleich einerseits für einen Teil der Verfahren prozedural bedingte und/oder durch Studien belegte Vorteile gegenüber einer Standardtherapie in Bezug auf die Krankenhausverweil- und teilweise auch auf die Katheterisierungsdauer geltend gemacht werden können, kann andererseits unter der Prämisse, dass die Beeinflussung der Symptomatik und ggf. damit verbunden der Lebensqualität vordringliches Ziel der Behandlung eines BPS ist, für keines der Verfahren letztendlich ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden.“

Das heißt, hier wird das Argument benutzt, es ist kein patientenrelevantes Ziel oder es ist nicht das, was wir eigentlich sehen wollen. So interpretiere ich jedenfalls diesen Satz. Auf jeden Fall sehen wir eine Zurücksetzung des Ziels der geringeren Morbidität. Wir sagen – das ist ein Kernpunkt, den wir ausdiskutieren sollten –, die Verfahren sind entwickelt worden, um bei gleichem Effekt dem Patienten eine Behandlungsmorbidität anzubieten, natürlich immer vorausgesetzt, sie tun das, was die TURP kann, abgesehen von der Masse der Ablation des Gewebes. Sie schreiben selbst zur Symptomatik auf Seite 1 des Berichts: Für 5 Verfahren wird kein Beleg für einen Unterschied gefunden. Das heißt, symptomatisch sind vergleichbar VLAP, HoLEP, TEAP und es liegt kein Beleg für einen Gruppenunterschied vor. Das heißt, Sie sagen, es gibt eine vergleichbare symptomatische Wirkung dieser Verfahren, die hier dargestellt sind. Die Lebensqualität für 7 Verfahren: entweder ein Vorteil gegenüber Standard als Beleg oder es wird kein Beleg für einen Unterschied gefunden, das heißt, entweder Vorteil oder vergleichbar. Das heißt, es ist an verschiedenen Stellen dargelegt, es gibt Verfahren, die

im Sinne der von Ihnen definierten patientenrelevanten Ziele vergleichbar sind. Das ist ja dokumentiert – das können wir zur Kenntnis nehmen –, das sind ja Ihre Daten. Das heißt, wenn man es zusammenfasst – Seite 1 –: Bei der Symptomatik können in 5 von 11, bei der Lebensqualität in 7 von 9, bei der Krankenhausverweildauer bei allen Verfahren und bei der Katheterisierungsdauer in 8 von 9 beurteilten Verfahren eine Vergleichbarkeit oder Überlegenheit, wie Sie es hier definieren, gegenüber Standard dokumentiert werden. So weit, so gut. Das war jetzt nur Krankenhausverweildauer und Katheterisierungsdauer. Ich denke, es ist unbestritten, dass zumindest die Katheterisierungsdauer für einen Mann ein Problem ist, nämlich muss ich überhaupt einen Katheter haben und wie lange muss ich den tragen. Das ist in Ihrem Methodenpapier auch so definiert. Da steht auf Seite 40:

„Die Inanspruchnahme von Leistungen des Gesundheitssystems kann ebenfalls eine patientenrelevante Größe sein.“

Das heißt, das sind patientenrelevante Größen, die wir gleichartig definieren müssen im Sinne der Verbesserung der Symptomatik und Lebensqualität, was wir, wenn wir das richtig interpretieren, als Hauptkriterien Ihrer Ziele gesehen haben.

Jetzt kommen wir zu den UEs.

PD Dr. Stefan Lange: Ich würde zunächst etwas sagen, weil es sonst zu kompliziert wird.

Der Satz, den Sie vorgelesen haben, erstreckt sich über sechseinhalb Zeilen. Er beschreibt genau das, was Sie gerade selbst beschrieben haben und gleichzeitig infrage stellen. Sie haben nämlich gesagt, vorausgesetzt die kommen zu den gleichen Ergebnissen wie die TURP. Genau das haben wir gemacht. Wir haben in der Methodik – wir wollen ja im Wesentlichen über die Methodik sprechen – die Zielgrößen aufgeschrieben. Da steht an erster Stelle die Beeinflussung der Symptomatik. Sonst würden die ja gar nicht zu Ihnen kommen. Wenn die nicht zu Ihnen kämen, dann würden die auch nicht einen Katheter bekommen müssen, weil sie plötzlich bluten, oder sie würden nicht impotent werden, weil da jemand an ihren Nerven herumarbeitet. Das ist mit der Prämisse gemeint. Das heißt, die verminderte Morbidität kann doch nur dann zum Zuge kommen, wenn wir wissen, dass man in Bezug auf die Symptomatik zu einem vergleichbaren Ergebnis kommt.

(Prof. Klaus Höfner: Da gibt es keinen Dissens!)

Ich glaube, das Problem ist, dass wir für den Aspekt der Vergleichbarkeit keine Problematisierung in allen Berichten, die es dazu gibt, und in den Primärstudien gefunden haben. Was heißt denn „vergleichbar“? Ist der IPSS noch 3, 2, 5, 10 Punkte schlechter? Das finden Sie nirgendwo. Wir können ja – Sie haben ja ein Papier zitiert – noch einmal darauf zurückkommen.

Jetzt komme ich zum zweiten Punkt. Sie haben nämlich ein bisschen unsere Tabelle im Bericht, wo wir die Ergebnisse zusammengefasst haben, missinterpretiert, dass es eben, wenn ein Pfeil nicht nach oben oder unten ist, sondern ein Doppelpfeil, nicht bedeutet, da gibt es keinen Unterschied, sondern es gibt keinen Beleg für einen Unterschied, was aber eben nicht gleichbedeutend ist mit, es gibt einen Beleg für einen nicht vorhandenen Unterschied. Dieser Beleg impliziert, dass man sich vorher darüber Gedanken macht, was Vergleichbarkeit ist. Das ist das, was wir mit dem Satz über sechseinhalb Zeilen beschreiben wollten. Für mich wäre es heute ein ganz wesentliches Ziel, dass wir da weiterkommen und dass Sie uns da eine Hilfe geben: Was heißt Vergleichbarkeit?

Moderator Dr. Thomas Kaiser: Wir sollten erst einmal dabei bleiben, wie es mit der Methodik aussieht. Wir werden ja auch auf Basis dieser Überlegungen zur Methodik, zum Beispiel auch was Vergleichbarkeit ist, gegebenenfalls zu Änderungen in Formulierungen des Fazits des Vorberichts oder Darstellung in der Tabelle kommen. Das ist durchaus möglich. Voraussetzung ist aber zunächst, dass man sich über die methodischen Dinge Gedanken macht. Wir sollten jetzt nicht das Fazit und auch nicht die Abwägung dieser Dinge gegeneinander diskutieren.

Prof. Klaus Höfner: Das, was in unserem Bericht steht, ist ja das Zitat, also das, was Sie selbst geschrieben haben. Es ist nichts hineininterpretiert, sondern es wird exakt zitiert, was in dem Bericht steht. Wenn man den Gesamtzusammenhang sieht, dann haben wir den Eindruck – vielleicht ist der falsch –, dass am Ende die niedrige Morbidität nicht den Stellenwert hat in der Interpretation. Insofern ist der Ausgangspunkt, dass man zunächst einmal das als vergleichbares patientenrelevantes Ziel identifiziert und dieses auch so interpretiert und so bis zum Ende durchdekliniert, also nicht die Einschränkung macht. Diesen Satz muss man als Einschränkung definieren. Dass Vorteile gegenüber Standard – ich verkürze das – in Bezug auf Krankenhausverweildauer und Katheterisierungsdauer geltend gemacht werden können, ist ein Zitat aus Ihrem Bericht. Dort steht ja, dass unter der Prämisse, dass die Beeinflussung der Symptomatik und gegebenenfalls damit verbunden der Lebensqualität vordringliches Ziel der Behandlung eines BPS ist, für keines der Verfahren letztendlich ein Zusatznutzen gegenüber einer Standardtherapie als belegt angesehen werden.

Moderator Dr. Thomas Kaiser: Ich denke, das ist ein Punkt, der erst einmal so verstanden wurde. Es ist etwas, was möglicherweise durch die Art der Intervention selbstevident ist. Ihre Anregung ist, das in der Schilderung etwas anders abzuwägen. Aber lassen Sie uns nicht am konkreten Fazit diskutieren, weil andere Aspekte dazu noch einfließen werden, zum Beispiel auch die Definition der Vergleichbarkeit. Aber diesen Punkt haben wir für eine Überarbeitung des Vorberichts aufgenommen.

Prof. Klaus Höfner: Ich komme dann nun zu den unerwünschten Ereignissen. Ich zitiere wieder Ihren Bericht, und zwar zu Tabelle 157: Für einen Unterschied zugunsten der untersuchten Verfahren gegenüber Standard ergaben sich Hinweise auf schwere Blutung bei 5 von

5 berichteten Methoden, Ejakulationsdysfunktion bei 5 von 5 berichteten Methoden, erektile Dysfunktion bei 2 von 2 berichteten Methoden, Strikturen bei einer von einer berichteten Methode. Da gibt es natürlich auch Hinweise auf den Unterschied zugunsten Standard. Das heißt, man sieht, dass bei der genaueren Betrachtung der unerwünschten Ereignisse – das sind Zitate Ihres Berichtes – ein Unterschied zugunsten der Verfahren gegenüber Standard definiert sind. Dann wird wieder eine Interpretation vorgenommen, die man als Leser nicht nachvollziehen kann. Auf Seite 389 steht, dass in der Gesamtbetrachtung der UEs für die Vergleiche mit Standard bezüglich sexueller Dysfunktionen und der behandlungsbedürftigen Blutungen Hinweise darauf vorliegen, dass die Prüfinerventionen der Standardtherapie überlegen sind. Am Rande bemerkt: Hier wird von Überlegenheit gesprochen, während in den meisten Fällen die Definition von Überlegenheit abgelehnt wird. Das ist ebenfalls eine Originalformulierung des Berichts. Das heißt, wir stellen fest, bei der Frage, wie diese UEs einzustufen sind - -¹² Ist schwere Blutung wichtiger, als wenn jemand eine Retention bekommt? Wir meinen ja, denn das sind die Hauptkriterien der UEs. Das heißt, wir empfinden keine Unterscheidung der UEs aus Patientensicht. Wir sagen also, Blutung ist ein ganz wichtiges Kriterium. Ich zitiere hier den BQS-Bericht. Im BQS-Bericht sind die meisten Komplikationen der TURP die Blutung, 3,63 %, 31 771 Patienten. Das heißt, wenn eine Methode dieses gerade nicht tut, dann ist das ein wichtiger Punkt. Die Impotenz und auch die Ejakulationsdysfunktion sind ebenfalls wichtige Punkte.

Moderator Dr. Thomas Kaiser: Ich möchte Sie kurz unterbrechen. Wir sollten zu den methodischen Aspekten kommen und nicht zur Zusammenschau. Die Zusammenschau wird ja hinterher die Folge sein. Wenn ich das richtig verstehe, dann haben Sie zwei Grundaussagen getroffen: Erstens. Das mit den Komplikationen ist auch wichtig. Es gehört also mit in die Betrachtung.

(Stefan Zenk: Als Zielgröße!)

– Als Zielgröße. Das ist ja auch genannt; ist ja aufgeführt.

(Stefan Zenk: Nicht in den Ergebnistabellen!)

– Ich sehe in den Ergebnistabellen Zielgrößen, unerwünschte Ereignisse. - Sie haben eben gesagt, Sie halten die Abwägung für nicht ganz gelungen.

(Dr. Richard Berges: Falsch gewichtet!)

Das heißt, Sie sagen, es ist zwar aufgeführt, es ist auch genannt, aber die Abwägung halten Sie für nicht gelungen. Aber das ist genau Gegenstand des Vorberichts.

¹² Abbrechen des Satzes

Zweitens. Sie sehen innerhalb des großen Komplexes Komplikationen, unerwünschte Ereignisse, in dem Sinne nichts Herausragendes, sondern Sie sagen, alles ist wichtig. Zumindest haben Sie vorhin diesen Satz so gesagt.

Prof. Klaus Höfner: Ich sehe bei den UEs schon Unterschiede. Ich sagte gerade: Wir haben ja Vorteile der untersuchten Verfahren gegenüber Standard und Nachteile gegenüber Standard, sodass wir am Ende eine Wertung vornehmen müssen, dass wir sagen, was im Sinne des patientenrelevanten Ziels wichtiger ist. Das ist eine Sache, zu der man hinterher in der Interpretation kommt. Das ist ein ganz wichtiger Punkt. Es ist eine Frage, die man jetzt sicherlich nicht entscheiden kann. Aber ich sage: Die Probleme Blutung und Impotenz haben sicherlich eine höhere Wertigkeit als das, was unten drunter steht.

PD Dr. Stefan Lange: Sie sagen also, man muss versuchen, eine Hierarchisierung bei den Komplikationen vorzunehmen?

(Prof. Klaus Höfner: Richtig!)

Dr. Kai Schorn: Herr Höfner hat Herrn Lange nonchalant zugestimmt, dass die Prämisse ist, wenn man die Nebenwirkungen und solche Schichten betrachtet, dass man die gleiche Effektivität von den Verfahren nachgewiesen hat bzw. die Symptomreduktion gleich ist. Da waren sich ja alle am Tisch einig. Die Frage ist, ob diese Prämisse aus Patientensicht überhaupt stimmt, ob nicht die Realität so aussieht, dass der Patient eine Abwägung vornimmt, indem er sagt, ich nehme möglicherweise eine etwas geringere Effektivität in Kauf bei deutlich geringeren Komplikationsrisiken. Die Frage ist, ob die Nutzenprämisse, wir müssen erst einmal verlangen, dass die Verfahren von ihrer Symptomreduktion und den Lebensqualitäts-Outcome her gleich sind, und dann können weitere Aspekte wie Krankenhausaufenthaltsdauer und Komplikationen mit einbezogen werden, die Realität widerspiegelt oder ob das mit den Patienten überhaupt rückgespiegelt ist, dass das für die Patienten wirklich ein relevanter Outcome ist. Was ist für den Patienten relevant? Kann er möglicherweise in seinem Abwägungsprozess eine geringere Symptomreduktion in Kauf nehmen, weil er bestimmte Nebenwirkungen hundertprozentig ausschließen möchte?

Moderator Dr. Thomas Kaiser: Ich sehe das von der globalen Betrachtung her zunächst als einen berechtigten Punkt. Da sind Sie natürlich sofort dabei, dass die Sham-Intervention mit geringster Komplikationsrate eingeht, aber ebenfalls symptomreduziert. Das ist also eine ganz schwierige Sache.

PD Dr. Stefan Lange: Herr Schorn, wir sind hier ganz nahe beieinander. Das entbindet Sie aber nicht davon, dass Sie definieren müssen, was Sie als noch akzeptable Unterlegenheit definieren.

(Dr. Kai Schorn: Richtig!)

Das ist etwas, was in der Literatur nicht passiert ist.

Dr. Kai Schorn: Habe ich Sie nicht richtig verstanden: In Ihrem Methodenpapier haben Sie den Anspruch, wenn das in der Literatur nicht definiert ist, zu versuchen, das mit Patienten zusammen zu definieren?

PD Dr. Stefan Lange: Nein, da haben Sie uns falsch verstanden.

Prof. Klaus Höfner: Das Problem, das wir hier haben, ist interessant. Da gibt es keine Studie, aber die Realität zeigt es uns jeden Tag. Die symptomatische Verbesserung eines Medikaments liegt im Durchschnitt bei 40 %, also weit unter dem, was wir mit allen Interventionen erreichen, egal welche Intervention. Das heißt, der Patient entscheidet ja und weiß auch, weil er das gesagt bekommt, ich nehme ein Medikament und erreiche einen symptomatischen Benefit von 40 %. Mit einer Therapie liege ich bei 80 % und bei (...) ¹³ bei 60 %. Das ist in etwa die Relation. Das heißt, das ist bereits entschieden. Das heißt, darüber müssen wir gar nicht reden, dass ein Patient akzeptiert in der Wahl eines Verfahrens mit einer geringeren Morbidität, dass er einen geringeren Benefit hat. Das passiert jeden Tag. Das Medikament ist nicht unser Ding jetzt hier, aber trotzdem ist das die Realität, die jeden Tag passiert. Das ist genau das, was er sagt. Das heißt, der Level ist bereits entschieden. Der Patient akzeptiert 40 %. Das können Sie überall nachlesen. 40 % Benefit ist das, was wir mit einem Medikament erreichen bei viel geringerer Morbidität.

Moderator Dr. Thomas Kaiser: Aber es widerspricht nicht der Aufgabe des Berichtes, zunächst zu beschreiben, was bezüglich Zusatznutzen oder Nichtzusatznutzen Gleichwertigkeit hinsichtlich Symptomen ist und was bezüglich Gleichwertigkeit Zusatznutzen und Komplikationen nachgewiesen ist. Das ist ja das Ziel des Berichts. Was die Entscheidungsträger daraus machen, ist ein anderer Punkt. Hier ist Ziel zu beschreiben, ob es – dafür fehlt uns leider die Definition – den Nachweis einer Gleichwertigkeit, den Nachweis eines noch akzeptablen Unterschiedes gibt. Das ist ja genau der Punkt. Es geht erst einmal darum, das zu beschreiben.

Dr. Richard Berges: Es gibt diesbezüglich sicherlich eine Wertung patientenseitig, denn sonst würden ja alle Patienten eine Operation wollen. Die würden auch nach nichts anderem fragen. Dazu gibt es auch Untersuchungen – wir haben das ja auch nachgereicht –, auch aus den USA, obwohl die Situation dort anders aussieht, weil nicht alle Patienten grundsätzlich schon medikamentös therapiert wurden, wie das leider in Deutschland der Fall ist. Es gibt aber auch noch Unterschiede bezüglich der Verfahren. Da verweise ich noch einmal auf die Grobeinteilung, sekundär ablativ und primär ablativ. Durch die technischen Unterschiede in diesen Verfahren sind Sachen systemimmanent und deswegen auch nicht gleich zu werten. Das ist das, was Sie mit dem Katheter ansprachen. Wenn ich als Patient von vornherein weiß, ich bekomme einen Katheter für diese Behandlung, weil es ein sekundär ablatives Verfahren

¹³ In der Aufzeichnung nicht zu verstehen

ist – das ist mir aber nicht so wichtig wie die Tatsache, dass ich keine Narkose und kein Krankenhaus brauche –, dann ist das ganz anders zu werten, als wenn ich ein primär ablatives Verfahren mit der TURP vergleiche und da nach Katheterisierungsdauer schaue. Das ist ja eine ganz andere Wertung. Das muss definiert sein.

Dr. Christoph von Zastrow: Herr Höfner hat ja die Zahlen der BQS wegen der Komplikationsrate zitiert. Die sind nur bedingt so zu interpretieren. Es wird nur erfasst, zum Beispiel bei Blutungen, wer Bluttransfusionen postoperativ bekam, aber es wurde nicht differenziert, wer schon präoperativ aus anderen Gründen eine Anämie hat. Das heißt, die Blutgabe war nicht aufgrund eines operationsbedingten Blutverlustes notwendig. Nach meiner Kenntnis ist nur die Hälfte der Bluttransfusionen wegen wirklich operationsbedingter Blutverluste notwendig und die andere Hälfte aufgrund anderer Umstände, die mit der Operation primär nichts zu tun haben.

Prof. Klaus Höfner: Ich habe das vorhin nicht so gesagt, aber wir müssen eine sehr viel höhere Rate einer signifikanten Blutung annehmen im Vergleich zu den Alternativen, weil ja hier nur 3,63 % transfusionspflichtige Blutungen waren. Es gibt also eine ganz gute Evidenz für die Verbesserung der Modifikation der UEs¹⁴. Es gibt noch keine Daten, aber das, was wir bisher wissen, ist, dass diese Verfahren sehr viel weniger Blutung erzeugen. Das ist also eine Zahl, die sehr ernst zu nehmen ist, weil es hier transfusionspflichtige Blutungen sind. Ich denke, die Anzahl der signifikanten Blutungen ist weit höher. Das ist auch das, was der Patient weiß. Das ist ja das, was wir immer wieder kommunizieren und wo der Patient sagt: Nein, das ist ein Blutbad. Ich möchte das nicht haben. Aber Sie haben natürlich Recht, dass nicht alles auf die Operationen bezogen ist.

PD Dr. Stefan Lange: Das ist ein ganz wichtiger Punkt, was es wiederum so schwierig macht, diese Daten außerhalb einer vergleichenden Studie zu interpretieren. Wenn das jetzt eine vergleichende Studie wäre, dann hätten alle die gleichen Ausgangsbedingungen. Wir könnten es genau interpretieren. Dann würden wir sehen, hier sind es nicht die 3 %, sondern 1,5 % oder 1,2 % und da 0,1 %, weil vielleicht auch einmal eine Komplikation eingetreten ist, weil ich zu viel koaguliert oder vaporisiert habe. Das kann man sich ja auch vorstellen. Das macht es eben so schwierig, solche Daten vernünftig zu interpretieren.

Prof. Klaus Höfner: Wenn ich 2 Gruppen von Patienten habe, dann muss man ja unterstellen, dass es in beiden Gruppen das gleiche Maß an nicht operationsbedingten Transfusionen gibt. Wir haben ja Daten zu Blutungen; das steht ja alles da. Wir haben schwere Blutungen (...)¹⁵

PD Dr. Stefan Lange: Richtig, vergleichende Daten, das ist ja nicht die Frage.

¹⁴ In der Aufzeichnung nicht deutlich zu verstehen

¹⁵ In der Aufzeichnung nicht zu verstehen

Moderator Dr. Thomas Kaiser: Es geht um die Registerdaten BQS. Da haben Sie eine Selektion von Patienten zu verschiedenen Verfahren. Das ist etwas ganz anderes. Darauf ist Herr Lange eingegangen.

PD Dr. Stefan Lange: Eben. Dazu brauche ich nichts mehr zu sagen.

Ich werde gerne darauf zurückkommen. Herr Höfner, ich kann das nicht akzeptieren. Das ist mir zu roh, zu sagen, da die 40 %, die akzeptieren 60 %, die akzeptieren 80 %. Jetzt müsste man noch versuchen, das in IPSS-Punkte umzurechnen und dann zu sagen, ob das wirklich ein irrelevanter Unterschied ist. Das hört sich ein bisschen an wie Abstimmung mit den Füßen. Das ist ein bisschen aus der Erfahrung heraus. Solche Daten kann man auch strukturiert gewinnen. In anderen Indikationen gibt es so etwas, dass man tatsächlich Patienten befragt – Herr Schorn hatte das gerade angesprochen –, aber das können wir nicht machen, sondern das müssen andere machen, was sie bereit sind, für einen bestimmten Zusatznutzenvorteil in Kauf zu nehmen an Unannehmlichkeiten, geringere Effektivität oder sonst etwas. Dazu gibt es relativ wenig. Sie haben die Arbeit von Barry genannt, die da sicherlich auch hilfreich ist. Aber wenn man sich diese genau anguckt, dann stellt man fest, dass das viel geringere Bereiche sind, die von Patienten offensichtlich schon empfunden werden. Jetzt kann man sich fragen, wie man das bewertet. Schon eine Veränderung, die dort beschrieben ist – es hängt zum Teil davon ab, von welchem Niveau diese Patienten ausgehen –, um 2 Punkte werden von Patienten als spürbar beschrieben. Das ist nicht viel. Wir kommen jetzt zu der zentralen Frage: Sind es 2 Punkte, 4 Punkte, 6 Punkte, 8 Punkte?

Moderator Dr. Thomas Kaiser: Aber ohne Abstimmung, sondern die Frage ist: Was ist die wissenschaftliche Evidenz dafür, dass 3 Punkte und nicht 2 Punkte?

Dr. Richard Berges: Sie zitierten ja bereits Barry und sehen auch, dass es sehr schwierig ist, eine ganz klare Abtrennung vorzunehmen, weil die Erfahrbarkeit der Verbesserung eben von der Höhe der Ausgangssymptome abhängig ist. Vieles davon liegt an dem Instrument IPSS. Das ist ja ein endlicher Score, der bei 35 Punkten aufhört. Es ist viel einfacher, einen Patienten im unteren Bereich nach oben und nach unten bewegen zu lassen als ganz oben und ganz unten – das hat schon einen limitierenden Faktor. Das ist eben ganz gut untersucht worden, dass ein Patient in dem starken symptomatischen Bereich, also über 20, schon sehr viele Punkte Verbesserung braucht, um diesen Unterschied erfahren zu können. Im mittleren Bereich ist das halt entsprechend etwas geringer. Im unteren, mildereren Bereich ist das sogar in dem Bereich, wo Sie sagen, 2 Punkte machen da schon etwas aus. Nur das sind nicht die Patienten, die wir typischerweise operieren, sondern die typischerweise operierten Patienten liegen im mittleren oder hoch symptomatischen Bereich, denn das ist ja die Indikation, die oft in diesen Studien verwandt wurde.

PD Dr. Stefan Lange: Da haben Sie völlig Recht. Für unsere Frage interessiert aber nicht der Ausgangswert, wo die Patienten herkommen, denn die werden ja durch beide Verfahren in genau diesen moderaten Bereich gebracht. Jetzt müssen wir im moderaten Bereich gucken,

was hier noch irrelevant ist. Das heißt, die kommen von 20, 22 auf 12, 13 herunter, und hier müssen wir fragen, was hier noch ein irrelevanter Unterschied ist, den der Patient praktisch nicht merkt oder aus bestimmten Gründen nicht merken will, weil er sagt, mir ist es schon recht, wenn ich nicht transfundiert werden oder keinen Katheter bekommen muss. Aber da brauchen wir die Aussage. Nach Barry sind das diese 2 Punkte; das ist nicht viel.

(Prof. Dr. Rolf Muschter: 3 Punkte stehen bei Barry!)

– Aber im unteren Bereich, also unter 20 – er hatte das so aufgeschlüsselt –, wurde 2 Punkte gesagt. In dem Bereich bewegen wir uns.

Moderator Dr. Thomas Kaiser: Es gibt jetzt ja mehrere Äußerungen zu 2, 3, 6 Punkten. Gibt es denn über die Barry-Arbeit hinaus eine Arbeit, die sich diesem Thema widmet, die Sie nennen können, die Sie in den Stellungnahmen genannt haben? Ich glaube, außer Barry ist in den Stellungnahmen nichts genannt.

Dr. Richard Berges: Man müsste wirklich in die Literatur einsteigen. Das ist sicherlich eine Wochenaufgabe, da nachzugucken. Es gibt eine ganze Reihe von Untersuchungen über die Erfahrbarkeit von Veränderungen beim IPSS. Diese sind vornehmlich in den WHO-Leitlinien – diese habe ich Ihnen geschickt – beleuchtet. Da müsste man noch einmal nachschauen. Das ist sicherlich ein tagesfüllendes Geschäft. Aber dazu gibt es mehr Daten als nur die Barry-Arbeit. Die Barry-Arbeit ist ein Beispiel dafür.

Moderator Dr. Thomas Kaiser: Das heißt, konkret können Sie keine benennen, aber in der WHO-Leitlinie ist darauf verwiesen, was es da eigentlich an Literatur gibt.

Dr. Richard Berges: Es gibt eine Fülle von Untersuchungen zum IPSS-Score und die Erfahrbarkeit der Veränderungen dazu.

PD Dr. Stefan Lange: Ich glaube, das können wir versuchen, auf uns zu nehmen. Die Problematik ist ja eigentlich da, also dass man das schon machen muss, um das einem Patienten auch sagen zu können: Dafür, dass du die und die Komplikationen mit einer geringeren Wahrscheinlichkeit oder gar nicht bekommst, musst du damit rechnen, dass du vielleicht so und so viele Score-Punkte schlechter abschneidest. Nach bisherigen Erfahrungen aus der Literatur kann man sagen, das ist für dich eigentlich gar nicht erfahrbar. So müsste man das doch eigentlich beschreiben. So ist auch unsere Vorgehensweise der Methodik, ohne das aufs Fazit eingehend zu verstehen. Das ist der erste Punkt. Deswegen kommen die Patienten. Sonst könnte man sie auch nach Sham operieren. Dann haben sie gar keine Blutungen und sie erfahren trotzdem eine Verbesserung in der Symptomatik. Das ist ja das Schöne dabei. Wenn wir das geschafft haben, dann können wir gucken, bei den Komplikationen gibt es gegebenenfalls den oder den Zusatznutzen oder auch – jetzt kommt wieder eine Hierarchisierung – Zusatzschaden, also Besserung schwere Blutungen, aber dafür vielleicht je nach Verfahren häufiger

Retentionen, was ja vielleicht auch nicht so schön ist. Das muss man dann sehen. Dann muss man gegebenenfalls hierarchisieren.

Prof. Klaus Höfner: Ich denke, Sie haben das Problem, weil Sie sagen, die Einteilung, die Sie vorhin formuliert haben, ist Ihnen zu grob. Wenn man sich jetzt auf einen Level einrichtet, was ja berechtigt kritisiert wird, zu sagen, es muss unbedingt sein, dass die Verfahren, wenn wir über die Komplikationsrate reden, den gleichen symptomatischen Benefit haben – das ist ein Ausgangspunkt –, dann sind wir zumindest auf der sicheren Seite. Dann können wir wirklich nur noch über die Komplikationen reden. Aber er hat natürlich schon Recht – deswegen habe ich das Beispiel gebracht –: Es ist natürlich klar, dass ein Patient sagt, ich akzeptiere durchaus einen geringeren Benefit an Symptomatik. Dafür gibt es keine Daten. Wir wissen nur, wenn wir das retrospektiv vergleichen, dass es in der Tat so ist, dass den schlechtesten symptomatischen Benefit – das ist bei den Studien klar dokumentiert – Medikamente erzeugen, und zwar bei einer anderen natürlich im Sinne von Blutungen und Interventionen geringeren Morbidität. Das ist die Realität, die wir haben – das ist nicht unser Thema –, aber das ist die Realität. Deswegen habe ich das angesprochen. Wenn Sie alles über ein Knie brechen, dann sehen Sie immer, wenn Sie es herunterrechnen, wie viel Prozent Besserung zum Ausgangswert denn da ist. Das ist natürlich auch für andere Verfahren gemacht worden. Was das Volumen, das die leisten, angeht – nicht in Punkten, sondern in Prozent –, liegen Sie bei 40 % bei den Medikamenten, bei etwa 60 bis 70 % bei den Interventionen und bei den Operationen liegen Sie am höchsten. Das ist die Realität. Im Sinne der Evidenz, wie Sie es verlangen, gibt es natürlich keine Daten, aber das ist die Realität.

Dr. Richard Berges: Ich möchte noch einmal auf die Erfahrbarkeit eingehen. Sie sehen das sehr gut an den Medikamentenstudien, die zum BPS durchgeführt worden sind. Sie haben in der Regel bei einem Alphablocker, der als standardmedikamentöse Therapie gilt, eine über den Placeboeffekt hinausgehende Verbesserung von gerade einmal 2 IPSS-Punkten. Selbst da ist es schon fraglich, ob der Patient diese beiden Punkte erfahren kann. Der Grund, warum er eine Besserung erfährt, ist der riesige Placeboeffekt von fast 80 % bei diesen Medikamenten. Dies ist auch der Grund, warum sich so hartnäckig die Phytotherapeuten halten, weil der Placeboeffekt bei dieser Geschichte so groß ist. Jetzt sagen wir, von allen unseren operativen Techniken sind wir um Größenordnungen besser in der Symptomreduktion, und versuchen jetzt, herauszubekommen, ob ein Verfahren besser ist als das andere. Ich glaube, das wird sehr schwierig, weil diese Grenzen für den Patienten wahrscheinlich nicht erfahrbar sind. Ergo bleibt nur noch übrig, nach dem zu gucken, was den Patienten in die anderen Therapien treibt, Laser, TUMT oder TUNA, weil er eben Angst vor der Operation hat. Das hat für die Patienten – das ist natürlich unterschiedlich – einen sehr großen Stellenwert. Das ist ja eine Abstimmung mit Füßen. Die Leute rennen ja dort hin, wo der Laser steht, nicht weil der Operateur so schlecht ist – das können sie gar nicht beurteilen –, sondern weil sie einfach Angst haben vor der Operation. Das ist von der Patientenseite her ein sehr wichtiger und maßgeblicher Aspekt. Der nächste Patient sagt, ich will nicht ins Krankenhaus und keine Narkose. Das

ist für diesen Patienten ein maßgeblicher Aspekt. Deswegen ist dem auch egal, ob er dann beim sekundär ablativen Verfahren einen Katheter tragen muss.

Moderator Dr. Thomas Kaiser: Aber die Aussage in dem Bericht ist ja nicht, dass die Überlegenheit bezüglich der Symptome Voraussetzung für den Nutzen ist. Sie haben gerade geschildert, da kann man keine Überlegenheit nachweisen. Das ist ja nicht die Voraussetzung, sondern die Frage ist, ob man zumindest auf die Gleichwertigkeit kommt. An diesem Problem hangeln wir uns gerade entlang. Das ist ja ein Unterschied.

Dr. Richard Berges: Ich glaube, das ist aber gezeigt worden – auch in Ihren Daten –, dass diese Gleichwertigkeit - ¹⁶ Sie kommen vielleicht zu einer statistischen Signifikanz, zum Beispiel 2 Punkte zugunsten der TURP, aber die Frage ist: Was ist das im Vergleich zu der Gesamtreduktion der Symptome, die Sie erreichen? Das ist ein ganz kleiner prozentualer Anteil. Das heißt, das ist für den Patient irrelevant. Der Patient wird, wenn Sie ihn fragen würden, welche Operation er denn hatte, Ihnen das anhand der Symptomverbesserung nicht sagen können.

Moderator Dr. Thomas Kaiser: Der mit 2 Punkten Unterschied nicht, aber der mit 10 Punkten schon.

Dr. Richard Berges: Wenn es den gäbe.

Moderator Dr. Thomas Kaiser: Den gibt es ja. Sie haben ja eine Mittelwertsbetrachtung.

PD Dr. Stefan Lange: Jetzt müssen wir aufpassen. Es geht schon um 2 Punkte. In Barry steht: 2 Punkte erfahren wir als ein „slightly improvement“, also als eine leichte Verbesserung. Wahrscheinlich können die das schon spüren. Das ist aber ein individuelles Kriterium, also wie sich der Patient verändert. Jetzt muss man einmal differenzieren, was ein irrelevanter Unterschied zwischen Gruppen ist. Da kann man verschiedenste Berechnungen anstellen; dazu möchte ich jetzt nicht kommen. Wir kommen nicht darum herum. Ich finde, Sie drücken sich ein bisschen darum. Ich nenne ein Beispiel: Wenn wir Arzneimittel betrachten und das Arzneimittel sagt, ich habe einen Vorteil bei Nebenwirkungen als Durchschnittsbehandlung von Patienten mit akutem Herzinfarkt, ich glaube, ich bin ähnlich gut, was die Verhinderung von 30-Tage-Sterblichkeit angeht, dann sagen die Behörden: Das ist ja schön und gut. Wir finden das auch toll, dass ihr weniger Hirnblutungen verursacht, aber ihr müsst uns in einer Studie nachweisen, dass ihr tatsächlich vergleichbar seid, was die 30-Tage-Sterblichkeit angeht. – Dann sind die gefordert, zu sagen, damit wir das überhaupt nachweisen können, müssen wir uns vorher Gedanken über Gleichwertigkeiten machen, damit wir das in einer vernünftigen statistischen Hypothese überhaupt in einer Studie überprüfen können. Genau das brauchen wir hier. Das heißt, Sie sagen, die 3, 4, 5 Punkte Unterschied sehen wir doch hier. Das ist doch nichts. Wir müssen das aber im Prinzip vorher definiert haben. Denn jetzt ist es

¹⁶ Abbrechen des Satzes

schwierig. Jetzt sind wir post hoc. Jetzt sehen wir die Ergebnisse. Eigentlich hätte man sich vorher darüber Gedanken machen müssen, damit ich es überhaupt nachweisen kann, was ist etwas, was ich noch als gleichwertig betrachte, was nicht bedeutet, dass die exakt gleich gut sein müssen, sondern wo man vielleicht ein bisschen mit einbezieht, dass es gegebenenfalls geringere Komplikationen bedeutet. Das ist ein Mangel, der etwas schwierig ist zu heilen, weil es in den Studien und in der Literatur nicht passiert ist. Wenn man sich das aber in den wenigen Studien, die das berichten, ansieht, dann stellt man fest, dass die Fallzahlplanung meistens auf einem Gruppenunterschied von etwa 2 bis 3 Punkten beruht. Das wiederum impliziert, dass es sich dabei um einen relevanten Unterschied handelt, weil der relevante Unterschied die Grundlage für die Fallzahlplanung ist. Das heißt, der irrelevante Unterschied muss eigentlich noch geringer sein.

Prof. Dr. Rolf Muschter: Es ist natürlich schwer, Ihnen evidenzgestützt eine Zahl zu nennen. Das ist nicht nur schwer, sondern unmöglich. Ich denke, das hätten wir längst getan, gäbe es die. Dennoch ist die Evidenz, wenn Sie sich die Zahl der Transurethralen Resektionen der Prostata in den USA ansehen, bis zur Einführung der Alphablocker im Prinzip ansteigend bzw. konstant, mit Einführung der Alphablocker stark rückläufig. Das heißt, hier hat es eine Nachfrage von Patienten gegeben, die in dem Moment, wo es ein Medikament gab, die medikamentöse und nicht die operative Therapie gewählt haben, obwohl ganz klar davon ausgegangen werden musste, dass das nicht annähernd die gleiche Größenordnung an Verbesserung erreicht. Sie mögen das angreifen. Das, was ich hier von mir gebe, ist natürlich keine wissenschaftliche Betrachtung, sondern nur eine logische. Aber das ist letztlich die Basis, auf der wir hier nur argumentieren können, in Ermangelung von Evidenz für die Fragestellung, ob 5 Punkte, 4 Punkte, 8 Punkte relevant ist oder nicht. Das ist sehr schwierig. Viele Patienten bleiben erfahrungsgemäß – das ist jetzt Level E; ganz schlechte Expertenmeinung – auf einem Medikament mit einer Besserung von 4 Punkten oder 3 Punkten, weil sie nach wie vor bei der Operation eine viel höhere Morbidität befürchten, auch wenn das vielleicht nur eine fiktive Morbidität ist, die im Einzelfall gar nicht greift. Das ist die Alltagserfahrung, die eigentlich jeder, der in der Praxis ist, macht. Wenn man bei Kongressen von Urologen eine Umfrage macht und fragt, was sie machen würden, dann heben die meisten bei der medikamentösen Therapie die Hand. Das ist die Problematik. Es ist eine subjektive Sache. Das ist das, was das Ganze hier letztlich ausdrückt. Dieser Grund ist individuell und nicht statistisch messbar. Sie können vielleicht herausfinden, wie viele Patienten so entscheiden, aber das ist für die andere Gruppe, die sagt, wir wollen kein invasives Verfahren, immer noch nicht relevant, auch wenn die Mehrheit sagt, das ist uns egal.

Dr. Richard Berges: Ich glaube, dass fast alle diese Studien so angelegt wurden, nicht eine Überlegenheit, sondern eine Gleichwertigkeit zu zeigen, zumindest was die Reduktion von Symptomen bei den klassischen Standardparametern Flow, IPSS, Restharn, Prostatavolumen angeht. Da ist immer versucht worden, eine Gleichwertigkeit zu zeigen. Gleichzeitig ist natürlich geschaut worden, wie es mit den Komplikationen aussieht. Wenn man sich die Gesamtzahl der Studien ansieht, dann stellt man fest, dass man fast überall in relativ gleichen

Größenordnungen liegt, was Ihre Frage zum Zusatznutzen gegenüber der TURP beantwortet. Das Einzige, wo wir einen Zusatznutzen sehen würden, wäre die Seite der Komplikationen. Das ist auch patientenseitig die Seite, die sie sozusagen in die Nachfrage treibt. Das muss man sicherlich herausarbeiten und gucken, ob man dort Unterschiede findet. Ich kann Ihnen am Beispiel der HE-TUMT oder der PLFT-TUMT als modernstes Verfahren unter diesen 3 Entwicklungsstufen nur sagen, das ist das, was viele Patienten wollen. Sie wollen nicht ins Krankenhaus, sie wollen direkt wieder nach Hause und sie wollen keine Narkose, ohne jetzt auf die Nischenpatienten zu schauen. Nischenpatienten gibt es da natürlich auch, zum Beispiel die schwer Herzkranken oder Macumarisierten. Für den anderen Patienten ist das aber keine Option, weil er sagt, ich will lieber morgen pinkeln können. Der rennt aber zum Laseroperateur, weil er dort weniger Angst hat.

Rebecca Deppisch: Zu den Nebenwirkungen: Sie sagen ja, dass eigentlich die Morbidität der zentrale Punkt ist, der auch für die Patienten relevant ist. Das Problem, das ich in den vorhandenen Studien sehe, ist, dass keine einzige dieser Studien unerwünschte Ereignisse definiert hat. Ab wann liegt denn beispielsweise ein Harnwegsinfekt vor? Wir haben häufig das Problem, dass eine Antibiotikaprophylaxe durchgeführt wird und da die Beschreibung fehlt, dass das in den Gruppen sehr uneinheitlich ist. Es fehlen also klare Definitionen für diese Studien der unerwünschten Ereignisse. Aus anderen Bereichen kennt man das nicht so. In Studien im Bereich der Kardiologie werden Endpunkte ganz klar definiert: Ein Herzinfarkt liegt vor, wenn ... Oder: Eine Reintervention ist, wenn das und das passiert. In sämtlichen Studien, die uns dazu vorlagen, fehlt das eigentlich. Somit ist das natürlich überhaupt nicht vergleichbar. Reden denn die Studien generell von dem Gleichen, wenn sie von einer Retention sprechen? Oder gibt es da vielleicht Definitionsunterschiede? Das ist überhaupt nicht nachvollziehbar. Darauf muss geachtet werden, und das haben wir sehr deutlich erwähnt.

Dr. Christoph von Zastrow: Es wurde vorhin gesagt, dass die Patienten abstimmen würden, indem sie bestimmte Methoden vorziehen. Zum Teil ist das aber auch beeinflusst durch die Art der Aufklärung. Ich weiß aus meiner Tätigkeit, dass es durchaus Anwender gibt, die die TUR dem Patienten mit Komplikationsraten schildern, die mindestens um den Faktor 10 über dem liegt, was allgemein üblich ist, sodass natürlich damit der Patient auf diese Weise gewarnt wird. Das muss man also regelrecht kritisch sehen.

Dr. Richard Berges: Das finden Sie überall. Grund dafür sind sicherlich auch pekuniäre Interessen. Das ist, glaube ich, nicht Gegenstand der Diskussion.

Wichtig finde ich, die Definition der Nebenwirkungen aufzugreifen. Ich glaube, das ist wirklich ein Problem, und das ist, wenn man sich viele Studien ansieht, nicht gut beschrieben. Da gehe ich vollkommen d'accord: Man liest sich durch den Methodenteil und sucht diese Sachen. Es gibt aber auch andere Untersuchungen, zum Beispiel De la Rosette-Untersuchungen, wo klar definiert ist, was eine Reintervention ist, zum Beispiel schon der Alphablocker und nicht nur eine Reoperation. Das gibt es auch. Man muss dort schon sehr genau hingucken und

eigentlich von Protokoll zu Protokoll hüpfen, was eine Sisyphusarbeit ist, um festzustellen, ob das in der randomisierten Studie – jetzt haben wir wieder den Zusatznutzen – definiert ist oder nicht. Ich kann nur die letzte Arbeit von Mattiasson zitieren, die ich Ihnen geschickt habe. Dort ist das in Teilen der Fall, aber – da gebe ich Ihnen Recht – am Punkt Hämaturie ist das schlecht beschrieben. Da wissen Sie nicht, ob das eine Makrohämaturie oder eine Mikrohä-maturie ist. Ich gehe aufgrund der Verfahrensweise davon aus, dass das eine Mikrohä-maturie ist. Die Relevanz dafür ist auch wieder eine Frage. Das ist die Schwierigkeit.

Dr. Klaus Höfner: Ich denke, das ist alles richtig. Nur eine Hierarchisierung wird das Problem lösen. Harnwegsinfekte entstehen bei jeder transurethralen Maßnahme. Insofern hat das einen besseren Stellenwert als beispielsweise die Blutung. Von daher muss man eine Hierar-chisierung vornehmen, wie Sie das vorgeschlagen haben. Dann muss man sehen, was die so genannten Top-of-the-list – (...) ¹⁷ Schwere Blutungen sind eigentlich dokumentiert. Das heißt, es ist nicht alles schlecht dokumentiert, sondern der Urologe ist vielleicht nur etwas unpräziser als der Kardiologe. Es ist jede Komplikationsrate wichtig, und die ist schlecht do-kumentiert. Die in der Klinik interessanten Dinge kann man aber nur durch eine Unterschei-dung ihrer Wertigkeit vornehmen. Dann, denke ich, wird man besser klarkommen.

Moderator Dr. Thomas Kaiser: Da könnte ja, wenn es denn berichtet worden wäre, das Kri-terium hilfreich sein, das bei Medikamentenstudien für die Bewertung von unerwünschten Ereignissen angewandt wird, also zum Beispiel schwerwiegendes unerwünschtes Ereignis, das so zusätzliche Kriterien zu einem unerwünschten Ereignis setzt, Notwendigkeit einer Krankenhausaufnahme, lebensbedrohlich usw. Das sind ja eigentlich übergeordnete Kriterien, die diese Hierarchisierung erleichtern würden. Ich würde mich sehr schwer tun, innerhalb zum Beispiel von 10 verschiedenen leichten unerwünschten Ereignissen eine Rangfolge zu machen. Das ist ja sehr schwierig. Aber von den tatsächlichen Krankheitswerten her, schwer-wiegend/nicht schwerwiegend, das kann gelingen, wenn es denn beschrieben ist. Das ist ja auch bei der Bewertung von Medikamenten ein übliches Vorgehen, was wir bei den Medika-mentenbewertungen auch so machen.

Ich würde jetzt gerne die beiden Tagesordnungspunkte 4 und 5 zusammenfassen. Nach mei-nem Verständnis war weitgehende Übereinstimmung dahingehend, dass man, wenn man den Zusatznutzen bewerten möchte, also nicht den Nutzen eines Verfahrens selber, sondern den Zusatznutzen im Vergleich zu was auch immer, also – lassen wir einmal den Begriff „Stan-dard“ weg – zum Beispiel zu einer etablierten Therapie –, um vergleichende Studien und in der Regel randomisierte vergleichende Studien nicht herunkommt. Gegebenenfalls können nicht randomisierte Studien ebenfalls hilfreiche Informationen liefern. Für die Bewertung des Zusatznutzens können nicht randomisierte Studien eigentlich nur dann eine Zusatzinformation liefern, wenn dort ein Aspekt gesehen, betrachtet wird, der einen dramatischen Effekt zeigt.

¹⁷ In der Aufzeichnung nicht zu verstehen

Jetzt war es so, dass wir gefragt haben, ob es solche Studien gibt, die einen Hinweis auf solch einen dramatischen Effekt liefern. Sie haben gesagt, das können wir ad hoc nicht sagen. Es wurde gesagt, eine Stellungnahme wird es sein können. Wir haben uns eben kurz besprochen. Es ist ja so, dass ein Protokoll zu der Erörterung geschrieben wird. Das wird ja auch zwei bis 3 Wochen dauern. Dann muss es allerdings auch mit dem Projekt weitergehen. Insofern wäre es hilfreich, wenn Sie uns innerhalb der nächsten 2 bis 3 Wochen solche Studien nennen könnten, die eine wissenschaftliche Evidenz haben, also sagen, zitieren und schicken: Diese Studie deutet auf einen derart dramatischen Effekt hin, dass es erforderlich ist, nicht randomisierte Studien einzubeziehen. Vielleicht ist das ein ganz guter Weg. Es ist ja schon eine besondere Situation mit dem kurzfristig getrennten Verfahren von der Erörterung nur zum Berichtsplan.

Zur Gleichwertigkeit und zu irrelevanten Unterschieden: Ich habe das so verstanden, dass das alle als sehr wichtig betrachten, also die Frage, was eigentlich Gleichwertigkeit ist. Das ist methodisch nicht ganz ohne. Es gibt an uns bekannter konkreter Evidenz zunächst einmal wenig – das ist die Barry-Arbeit –, aber es gibt möglicherweise in Sekundärliteratur noch Hinweise auf andere Arbeiten, die es noch einmal zu sichten gilt, um zu sehen, was dort als irrelevanter Unterschied oder klinisch relevanter Unterschied beschrieben ist, was ja etwas anderes ist. Darauf aufbauend müsste man sich die Bewertung im Vorbericht noch einmal vornehmen und fragen: Zeigt das, was wir in den randomisierten kontrollierten Studien an Unterschieden bezüglich des IPSS-Scores gesehen haben, einen Nachweis eines irrelevanten Unterschiedes, einen Nachweis eines relevanten Unterschiedes, oder wie ist das? Diese Wertung muss also mit hineingenommen werden.

Das wäre meine Zusammenfassung der Diskussion über die Tagesordnungspunkte 4 und 5.

Dr. Richard Berges: Ich wollte nur darauf hinweisen, dass, auch wenn wir die eklatanten Unterschiede benennen, dies nicht die klassischen UEs sein müssen, sondern auch die eklatanten Unterschiede „Klinik ja/nein“, „Narkose ja/nein“. Denn das sind auch eklatante Unterschiede, die in diesem Verfahren liegen.

Moderator Dr. Thomas Kaiser: Wobei das dann nicht in einer Studie bewertet werden muss, sondern das ist ja verfahrensinhärent.

Prof. Klaus Höfner: Eine Bemerkung, die noch nicht berücksichtigt ist – wir kamen einfach nicht dazu, aber das gehört mit dazu –, ist der Parameter zur Messung der Lebensqualität. Wir stellen im Bericht fest, dass der IPSS-1-Item-Score zur Lebensqualität berechtigt kritisiert wird. Er gibt neuere Instrumente. Wir haben das so auch argumentiert. Das, was uns nicht schlüssig erscheint, ist letztlich eine Gesamtaussage – das sollte man bei der Revision vielleicht einmal überprüfen –, dass man sagt, man kann die ganzen Ergebnisse, die uns mit dem 1-Item-Score mitgeteilt werden, deswegen nicht interpretieren, weil der Parameter 1-Item-Score für uns nicht ausreichend ist. Das sollte man noch zu Protokoll geben. Das ist eine Aussage, die wir so nicht gerne sehen. Wir sind der Meinung, dass zu dem Zeitpunkt – das ist

auch bis jetzt nicht hinterfragt - -¹⁸ Es muss erst einmal ein Score entwickelt werden, der in dieser Einfachheit und Globalität so komplex die Dinge erfasst wie der 1-Item-Score. Es gibt sicherlich andere Bereiche, die man da untersuchen kann. Die Gesamtinterpretation ist uns etwas aufgestoßen, weil die Kritik an diesem 1-Item-Score dazu geführt hat, die Ergebnisse zu schmälern. Das sollte man noch aufnehmen.

Prof. Dr. Rolf Muschter: Einen Schritt zurück zu Ihrer Zusammenfassung Punkt 1: Sie sagten, wir hätten noch keine dramatischen Unterschiede in den Studien schildern können. Das ist auch korrekt. Aber ich möchte einen dramatischen Unterschied beispielhaft nennen: Allein die Existenz von Studien, die die Holmium-Laser-Enukleation gegen die offene Operation untersucht, ist so ein dramatischer Unterschied. Die sind ja sogar randomisiert. Ich vergleiche hier einen offenen Eingriff mit einem durchschnittlich zehntägigen Krankenhausaufenthalt mit einer nahezu ambulant durchführbaren Methode, zumindest in manchen Studien. Ich mache es absichtlich ein bisschen dramatischer, als die Wirklichkeit sein wird. Alleine dort haben wir einen sehr dramatischen Unterschied. Offen versus endoskopisch, das ist für Patienten ganz sicher erfahrbar und ganz sicher ein Grund, sich für ein solches Verfahren zu entscheiden, wenn sie im klinischen Outcome noch vergleichbar sind, also vergleichbar mit allen Einschränkungen, die Sie in der Zusammenfassung genannt haben.

Dr. Volker Rohde: Ich wollte kurz auf die Lebensqualität Bezug nehmen. Mitte der 90er-Jahre hat es einen so genannten Minimalkonsens gegeben, der mit diesen 3 Säulen definiert worden ist, im Wesentlichen der somatischen, psychischen Komponente. Diese Komponenten spiegeln sich in dieser einzelnen Frage nicht wider. Das heißt, wenn Sie Lebensqualität abbilden wollen, müssen Sie verschiedene Dimensionen von Lebensqualität abbilden. Das leistet eine Frage zweifelsohne nicht. Ich denke, es ist die Frage zu stellen, ob das nicht eher nur Zufriedenheit ausdrückt. Der IPSS selber ist ja ein Abfragen von Symptomen und spiegelt auch nicht die Lebensqualität unter wissenschaftlicher Definition von Lebensqualität wider; das muss man ganz klar sagen. Das heißt, die Instrumente, um Lebensqualität als nur dimensionales Konstrukt abzubilden, gibt es nicht. Das heißt, da haben wir sicherlich ein Definitionsproblem von Lebensqualität innerhalb des IPSS-Fragebogens. Das ist, denke ich, nicht ganz ohne.

Prof. Klaus Höfner: Das ist richtig. Wir erkennen natürlich, dass es sehr viel komplexere Instrumente gibt, die von Mitte der 90er-Jahre bis heute weiterentwickelt worden sind. Aber es ist trotz dieser gesamten Scores kein Score erkennbar, auch heute nicht, der für uns eine optimale Beschreibung der Lebensqualität eines Patienten zeigt. Deswegen sagen wir: Wenn diese Evidenz nicht besteht und wir heute kein anderes Instrument haben, das uns sagt, dass es besser ist als das, was wir hatten, dann ist die Gesamtinterpretation überzogen. Ich kann nicht sagen, der Score ist aus heutiger Sicht auch in der Entwicklung der ganzen Instrumente unzu-

¹⁸ Abbrechen des Satzes

reichend, weswegen wir alles das nicht akzeptieren oder derartig einschränken, dass wir sagen, das ist eigentlich eine Null-Aussage. Das ist zu viel.

Dr. Volker Rohde: Dann darf man natürlich die Lebensqualität als Parameter in der Beurteilung der ganzen zu vergleichenden Technologien überhaupt nicht zur Anwendung bringen, was hier immer wieder als patientenrelevantes Kriterium aufgeführt worden ist. Denn wenn wir diese Instrumente nicht als Instrumente zur Erfassung von Lebensqualität akzeptieren, dann haben wir keine wissenschaftlichen Daten, mit denen wir das beurteilen können.

Dr. Richard Berges: Herr Rohde, ich muss Ihnen widersprechen. Das geht jetzt in eine ganz falsche Richtung. Wir diskutieren jetzt plötzlich über die multidimensionale Erfassung von Lebensqualität oder gesundheitsbezogener Lebensqualität. Der IPSS hat nie den Anspruch gehabt, genau wie alle anderen Scores, die auch für Miktionsstörungen, also Beschwerden beim Wasserlassen, entwickelt wurden – davon gibt es viele –, multidimensional Lebensqualität zu messen, sondern er misst genau das – das ist in allen Sprachen multipel validiert worden –, was er messen soll, nämlich Miktionsstörungen. Nur das misst er und nur das gibt er vor. Nur deswegen ist er, weil er praktikabel und einfach ist, in den ganzen Untersuchungen angewendet worden, da man ein Instrument brauchte, mit dem man validiert nachvollziehbar und nachprüfbar die Veränderung von Miktionsstörungen messen wollte. Dem angehängt ist ein Globalscore. Auch das ist absolut üblich in der Lebensqualitätsforschung – das macht jeder große multidimensionale Score inzwischen auch –, dass man eine Globalfrage ans Ende hängt, um die ganze Sache in einen Zusammenhang zu bringen. Das ist validiert und richtig und auch vernünftig. Ich meine, es ist höchst unvernünftig, dieses Instrument in seiner Form, wie es vorliegt, zu diskutieren. Das bringt überhaupt nichts. Die Untersuchungen sind alle so gemacht worden, sowohl die vergleichenden als auch die offenen Untersuchungen, um diesen Unterschied zu messen. Dazu gibt es hervorragende Literatur, dass das auch klappt. Der Score misst, was er messen will, nicht multidimensionale Lebensqualität, sondern Miktionsstörungen. Das ist das, was der Urologe behandelt.

Moderator Dr. Thomas Kaiser: Ich sehe zusammenfassend überhaupt keinen Widerspruch, sondern es gibt eben offensichtlich in den Studien berichtete Scores, die keine Lebensqualitätsinstrumente sind, sondern die etwas für sich Abgeschlossenes messen, nämlich Symptomscores. Das kennen wir aus ganz anderen Bereichen ja auch. Bei Asthma ist nicht jeder Symptomscore ein Lebensqualitätsscore, sondern ein Lebensqualitätsscore hat bestimmte methodische Voraussetzungen. Nichts anderes hat Herr Rohde gesagt. Jeden in den einzelnen Studien berichteten Score muss man sich dahingehend anschauen, was er berichtet. Ist es ein Bericht über Lebensqualität mit all den Voraussetzungen oder ein Bericht zum Beispiel über ein Symptom? Nichts anderes muss man machen.

Prof. Klaus Höfner: Ich muss widersprechen. So ist es nicht. Es ist ja eine Frage des IPSS. Das ist dezidiert als Messinstrument der Lebensqualität akzeptiert.

(Moderator Dr. Thomas Kaiser: Von wem?)

– Von Leitlinienkommissionen etc. Es sind Expertengremien, wie auch immer. Das kann man natürlich kritisieren. – Das heißt, wir reden hier über die Frage Lebensqualität bei Miktions-symptomen. Jetzt ist die Frage, die Sie stellen – die ist nicht untersucht –, ob das eine Erfas-sung der Symptomatik oder der Lebensqualität ist. Dafür gibt es keine Evidenz. Es gibt auch keine Evidenz, die sagt, das ist kein Messinstrument für die Lebensqualität. Es ist das eine wie das andere richtig. Es ist nicht multidimensional.

Moderator Dr. Thomas Kaiser: Ich denke, wir sollten den Punkt abschließen. Es gibt klare Übereinkunft zwischen Methodikern, Forschern zu Symptomscores und Lebensqualität, was Lebensqualität misst, was Symptome misst. Dahingehend werden alle diese Scores unter-sucht, ob es Einzelfragen oder mehrere Fragen sind. Wenn die übereinstimmenden methodi-schen Erkenntnisse als Voraussetzung für eine tatsächlich valide Messung von Lebensqualität gegeben sind, dann kann man das betrachten. Wenn es nicht gegeben ist, dann kann man es nicht betrachten. Daran ändert auch nichts, dass andere Leute das für Lebensqualität halten.

(Prof. Klaus Höfner: Der Beweis ist nicht erbracht, dass das nicht die Lebensqualität ist!)

PD Dr. Stefan Lange: Hier muss man überlegen, wer in der Beweispflicht ist. Menschen, die sich mit der Messung der Lebensqualität beschäftigen, tun das nicht erst seit gestern, sondern schon ein bisschen länger. Das heißt, da gibt es vielfältige Arbeitsgruppen und Wissenschaft-ler zu verschiedensten Krankheiten. Die haben sich darüber schon sehr viele Gedanken ge-macht. Inzwischen gibt es auch Guidelines von der FDA, also von Gesundheitsbehörden. Ich denke, wir werden das vielleicht zum Anlass nehmen, zu versuchen, darauf zu rekurrieren, denn dann hat man es auf einer ganz klaren Basis. Lebensqualität ist nicht das, worauf sich Experten einigen, sondern Lebensqualität ist das, wo der Patient extrem im Fokus steht. Da muss man gucken, wie die Instrumente auf diesen Punkt, dass es nämlich den Patienten be-trachtet und was dieser erlebt, entwickelt und validiert wurden. Wenn man etwas Neues rein-bringt, zum Beispiel wenn die Urologen, Leitliniengruppen, die 1-Item-Frage des IPSS als Frage der Lebensqualität betrachten, dann kann man nicht sagen, die anderen müssen bewei-sen, dass es nicht Lebensqualität ist, sondern man muss selber beweisen, dass das etwas Ver-nünftiges ist. Die Validierung liegt bei demjenigen, der ein Instrument reinbringt.

Prof. Klaus Höfner: Man muss das kommentieren. Wir reden ja nicht nur über einen Haufen illustrier Urologen, sondern über den Standard, der in sämtlichen Leitlinien, Empfehlungen existiert, die diesen IPSS-Score, diese eine Frage als Lebensqualitätsindex definiert. Das ist in diese Richtung auch immer wieder validiert worden. Da gibt es Tonnen von Daten. Das heißt, mit anderen multidimensionalen Instrumenten muss überprüft werden, ob das wirklich so ist, wie diese Leute das seit 15 Jahren – seit 15 Jahren existiert dieser Score – annehmen. Wenn wir keinen Gegenbeweis haben, dann ist es legitim, in den Studien diesen Parameter zu ak-zeptieren, der weltweit als solcher akzeptiert ist. Das sind nicht nur die deutschen Urologen, sondern auch die ganzen Europäer, Amerikaner etc. Die sind überall so definiert. Warum sa-gen wir, das ist nicht multidimensional validiert wie andere, zum Beispiel Depressionen etc.,

und deswegen kein Parameter zur Messung der Lebensqualität? Das ist ja das, was jetzt passiert.

PD Dr. Stefan Lange: Ich komme auf das Auto Beispiel zurück: Wenn ich CO₂-Emissionen reduzieren will, dann muss ich schon gucken, dass es CO₂ und nicht irgendetwas anderes ist. Wir werden versuchen, diesen Punkt entsprechend zu würdigen. Man kann diese Position darstellen, aber muss auch akzeptieren, dass es eine Wissenschaft gibt, die sich exklusiv mit Lebensqualitätsforschung auseinandersetzt. Diesen Standpunkt muss man auch darstellen.

Stefan Zenk: Es gibt eine Arbeit von Venrooij, die sagt: „while symptom score and quality of life correlated well“.

Dr. Richard Berges: Wenn Herr Lange die Literatur durchforstet bezüglich der Erfahrbarkeit des IPSS-Score, dann wird er auch auf die Korrelationsarbeiten stoßen, die mit multidimensionalen Messinstrumenten der Lebensqualitätsforscher den IPSS korreliert haben. Diese Hausaufgaben sind gemacht worden.

Dr. Guido Skipka: Was ich aus der Diskussion um den 1-Item-Score herausgehört habe, ist, dass nicht so ganz klar ist, was da letztendlich gemessen wird, ob das mehr in Richtung Lebensqualität geht oder Teile von Symptomatik hat. Das ist für mich, der ich an der Berichterstellung beteiligt bin, auch gar nicht so wesentlich. Für mich wäre wesentlich – darauf hätte ich gerne eine Antwort –: Ist die 1-Item-Frage patientenrelevant? Ich vermute, dass Sie hier zustimmen werden. Das Zweite ist: Wie würde man das mit den anderen Endpunkten gewichten? Wir haben ja heute schon darüber geredet, dass das nicht alles gleich zu behandeln ist. Wie würde man diesen Score dann einsortieren, wenn man so etwas wie eine Hierarchisierung anstrebt?

PD Dr. Stefan Lange: Lebensqualität vor Symptome oder umgekehrt? Oder sind Lebensqualität und Komplikationen zusammen?

Dr. Richard Berges: Man muss es differenzieren. Lebensqualität ist sicherlich eines der wesentlichen Aspekte der Behandlung an sich, wobei hier eine ganz speziell gesundheitsbezogene Lebensqualität die Miktionsstörungen betreffend. Es wird ja gefragt: Wie würden Sie sich fühlen, wenn sich Ihr Zustand diesbezüglich nicht mehr ändert? Das ist schon eine eingeschränkte Aussage, nicht die globale Lebensqualitätsfrage, sondern die miktionsbezogene. Das heißt, das steht im Vordergrund. Dann müssen Sie den Score selber auffächern, der sich in obstruktive und irritative Symptome unterteilen lässt. Hier ist ganz klar – auch das werden Sie finden, wenn Sie in medias res bei den Scores gehen –, dass die irritativen Sachen natürlich die sind, die den Patienten viel mehr stören als die obstruktiven Teile des Scores. Strahlverbesserung ist zum Beispiel weniger zu bewerten als Nykturieverbesserung. Das ist auch multipel untersucht worden und auch klar.

PD Dr. Stefan Lange: Jetzt sprechen Sie wieder über IPSS. Herr Skipka sprach aber über Lebensqualität.

Dr. Richard Berges: Ich sage ja: Der erste Punkt ist, was ich mit Lebensqualität global auf diese Symptome bezogen mache. Der zweite Punkt ist die Aufdifferenzierung des IPSS, wobei das noch viel schwieriger ist als die Gesamtreduktion. Deswegen ist es häufig in den Studien nur als Gesamtreduktion des IPSS gewichtet worden.

Prof. Klaus Höfner: Es ist niemals versucht worden, zwischen beiden eine Hierarchisierung vorzunehmen. Das sind, wie wir meinen, 2 verschiedene Dinge. Ob der 1-Item-Score zur Lebensqualität nach Definition von Herrn Rohde das wirklich so misst, ist eine andere Frage, aber er misst zumindest nach unserer Meinung keine Symptomatik. Das heißt, wir haben einen zusätzlichen Parameter. Insofern sollte man die gegenüberstellen und einfach getrennt diskutieren, wie es hier gemacht ist. Wir würden keine Unterscheidung vornehmen. Das ist ja das, was Sie gefragt haben.

Dr. Guido Skipka: Ich meinte jetzt nicht nur die 1-Item-Frage innerhalb des ganzen IPSS-Score-Systems, sondern zum Beispiel auch im Vergleich zur Krankenhausverweildauer oder dazu, was wir heute zur Morbidität zusammengefasst haben. Wir müssen es ja irgendwie einordnen.

Prof. Klaus Höfner: Das eine ist der Benefit und das andere ist die Negativstory. Man kann jetzt nicht sagen, das eine ist wichtiger. Wir wollen nicht das Maximum an Wirkungen haben und dafür eine Riesenmenge an Komplikationen akzeptieren. Das muss man gegenüberstellen und sehr gut abwägen. Da gibt es viele Probleme. Das haben wir ja heute schon diskutiert. Das heißt, wir können nicht sagen, Symptomatik ist wesentlicher als niedrige Komplikationen oder umgekehrt. Ich finde, da muss man sich irgendwo einpendeln, sodass man hier sicherlich keine Wertung vornehmen kann. Klar ist, man sollte eine Wertigkeit der UEs vornehmen. Das ist wichtig.

(Moderator Dr. Thomas Kaiser: Schwerwiegend/nicht schwerwiegend?)

– Ja.

Moderator Dr. Thomas Kaiser: Dazu eine konkrete Nachfrage: Gibt es für den Einzelscore Untersuchungen bezüglich eines minimal wichtigen Unterschieds? Sie sagen ja, den wollen Sie separat betrachten. Wenn ich das separat betrachte, dann muss ich für diese separate Betrachtung mir auch Gedanken machen, was ein relevanter Unterschied an Scorepunkten ist.

Dr. Richard Berges: Das muss man nachgucken. Das kann ich nicht beantworten.

Tagesordnungspunkt 6:

Verschiedenes

Moderator Dr. Thomas Kaiser: Gibt es über das in dieser länger als gedachten Erörterung hinaus Gesagte noch Irgendetwas, was diskussionswürdig zum Berichtsplan erscheint? – Das ist nicht der Fall.

Ich bedanke mich für die sehr konstruktive Diskussion. Ich glaube, dass das für den Bericht sehr hilfreich ist. Auch Ihnen, Herr Muschter, danke ich, dass Sie wiedergekommen sind und teilgenommen haben. Wir würden uns, wie Herr Lange es bereits gesagt hat, mit den Ausnahmeregelungen nicht so schwertun, wenn uns das Leben mit den Erörterungen nicht so schwer gemacht werden würde. Wir haben das Ziel, vernünftig zu diskutieren; andere haben das Ziel leider nicht. Sie hatten das Ziel heute, wofür ich mich ausdrücklich bedanken möchte.

Den Fahrplan haben wir im Groben mitgeteilt. Es wird jetzt eine Veröffentlichung des Protokolls und eine Würdigung der Stellungnahmen, einen dann überarbeiteten Berichtsplan und daran anschließend einen überarbeiteten Vorbericht, zu dem wieder Gelegenheit zur Stellungnahme besteht, geben.

PD Dr. Stefan Lange: Herr Kaiser hatte den Zeitraum 2 bis 3 Wochen genannt. Wir können den Zeitraum ja auf 3 Wochen festlegen. Ich bitte Sie, uns Literatur zu gegebenenfalls dramatischen Effekten, CCTs, Registerstudien, zu liefern. Wenn Ihnen bei der Gelegenheit noch etwas auffällt, zum Beispiel die 1-Item-Frage des IPSS, wären wir sehr dankbar, wenn Sie uns dazu methodische Literatur zur Verfügung stellen könnten. Das nehmen wir gerne entgegen.

Moderator Dr. Thomas Kaiser: Gute Heimfahrt.