

# Entscheidungsanalytische Modellierungen aus der Perspektive eines Modellierers

IQWiG im Dialog, 16.06.2023, Köln  
"Wie bewertet man die Qualität von KNB's?"

*Uwe Siebert, MPH, MSc, ScD<sup>1,2</sup>*


*Professor of Public Health, Medical Decision Making and Health Technology Assessment (UMIT TIROL)  
Adjunct Professor of Epidemiology and Health Policy & Management (Harvard University)*

<sup>1</sup> UMIT TIROL - University for Health Sciences and Technology

<sup>2</sup> Harvard Chan School of Public Health, and Massachusetts General Hospital/Harvard Medical School

# Uwe Siebert, MD, MPH, MSc, ScD



- Chair & Professor, Dept. of Public Health, Health Services Research and Health Technology Assessment, UMIT TIROL, Austria
- Adjunct Professor of Epidemiology and Health Policy & Management, Harvard Chan School of Public Health, Boston
- Cardiovascular Research Program, Institute for Technology Assessment and Dept. of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston
- Former President, Society for Medical Decision Making (SMDM)
- **Follow:**  @UweSiebert9, LinkedIn: [uwe-siebert9](#), [www.htads.org](http://www.htads.org)
- **Contact:** [uwe.siebert@umit-tirol.at](mailto:uwe.siebert@umit-tirol.at)

*Drop email with  
questions after talk*

## Research Interests:

- *Causal Modeling*: Combining decision-analytic modeling for health interventions with causal inference methods to estimate model parameters from trials and RWD
- Related concepts: **Target Trial Emulation**, Machine Learning/Causal AI
- Incremental **Trade-off Assessment**

# Übersicht

- Entscheidungsanalytisches Modellierungsframework
  - Policy Question, Auswahl Modelltyp, Evidenzsynthese, Parametervalidität, Trade-offs, Validierung
- Anwendungsbereiche
- Methoden- und Reporting-Standards
- Validität und Aussagekraft anhand von Fallbeispielen
- Kausalität in der Modellierung

*\*Stern-Folien: Material zum Nachschlagen, wird im Vortrag nicht vorgestellt*

## \*Model Types

- Decision tree models
  - Simple problem, short and fixed time horizon (e.g., 7-day mortality after surgery)
- State-transition cohort (Markov) models
  - Series of health states and transitions over time
  - With a manageable number of states → cohort simulation
- State-transition individual-level models (microsimulation)
  - Level of individual behavioural entity (e.g. virus, patient, family)
  - Number of health states is not manageable → model individuals rather than the entire cohort


# Abgrenzung

- Keine tiefergehende Beurteilung zur politischen Rolle der KNB im AMNOG-Verfahren
- Keine generelle Abhandlung von KNBs / ökonomischen Evaluationen
- Keine Diskussion von statistischen Modellen zur Prädiktion, Prognose oder Projektion
- Fokus: **entscheidungsanalytische Modellierungen** und **Voraussetzungen für deren Validität und Aussagekraft**

# IQWiG im Dialog 2023

## "Wie bewertet man die Qualität von KNBs?"

### IQWiG im Dialog am 16. Juni: "Wie bewertet man die Qualität von KNBs?" - Die Anmeldefrist ist inzwischen abgelaufen

In Kosten-Nutzen-Bewertungen (KNBs) wird der gesundheitliche  Nutzen medizinischer Maßnahmen den Kosten gegenübergestellt. Grundlage für KNBs sind in der Regel entscheidungsanalytische Modellierungen, die es ermöglichen, einen längeren Zeithorizont zu betrachten als in den klinischen Studien, die Nutzenbewertungen zugrunde liegen. Im Rahmen unserer diesjährigen IQWiG-im-Dialog-Veranstaltung wollen wir gemeinsam mit Expertinnen und Experten von Universitäten, Forschungseinrichtungen, Industrie und HTA-Institutionen diskutieren, welche Kriterien für die Bewertung der Aussagekraft von gesundheitsökonomischen Modellen bedeutsam sind.

[zum Programm](#)  
[zu den Abstracts](#)

<https://www.iqwig.de/veranstaltungen/iqwig-im-dialog/>

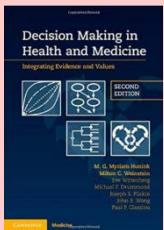
# \*Literatur zu Modellierung



U. Siebert. Kapitel 20 - Transparente Entscheidungen in Public Health mittels systematischer Entscheidungsanalyse. In: Schwartz et al. (Hrsg.): Public Health. Gesundheit und Gesundheitswesen, 2022, Kap. 57.2, S. 872-889.



U. Siebert, B. Jahn, N. Mühlberger, F.-U. Fricke, O. Schöffski. Kapitel 3 - Entscheidungsanalyse und Modellierungen. In Schöffski et al. (Hrsg.): Gesundheitsökonomische Evaluation.



M.G. Hunink, M.C., Weinstein, E. Wittenberg, M.F. Drummond, J.S. Pliskin, J.B. Wong, P.P. Glaziou. Decision Making in Health and Medicine. Integrating Evidence and Values.

## Editorial

Uwe Siebert<sup>1,2</sup>  
<sup>1</sup>Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA  
<sup>2</sup>Bavarian Public Health Research and Coordinating Centre, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University Munich, Germany

*When should decision-analytic modeling be used in the economic evaluation of health care?*

U. Siebert. When should decision-analytic modeling be used in the economic evaluation of health care? European Journal of Health Economics 2003;4(3):143-150.

# \*Literatur zu Modellierung & Kausalinferenz

Kühne F, Schomaker M, Stojkov I, Jahn B, Conrads-Frank A, Siebert S, Sroczynski G, Puntischer S, Schmid D, Schnell-Inderst P, Siebert U. Causal evidence in health decision making: methodological approaches of causal inference and health decision science. HTA Report Vol. 509-1. GMS German Medical Science — an Interdisciplinary Journal 2022;20:Doc12. doi: 10.3205/000314. Open Access: <https://www.egms.de/static/en/journals/gms/2022-20/000314.shtml>

Kuehne F, Hallsson L, Arvandi M, Puntischer S, Jahn B, Sroczynski G, Siebert U. Vergleich der Effektivität von multiplen dynamischen Behandlungsstrategien unter Nutzung der Target-Trial-Emulierung: Kontrafaktischer Ansatz zur Kausalinferenz aus Real-World-Daten [Comparing the effectiveness of multiple dynamic treatment strategies using target trial emulation: A counterfactual approach to causal inference from real-world data]. Präz Gesunderheitsf. 2023:1–11. German. doi: 10.1007/s11553-023-01033-8. Open Access: [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10259361/pdf/11553\\_2023\\_Article\\_1033.pdf](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10259361/pdf/11553_2023_Article_1033.pdf)



# Entscheidungsfindung im Gesundheitswesen

Soll eine bestimmte Gesundheitsförderungs-, Präventions-, Diagnose-, Therapie-, Reha- oder Pflegemaßnahme bei einem bestimmten Menschen durchgeführt werden bzw. im Gesundheitssystem eingeführt werden? Und wenn ja, welche Variante davon? Wie hoch darf der Preis dafür sein?



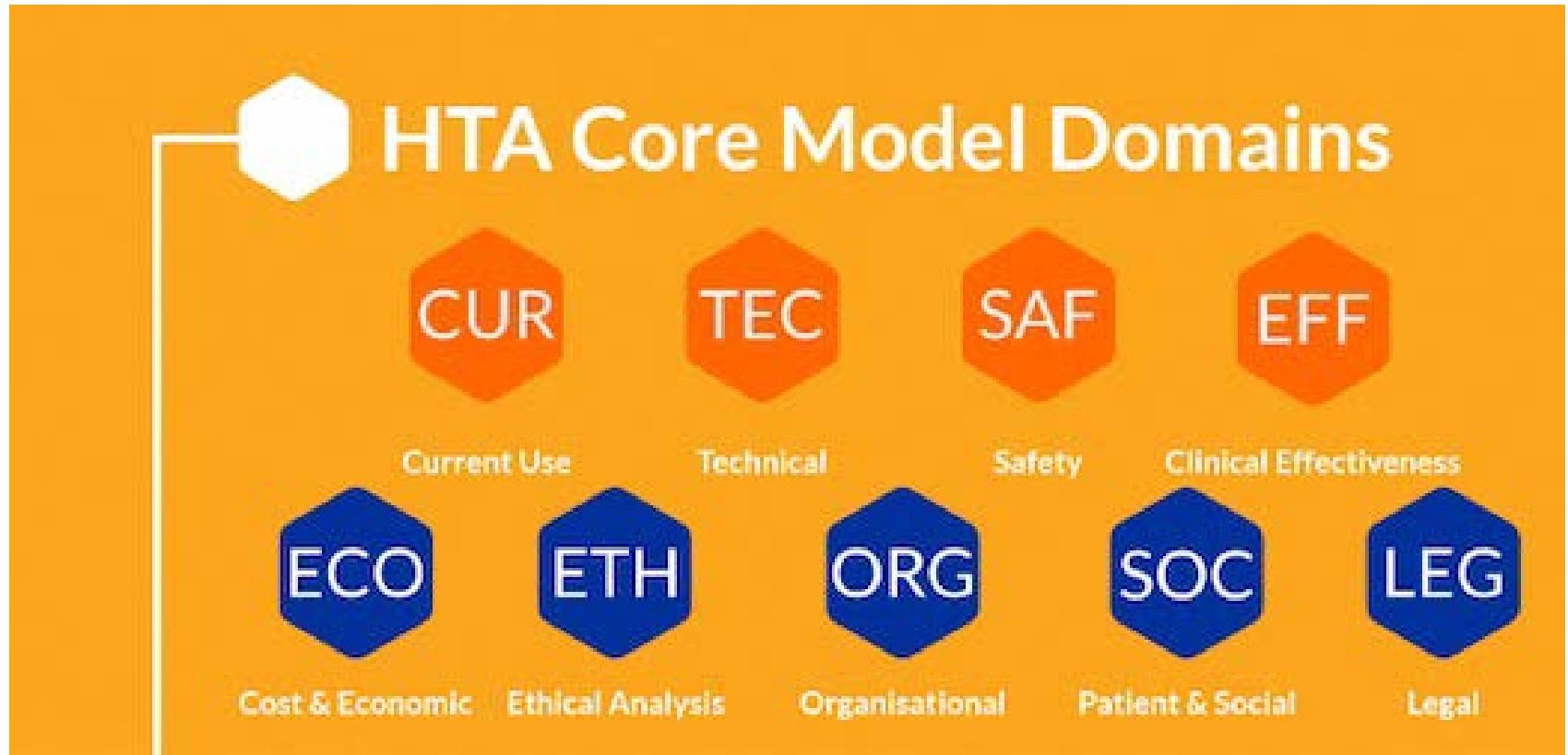
Alternative roads,  
potential/counterfactual outcomes



# Abwägungen in Health Decision Science



# HTA Domains



[www.eunetha.eu](http://www.eunetha.eu)

# Definition Entscheidungsanalyse

Entscheidungsanalyse / entscheidungsanalytische Modellierung

- ist ein **systematischer**, **expliziter** und **quantitativer** Ansatz zur Entscheidungsfindung unter **Unsicherheit**,
- **kombiniert Evidenz** für verschiedene Estimands und Outcomes aus verschiedenen Evidenzquellen und Studientypen
- bewertet Nutzen, Schaden, Ressourcen (und ggf. Verteilungen) – „**valuing**“
- um **kausale Schlussfolgerungen** abzuleiten für die Unterstützung von Entscheidungen und Abwägungen in Klinik, Policy und Forschung

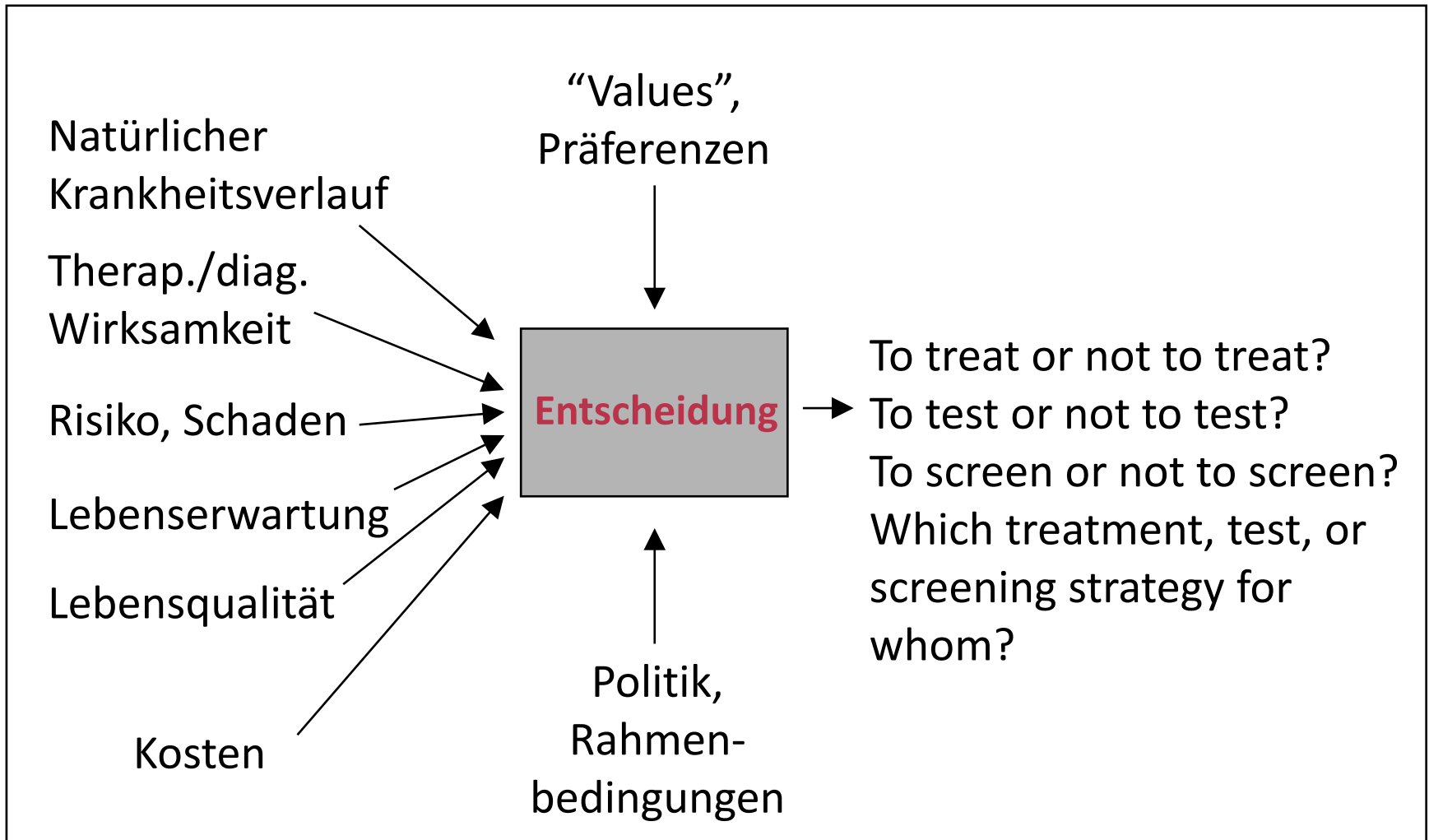
Synonyme (engl.): *decision analysis, decision-analytic modeling, decision modeling, mathematical modeling, simulation*

# Ziel der Entscheidungsanalyse

Ziel: Auswahl der „optimalen Handlungsstrategie“ nach Gewichtung medizinischer **Nutzen**, **Schäden**, **Kosten** und anderer Aspekte der verschiedenen **Alternativen**

*Entscheidungsanalyse ist nicht das Verfahren zur Schätzung von Parametern, sondern zur **Berechnung/Simulation** der Ergebnisse unter Verwendung einer den Krankheitsverlauf abbildenden Modellstruktur sowie der Modellparameter und deren Unsicherheiten*

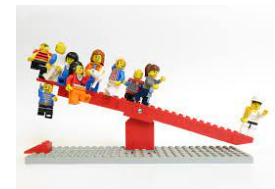
# Entscheidungsanalyse



# Purposes of Decision Analysis

## Type 1 Decision: *How should we act based on the current evidence?*

- Benefit-harm analysis
  - informing clinical guidelines and personalized medical decision making
- Cost-effectiveness analysis
  - informing reimbursement decision making
- Distributional decision analysis
  - including health inequality for social decision making



## Type 2 Decision: *Should we gather further evidence?*

- Value-of-information analysis
  - guiding future research prioritization



# \*Incorporating health inequality concerns?

# \*When is enough evidence enough?

TUTORIAL

## Distributional Cost-Effectiveness Analysis: A Tutorial

Miqdad Asaria, MSc, Susan Griffin, PhD, Richard Cookson, PhD

*Distributional cost-effectiveness analysis (DCEA) is a framework for incorporating health inequality concerns into the economic evaluation of health sector interventions. In this tutorial, we describe the technical details of how to conduct DCEA, using an illustrative example comparing alternative ways of implementing the National Health Service (NHS) Bowel Cancer Screening Programme (BCSP). The 2 key stages in DCEA are 1) modeling social distributions of health associated with different interventions, and 2) evaluating social distributions of health with respect to the dual objectives of improving total population*

*health and reducing unfair health inequality. As well as describing the technical methods used, we also identify the data requirements and the social value judgments that have to be made. Finally, we demonstrate the use of sensitivity analyses to explore the impacts of alternative modeling assumptions and social value judgments. Key words: cost-effectiveness analysis; economic evaluation; efficiency; equality; equity; fairness; health distribution; health inequality; inequality measures; opportunity cost; social value judgments; social welfare functions; tradeoff. (Med Decis Making 2016;36:8-19)*

### INTRODUCTION

When designing and prioritizing interventions, health care decision makers often have concerns about reducing unfair health inequality as well as improving total population health. However, the economic evaluation of such interventions is typically conducted using methods of cost-effectiveness analysis (CEA), which focus exclusively on maximizing total

population health. These standard methods of CEA do not provide decision makers with information about the health inequality impacts of the interventions evaluated, or the nature and size of any tradeoffs between improving total population health and reducing unfair health inequality.

To address these shortcomings, we have developed a framework for incorporating health inequality impacts into CEA, which we call *distributional cost-effectiveness analysis* (DCEA).<sup>1</sup> DCEA is suitable for health sector decisions concerning the design

Asaria, Griffin, Cookson, MDM 2016

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ) (2013) 107, 575–584



ELSEVIER  
URBAN & FISCHER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

Journal homepage: <http://journals.elsevier.de/zefq>



SCHWERPUNKT

When is enough evidence enough? – Using systematic decision analysis and value-of-information analysis to determine the need for further evidence



“Wann ist genug Evidenz genug? – Der Einsatz systematischer Entscheidungsanalyse und Value-of-Information Analyse zur Bestimmung des Bedarfs an zusätzlicher Evidenz”

Uwe Siebert<sup>1,2,3,4,\*</sup>, Ursula Rochau<sup>1,2</sup>, Karl Claxton<sup>5,6</sup>

<sup>1</sup> Department of Public Health and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria

<sup>2</sup> Area Health Technology Assessment and Bioinformatics, ONCOTYROL - Center for Personalized Cancer Medicine, Innsbruck, Austria

<sup>3</sup> Center for Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA

<sup>4</sup> Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

<sup>5</sup> Centre for Health Economics, University of York, UK

<sup>6</sup> Department of Economics and Related Studies, University of York, UK

Submitted/eingegangen 1 October 2013; revised/überarbeitet 18 October 2013; accepted/akzeptiert 18 October 2013

Siebert, Rochau, Claxton, ZEFQ 2013

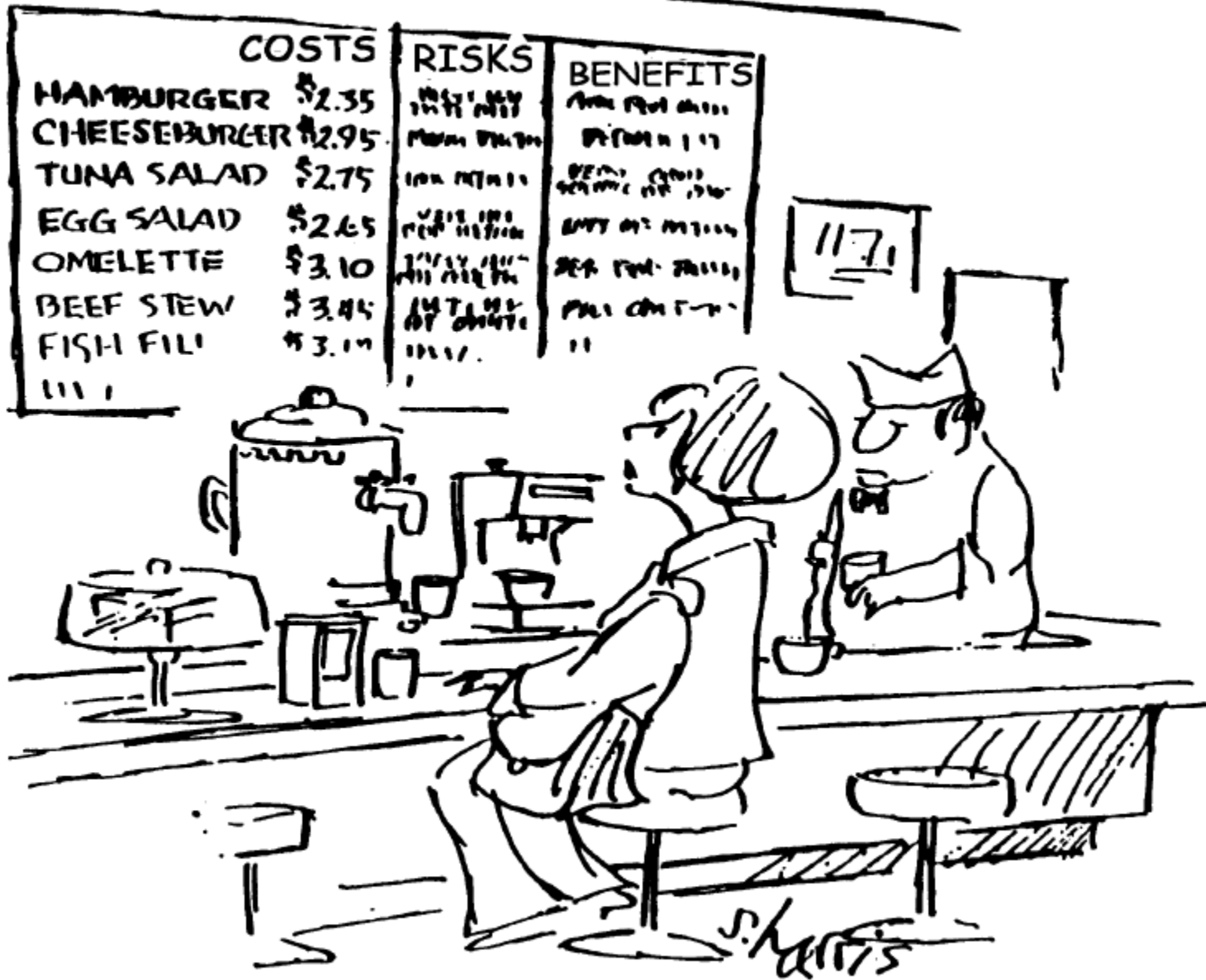
# Prämisse Entscheidungsanalyse

***The decision must be made***

*“Clinical decisions and health care policy decisions **must be made** whether the clinical circumstances are obvious or complex. Even choosing not to perform a diagnostic test, not to intervene, or not to reimburse a health technology is a decision with consequences that will be experienced by the patient.”*

Siebert: U. When should decision-analytic modeling be used in the economic evaluation of health care? Eur J Health Econom 2003; 4:143–150





# Definition "Model"

The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Task Force on Good Research Practices – Modeling Studies:

*"[...] an **analytic methodology** that accounts for **events over time and across populations**, that is based on data drawn from primary and/or secondary sources, and whose purpose is to **estimate the effects of an intervention on valued health consequences and costs.**"*

[Source: Weinstein et al., Value in Health 2003]

# Definition "Model"

Buxton and colleagues defined models in scientific disciplines:

*"Models [...] are a way of representing the complexity of the real world in a more simple and comprehensible form"*

*[Buxton, Health Economics. 1997]*

# \*Definition "Model"

## Modeling Good Research Practices - Overview:

"[...] Models are **essentially communication tools** that allow the **complexity** of a given system to be reduced to its essential elements. As such, models represent a **simplification of reality** and modeling is necessarily a reductionist methodology. "

[Caro et al. 2012: Modeling Good Research Practices – Overview  
<http://www.ispor.org/workpaper/Modeling-Good-Research-Practices-Overview.asp>]

# Purposes of Modeling in Economic Evaluation

## Editorial

Eur J Health Econom 2003 · 4:143–150  
DOI 10.1007/s10198-003-0205-2  
Published online: 10. September 2003  
© Springer-Verlag 2003

Uwe Siebert<sup>1,2</sup>

<sup>1</sup> Institute for Technology Assessment and Department of Radiology, Massachusetts, General Hospital, Harvard Medical School, Boston, MA, USA

<sup>2</sup> Bavarian Public Health Research and Coordinating Center, Institute of Medical Informatics, Biometry, and Epidemiology, Ludwig Maximilian University Munich, Germany

***When should decision-analytic modeling be used in the economic evaluation of health care?***

*Siebert, Eur J Health Econom 2003*

# Wozu EA Modelle?

## Synthese

- **Kombination** von Ergebnissen aus Prävalenz-, Diagnose-, Wirksamkeits- und NWstudien
- Verknüpfung **intermediärer** Ergebnisse aus RCTs mit **Langzeitergebnissen** aus Beobachtungsstudien & RWD
- Synthese von **Lebenszeit** und **Lebensqualität** (QALYs)
- Synthese von **Kosten** und **Effektivität** (Kosten pro gewonnenes QALY)

## Transfer

- **Zeitliche** Entwicklungen in Epidemiologie und Versorgung
- **Geographische** Unterschiede der Epidemiologie
- Unterschiede nationaler **Versorgungsstandards**
- Unterschiede bei der **Patientencompliance**
- Health Technology Assessment (**nationale Versorgungsspezifika**)

# Modelltypen

- Entscheidungsbäume (Decision tree models)
  - Einfache Probleme, kurzer fixer Zeithorizont
- Zustandsübergangsmodelle (State-transition models)
  - Krankheit durch Zustände und Übergänge abbildbar, keine Interaktionen
  - Kohortensimulation (Markov-Modelle) – Zustände überschaubar
  - Individuenbasierte (Mikro-)Simulation – „Zustandsexplosion“
- Diskrete Ereignissimulation (Discrete Event Simulation)
  - Limitierte Ressourcen, Warteschlangen
- Übertragungsmodelle (Transmission models)
  - Verbreitung akuter Infektionskrankheiten, Pandemien
- Weitere ...

# \*Model Types

- **Decision tree models**
  - Simple problem, short and fixed time horizon (e.g., 7-day mortality after surgery)
- **State-transition cohort (Markov) models**
  - Series of health states and transitions over time
  - With a manageable number of states → cohort simulation
- **State-transition individual-level models (microsimulation)**
  - Level of individual behavioural entity (e.g. virus, patient, family)
  - Number of health states is not manageable → model individuals rather than the entire cohort



# \*Model Types

- **Discrete event simulation models**
  - Queuing problems, waiting lines (e.g., organ transplantations)
- **Dynamic transmission models**
  - Modeling spread of disease over time (e.g., acute infectious diseases)
- **Other model types**
  - Agent-based models (modeling incentives and behavior)
  - System dynamics models (epidemiology, health care systems, ...)
  - Modeling substructures (molecules, cells, organs, ...)

# ISPOR-SMDM TF: Model Selection (I)

“For **relatively simple models**, or decision problems with special characteristics (e.g. very short time horizons, complex value structures) a **decision tree** may be appropriate.”

“If the conceptualization involves representing the disease or treatment process as a **series of health states, state-transition models** are often appropriate [...]. Their primary disadvantage, the Markovian assumption that transition probabilities do not depend on past history, can be addressed by increasing the number of states. **Individual-based state transition models** (termed “microsimulations”), which do not require this assumption, are an alternative **when the number of states grows too large.**”

[Caro et al. **Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force -1**. Freely available from:

<http://mdm.sagepub.com/content/32/5/667.full> and <http://www.ispor.org/workpaper/modeling-good-research-practices-overview.asp>. Medical Decision Making 2012; 32: 667-77/ Value in Health 2012; 15: 796-803]

# ISPOR-SMDM TF: Model Selection (II)

“When the disease or treatment process includes **interactions between individuals**, the modeling methods should be able to represent and evaluate the effects of those interactions (**dynamic transmission models, ...**)”

“When the **decision problem involves resource constraints**, the modeling method should be able to represent and evaluate the effects of those constraints (**discrete event simulation, ...**)”

“For some decision problems, combinations of model types, **hybrid models**, and other modeling methodologies are appropriate.”

[Caro et al. **Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force -1**. Freely available from: <http://mdm.sagepub.com/content/32/5/667.full> and <http://www.ispor.org/workpaper/modeling-good-research-practices-overview.asp>. Medical Decision Making 2012; 32: 667-77/ Value in Health 2012; 15: 796-803]

# \*Taxonomie & Modellauswahl

- Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004;9(2):110-118
- Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006;15:1295-1310
- Stahl JE. Modelling methods for pharmacoeconomics and health technology assessment. An overview and guide. *Pharmacoeconomics* 2008;26(2):141-148
- Kim SY, Goldie SJ. Cost-effectiveness analyses of vaccination programmes : a focused review of modelling approaches. *Pharmacoeconomics*. 2008;26(3):191-215
- Cooper K, Brailsford SC, Davies R. Choice of Modelling Technique for Evaluating Health Care Interventions. *The Journal of the Operational Research Society* 2007;58(2):168-176.
- Roberts M, Russell LB, Paltiel D, Chambers M, McEwan P, Krahn M, On Behalf of the ISPOR-SMDM Modeling Good Research Practices Task Force. Conceptualizing a Model. A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. *Med Decis Making* 2012;32:678–689

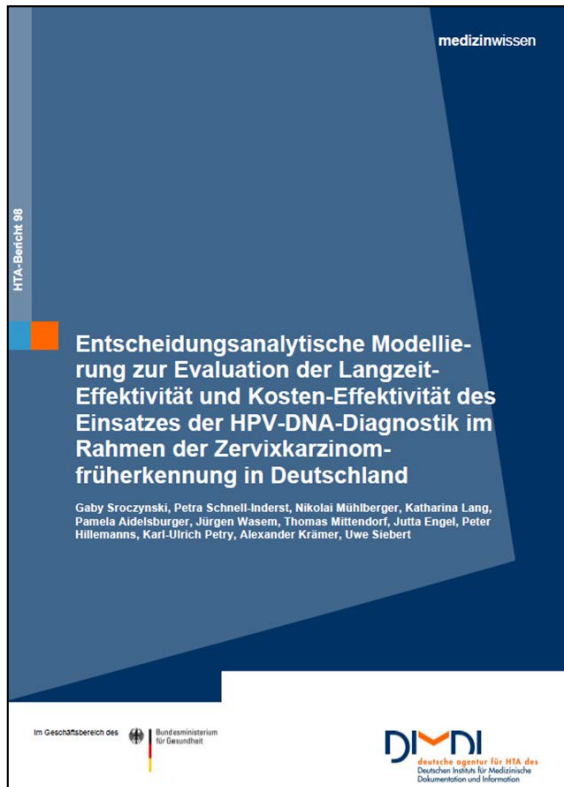
# Anwendungsbereiche



# Einsatzgebiete

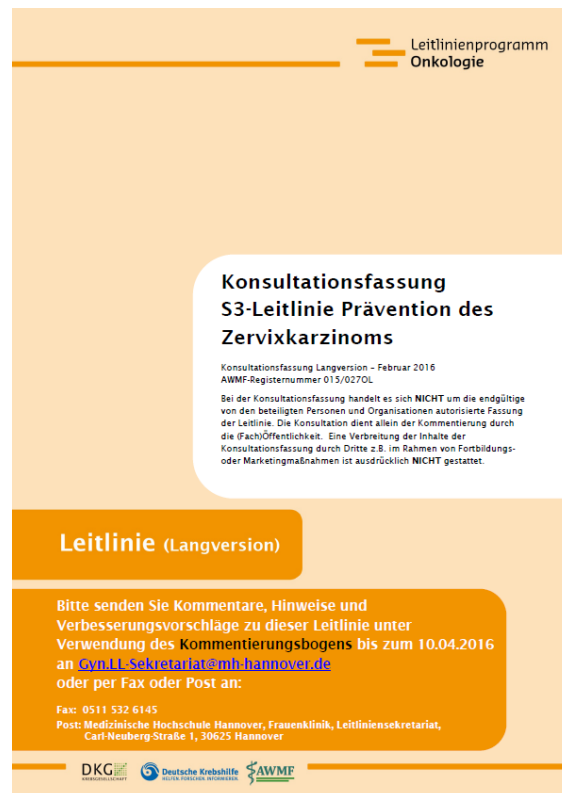
- Health Technology Assessment
  - Nutzen-Schaden-Balance
  - Kosteneffektivität, Preisfindung
  - Verteilungsgerechtigkeit und soziale Aspekte
- Klinische Leitlinien
  - Vorwiegend Nutzen-Schaden-Balance
- Patientenorientierte / partizipative Entscheidungsfindung
  - Individuelle Nutzen-Schaden-Balance
  - Berücksichtigung individueller Risiken und Präferenzen (z.B. Patienteninformationsbroschüren)

# HTA-Berichte



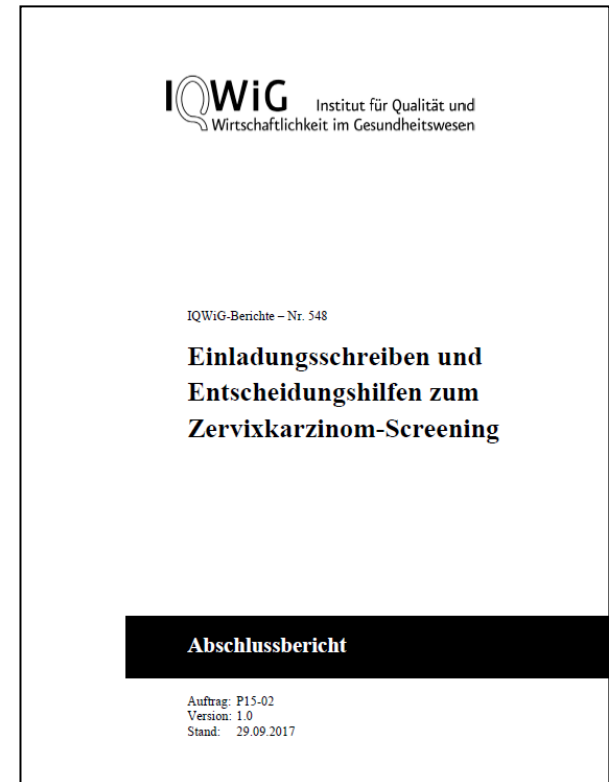
[http://portal.dimdi.de/de/hta/hta\\_berichte/hta234\\_bericht\\_de.pdf](http://portal.dimdi.de/de/hta/hta_berichte/hta234_bericht_de.pdf)

# S3-Leitlinien



<http://www.leitlinienprogramm-onkologie.de/leitlinien/zervixkarzinom-praevention/>

# Entscheidungshilfen



[https://www.iqwig.de/download/P15-02\\_Einladungsschreiben-und-Entscheidungshilfen-zum-Zervixkarzinom-Screening.pdf](https://www.iqwig.de/download/P15-02_Einladungsschreiben-und-Entscheidungshilfen-zum-Zervixkarzinom-Screening.pdf)

# \*S3-LL Ergebnisse als Faktenbox Nutzen, Schaden

Konsequenzen: (pro 10.000 Frauen) 1. Zahl: absoluter Wert 2. Zahl: Inkrement*	Kein Screening (Strategie 1)	5-jähr. Screening mit HPV+Pap Kotestung vs. Kein Screening (Strategie 21)	3-jähr. Screening mit HPV/Pap Triage vs. 5-jähr. HPV+Pap Kotestung (Strategie 26)	2-jähr. Screening mit HPV/Pap Triage vs. 3-jähr. HPV/Pap Triage (Strategie 25)
Diagnostizierte Zervix- karzinominzidenz (lebenslang)	314	50 264	25 25	13 12
Zervixkarzinomtodesfälle (lebenslang)	119	13 106	5 7	2 3
Lebensjahre	679.066	681.402 2.337	681.552 150	681.611 59
Positive Primärscreening- ergebnisse	0	10.273 10.273	13.118 2.845	17.867 4.749
Anzahl der Kolposkopien	0	6.021 6.021	7.830 1.809	9.832 2.002
Anzahl der Konisationen <CIN3	0	650 650	777 127	874 98
Anzahl aller Konisationen	0	859 859	970 111	1042 72

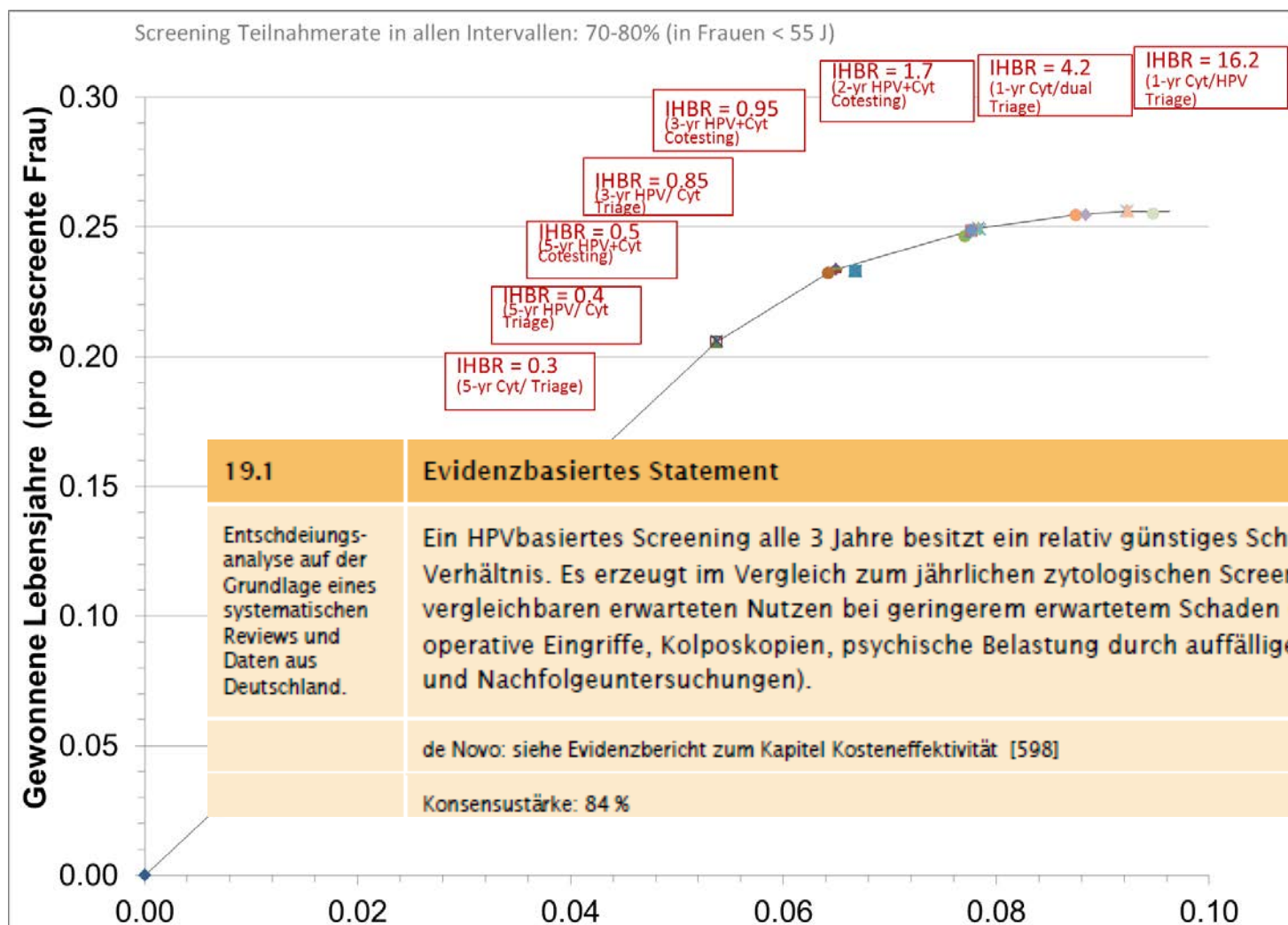
Nutzen

Schaden



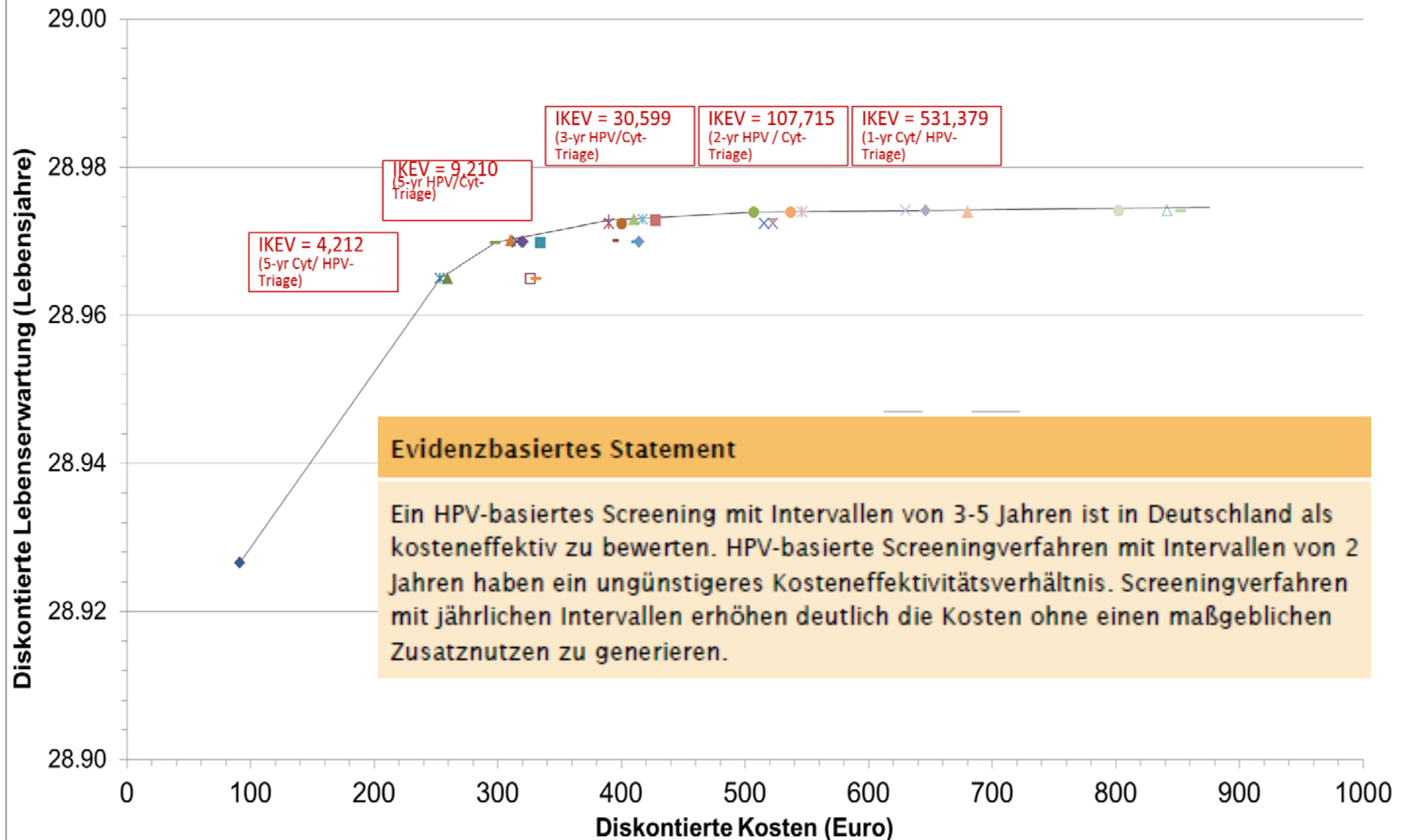
# Ergebnisse: Schaden-Nutzen

## Konisationen < CIN3 pro gewonnenes Lebensjahr



# Ergebnisse: Kosteneffektivität Euro pro gewonnenes Lebensjahr

Screening Teilnehmerate in allen Intervallen: 70-80% (in Frauen < 55 J)



# IQWiG Broschüre Entscheidungshilfe

## **Gebärmutterhalskrebs- Früherkennung**

Das Angebot der gesetzlichen  
Krankenversicherung

**Eine Entscheidungshilfe für  
Frauen zwischen 20 und 34 Jahren**

# \*IQWiG Broschüre Entscheidungshilfe

## Unterstützung für Ihre Entscheidung: Gebärmutterhalskrebs-Früherkennung

Frauen bewerten die Vor- und Nachteile von Früherkennungsuntersuchungen für sich unterschiedlich. Die **untenstehende Tabelle** fasst noch einmal wesentliche Vor- und Nachteile der Früherkennung von Gebärmutterhalskrebs zusammen. So können Sie abwägen, ob Sie teilnehmen möchten.

	Keine Früherkennung	Regelmäßige Früherkennung (lebenslang)
Wie viele Frauen erkranken an Gebärmutterhalskrebs?	<b>Nicht gegen HPV geimpft:</b> Etwa <b>30 von 1000 Frauen</b> erkranken an Gebärmutterhalskrebs. <b>Gegen HPV geimpft:</b> Etwa <b>10 von 1000 Frauen</b> erkranken an Gebärmutterhalskrebs.	Weniger als <b>1 von 1000 Frauen</b> erkranken an Gebärmutterhalskrebs. Dadurch werden auch belastende Behandlungen vermieden.
Wie viele Frauen sterben an Gebärmutterhalskrebs?	Etwa <b>12 von 1000 nicht gegen HPV geimpfte Frauen</b> sterben an Gebärmutterhalskrebs. Etwa <b>4 von 1000 gegen HPV geimpfte Frauen</b> sterben an Gebärmutterhalskrebs.	Weniger als <b>1 von 1000 Frauen</b> sterben an Gebärmutterhalskrebs.

Wie oft kommt es zu Behandlungen von harmlosen Vorstufen?	Bei Frauen, <b>die nicht zur Früherkennung</b> gehen, können Vorstufen nur dann entdeckt und behandelt werden, wenn sie aus anderen Gründen einen Pap-Test machen lassen. Wie häufig dabei harmlose Vorstufen behandelt werden, lässt sich nicht genau sagen.	Bei etwa <b>110 bis 120 von 1000 nicht gegen HPV geimpften Frauen</b> werden irgendwann Dysplasien durch eine Konisation behandelt. Bei etwa <b>40 von 1000 gegen HPV geimpften Frauen</b> werden irgendwann Dysplasien durch eine Konisation behandelt. Bei einem Teil von ihnen wäre die Dysplasie nie zu Krebs geworden.
Wie hoch ist das Risiko für Schwangerschafts-Komplikationen durch die Behandlung von Dysplasien?	Konisationen sind selten bei Frauen, die nicht zur Früherkennung gehen. Denn Vorstufen können nur dann erkannt werden, wenn sie den Pap-Test aus anderen Gründen machen lassen.	Bei etwa <b>3 von 100 Schwangeren</b> , die vorher eine größere Konisation hatten, kommt es infolge des Eingriffs zu einer Frühgeburt. Gewebeschonende Eingriffe verringern dieses Risiko deutlich.
Wie zuverlässig sind die Untersuchungen?	Gebärmutterhalskrebs wird meist erst dann festgestellt, wenn er fortgeschritten ist.	Die Früherkennung ist bei regelmäßiger Teilnahme recht zuverlässig. Dennoch können in seltenen Fällen Vorstufen und Tumore übersehen werden.

# Standards und Guidelines für entscheidungsanalytische Modellierungen

# Duale Validität

Besonderheit bei EA Modellierungen im Vergleich zu anderen Studientypen:

**Validität und Aussagekraft** der Modellierungsergebnisse hängt ab von 2 Faktoren:

1. Korrektheit der **Modellstruktur**

- „Biologie des Modells, beinhaltet immer auch **Expertenwissen**

2. Adäquate **Parameterwahl**

- Adäquate statistische Analyse(selbst oder in Literatur), Adjustierung, bedingte Schätzer, kausale Inferenz etc.

Bei der Beurteilung der Modellvalidität müssen **beide Aspekte im Zusammenhang** betrachtet werden

→ Checklisten sind notwendig aber nicht hinreichend

# International Guidelines

- **Framework Quality Assessment:** Assessing quality in decision analytic cost-effectiveness models. Sculpher et al., 2000
- **HTA Key Principles:** Key principles for the improved conduct of health technology assessments for resource allocation decisions. International Working Group for HTA Advancement, 2010
- **Comprehensive Methodological Guidelines:** ISPOR-SMDM Modeling Good Research Practices Task Force, 2012
- **Reporting:** Consolidated health economic evaluation reporting standards (CHEERS 2022 statement), Husereau et al., 2022
- **Certainty Assessment:** Brozek et al., GRADE Guideline 30: Assessment of certainty of modeled evidence, JCE, 2021

# \*Key Principles of HTA

*International Journal of Technology Assessment in Health Care*, 24:3 (2008), 244–258.  
Copyright © 2008 Cambridge University Press. Printed in the U.S.A.  
doi:10.1017/S0266462308080343

## Key principles for the improved conduct of health technology assessments for resource allocation decisions

**Michael F. Drummond**

*University of York*

**J. Sanford Schwartz**

*University of Pennsylvania*

**Bengt Jönsson**

*Stockholm School of Economics*

**Bryan R. Luce**

*United BioSource Corporation*

**Peter J. Neumann**

*Tufts University*

**Uwe Siebert**

*UMIT—University for Health Sciences, Medical Informatics and Technology*

**Sean D. Sullivan**

*University of Washington*

Health technology assessment (HTA) is a dynamic, rapidly evolving process, embracing different types of assessments that inform real-world decisions about the value (i.e., benefits, risks, and costs) of new and existing technologies. Historically, most HTA agencies have focused on producing high quality assessment reports that can be used by a range of decision makers. However, increasingly organizations are undertaking or

*Drummond et al. 2008, IJTAHC*



# \*Key HTA Principles

## STRUCTURE OF HTA

1. Goal and scope explicit & relevant
2. Unbiased & transparent
3. All relevant technologies
4. Clear priority setting for topics

## METHODS OF HTA

5. Approp. cost & benefit methods
6. Wide range evidence & outcomes
7. Full societal perspective
8. Explicitly characterize uncertainty
9. Generalizability & Transferability

## PROCESSES FOR CONDUCT

10. Engage all stakeholders
11. Seek all available data
12. Monitor implementation of findings

## USE IN DECISION MAKING

13. Should be timely
14. Appropriate communication
15. Clear link HTA–decision making

*Drummond et al. 2008, Int J Technol Assessment Health Care*

# Task Force Co-Chairs

## ISPOR-SMDM Good Research Practice 2010-2012



**J. Jaime Caro** MDCM, FRCPC, FACP. Adjunct Professor Medicine & Epidemiology & Biostatistics, McGill University, Montreal, Quebec, Senior Vice President, Health Economics, UBC, Lexington, MA, USA



**Andrew Briggs** DPhil, Lindsay Chair in Health Policy & Economic Evaluation, Public Health & Health Policy, University of Glasgow, Glasgow, Scotland UK



**Karen Kuntz** ScD, SMDM Issues in Methodology Working Group Chair, and Professor, University of Minnesota School of Public Health, Minneapolis, MN USA



**Uwe Siebert** MD, MPH, MSc, ScD, SMDM Board representative-Issues in Methodology Working Group Co-Chair, Professor of Public Health, UMIT – Univ. of Health Sciences, Medical Informatics & Technology, Hall i.T., Austria

\*

**Modeling Good Research  
Practices—Overview: A Report of the  
ISPOR-SMDM Modeling Good Research  
Practices Task Force–1**

*J. Jaime Caro, MDCM, FRCPC, FACP, Andrew H. Briggs, DPhil,  
Uwe Siebert, MD, MPH, MSc, ScD, Karen M. Kuntz, ScD,  
On Behalf of the ISPOR-SMDM Modeling Good Research Practices Task Force*

**Webinars:**

[https://www.ispor.org/education/Educational\\_Webinars.asp](https://www.ispor.org/education/Educational_Webinars.asp)

**URL task force reports:**

<https://www.ispor.org/heor-resources/good-practices-for-outcomes-research/article/modeling-good-research-practices---overview>

*Caro JJ, Briggs AH, Siebert U, Kuntz KM, on Behalf of the ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force -1. Value in Health 2012;**15**:796-803.*

*Caro JJ, Briggs AH, Siebert U, Kuntz KM, on behalf of the ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force -1. Medical Decision Making 2012;**32**(5):667-77.*


# \*Certainty of Evidence (GRADE)

- Domains determining certainty of evidence

- risk of bias
- Indirectness
- Inconsistency
- Imprecision
- reporting bias
- magnitude of effect
- dose-response relation
- direction of residual confounding

- Recommendations

- develop a model de novo
- identify best existing model
- use outputs from multiple models.



**HHS Public Access**  
Author manuscript  
*J Clin Epidemiol.* Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:  
*J Clin Epidemiol.* 2021 January ; 129: 138–150. doi:10.1016/j.jclinepi.2020.09.018.

**GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence—An overview in the context of health decision-making**

Jan L. Brozek<sup>a,b,c,\*</sup>, Carlos Canelo-Aybar<sup>d,e,1</sup>, Elie A. Akf<sup>f</sup>, James M. Bowen<sup>a,g</sup>, John Bucher<sup>h</sup>, Weihsueh A. Chiu<sup>i</sup>, Mark Cronini<sup>j</sup>, Benjamin Djulbegovic<sup>k</sup>, Maicon Falavigna<sup>l</sup>, Gordon H. Guyatt<sup>a,b,c</sup>, Ami A. Gordon<sup>m</sup>, Michele Hilton Boon<sup>n</sup>, Raymond C.W. Hutubessy<sup>o</sup>, Manuela A. Joore<sup>p</sup>, Vittal Katikireddi<sup>q</sup>, Judy LaKind<sup>r,s</sup>, Miranda Langendam<sup>t</sup>, Veena Manja<sup>a,t,u</sup>, Kristen Magnuson<sup>m</sup>, Alexander G. Mathioudakis<sup>v</sup>, Joerg Meerpoel<sup>w,x</sup>, Dominik Mertz<sup>a</sup>, Roman Mezencev<sup>y</sup>, Rebecca Morgan<sup>a</sup>, Gian Paolo Morgano<sup>a,c</sup>, Reem Mustafa<sup>a,z</sup>, Martin O'Flaherty<sup>aa</sup>, Grace Patlewicz<sup>ab</sup>, John J. Riva<sup>c,ac</sup>, Margarita Posso<sup>e</sup>, Andrew Rooney<sup>h</sup>, Paul M. Schlosser<sup>y</sup>, Lisa Schwartz<sup>a</sup>, Ian Shemilt<sup>ad</sup>, Jean-Eric Tarride<sup>aa,ae</sup>, Kristina A. Thayer<sup>af</sup>, Katya Tsaoun<sup>af</sup>, Luke Vale<sup>ag</sup>, John Wambough<sup>ab</sup>, Jessica Wignall<sup>m</sup>, Ashley Williams<sup>m</sup>, Feng Xie<sup>a</sup>, Yuan Zhang<sup>a,ah</sup>, Holger J. Schünemann<sup>a,b,c</sup>, GRADE Working Group

<sup>a</sup>Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada  
<sup>b</sup>Department of Medicine, McMaster University, Hamilton, Ontario, Canada  
<sup>c</sup>McMaster GRADE Centre & Michael DeGroot Cochrane Canada Centre, McMaster University, Hamilton, Ontario, Canada  
<sup>d</sup>Department of Paediatrics, Obstetrics and Gynaecology, Preventive Medicine, and Public Health. PhD Programme in Methodology of Biomedical Research and Public Health. Universitat Autònoma de Barcelona, Bellaterra, Spain  
<sup>e</sup>Iberoamerican Cochrane Center, Biomedical Research Institute (IIB Sant Pau-CIBERESP), Barcelona, Spain  
<sup>f</sup>Department of Internal Medicine, American University of Beirut, Beirut, Lebanon  
<sup>g</sup>Toronto Health Economics and Technology Assessment (THETA) Collaborative, Toronto, Ontario, Canada  
<sup>h</sup>National Toxicology Program, National Institute of Environmental Health Sciences, Durham, NC, USA

# Kriterien für die Validität von Modellen und deren Aussagekraft



# Glaubwürdigkeit und Verlässlichkeit eines Modells

„**Trust** and **confidence** are critical to the success of health care models. There are two main methods for achieving this: **transparency** (people can see how the model is built) and **validation** (how well it reproduces reality). [...]

Validation involves **face validity** (wherein experts evaluate model structure, data sources, assumptions, and results), **verification or internal validity** (check accuracy of coding), **cross validity** (comparison of results with other models analyzing same problem), **external validity** (comparing model results to real-world results), and **predictive validity** (comparing model results with prospectively observed events).”

Eddy D et al. Model Transparency and Validation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Medical Decision Making 2012;32:733-43.

# Kriterien für die Validität von Modellen und deren Aussagekraft

- Klare Formulierung der alternativen Strategien (kausale, kontrafaktische Forschungsfrage) PICOST
- Adäquater Modelltyp und adäquate Modellstruktur
- Kausal interpretierbare Inputparameter
- Explizite Beschreibung der Annahmen
- Erfolgreiche Durchführung von Kalibrierungen und Validierungen
- Unsicherheitsanalysen
- Adäquates Reporting der Ergebnisse und Trade-offs

# Entwicklung der Modellierung beim IQWiG

- Insgesamt ist **Modellierung die Ausnahme** (s. Vortrag Dr. Anja Schwalm)
- Wird sich das ändern angesichts mancher neuen hochpreisigen Medikamente?
- Erster Einsatz 2013 KNB bei Vielfalt der Antidepressiva mit sehr unterschiedlichen Preisen  
→ **Erstattungspreise**
- Jüngster Einsatz 2022 bei Nutzen-Schaden-Bewertung Mammografie-Screening  
→ **Auswahl optimaler Screeningvarianten**



# Erste IQWiG KNB Antidepressiva

Mein Fokus heute:  
Verwendung der  
Ergebnisse zur  
Bestimmung des  
Erstattungspreises

## **Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin im Vergleich zu weiteren verordnungsfähigen medika- mentösen Behandlungen**

### **Abschlussbericht**

Auftrag: G09-01  
Version: 1.0  
Stand: 03.09.2013

[https://www.iqwig.de/download/g09-01\\_abschlussbericht\\_kosten-nutzen-bewertung-von-venlafaxin-duloxetin.pdf](https://www.iqwig.de/download/g09-01_abschlussbericht_kosten-nutzen-bewertung-von-venlafaxin-duloxetin.pdf)

# Markov-Modell Antidepressiva (mit Erst- und Zweitlinientherapie)

Abschlussbericht G09-01

Version 1.0

Kosten-Nutzen-Bewertung von Venlafaxin, Duloxetin, Bupropion und Mirtazapin

03.09.2013

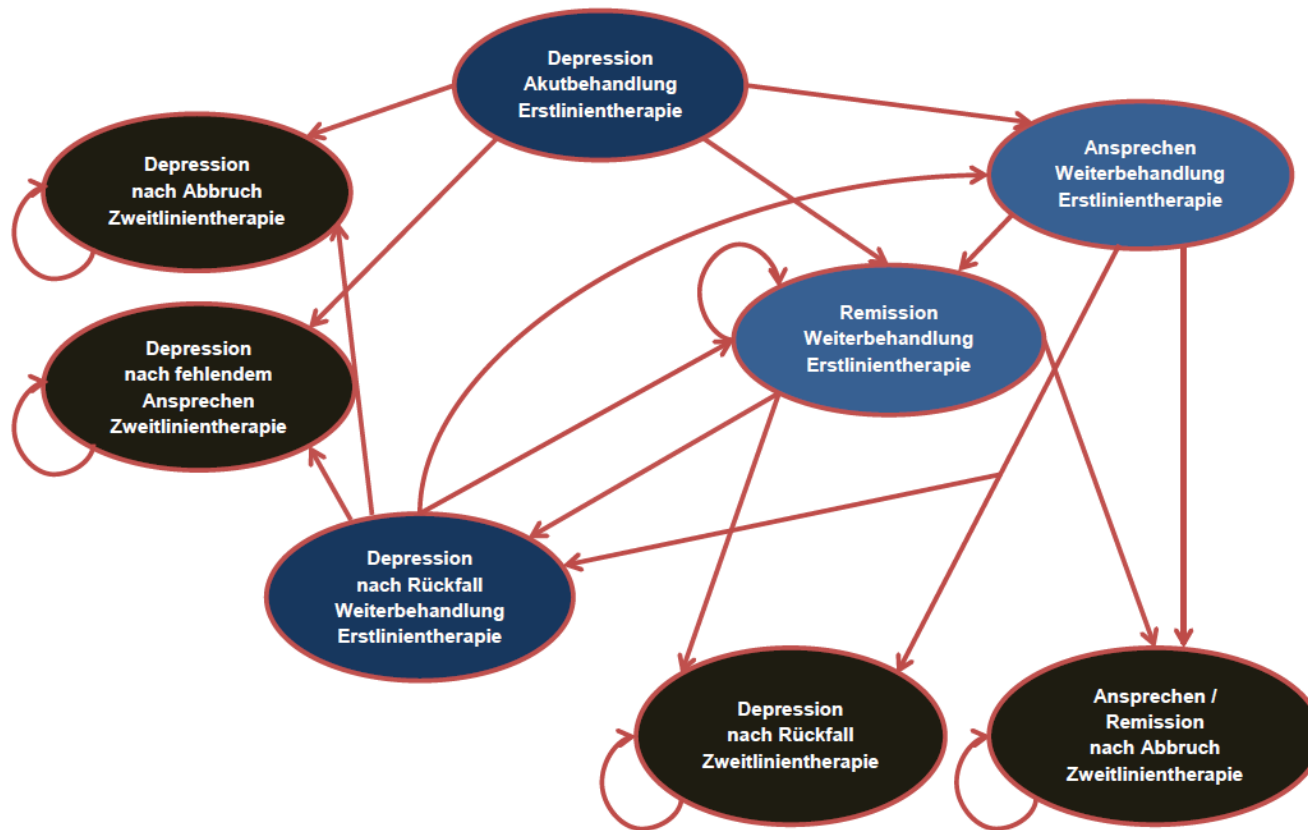
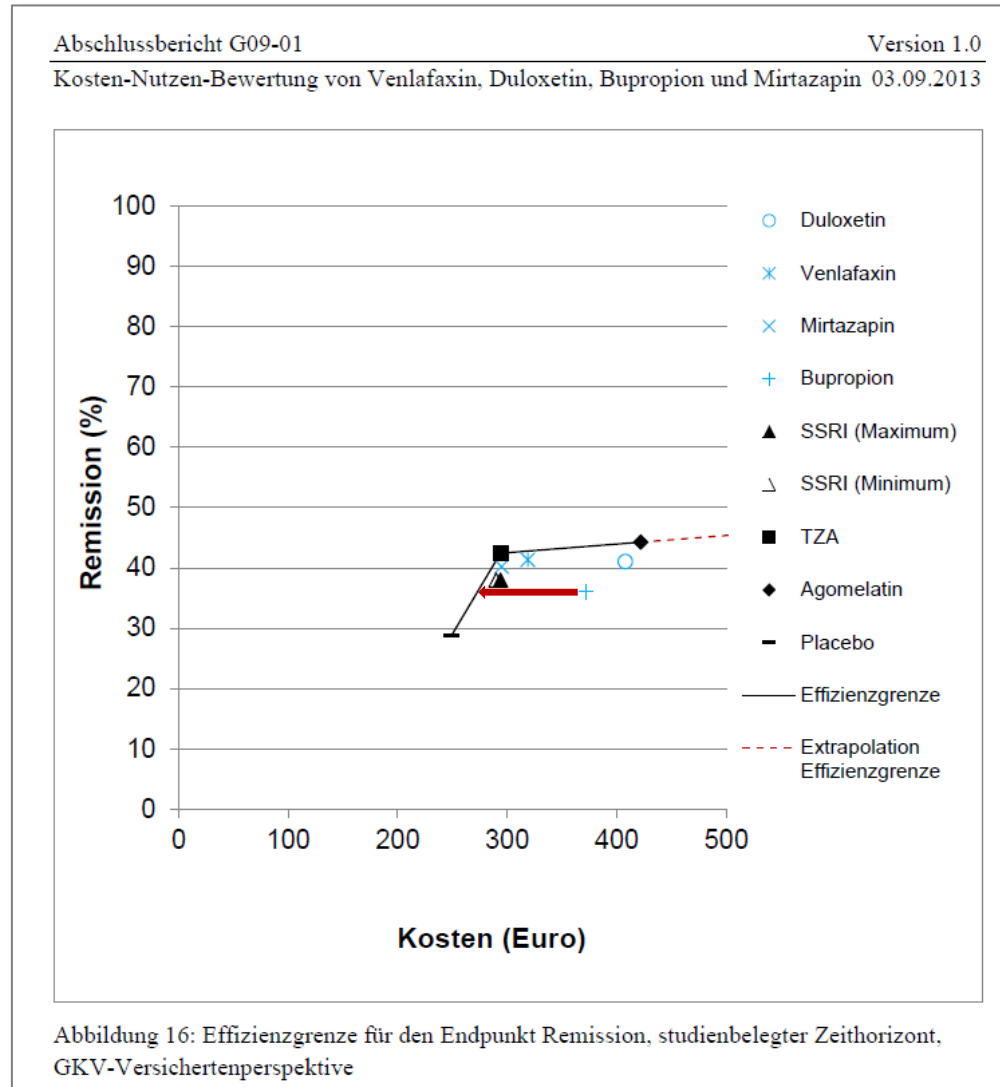



Abbildung 14: Blasendiagramm Modellstruktur: komplexe Darstellung

# Effizienzgrenze Remission Antidepressiva



# Zusatznutzenbereinigte Erstattungspreise Antidepressiva

Tabelle 100: Übersicht über die zusatznutzenbereinigten Erstattungspreise auf Grundlage der Basisfallanalysen und der probabilistischen Sensitivitätsanalysen

Prüfsubstanz	Basispreis (€)	EP Remission (€)	EP IQR PSA Remission (€)	EP Ansprechen (€)	EP IQR PSA Ansprechen (€)
<b>Bupropion</b>	104,88 	2,93	0–10,32	1,48	0–8,29
<b>Duloxetin</b>	241,18	30,66	22,94–69,66	9,30	0,35–21,95
<b>Mirtazapin</b>	46,46	31,66	20,68–44,90	24,28	14,29–35,96
<b>Venlafaxin</b>	92,57	42,99	35,33–83,04	40,91	31,22–54,25

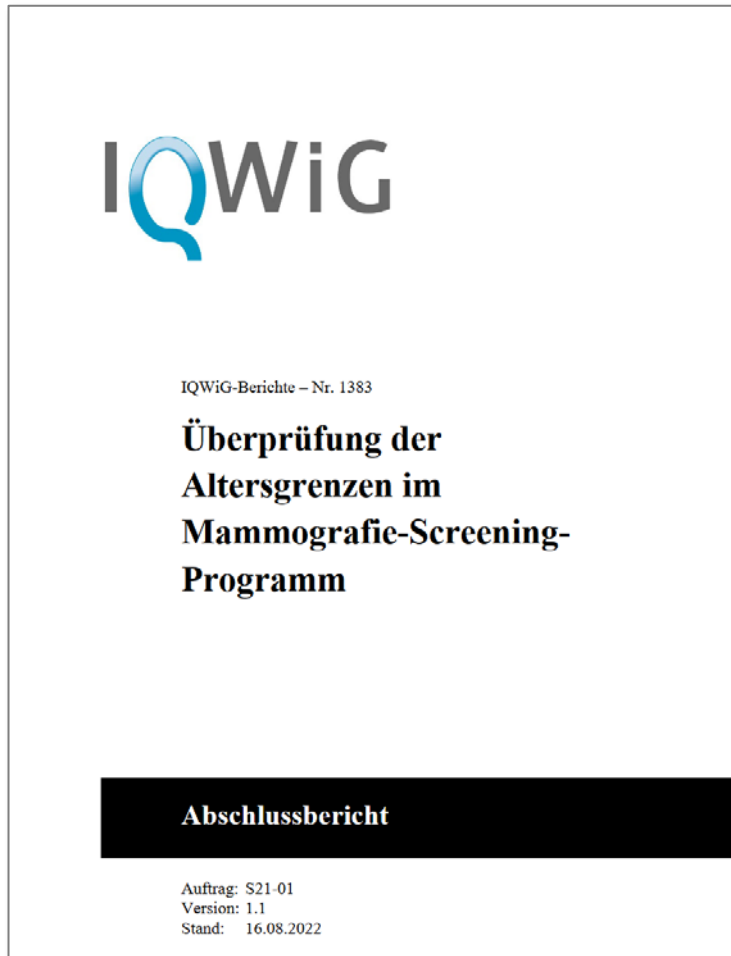
EP: Erstattungspreis, IQR: Interquartilsregion, PSA: probabilistische Sensitivitätsanalyse

# Effizienzgrenzenansatz

## Offene Fragen:

- **Interpretierbarkeit der ICERs**
  - EUR pro % Ansprechrate
  - Wie vergleichen mit EUR pro % Langzeitansprechrate (SVR) bei chronischer Hepatitis C?
- **Multiple Effizienzgrenzen**
  - Pro Outcome eine Effizienzgrenze
  - Keine Methode zur Integration multipler Effizienzgrenzen
  - Aber: Methoden zur Integration von Outcomes existieren z.B. EUR pro gewonnenes QALY

# IQWiG Nutzen-Schaden-Modellierung: Altersgrenzen Mammografie-Screening



## Überprüfung der Altersgrenzen im Mammographie-Screening-Programm (S21-01) – Modellierungsstudie

09. Mai.2022

**Autoren:** Gaby Sroczynski, Lára R. Hallsson, Nikolai Mühlberger, Felicitas Kühne, Beate Jahn, Uwe Siebert

**Institution:** UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technologie

**Projekttitel:** „Überprüfung der Altersgrenzen im Mammographie-Screening-Programm“ (S21-01) – Modellierungsstudie

Erstellt für das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)

Mein Fokus heute:  
Konzept und Validität  
der Ergebnisse

# Entscheidungsanalytisches Framework

- **Population und Perspektive:** Frauen in Deutschland ab 45 Jahren, für die eine Mammographie in Frage kommt
- **Zeithorizont:** lebenslang
- **Strategien:** verschiedene Start- und Stop-Altersgrenzen und Screeningintervalle
- **Outcomes:**
  - **Nutzen:** Reduktion von DCIS/Brustkrebs/krebsbedingter Tod, gewonnene Lebensjahre/QALYs
  - **Schaden:** Zusätzliche Mammografien, positive/falsch-positive Mammografien, Überdiagnosen
  - **Trade-offs:** Inkrementelle Schaden-Nutzen-Verhältnisse (ISNV).
- **Modelltyp:** Markov Zustand-Übergangsmodell
- **Simulation:** Kohortensimulation

# Beschreibung alternative Strategien

## 4.1.3. Evaluierte Strategien

Folgende Strategien werden in der vorliegenden entscheidungsanalytischen Studie verglichen, wobei der derzeitige Screening-Standard in Deutschland als Referenzstrategie dient.

- 1) Kein Screening
- 2) Referenzstrategie: Mammographie, alle 2 Jahre im Alter 50 - 69 Jahre (Standard in Deutschland)
- 3) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 69 Jahre
- 4) Mammographie alle zwei Jahre im Alter 45 - 69 Jahre
- 5) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 74 Jahre
- 6) Mammographie alle zwei Jahre im Alter 45 - 74 Jahre
- 7) Mammographie, jährlich im Alter 45 - 49 Jahre und alle zwei Jahre im Alter 50 - 79 Jahre
- 8) Mammographie alle zwei Jahre im Alter 45 - 79 Jahre
- 9) Mammographie, jährlich im Alter 45 - 49 Jahre, alle zwei Jahre im Alter 50 - 69 Jahre und alle drei Jahre im Alter 70 - 74 Jahre
- 10) Mammographie, jährlich im Alter 45 - 49 Jahre, alle zwei Jahre im Alter 50 - 69 Jahre und alle drei Jahre im Alter 70 - 79 Jahre
- 11) Mammographie alle zwei Jahre im Alter 50 - 74 Jahre
- 12) Mammographie alle zwei Jahre im Alter 50 - 79 Jahre
- 13) Mammographie alle zwei Jahre im Alter 50 - 69 und alle drei Jahre im Alter 70 - 74 Jahre
- 14) Mammographie alle zwei Jahre im Alter 50 - 69 und alle drei Jahre im Alter 70 - 79 Jahre



# Beschreibung der Outcomes/Endpunkte (disaggregiert und aggregiert)

## 4.1.6. Zielparameter (Endpunkte)

Es werden verschiedene Zielparameter, die den langzeitigen Nutzen oder möglichen Schaden darstellen, in den Analysen berechnet:

1. Lebenszeitrisiko für entdeckte invasive Brustkrebskarzinome (BK) (in %)
2. Lebenszeitrisiko für entdeckte ductale Carcinoma in situ (DCIS) (in %)
3. Lebenszeitrisiko, an entdecktem invasivem Brustkrebskarzinom zu versterben (in %)
4. Verbleibende Restlebenserwartung (in Jahre) pro 100 Frauen
5. Verbleibende lebensqualitätsadjustierte Restlebenserwartung (in qualitätsadjustierte Lebensjahre (QALY)) pro 100 Frauen
6. Absolute Anzahl von Screeningbefunden (Mammographien) pro 100 Frauen
7. Absolute Anzahl positiver und falsch-positiver Screeningbefunde (Mammographien) pro 100 Frauen
8. Lebenszeitrisiko für eine Überdiagnose (DCIS und Brustkrebs) (in %)
9. Number-Needed-To-Screen (NNS), um einen Brustkrebs-bedingten Todesfall zu vermeiden

# Begründung Modelltyp ISPOR-SMDM TF

“If the conceptualization involves representing the disease or treatment process as a **series of health states, state-transition models** are often appropriate [...]. Their primary disadvantage, the Markovian assumption that transition probabilities do not depend on past history, can be addressed by increasing the number of states. **Individual-based state transition models** (termed “microsimulations”), which do not require this assumption, are an alternative **when the number of states grows too large.**”

[Caro et al. **Modeling Good Research Practices - Overview: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force -1**. Freely available from:

<http://mdm.sagepub.com/content/32/5/667.full> and <http://www.ispor.org/workpaper/modeling-good-research-practices-overview.asp>. Medical Decision Making 2012; 32: 667-77/ Value in Health 2012; 15: 796-803]

# Begründung des Modelltyps

Für die Evaluation klinischer Langzeitkonsequenzen einer Brustkrebs-Früherkennungsuntersuchung für Frauen zwischen 45 und 49 Jahren, 70 und 74 sowie zwischen 75 und 79 Jahren wurde ein entscheidungsanalytisches Markov-Zustands-Übergangs-Modell (4-9) entwickelt und für den deutschen klinischen und epidemiologischen Kontext angepasst, kalibriert und validiert. Ein ...

Das Markov-Zustands-Übergangs-Modell kann den Krankheitsverlauf des Brustkrebses widerspiegeln, da die Krankheit verschiedenen wohldefinierten, histologischen und klinischen Zuständen mit spezifischen Übergangs- und Ereigniswahrscheinlichkeiten folgt (7, 8, 12, 13). Außerdem kann ein Markov-Modell die wiederkehrenden Ereignisse des Screenings bis zum finalen Endpunkt abbilden, welcher für die Evaluation von Screening-Programmen erforderlich ist. Dieser Modelltyp wurde daher für die hier beschriebene Analyse gewählt. Da die Anzahl der Gesundheitszustände überschaubar ist, wurde für die Analyse des Modells das Verfahren der Kohortensimulation gewählt. (13).

[https://www.iqwig.de/download/s21-01\\_altersgrenzen-im-mammografie-screening-programm\\_abschlussbericht\\_v1-1.pdf](https://www.iqwig.de/download/s21-01_altersgrenzen-im-mammografie-screening-programm_abschlussbericht_v1-1.pdf)

# Transparente Beschreibung Modellstruktur: Graphisch und als Text

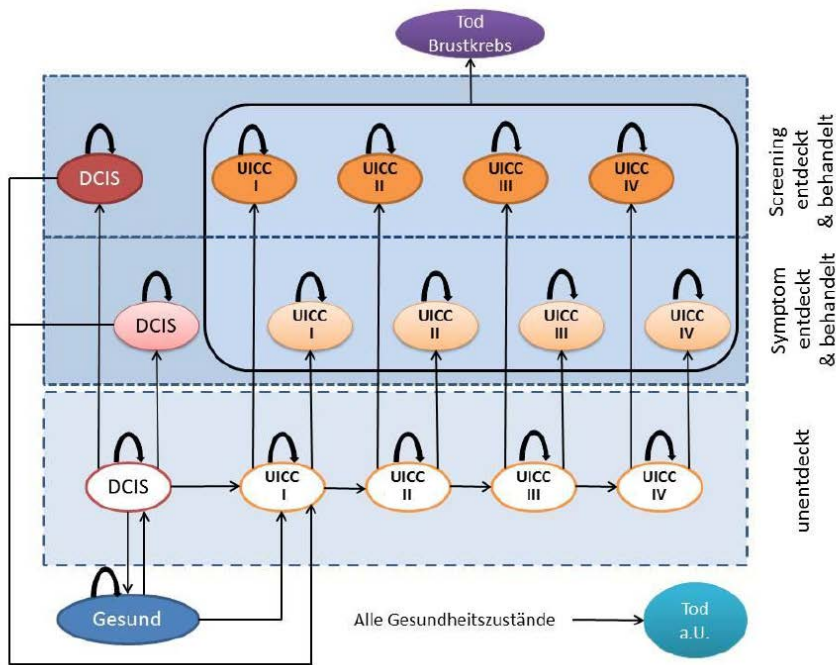


Abbildung 1 Natürlicher Krankheitsverlauf von Brustkrebs im Markov Zustand-Übergangs-Modell.

DCIS - Ductales Carcinoma in situ, UICC - Klassifikation der Union for International Cancer Control, a. U. andere Ursachen.

Jede Blase steht für einen Gesundheitszustand. Jeder Pfeil steht für mögliche Übergänge zwischen Gesundheitszuständen, die jedes Jahr mit bestimmten Übergangswahrscheinlichkeiten auftreten können. Alle Frauen beginnen im gesunden Zustand. Im Laufe der Zeit können die Frauen unentdecktes DCIS oder unentdeckten Brustkrebs im Stadium UICC I entwickeln. Unentdecktes DCIS kann zum Zustand Gesund regredieren. Unentdeckter Brustkrebs im Stadium UICC I kann zu fortgeschrittenen Stadien UICC II-IV fortschreiten. Unentdecktes DCIS und unentdeckter Krebs kann in jedem Stadium durch Symptome oder Screening diagnostiziert werden. Frauen mit diagnostiziertem DCIS oder diagnostiziertem Krebs (durch Symptome oder Screening) gehen in den jeweiligen diagnostizierten Gesundheitszustand und erhalten eine der Diagnose entsprechende Behandlung. Frauen mit diagnostiziertem DCIS können fortschreiten zu einem Zustand mit unentdecktem Brustkrebs im Stadium UICC I. Frauen mit diagnostiziertem Brustkrebs können an Brustkrebs versterben, wobei die Sterbewahrscheinlichkeit und somit die Zeit bis zum Tod vom initial entdeckten Tumorstadium bestimmt wird. Alle Frauen können aus einem beliebigen Zustand heraus an anderen Ursachen sterben entsprechend der alters- und geschlechtsspezifischen Sterblichkeit in Deutschland.

[https://www.iqwig.de/download/s21-01\\_altersgrenzen-im-mammografie-screening-programm\\_abschlussbericht\\_v1-1.pdf](https://www.iqwig.de/download/s21-01_altersgrenzen-im-mammografie-screening-programm_abschlussbericht_v1-1.pdf)

# Transparente Beschreibung Modellparameter: Progression, Mortalität, Klein. Variablen, Testgenauigkeit, Lebensqualität, ...

**Tabelle 1** Modellparameter des natürlichen Krankheitsverlaufs.

Übergang von	nach	Alter (Jahre)	Jährliche Übergangswahrscheinlichkeit	Quelle			
Gesund	DCIS unentdeckt	0-12	0	kalibriert <sup>1</sup>			
		13-29	0,000068				
		30-39	0,002047				
		40-45	0,004536				
		45-49	0,004353				
		50-54	0,004288				
		55-59	0,003184				
		60-64	0,002534				
		65+	0,000122				
		Gesund	UICC I unentdeckt		0-14	0	kalibriert <sup>2</sup>
15-19	0,000001						
20-24	0,000018						
25-29	0,000184						
30-34	0,000449						
35-39	0,000663						
40-44	0,001402						
45-49	0,001611						
50-54	0,002361						
55-59	0,002802						
60-64	0,002910						
65+	0,002993						
	0,003311						
DCIS unentdeckt	Gesund				0,068184	Schiller-Fruehwirth 2017 (20), kalibriert <sup>2</sup>	
				DCIS symptomatisch entdeckt <sup>2</sup>	0,009446		
		UICC I unentdeckt	0,123799				
UICC I unentdeckt	UICC I symptomatisch entdeckt <sup>2</sup>		0,163963				
		UICC II unentdeckt <sup>2</sup>	0,339696				
		UICC II symptomatisch entdeckt <sup>2</sup>	0,394884				
UICC II unentdeckt	UICC III symptomatisch entdeckt <sup>2</sup>		0,391166				
		UICC III symptomatisch entdeckt <sup>2</sup>	0,528699				
		UICC IV unentdeckt <sup>2</sup>	0,619020				
UICC IV unentdeckt	UICC IV symptomatisch entdeckt <sup>2</sup>		0,829060				

UICC - Union for International Cancer Control classification, DCIS - Ductales Carcinoma in situ

**Tabelle 2** Relative Überlebenschancen für Patientinnen mit entdecktem Brustkrebs.

Jahr nach Erstdiagnose	Relative Überlebenschancen für Patientinnen nach Erstdiagnose von entdecktem Brustkrebs (22)			
	UICC I*	UICC II*	UICC III*	UICC IV*
	1	1	1	1
<b>1. Jahr</b>	1,000	0,998	0,983	0,739
<b>2. Jahr</b>	1,000	0,990	0,960	0,783
<b>3. Jahr</b>	1,000	0,979	0,954	0,763
<b>4. Jahr</b>	1,000	0,979	0,958	0,790
<b>5. Jahr</b>	1,000	0,977	0,961	0,811
<b>6. Jahr</b>	1,000	0,979	0,960	0,834
<b>7. Jahr</b>	1,000	0,978	0,962	0,852
<b>8. Jahr</b>	0,997	0,980	0,969	0,856
<b>9. Jahr</b>	0,997	0,980	0,967	0,884
<b>10. Jahr</b>	0,993	0,978	0,966	0,882
<b>11. Jahr</b>	0,994	0,982	0,970	0,896
<b>12. Jahr</b>	0,994	0,977	0,967	0,967
<b>13. Jahr</b>	0,990	0,980	0,977	0,905
<b>14. Jahr</b>	0,991	0,979	0,974	0,914
<b>15. Jahr</b>	0,991	0,981	0,978	0,948

UICC - Union for International Cancer Control classification  
\* Daten wurden auf UICC transferiert

**Tabelle 4** Testgenauigkeit (Sensitivität und Spezifität) der Mammographie in Abhängigkeit von Brustdichte-Kategorie, Alter und Screeningintervall (27).

Brustdichte-Kategorie	Altersgruppe (Jahre)	Intervall <sup>1</sup>	Sensitivität (invasiver Krebs)	Sensitivität (DCIS)	Spezifität		
ACR A	40-49	Erste	0,921	0,954	0,872		
			Jährlich	0,806	0,919	0,930	
		Zweijährlich	0,881	0,942	0,925		
			Erste	0,948	0,955	0,903	
			Jährlich	0,868	0,921	0,948	
	ACR B	50-64	Zweijährlich	0,921	0,943	0,944	
				Erste	0,963	0,955	0,916
			Jährlich	0,903	0,922	0,955	
				Erste	0,943	0,944	0,952
				Jährlich	0,894	0,948	0,797
ACR C		40-49	Erste	0,751	0,911	0,884	
				Jährlich	0,844	0,935	0,876
			Zweijährlich	0,930	0,949	0,843	
				Erste	0,826	0,912	0,912
				Jährlich	0,895	0,937	0,906
	ACR D	50-64	Erste	0,950	0,950	0,863	
				Jährlich	0,871	0,913	0,924
			Zweijährlich	0,924	0,937	0,919	
				Erste	0,817	0,964	0,760
				Jährlich	0,615	0,937	0,860
ACR D		40-49	Erste	0,740	0,955	0,851	
				Jährlich	0,818	0,956	0,812
			Zweijährlich	0,909	0,965	0,836	
				Erste	0,782	0,938	0,908
				Jährlich	0,865	0,956	0,901
	ACR D	50-64	Erste	0,746	0,943	0,815	
				Jährlich	0,512	0,902	0,895
			Zweijährlich	0,852	0,939	0,888	
				Erste	0,822	0,944	0,857
				Jährlich	0,623	0,904	0,921
ACR D		≥65	Zweijährlich	0,747	0,930	0,915	
				Erste	0,868	0,944	0,876
			Jährlich	0,702	0,904	0,932	
				Zweijährlich	0,808	0,931	0,927
				Jährlich			

ACR - American College of Radiology, DCIS - Ductales Carcinoma in situ  
ACR A: fast vollständig fettig (< 25% drüsenförmig)

**Tabelle 3** Altersspezifischer Anteil (in %) der Brustdichte-Kategorien in der weiblichen Bevölkerung in Deutschland (24-26).

Alter (Jahre)	Brustdichte-Kategorie			
	ACR A (%)	ACR B (%)	ACR C (%)	ACR D (%)
45-49§	4,8	31,9	55,3	8,0
50-54	6,1	38,7	48,3	6,9
55-59	7,4	45,5	41,3	5,8
60-64	8,6	48,5	37,9	5,0
65-69	9,45	50,0	36,1	4,45
70+	10,0	50,85	35,0	4,15

ACR - American College of Radiology  
ACR A: fast vollständig fettig (< 25% drüsenförmig)  
ACR B: verstreute fibroglanduläre Dichten (25-50% drüsenförmig)  
ACR C: ungleichmäßig dicht (51-75% drüsenförmig)  
ACR D: extrem dicht (> 75% drüsenförmig)  
§ berechnet aus den absoluten Differenzen der Alterskategorie 50-54 vs. 55-59 in der jeweiligen ACR Kategorie

**Tabelle 5** Relatives Risiko für Brustkrebs in Abhängigkeit von Alter und Brustdichte-Kategorie (24).

Altersgruppe (Jahre)	Brustdichte-Kategorie			
	ACR A	ACR B	ACR C	ACR D
40-49§	0,376	0,783	1,212	1,573
50-59	0,388	0,807	1,251	1,623
60-64	0,400	0,832	1,291	1,675
65-69	0,581	0,885	1,228	1,283
70+	0,600	0,914	1,268	1,325

ACR - American College of Radiology  
ACR A: fast fettig (< 25% drüsenförmig)  
ACR B: verstreute fibroglanduläre Dichten (25-50% drüsenförmig)  
ACR C: ungleichmäßig dicht (51-75% drüsenförmig)  
ACR D: extrem dicht (> 75% drüsenförmig)  
§ berechnet mit relativen Gewichten der Alterskategorien (50-54) zu (55-59): Faktor 0,376 - 1,573

**Tabelle 6** Nutzwerte für die gesundheitsbezogene Lebensqualität.

Lebensqualität deutsche weibliche Bevölkerung (EQ-5D)			
Alter (Jahre)	Nutzwert	Quelle	
0-17	1	(28)	
18-24	0,950		
25-34	0,949		
35-44	0,943		
45-54	0,908		
55-64	0,881		
65-74	0,838		
75+	0,771		
Nutzwert-Dekrement für Screening (TTO)			
Mammographie	Nutzwert Dekrement	Dauer	Dekrement über 1 Jahr
Mammographie	0,196	2 Wochen	0,0075 (24)
Positiver Befund	0,105	5 Wochen	0,0100 (31)

[https://www.iqwig.de/download/s21-01\\_altersgrenzen-im-mammografie-screening-programm\\_abschlussbericht\\_v1-1.pdf](https://www.iqwig.de/download/s21-01_altersgrenzen-im-mammografie-screening-programm_abschlussbericht_v1-1.pdf)



# Kalibrierung

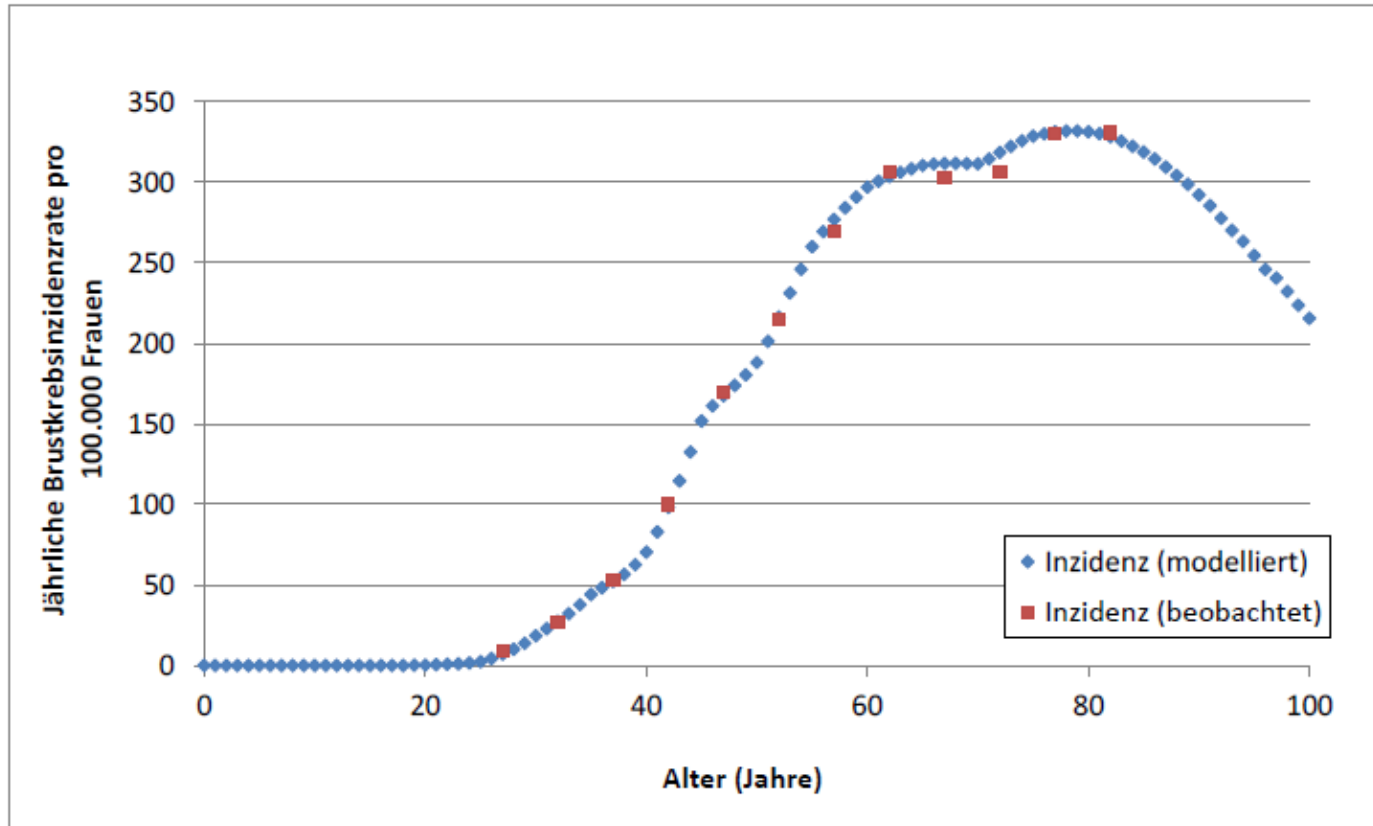
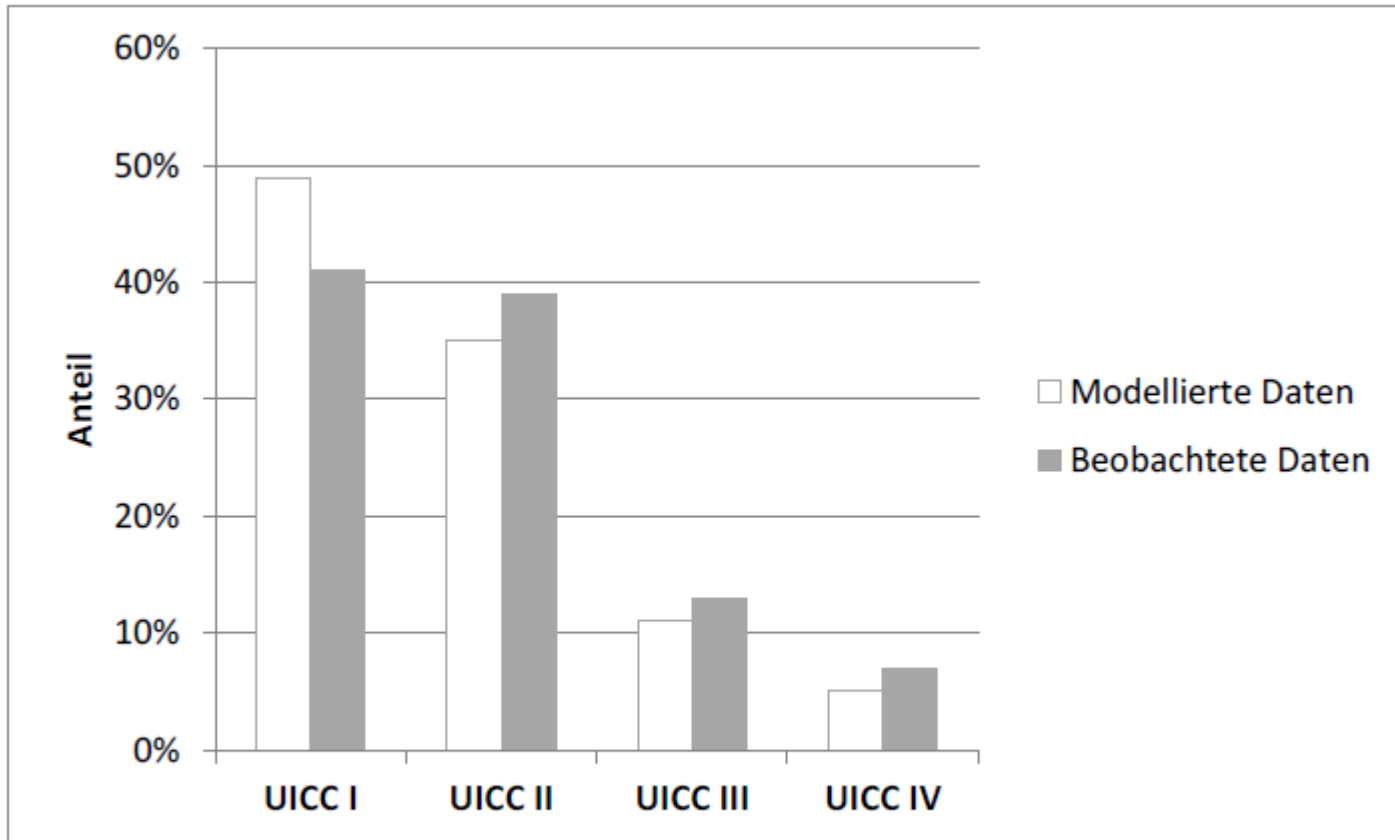


Abbildung 3 Kalibrierungsergebnis: Altersspezifische jährliche Inzidenzrate von symptomatisch entdecktem Brustkrebs pro 100.000 Frauen.

# Validierung



**Abbildung 5 Ergebnisse der externen Validierung: UICC-Stadienverteilung entdeckter Brustkrebsfälle.**

# Ergebnisdarstellung multiple Outcomes

Tabelle 10 Basisfallanalyse: Restlebenserwartung und für die lebensqualitätsadjustierte Restlebenserwartung pro 100 Frauen.

Strategie	Restlebenserwartung pro 100 Frauen (nach dem 45. Lebensjahr)			Lebensqualitätsadjustierte Restlebenserwartung pro 100 Frauen (nach dem 45. Lebensjahr)		
	Lebensjahre	Inkrementelle Lebensjahre	Inkrementelle Lebensjahre	QALY	Inkrementelle QALY	Inkrementelle QALY
		vs. Kein Screening	vs. Referenz		vs. Kein Screening	vs. Referenz
Kein Screening	3.847,7	---	---	3.310,0	---	---
Screening: Alter 50-69 J, 2 J (Ref)	3.870,0	22,4	---	3.322,5	12,5	---
Screening: Alter 50-69 J, 2 J; 70-74 J, 3 J	3.871,2	23,5	1,1	3.322,9	13,0	0,5
Screening: Alter 50-74 J, 2 J	3.872,5	24,8	2,4	3.323,1	13,1	0,6
Screening: Alter 50-69 J, 2 J; 70-79 J, 3 J	3.872,6	25,0	2,6	3.323,2	13,3	0,8
Screening: Alter 50-79 J, 2 J	3.873,4	25,7	3,3	3.322,9	13,0	0,4
Screening: Alter 45-69 J, 2 J	3.876,0	28,4	6,0	3.325,6	15,6	3,1
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J	3.877,1	29,4	7,0	3.325,0	15,0	2,5
<b>Screening: Alter 45-74 J, 2 J</b>	3.877,7	30,0	7,6	<b>3.326,0</b>	<b>16,0</b>	<b>3,5</b>
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-74 J, 3 J	3.878,1	30,4	8,1	3.325,4	15,4	2,9
Screening: Alter 45-49 J, 1 J; 50-74 J, 2 J	3.878,7	31,0	8,6	3.325,4	15,4	2,9
Screening: Alter 45-79 J, 2 J	3.879,0	31,4	9,0	3.325,8	15,8	3,3
Screening: Alter 45-49 J, 1 J; 50-69 J, 2 J; 70-79 J, 3 J	3.879,4	31,7	9,3	3.325,6	15,6	3,1
<b>Screening: Alter 45-49 J, 1 J; 50-79 J, 2 J</b>	<b>3.880,0</b>	<b>32,4</b>	<b>10,0</b>	3.325,2	15,2	2,7

J – Jahre; QALY – qualitätsadjustiertes Lebensjahr (quality-adjusted life year), Ref - Referenzstrategie

Fett gedruckt: die Strategien mit höchstem Nutzenzuwachs hinsichtlich Restlebenserwartung und lebensqualitätsadjustierte Restlebenserwartung.



# Ergebnisdarstellung Trade-offs

Stufenweise inkrementelle Betrachtungsweise, Effizienzgrenzen

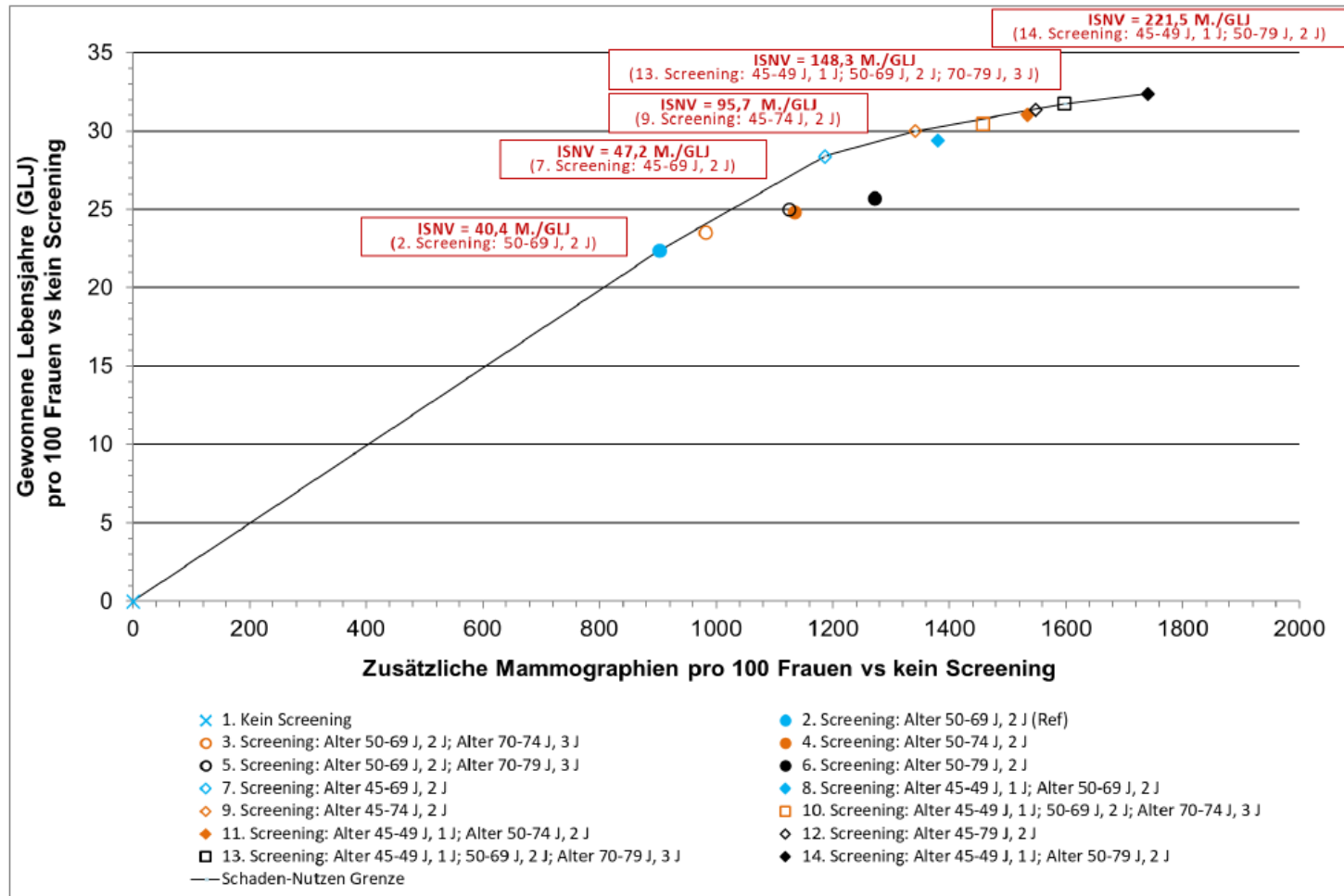


Abbildung 7 Ergebnisse der Basisfallanalyse: Gesamtanzahl an zusätzlichen Mammographien versus gewonnene Lebensjahre.

ISNV - inkrementelles Schaden-Nutzen-Verhältnis; J - Jahr; M/GLJ - Mammographie pro gewonnenes Lebensjahr

[https://www.iqwig.de/download/s21-01\\_altersgrenzen-im-mammografie-screening-programm\\_abschlussbericht\\_v1-1.pdf](https://www.iqwig.de/download/s21-01_altersgrenzen-im-mammografie-screening-programm_abschlussbericht_v1-1.pdf)

# Unsicherheitsanalysen

- Im Bericht auf 26 Seiten Ergebnisse der Sensitivitätsanalysen (hier nicht dargestellt)

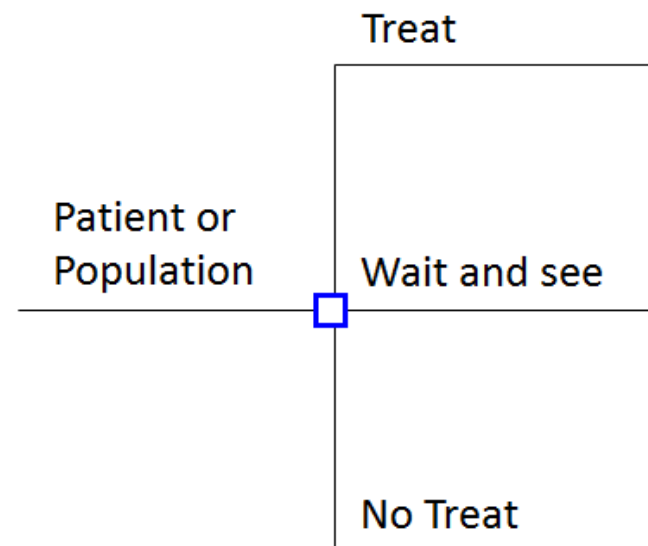
# Kausale Modellierung



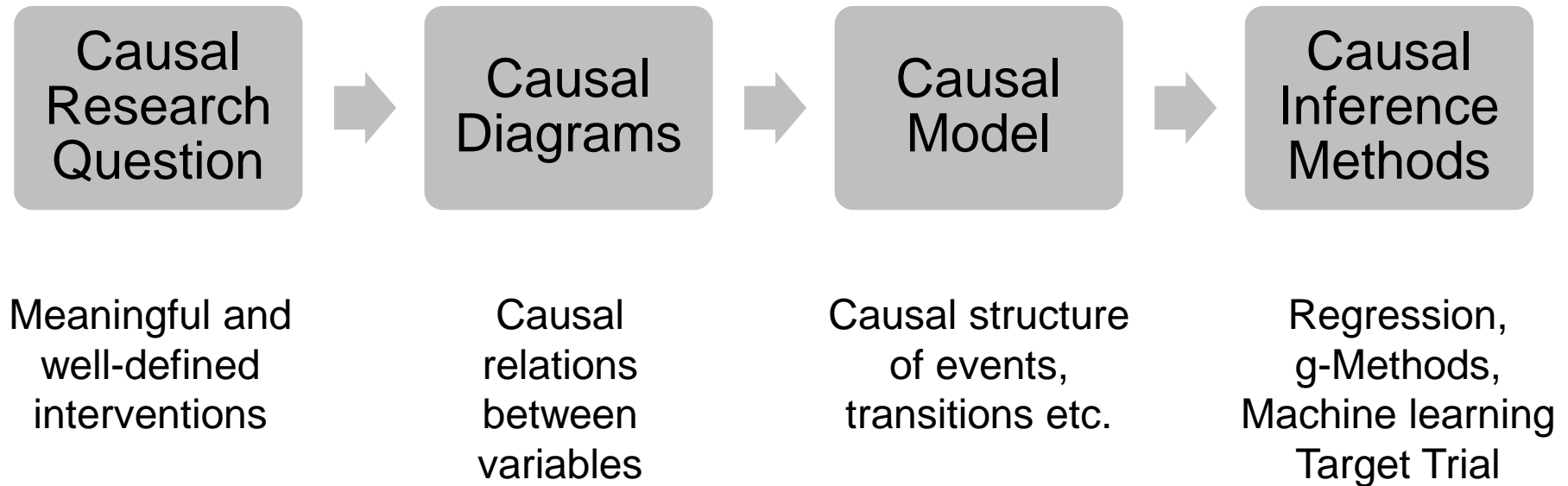
# Kausalität in Entscheidungsanalysen

- Entscheidungsanalytische Modelle müssen per se kausal interpretierbar sein
  - im Unterschied zu Prädiktionsmodellen
- Entscheidung = **What if** = kausal/kontrafaktisch
- Die Zielpopulation durchläuft jeden alternativen Handlungsarm

- **Entscheidungsknoten**



# Ingredients of Causal Decision Modeling



# Interface Modeling and Causal Inference

## Abstract

**Objectives:** Public health decision making is a complex process based on thorough and comprehensive health technology assessments involving the comparison of different strategies, values and tradeoffs under uncertainty. This process must be based on best available evidence and plausible assumptions. Causal inference and health decision science are two methodological approaches providing information to help guide decision making in health care. Both approaches are quantitative methods that use statistical and modeling techniques and simplifying assumptions to mimic the complexity of the real world. We intend to review and lay out both disciplines with their aims, strengths and limitations based on a combination of textbook knowledge and expert experience.

**Methods:** To help understanding and differentiating the methodological approaches of causal inference and health decision science, we reviewed both methods with the focus on aims, research questions, methods, assumptions, limitations and challenges, and software. For each methodological approach, we established a group of four experts from our own working group to carefully review and summarize each method, followed by structured discussion rounds and written reviews, in which the experts from all disciplines including HTA and medicine were involved. The entire expert group discussed objectives, strengths and limitations of both methodological areas, and potential synergies. Finally, we derived recommendations for further research and provide a brief outlook on future trends.

**Results:** Causal inference methods aim for drawing causal conclusions from empirical data on the relationship of pre-specified interventions on a specific target outcome and apply a counterfactual framework and statistical techniques to derive causal effects of exposures or interventions from empirical data. Causal inference is based on a causal diagram, more precisely a directed acyclic graph (DAG), which encodes the assumed causal relations between variables. Depending on the method used, traditional statistical methods are applied to derive valid causal effects. Health decision science is a discipline that aims at supporting health interventions. The main objectives of a decision based on empirical data and studies, for example prevalence studies, observational routine effectiveness studies, and costs. It involves decision analysis, a systematic, explicit and quantitative framework to guide decisions under uncertainty. Decision

Felicitas Kühne<sup>1</sup>  
 Michael Schomaker<sup>1,2</sup>  
 Igor Stojkov<sup>1</sup>  
 Beate Jahn<sup>1,3</sup>  
 Annette Conrads-Frank<sup>1</sup>  
 Silke Siebert<sup>1</sup>  
 Gaby Sroczyński<sup>1</sup>  
 Sibylle Puntschner<sup>1</sup>  
 Daniela Schmid<sup>1</sup>  
 Petra Schnell-Inderst<sup>1</sup>  
 Uwe Siebert<sup>1,3,4,5</sup>

1 Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT TIROL - University for Health Sciences, Medical Informatics and Technology, Hall T.T., Austria

2 Centre for Infectious Disease Epidemiology & Research, University of Cape Town, South Africa

3 Division of Health Technology Assessment, ONCOTYROL - Center for Personalized Cancer Medicine, Innsbruck, Austria

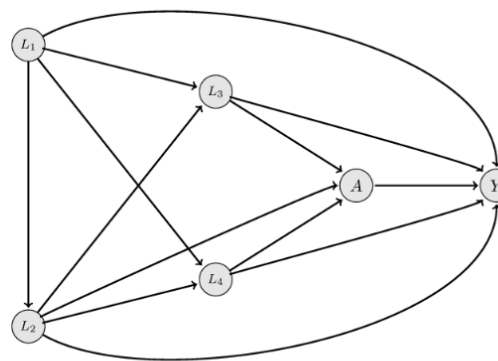
4 Center for Health Decision Science, Departments of Epidemiology and Health Policy & Management, Harvard T.H. Chan School of Public Health, Boston, MA, USA

159 references



GMS German Medical Science 2022, Vol. 20, ISSN 1612-3174

1/24



Y = Death; A = cancer treatment; L<sub>1</sub> = age; L<sub>2</sub> = SES; L<sub>3</sub> = comorbidities; L<sub>4</sub> = clinical stage

Figure 2: Directed acyclic graph for cancer

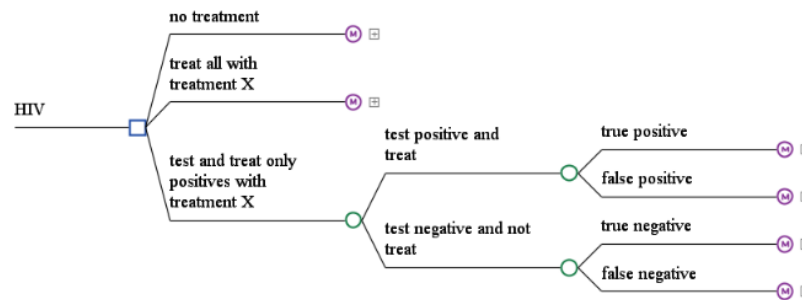


Figure 3: Potential decision tree as start of decision problem in HIV-positive patients

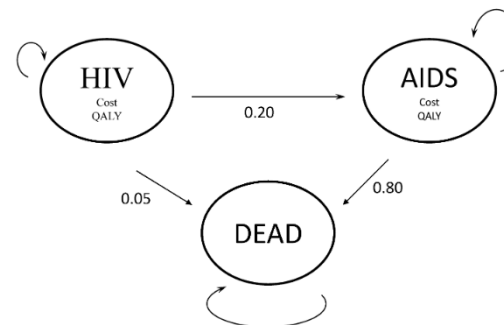


Figure 4: State transition diagram

Kuehne et al., GMS, 2022

# Best Practice Causal Inference

## ISPOR Good Research Practices for Retrospective Databases Analysis Task Force -- Part I-III

Volume 12 • Number 8 • 2009  
VALUE IN HEALTH

Volume 12 • Number 8 • 2009  
VALUE IN HEALTH

Volume 12 • Number 8 • 2009  
VALUE IN HEALTH

### Good Research Practices for Comparative Effectiveness Research: Analytic Methods to Improve Causal Inference from Nonrandomized Studies of Treatment Effects Using Secondary Data Sources: The ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part III

Michael L. Johnson, PhD,<sup>1</sup> William Crown, PhD,<sup>2</sup> Bradley C. Martin, PharmD, PhD,<sup>3</sup> Colin R. Dormuth, MA, ScD,<sup>4</sup> Uwe Siebert, MD, MPH, MSc, ScD<sup>5</sup>

<sup>1</sup>University of Houston, College of Pharmacy, Department of Clinical Sciences and Administration, Houston, TX, USA; Houston Center for Quality of Care and Utilization Studies, Department of Veteran Affairs, Michael E. DeBakey VA Medical Center, Houston, TX, USA; <sup>2</sup>St. Joseph, William, MA, USA; <sup>3</sup>Division of Pharmaceutical Evaluation and Policy, College of Pharmacy, University of Arkansas for Medical Sciences, Little Rock, AR, USA; <sup>4</sup>Department of Anesthesiology, Pharmacology & Therapeutics, University of British Columbia; Pharmacopidemiology Group, Therapeutics Initiative, Vancouver, BC, Canada; <sup>5</sup>Department of Public Health, Information Systems and Health Technology Assessment, UMIT, University of Health Sciences, Medical Informatics and Technology, Hiltl IT, Austria; Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA, and Department of Health Policy and Management, Harvard School of Public Health, Boston, MA, USA

#### ABSTRACT

**Objectives:** Most contemporary epidemiologic studies require complex analytical methods to adjust for bias and confounding. New methods are constantly being developed, and older more established methods are yet appropriate. Careful application of statistical analysis techniques can improve causal inference of comparative treatment effects from nonrandomized studies using secondary databases. A Task Force was formed to offer a review of the more recent developments in statistical control of confounding.

**Methods:** The Task Force was commissioned and a chair was selected by the ISPOR Board of Directors in October 2007. This Report, the third in this issue of the journal, addressed methods to improve causal inference of treatment effects for nonrandomized studies.

**Results:** The Task Force Report recommends general analytic techniques and specific best practices where consensus is reached including: use of

stratification analysis before multivariable modeling, multivariable regression including model performance and diagnostic testing, propensity scoring, instrumental variable, and structural modeling techniques including marginal structural models, where appropriate for secondary data. Sensitivity analyses and discussion of extent of residual confounding are discussed.

**Conclusions:** Valid findings of causal therapeutic benefits can be produced from nonrandomized studies using an array of state-of-the-art analytic techniques. Improving the quality and uniformity of these studies will improve the value to patients, physicians, and policymakers worldwide.

**Keywords:** causal inference, comparative effectiveness, nonrandomized studies, research methods, secondary databases.

#### Background to the Task Force

In September 2007, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Health Science Policy Council recommended that the issue of establishing a Task Force to recommend Good Research Practices for Designing and Analyzing Retrospective Databases be considered by the ISPOR Board of Directors. The Council's recommendations concerning this new Task Force were to keep an overarching view toward the need to ensure internal validity and improve causal inference from observational studies, review prior work from past and ongoing ISPOR task forces and other initiatives to establish baseline standards from which to set an agenda for work. The ISPOR Board of Directors approved the creation of the Task Force in October 2007. Task Force leadership and reviewer groups were finalized by December 2007, and the first teleconference took place in January 2008.

**Address correspondence to:** Michael L. Johnson, University of Houston, College of Pharmacy, Department of Clinical Sciences and Administration, Houston, TX 77030, USA. E-mail: mikejohnson@uh.edu  
10.1111/j.1524-4733.2009.00602.x

Task Force members were experienced in medicine, epidemiology, biostatistics, public health, health economics, and pharmacy sciences, and were drawn from industry, academia, and as advisors to governments. The members came from the UK, Germany, Austria, Canada, and the United States.

Beginning in January 2008, the Task Force conducted monthly teleconferences to develop core assumptions and an outline before preparing a draft report. A face-to-face meeting took place in October 2008, to develop the draft, and three forums took place at the ISPOR Meetings to develop consensus for the final draft reports. The draft reports were posted on the ISPOR website in May 2009 and the task forces' reviewer group and ISPOR general membership were invited to submit their comments for a 2 week reviewer period. In total, 38 responses were received. All comments received were posted to the ISPOR website and presented for discussion at the Task Force forum during the ISPOR 12th Annual International Meeting in May 2009. Comments and feedback from the forum and reviewer and membership responses were considered and acknowledged in the final reports. Once consensus was reached, the manuscript was submitted to *Value in Health*.

© 2009, International Society for Pharmacoeconomics and Outcomes Research (ISPOR) 1098-3015/09/12 08-12

1

1. Berger et al. ... defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources ... Part I., Value in Health 2009
2. Cox et al., ... approaches to mitigate bias and confounding in the design of nonrandomized studies of treatment effects using secondary data sources ... Part II, Value in Health 2009
3. Johnson et al., ... analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources ... Part III, Value in Health 2009

# Best Practice Causal Inference

## ISPOR Good Research Practices for Retrospective Databases Analysis Task Force -- Part I-III

### Recommendation Part II (Cox et al, ViH 2009):

- "[...] the need for data analysis plan with **causal diagrams**; detailed attention to classification bias in definition of exposure and clinical outcome; careful and appropriate use of restriction; extreme care to **identify and control for confounding factors**, including **time-dependent confounding**."



# Best Practice Causal Inference

## ISPOR Good Research Practices for Retrospective Databases Analysis Task Force -- Part I-III

### Recommendations Part III (Johnson et al, ViH 2009):

- "All factors that are theoretically related to outcome or treatment selection should be included **despite statistical significance** at traditional levels of significance."
- "In the presence of **time-varying confounding**, standard statistical methods may be biased, and alternative methods such as **marginal structural models or g-estimation** should be examined."

# UKPDS Outcomes Model

Diabetologia (2013) 56:1925–1933  
DOI 10.1007/s00125-013-2940-y

ARTICLE

## UKPDS Outcomes Model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82

A. J. Hayes · J. Leal · A. M. Gray · R. R. Holman · P. M. Clarke

Received: 2 January 2013 / Accepted: 24 April 2013 / Published online: 22 June 2013  
© Springer-Verlag Berlin Heidelberg 2013

### Abstract

**Aims/hypothesis** The aim of this project was to build a new version of the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS-OM1), a patient-level simulation tool for predicting lifetime health outcomes of people with type 2 diabetes mellitus.

**Methods** Data from 5,102 UKPDS patients from the 20 year trial and the 4,031 survivors entering the 10 year post-trial monitoring period were used to derive parametric proportional hazards models predicting absolute risk of diabetes complications and death. We re-estimated the seven original event equations and estimated new equations for diabetic ulcer and some second events. The additional data permitted inclusion of new risk factor predictors such as estimated GFR. We also developed four new equations for all-cause mortality. Internal validation of model predictions of cumulative

incidence of all events and death was carried out and a contemporary patient-level dataset was used to compare 10 year predictions from the original and the new models.

**Results** Model equations were based on a median 17.6 years of follow-up and up to 89,760 patient-years of data, providing double the number of events, greater precision and a larger number of significant covariates. The new model, UKPDS-OM2, is internally valid over 25 years and predicts event rates for complications, which are lower than those from the existing model.

**Conclusions/interpretation** The new UKPDS-OM2 has significant advantages over the existing model, as it captures more outcomes, is based on longer follow-up data, and more comprehensively captures the progression of diabetes. Its use will permit detailed and reliable lifetime simulations of key health outcomes in people with type 2 diabetes mellitus.

# UKPDS Outcomes Model

**Start: define the following patient characteristics:**

**Age at diagnosis**, ethnicity, sex, BMI, HbA<sub>1c</sub>, total: HDL cholesterol (Lipids), blood pressure, smoking status, atrial fibrillation at diagnosis, PVD

**History of diabetes-related events:** IHD, CHF, blindness, amputation, renal failure, MI, stroke

**Update patient risk factors using risk factor equations:**

HbA <sub>1c</sub>	Eq.11
Blood pressure	Eq. 12
Total: HDL cholesterol	Eq. 13
Smoking	Eq. 14

**Update** history of diabetes-related events

**Calculate** life years & QALYs

Yes

No

**Dead?**

**Commence model cycle**

**Randomly order and run event equations:**

Ischaemic heart disease (IHD)	Eq. 1
Myocardial infarction (MI)	Eq. 2
Congestive heart failure (CHF)	Eq. 3
Stroke	Eq. 4
Amputation	Eq. 5
Blindness	Eq. 6
Renal failure	Eq. 7
Diabetes-related mortality	Eq. 8
<i>(conditional on CHF, amputation @ renal Failure, MI or stroke having occurred)</i>	Eq. 9
Other mortality	Eq. 10

# \*UKPDS Outcomes Model

- estimation of first occurrence of diabetes related complications (MI, other IHD, stroke, heart failure, amputation, renal failure, blindness)
- equations were fitted to observed events in UKPDS population

- diabetes-related complications: proportional hazards Weibull regression model

$$h(t|x_{tj}) = h_0(t) \exp(x_{tj}\beta_j); \quad h_0(t) = \lambda \gamma t^{\gamma-1}$$

- probability that first MI, IHD, CHF amputation or renal failure would be fatal: logistic regression

$$p_t = (1 + \exp(-x_{tj}\beta_j))$$

- risk of diabetes-related/non-diabetes related mortality: Gompertz regression function

$$h(t|x_{tj}) = h_0(t) \exp(x_{tj}\beta_j); \quad h_0(t) = \lambda \exp(\varphi t)$$

- risk factor progression: random effects panel data

$$RF_{it} = \alpha + \beta_j x_{itj} + \mu_i + \nu_{it}$$

- probabilistic discrete-time illness death model with annual cycles

$$H(t|x_{tj}) = \exp(\beta_0 + x_{tj}\beta_j)t^\gamma$$

$$1 - \exp(H(t|x_{tj}) - H(t+1|x_{tj}))$$

# UKPDS Outcomes Model

**Table 2.** Sample size, functional form, parameters and beta coefficients (SEs) for seven equations to estimate the probability of diabetes-related complications

	Eq. 1	Eq. 2	Eq. 3	Eq. 4	Eq. 5	Eq. 6	Eq. 7
Complication	IHD	MI	CHF	STROKE	AMP	BLIND	RENAL
No. of subjects	3612	3642	3607	3607	3642	3642	3642
Functional form	Weibull	Weibull	Weibull	Weibull	Weibull	Weibull	Weibull
Parameters	Estimate of coefficient (SE)						
$\lambda$	-5.310 (0.174)	-4.977 (0.160)	-8.018 (0.408)	-7.163 (0.342)	-8.718 (0.613)	-6.464 (0.326)	-10.016 (0.939)
$\rho$	1.150 (0.067)	1.257 (0.060)	1.711 (0.158)	1.497 (0.126)	1.451 (0.232)	1.154 (0.121)	1.865 (0.387)
AGE	0.031 (0.008)	0.055 (0.006)	0.093 (0.016)	0.085 (0.014)		0.069 (0.014)	
FEMALE	-0.471 (0.143)	-0.826 (0.103)		-0.516 (0.171)			
AC		-1.312 (0.341)					
SMOK		0.346 (0.097)		0.355 (0.179)			
BMI			0.066 (0.017)				
HBA1C	0.125 (0.035)	0.118 (0.025)	0.157 (0.057)	0.128 (0.042)	0.435 (0.066)	0.221 (0.050)	
SBP	0.098 (0.037)	0.101 (0.026)	0.114 (0.056)	0.276 (0.042)	0.228 (0.075)		0.404 (0.106)
TOTAL:HDL				0.113 (0.025)			
Ln (TOTAL:HDL)	1.498 (0.202)	1.190 (0.169)					
PVD					2.436 (0.521)		
ATRFIB				1.428 (0.472)			
IHD		0.914 (0.150)					
CHF		1.558 (0.202)		1.742 (0.287)			
BLIND					1.812 (0.462)		2.082 (0.551)

# EMA ICH e9 Addendum - Estimands

[https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/draft-ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical_en.pdf)

1 30 August 2017  
2 EMA/CHMP/ICH/436221/2017  
3 Committee for Human Medicinal Products

- 4 ICH E9 (R1) addendum on estimands and sensitivity  
5 analysis in clinical trials to the guideline on statistical  
6 principles for clinical trials  
7 Step 2b

Transmission to CHMP	July 2017
Adoption by CHMP for release for consultation	20 July 2017
Start of consultation	31 August 2017
End of consultation (deadline for comments)	28 February 2018

8

9

Comments should be provided using this [template](#). The completed comments form should be sent to [ich@ema.europa.eu](mailto:ich@ema.europa.eu)

10

11



# Kausale Modellierung

Zwei Beispiele, bei denen die Wahl des kausalen Estimands eine entscheidende Rolle spielt:

- Behandlungswechsel
- Compliance

*Extrem relevant für entscheidungsanalytische Kosteneffektivitätsmodellierungen, viele Entscheidungen sind gekippt.*

# \*Kausale Modellierung

## Behandlungswechsel

- Forschungsfrage für Erstattung: Welt mit vs. ohne neues Medikament
- In RCT aus ethischen Gründen Entblindung und Behandlungswechsel nach Progression zugelassen
- => Studiendesign kann Forschungsfrage nicht beantworten
- Wahl zwischen 2 Übeln:
  - (1) ITT-Analyse unbiased, aber falscher Estimand
  - (2) Kausale (nicht-naive) Per-Protocol-Analyse korrekter Estimand, aber risk of bias



# \*Kausale Modellierung

## Compliance

- Unterschiedliche Arten von Participation, Compliance und Adherence sind zu unterscheiden!
- In HTA interessiert of der „kontrollierte“ Effekt wie in RCTs und nicht in der Real-world (geringere Unsicherheit)
  - Kann in Modellierungen simuliert werden, z.B., Analyse für am Screening teilnehmende Personen
  - Für klinische Guidelines korrekter Ansatz, da Personen selbst wählen, ob sie teilnehmen oder nicht
  - Für KNB problematisch, wenn fixe Kosten höher sind als variable (Screeningimplementation)
  - Bsp.: Dickdarmkrebsscreening: Nationales Screening Komitee Österreich ließ in Empfehlungen die Wahl zwischen Koloskopie oder Blutstuhltests

# \*Using Causal Estimands as Model Inputs

## Types of strategies specified in the ICH e9 Addendum:

1. Treatment-policy: outcome collected regardless of treatment switching, compliance etc.

ITT analysis (no adjustment)

→ causal policy effect

2. Hypothetical strategy: what if the patients did not switch, were compliant etc.

Causal (non-naïve) per-protocol analysis (adjusting for post-baseline variables, noncompliance etc.)

→ counterfactual (causal) outcome

today using target trial emulation with “cloning”

Further strategies (composite, ...)

# Estimand Depends on Statistical Method



ELSEVIER



Journal of Clinical Epidemiology 152 (2022) 1–13

**Journal of  
Clinical  
Epidemiology**

## ORIGINAL ARTICLE

### Causal analyses with target trial emulation for real-world evidence removed large self-inflicted biases: systematic bias assessment of ovarian cancer treatment effectiveness

Felicitas Kuehne<sup>a</sup>, Marjan Arvandi<sup>a</sup>, Lisa M. Hess<sup>b</sup>, Douglas E. Faries<sup>b</sup>,  
Raffaella Matteucci Gothe<sup>a</sup>, Holger Gothe<sup>a,c</sup>, Julie Beyrer<sup>b</sup>, Alain Gustave Zeimet<sup>d</sup>,  
Igor Stojkov<sup>a</sup>, Nikolai Mühlberger<sup>a</sup>, Willi Oberaigner<sup>a,e</sup>, Christian Marth<sup>d</sup>, Uwe Siebert<sup>a,f,g,\*</sup>

<sup>a</sup>Department of Public Health, Health Services Research and Health Technology Assessment, Institute of Public Health, Medical Decision Making and Health Technology Assessment, UMIT TIROL – University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria

<sup>b</sup>Eli Lilly and Company, Indianapolis, IN, USA

<sup>c</sup>Chair of Health Sciences/Public Health, Medical Faculty “Carl Gustav Carus”, Technical University Dresden, Dresden, Germany

<sup>d</sup>Department of Obstetrics and Gynecology, Innsbruck Medical University, Innsbruck, Austria

<sup>e</sup>Institute for Clinical Epidemiology, Cancer Registry Tyrol, Tirol Kliniken, Innsbruck, Austria

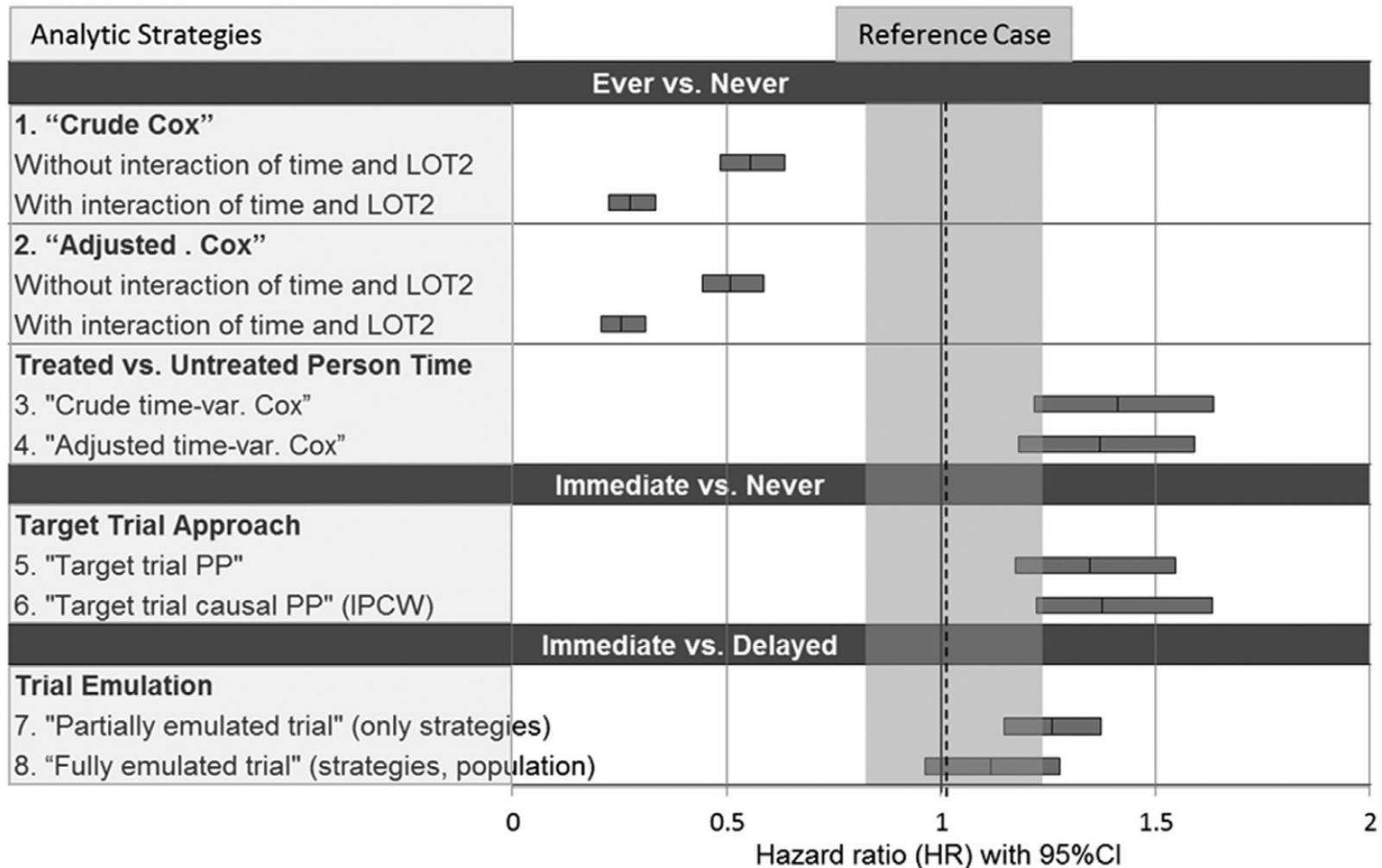
<sup>f</sup>Center for Health Decision Science and Departments of Epidemiology and Health Policy & Management, Harvard T.H. Chan School of Public Health, Boston, MA, USA

<sup>g</sup>Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Accepted 3 October 2022; Published online xxxx

Kuehne et al., JCE, 2022

# Comparative Bias Assessment



# A Scientist's Way to Address Bias in Modeling Studies

## Assumptions of the model

As in prior analyses,<sup>8 11</sup> when data were controversial or incomplete, we attempted to bias model assumptions against antiviral treatment. Although many studies suggest histological improvement or reduced likelihood of hepatocellular carcinoma in relapse patients,<sup>38</sup> we assumed no long term benefit from antiviral treatment relapse. We assumed that spontaneous or treatment induced loss of viraemia greatly reduces but does not eliminate the risk of developing progressive liver disease.<sup>39</sup> Lastly, serial liver biopsies, which would have

Siebert et al., Gut, 2003

# Rolle von AI in EA Modellierungen



# Modeling supported by AI

chatGPT:

- **Which decision-analytic model type** should be used for the assessment of the cost effectiveness of antiviral treatment in patients with chronic hepatitis C?

# Exactly same answer as from NICE



Which decision-analytic model type should be used for the assessment of the cost effectiveness of antiviral treatment in patients with chronic hepatitis C?



The decision-analytic model type commonly used for the assessment of the cost-effectiveness of antiviral treatment in patients with chronic hepatitis C is a Markov model. Markov models are widely employed in health economic evaluations due to their ability to capture the dynamic nature of disease progression over time and evaluate long-term outcomes.



In a Markov model for chronic hepatitis C, the disease progression is represented as a series of discrete health states, such as stages of liver disease (e.g., chronic hepatitis, compensated cirrhosis, decompensated cirrhosis, hepatocellular carcinoma) or treatment response categories. Transitions between these health states occur probabilistically and are influenced by various factors such as treatment efficacy, adverse events, and natural history of the disease. The model typically considers costs associated with healthcare resource utilization (e.g., medication, laboratory tests, hospitalization) and health outcomes (e.g., quality-adjusted life-years or life-years).

By simulating the progression of patients through different health states over time, Markov models allow for the comparison of alternative treatment strategies and the estimation of their long-term costs and health benefits. These models enable decision-makers to assess the cost-effectiveness of antiviral treatment options, inform resource allocation decisions, and evaluate the value of interventions in improving patient outcomes.

It's worth noting that other modeling approaches, such as individual-based models or microsimulation models, can also be used in the assessment of chronic hepatitis C treatment cost-effectiveness. However, Markov models remain a widely utilized and well-established approach in this context due to their simplicity, transparency, and ability to provide valuable insights into the long-term implications of treatment strategies.



# Modeling supported by AI

chatGPT:

- **Which Markov states** should be included in a valid decision-analytic Markov model to assess the cost-effectiveness of antiviral treatment in patients with chronic hepatitis C?

# Not quite there yet, but close

UW

Which Markov states should be included in a valid decision-analytic Markov model to assess the cost-effectiveness of antiviral treatment in patients with chronic hepatitis C?



When constructing a decision-analytic Markov model to assess the cost-effectiveness of antiviral treatment in patients with chronic hepatitis C, the specific Markov states to include will depend on the details of the model and the assumptions made. However, here are some typical Markov states that are often included in such models:



1. Susceptible: This state represents individuals who are at risk of contracting hepatitis C but have not been infected yet.
2. Acute Hepatitis C: This state represents individuals who have recently contracted hepatitis C and are in the acute phase of the infection.
3. Chronic Hepatitis C: This state represents individuals who have progressed from the acute phase to the chronic phase of hepatitis C infection. This is the target population for antiviral treatment.
4. Compensated Cirrhosis: This state represents individuals who have developed liver cirrhosis but have not yet experienced severe complications or decompensation.
5. Decompensated Cirrhosis: This state represents individuals with cirrhosis who have experienced complications such as portal hypertension, ascites, hepatic encephalopathy, or variceal bleeding.
6. Hepatocellular Carcinoma (HCC): This state represents individuals who have developed liver cancer as a result of chronic hepatitis C infection.
7. Liver Transplantation: This state represents individuals who have undergone liver transplantation due to end-stage liver disease caused by hepatitis C.
8. Death: This absorbing state represents individuals who have died from any cause.

These are some common Markov states used in hepatitis C models. The transition probabilities between these states will depend on factors such as the natural history of the disease, treatment efficacy, adverse events, and other relevant parameters. It is important to note that the actual model design may vary depending on the specific research question, available data, and modeling assumptions.

# Modeling supported by AI

chatGPT:

- What are valid **point estimates and 95% confidence intervals of model parameters** for a decision-analytic Markov model for the assessment of the cost effectiveness of antiviral treatment in patients with chronic hepatitis C?

# Definitely beyond capacity

UW

What are valid point estimates and 95% confidence intervals of model parameters for a decision-analytic Markov model for the assessment of the cost effectiveness of antiviral treatment in patients with chronic hepatitis C?



The point estimates and 95% confidence intervals of model parameters for a decision-analytic Markov model assessing the cost-effectiveness of antiviral treatment in patients with chronic hepatitis C can vary depending on the specific data sources, model assumptions, and country context. However, here are examples of commonly used parameters along with hypothetical point estimates and their corresponding 95% confidence intervals:



#### 1. Transition Probabilities:

- Example: Transition probability from chronic hepatitis C to cirrhosis: 0.05 (95% CI: 0.04-0.06)
- Example: Transition probability from compensated cirrhosis to decompensated cirrhosis: 0.10 (95% CI: 0.08-0.12)

#### 2. Treatment Efficacy:

- Example: Antiviral treatment response rate (SVR): 0.80 (95% CI: 0.75-0.85)
- Example: Treatment response rate in patients with cirrhosis: 0.60 (95% CI: 0.55-0.65)

#### 3. Adverse Event Rates:

- Example: Rate of fatigue during antiviral treatment: 0.30 (95% CI: 0.25-0.35)
- Example: Rate of anemia requiring intervention: 0.15 (95% CI: 0.10-0.20)

#### 4. Costs:

- Example: Cost of antiviral treatment per patient: \$30,000 (95% CI: \$28,000-\$32,000)
- Example: Annual cost of managing decompensated cirrhosis: \$10,000 (95% CI: \$9,000-\$11,000)

#### 5. Utilities:

- Example: Utility value for chronic hepatitis C: 0.75 (95% CI: 0.70-0.80)
- Example: Utility value for liver transplantation: 0.90 (95% CI: 0.85-0.95)

Please note that these values are for **illustrative purposes only** and should not be considered as actual estimates for any specific model. The point estimates and confidence intervals of model parameters **should be derived from empirical evidence, such as published studies,** clinical trial data, or country-specific databases, to ensure the accuracy and validity of the model's results. It's essential to consult relevant literature or engage with experts in the field to obtain the most appropriate and up-to-date estimates for the specific context in which the model is being applied.

# Want to hear more about Modeling?



The poster features the SMDM logo at the top left with the tagline 'Society for Medical Decision Making' and 'Better Health Through Better Decisions.' A yellow diagonal banner on the right says 'SAVE THE DATE'. The main text reads: '18th Biennial European Conference', '21-23 May 2023 Berlin, Germany', and 'Linking Research to Evidence-Based Action for Patients, Providers and Policy Decision Makers'. The central image shows the Brandenburg Gate in Berlin at night. At the bottom, there are five small inset images of various cityscapes and landmarks. The website 'www.smdm.org' and email 'info@smdm.org' are listed at the bottom.



## Society for Medical Decision Making 18<sup>th</sup> Biennial European Conference

Conference Co-Chairs:

Beate Jahn, PhD,  
Silke Siebert, MD,  
Tobias Kurth, MD, MSc, ScD,  
Uwe Siebert, MD, MPH, MSc, ScD

Visit [www.smdm.org](http://www.smdm.org) for more information.

Conference Venue:

Langenbeck Virchow Haus, Berlin

# Want to hear more about Modeling?

→ See our HTADS Continuing Education Program



## Public Health, Health Services Research and HTA

Home > Departments > Public Health, Health Services Research and Health Technology Assessment > HTADS Continuing Education Program

### Department-Team

#### News & Highlights

Institute of Public Health, Medical Decision Making and HTA

Institute for Management and Economics in Healthcare

Institute of Quality and Efficiency in Medicine

Division for Quantitative Methods in Public Health and Health Services Research

Division for Management in Health and Sport Tourism

Research Unit for Quality and Ethics in Health Care

Research Unit for Diabetes Epidemiology

Center for Statistical Consulting and Continuing Education

#### HTADS Continuing Education Program

Scientific Reporting and Writing - ONLINE

### HTADS Continuing Education Program

#### What is the Continuing Education Program on Health Technology Assessment & Decision Sciences (HTADS)?

Health Technology Assessment (HTA) has been defined by the International Network of Agencies for HTA (INAHTA) as "a multidisciplinary field of policy analysis, studying the medical, economic, social, and ethical implications of development, diffusion and use of health technologies (e.g., drugs, devices, surgical procedures, prevention techniques)".  
In conducting HTA, the discipline of decision sciences has become increasingly relevant.

Decision Sciences (DS) is the application of explicit and quantitative methods to analyze decisions under conditions of uncertainty (e.g., meta-analysis, decision-analytic modeling, cost-effectiveness analysis).

In recent years, HTA and DS have become very important to health care policymakers.

In order to keep pace with these developments, the UMIT-HTADS Program was designed to provide excellent quality education and comprehensive training in the key issues of HTA and DS for anyone involved in the health sector.

#### HTADS Program Director:



#### Contact

Univ.-Prof. Dr. Uwe Siebert,  
MPH, MSc

Head of Department / Head of Institute

[public-health@umit-tirol.at](mailto:public-health@umit-tirol.at)

t +43 50-8648-3930

f +43 50-8648-673930

HTADS Continuing Education Program

03 – 05 April 2024

3-DAY CERTIFIED COURSE

Modeling Approaches for HTA

A Practical Hands-on Workshop



[www.htads.org](http://www.htads.org)



[www.htads.org](http://www.htads.org)



[https://youtu.be/QHEXTI\\_L9dw](https://youtu.be/QHEXTI_L9dw)

# HTADS Continuing Education Program Certified Courses

Information: [htads@umit-tirol.at](mailto:htads@umit-tirol.at)



All in-person courses are either held in Hall in Tirol, Austria, or offered online

[www.htads.org](http://www.htads.org)

- Doctoral Program in Health Technology Assessment
- Doctoral Program in Public Health





# **HTADS Continuing Education Program Certified Courses**

**Information: [htads@umit-tirol.at](mailto:htads@umit-tirol.at)**

**[www.htads.org](http://www.htads.org)**

<i>28 – 30 Sept 2023</i>	<i>Introduction to Systematic Reviews and Meta-Analysis - ONLINE</i>
<i>16 – 18 Nov 2023</i>	<i>Scientific Reporting and Writing, Hall in Tirol</i>
<i>29 – 31 Jan 2024</i>	<i>Introduction to HTA and Health Economics – ONLINE</i>
<i>12 – 16 Feb 2024</i>	<i>Winter School in Clinical Epidemiology, Hall in Tirol</i>
<i>18 – 22 Mar 2024</i>	<i>Causal Inference in Observational Studies and Clinical Trials Affected by Treatment Switching: A Practical Hands-on Workshop - ONLINE</i>
<i>03 – 05 Apr 2024</i>	<i>Modeling Approaches for HTA: A Practical Hands-on Workshop, Hall in Tirol</i>
<i>23 – 25 May 2024</i>	<i>Advanced Systematic Reviews and Meta-Analysis - ONLINE</i>

***All courses are either held online or in Hall in Tirol, Austria***



***Doctoral Program in Health Technology Assessment  
Doctoral Program in Public Health***

# \*HTADS – Info Letter

If you would like to receive additional material on *causal inference, health decision science and HTA*, please send an email to:

[htads@umit-tirol.at](mailto:htads@umit-tirol.at)

Subject: Subscribe HTADS

or visit:

[www.htads.org](http://www.htads.org)



Health Technology Assessment  
and Decision Science

Education.  
Science.  
Competence.

Get information now!  
[www.umit-tirol.at/htads](http://www.umit-tirol.at/htads)

HTADS – Continuing Education Program on Health  
Technology Assessment & Decision Science

An online international university program

[www.umit-tirol.at](http://www.umit-tirol.at)



