

Surrogatendpunkte als Parameter der Nutzenbewertung

Use of Surrogate end points in HTA

Abstract

The different actors involved in health system decision-making and regulation have to deal with the question which are valid parameters to assess the health value of health technologies.

So called surrogate endpoints represent in the best case preliminary steps in the casual chain leading to the relevant outcome (e. g. mortality, morbidity) and are not usually directly perceptible by patients. Surrogate endpoints are not only used in trials of pharmaceuticals but also in studies of other technologies. Their use in the assessment of the benefit of a health technology is however problematic.

In this report we intend to answer the following research questions: Which criteria need to be fulfilled for a surrogate parameter to be considered a valid endpoint? Which methods have been described in the literature for the assessment of the validity of surrogate endpoints? Which methodological recommendations concerning the use of surrogate endpoints have been made by international HTA agencies? Which place has been given to surrogate endpoints in international and German HTA reports?

For this purpose, we choose three different approaches. Firstly, we conduct a review of the methodological literature dealing with the issue of surrogate endpoints and their validation. Secondly, we analyse current methodological guidelines of HTA agencies members of the International network of agencies for Health Technology Assessment (INAHTA) as well as of agencies concerned with assessments for reimbursement purposes. Finally, we analyse the outcome parameter used in a sample of HTA reports available for the public.

The analysis of methodological guidelines shows a very cautious position of HTA institutions regarding the use of surrogate endpoints in technology assessment. Surrogate endpoints have not been prominently used in HTA reports. None of the analysed reports based its conclusions solely on the results of surrogate endpoints. The analysis of German HTA reports shows a similar pattern.

The validation of a surrogate endpoint requires extensive research, including randomized controlled trials (RCT) assessing clinical relevant endpoints. The validity of a surrogate parameter is rather technology-specific than disease-specific. Thus – even in the case of apparently similar technologies – it is necessary to validate the surrogate for every single technology (i. e. for every single active agent).

The use of surrogate endpoints in the assessment of the benefit of health technologies is still to be seen very critically.

Sandra Mangiapane¹
Marcial Velasco
Garrido¹

¹ Technische Universität Berlin,
Berlin, Deutschland

Zusammenfassung

Die Frage nach den Ergebnisparametern, die für eine valide Bewertung des Nutzens medizinischer Technologien verwendet werden können, beschäftigt alle an Entscheidungsfindungen und an der Regulierung im Gesundheitssystem beteiligten Akteure.

Während die klinisch relevanten Endpunkte jene sind, die für den Patienten belangvoll sind (z. B. Morbidität, Mortalität), stellen Surrogatendpunkte im besten Fall Vorstufen der eigentlichen klinisch relevanten Endpunkte dar, die für den Patienten in der Regel nicht unmittelbar

spürbar sind. Surrogatparameter werden nicht nur in Studien über die Wirksamkeit von Arzneimitteln eingesetzt, sondern auch in Studien zu anderen Technologien. Der Einsatz von Surrogatendpunkten in der Bewertung des Nutzens von Gesundheitstechnologien ist jedoch problematisch.

In diesem Bericht wird folgenden Fragen nachgegangen:

- Welche Kriterien muss ein Surrogatparameter erfüllen, um als valider Endpunkt angesehen werden zu können?
- Welche Methoden werden zur Validierung von Surrogatendpunkten in der Literatur diskutiert?
- Welche methodischen Vorgaben machen internationale Agenturen aus dem Bereich Health Technology Assessment (HTA) oder Arzneimittelnutzenbewertung hinsichtlich des Einsatzes von Surrogatendpunkten?
- Welchen Stellenwert haben Surrogatendpunkte in HTA-Berichten internationaler HTA-Agenturen und in Berichten, die in Deutschland produziert werden?

Entsprechend dieser Fragestellungen werden drei verschiedene methodische Herangehensweisen gewählt: Reviews der einschlägigen methodischen Literatur zu Surrogatendpunkten und ihre Validierung, Analyse der aktuellen Methodenpapiere von HTA-Institutionen, Analyse von abgeschlossenen sowie öffentlich zugänglichen internationalen und deutschen HTA-Berichten.

Zusammenfassend zeigen die Empfehlungen der hier berücksichtigten Institutionen dahingehend eine kritische Einstellung zur Verwendung von Surrogatendpunkten in HTA. Es lässt sich zudem feststellen, dass Surrogatendpunkte einen geringen Stellenwert in HTA-Berichten haben. In keinem der untersuchten HTA-Berichte wird die Bewertung von nicht diagnostischen Technologien ausschließlich auf die Ergebnisse von Surrogatendpunkten gestützt. Die Ergebnisse der Analyse deutscher HTA-Berichte sind annähernd identisch zu denen internationaler.

Die Validierung eines Surrogatendpunkts erfordert ausführliche Forschung, einschließlich der Durchführung von randomisierten kontrollierten Studien (RCT) mit klinisch relevanten Endpunkten. Die Validität eines Surrogatendpunkts ist eher technologie- als krankheitsspezifisch, so dass die Ergebnisse der Validierung eines Surrogatendpunkts nicht auf andere Technologien übertragen werden können (auch nicht bei einem angeblich ähnlichen Wirkmechanismus). Um ein Höchstmaß an Sicherheit zu erreichen, muss die Validität eines Surrogatendpunkts bei jeder Technologie bzw. jedem Wirkstoff einzeln geprüft werden.

Nach wie vor ist der Einsatz von Surrogatendpunkten bei der Nutzenbewertung sehr kritisch zu betrachten.

Schlüsselwörter: Surrogatendpunkte, Surrogatparameter, HTA, Nutzenbewertung

Executive Summary

1. Health policy background

The issue which are valid and acceptable parameters for the assessment of the health benefit of the application of health technologies is recurrently discussed among the different actors and stakeholders in the health system. In Germany, with the establishment of the Institute for Quality and Efficiency in Health Care (IQWiG) in 2004 the discussions on the methods for health technology assessment in general and on the use of surrogate endpoints in particular have recently gained actuality und publicity.

2. Scientific background

So called surrogate endpoints represent in the best case preliminary steps in the casual chain leading to the relevant outcome (e. g. mortality, morbidity) and are not usually directly perceptible by patients. Characteristics of surrogate endpoints are:

- They are measured in lieu of the actually relevant outcome of interest.
- They are usually biochemical markers, physiological parameters or subclinical endpoints which for the patient are not directly perceptible. However, they are correlated with relevant clinical endpoints (e. g. high blood pressure is associated with higher risk of stroke, high LDL-cholesterine (LDL = Low density lipoprotein) is a risk factor for a heart attack, the CD4-cell count is associated with AIDS mortality).
- Changes in the surrogate are easier to observe than changes in the related relevant endpoint (i. e. the occur earlier and more commonly).
- Surrogates are sometimes named intermediate or intermediary outcomes, since they represent an intermediate step in the casual chain leading to the clinical relevant endpoint.
- Surrogate parameters are statistically associated with the clinical relevant outcome and have prognostic power.
- The association between the surrogate and the relevant endpoint is plausible from a biological and pathophysiological point of view.

Surrogate endpoints are not only used in trials of pharmaceuticals but also in studies of other clinical technologies. Parameters with an intermediary character are also applied in the field of community and public health interventions.

Their use in the assessment of the benefit of a health technology is however problematic. In the past, reliance on surrogate outcomes has led to false conclusions concerning the effects of a technology on the relevant health outcome. In many situations, relying on the strong correlation observed between surrogate and relevant endpoint to find an intervention has had fatal consequences (i. e. positive effects on the surrogate but increased mortality

with the intervention in question). The problematic is known since around 30 years. A classical example of the potential for fatal consequences as a result of reliance on surrogate is the case of class I antiarrhythmic drugs. Some drugs have been removed from market after the observation of an increased mortality or morbidity with their use, contrary to the expectations raised by the observation of positive effects on a surrogate endpoint. In other occasions, reliance on surrogate has also led to withholding effective therapies. For example, for many years betablockers – due to the bradycardic effect – were considered to be contraindicated in patients with heart failure, since following pathophysiological reasoning a reduction of the heart rate was thought to have deleterious effects in this patients.

3. Research questions

1. Which criteria need to be fulfilled for a surrogate parameter to be considered a valid endpoint?
2. Which methods have been described in the literature for the assessment of the validity of surrogate endpoints?
3. Which methodological recommendations concerning the use of surrogate endpoints have been made by international HTA agencies?
4. Which place has been given to surrogate endpoints in international and German HTA reports?

4. Methods

According to the above mentioned research questions, we follow different methodological approaches.

In order to answer research questions 1 and 2, related to the concepts and the methods of surrogate validation we conduct a systematic review of methodological papers. Electronic databases are searched with the following terms:

SURROGATE END POINT; SURROGATE END POINTS; SURROGATE ENDPOINT; SURROGATE ENDPOINTS; ENDPOINT, SURROGATE; ENDPOINTS, SURROGATE; ENDPOINT, SURROGATE; END POINTS, SURROGATE; BIOLOGICAL MARKER; BIOLOGICAL MARKERS; VALIDATION; STATISTICS; BIOMETRY; DECISION SUPPORT TECHNIQUES; ENDPOINT DETERMINATION; CAUSALITY

The methodological literature is summarised in a narrative review consisting of two parts: 1. an overview of the criteria to be fulfilled by a surrogate in order to be considered acceptable. 2. an overview of the statistical methods proposed in the literature for the validation of surrogates. In order to answer research question 3 we analyse the methodological guidelines and recommendations of HTA agencies, member of the International network of agencies for Health Technology Assessment (INAHTA), and of agencies involved in pharmaceutical pricing and reimbursement decisions.

In order to answer research question 4 we analyse a random sample of HTA reports from the HTA database. We extracted the type of outcome parameter used and

reported in the HTA reports. We analyse in the same way the full sample of HTA reports procured in Germany and registered in the database of the German Agency for HTA.

5. Results

5.1 Literature review

The literature search yields a total of N=1,109 hits. After checking title and abstract, n=2 duplicates and n=1,007 references lacking references are excluded. A total of n=100 papers is retrieved for more detailed analysis. At the end n=25 methodological papers are summarised in the review on criteria and validation methods.

The criteria that a surrogate parameter need to fulfil in order to be recognized an acceptable and valid endpoint can be summarised as follows:

- **Biological plausibility:** There is evidence from animal models and epidemiological studies of a causal relationship between the surrogate parameter and the clinical relevant endpoint. The surrogate is part of the pathophysiological causal path leading to the health outcome.
- **Magnitude of the association between surrogate and relevant endpoint:** Epidemiological evidence has shown repeatedly and consistently that changes in the surrogate are qualitative and quantitative associated with changes in the relevant health outcome.
- **Evidence of effect from randomized controlled trials (RCT):** There is evidence from RCT showing that the changes induced by an intervention in the surrogate lead to changes in the relevant outcome in the same direction. The effect of the intervention is fully captured by the surrogate. Even in the case of very similar active principles, the mechanism of action may differ. Thus, the transferability of conclusions on the validity of a surrogate from one technology to another needs to be carefully assessed.

In the full report, we summarise the different statistical methods discussed in the literature for the validation of surrogate endpoints. In summary, we conclude that there is no goldstandard for the validation of surrogate endpoints. Since the generalisation of results from single studies is more prone to produce fallacies, approaches summarising results from several studies (i. e. meta-analysis) are preferred.

5.2 Analysis of methodological guidance from HTA agencies

A total of 23 methodological papers from 14 INAHTA members (eleven countries) is identified. In addition, eleven further methodological guidelines from agencies involved in pricing and reimbursement decisions are found. We extract their recommendations concerning the selection of outcome parameters in general and the use of surrogate endpoints in particular.

A total of 13 from 23 analysed INAHTA member methodological papers' and seven of eleven from "fourth-hurdle agencies" provide information on how to choose outcome parameters for the assessment. All institutions agree, that patient relevant outcome parameters are strongly preferred in the assessment of the benefit of a health technology. All agencies underline that hard outcome parameters are to be preferred to surrogate endpoints. Nevertheless, the majority of agencies describes that under some circumstances surrogate endpoints may exceptionally be accepted – provided the validity of the surrogate is well established. In order to accept a surrogate, HTA agencies require the presentation of evidence which supports the causal relationship between surrogate and clinical relevant endpoint.

None of the methodological guidance papers from HTA agencies provided a list of well established/generally accepted surrogate endpoints.

5.3 Survey of HTA reports

A total of 140 HTA reports from INAHTA members and of 131 HTA reports from German institutions is analysed.

The reports cover different types of technologies, although the assessment of medical and surgical interventions represent the majority. A prospective description (e. g. in the research questions or in the methods section) of the outcome parameters in which the assessment would be based is present in less than half of the analysed HTA reports. Surrogate endpoints are extracted and reported in 87 (62%) HTA reports from the HTA database. Almost all HTA reports include also a clinical and patient relevant outcome. Only five reports use exclusively surrogate parameters, all of them assessing a diagnostic technology and being the surrogate test characteristics.

Similar results are obtained for the sample of German HTA-reports. Approximately one third of the German HTA reports describe to assess benefits and risks, or effectiveness and safety of the technology, without further describing how these terms were operationalised into outcome parameters. Surrogate endpoints are extracted and reported in 74 (56%) German HTA reports. Almost all German HTA reports also consider a clinical and patient relevant parameter. Only six reports exclusively use surrogate parameters, all of them assessing a diagnostic technology and being the surrogate test characteristics.

6 Discussion

6.1 Methods of the report

According to the different nature of our research questions, we follow several approaches in this report.

The literature review represents a good overview of the field. Besides original works, we also identify three recent systematic reviews which summarise additional methodological papers.

We also provide a representative overview of the methodological guidance regarding surrogate endpoints from HTA agencies worldwide.

In addition, we also provide a representative picture of the actual consideration of surrogate endpoints in HTA reports from international and German HTA agencies. The representative survey allows to understand the value surrogate endpoints have in the field of HTA.

6.2 Results of the report

In order to be considered valid and acceptable a surrogate needs to fulfil several criteria. Thus, favourable results from statistical validation approaches are not a sufficient condition to conclude on the validity of a surrogate endpoint. Information on biological and pathophysiological factors is also required. In addition, the validity of a surrogate is to be seen as technology-specific. Whether a surrogate is able to capture the full effect of a technology depends on the mechanism of action of the technology in question. This irrespective of whether a strong and consistent association between surrogate and relevant health outcome has been well established.

In summary, HTA agencies show a cautious position regarding the reliability of surrogate endpoints in HTA. Clinical relevant endpoints such as mortality and morbidity are preferred, since they allow a sound assessment of effectiveness and safety of the interventions. The considerations and recommendations provided by HTA agencies regarding the use of surrogates are conform to the discussions being held in the theoretical and methodological literature.

The quantitative analysis of the sample of HTA reports allows to conclude that surrogate endpoints have not been prominently used in HTA reports. The results are similar for the international sample and for the German sample, indicating that the handling of surrogate endpoints in Germany is not more stringent than in the international context.

7. Conclusions

The validation of a surrogate endpoint requires extensive research, including RCT assessing clinical relevant endpoints. The validity of a surrogate parameter is rather technology-specific than disease-specific. Thus – even in the case of apparently similar technologies – it is necessary to validate the surrogate for every single technology (i. e. for every single active agent), in order to avoid false conclusions potentially leading to fatal consequences. The use of surrogate endpoints in the assessment of the benefit of health technologies is still to be seen very critically.

Kurzfassung

1. Gesundheitspolitischer Hintergrund

Die Diskussionen, welche Ergebnisparameter für eine valide Bewertung des Nutzens medizinischer Technologien verwendet werden können, sind spätestens seit der Gründung des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) in Deutschland wieder aufgelebt. Diese Frage beschäftigt alle Akteure, die an die Entscheidungsfindung und Regulierung im Gesundheitssystem beteiligt sind.

2. Wissenschaftlicher Hintergrund

Während klinisch relevante Endpunkte Einschnitte bezeichnen, die für den Patienten bedeutend sind (z. B. Morbidität, Mortalität), stellen Surrogatendpunkte im besten Fall Vorstufen der eigentlichen klinisch relevanten Endpunkte dar, die für den Patienten in der Regel nicht unmittelbar spürbar sind. Surrogatendpunkte weisen folgende Charakteristika auf:

- Es handelt sich um Zielgrößen, die anstelle von (*in lieu*) den eigentlichen Endpunkten gemessen werden.
- Es handelt sich häufig um physiologische oder biochemische Marker bzw. subklinische Endpunkte, die für den Patienten in der Regel nicht unmittelbar von Bedeutung bzw. in irgendeiner Weise bemerkbar sind, jedoch stellvertretend für wichtige klinische Endpunkte stehen können (z. B. erhöhter Blutdruck verstärkt das Risiko für Schlaganfall, erhöhte LDL-Cholesterinwerte (LDL = Low density lipoprotein) sind ein Risikofaktor für Myokardinfarkt, der Anzahl der CD4-Zellen für AIDS-Mortalität (AIDS = Acquired Immuno Deficiency Syndrome)).
- Veränderungen dieser Parameter sind im Vergleich zum Auftreten der relevanten Endpunkte (Morbidität, Tod) häufiger und nach kürzerer Beobachtungszeit feststellbar.
- Sie stellen Zwischenschritte auf dem Weg von der Erkrankung hin zum klinischen Endpunkt dar und werden deshalb mitunter auch als „intermediate“ oder „intermediary outcomes“ bezeichnet.
- Es gibt einen (statistisch nachweisbaren) Zusammenhang zwischen dem Surrogatendpunkt und dem eigentlichen klinischen Endpunkt im Sinn einer guten Prognosekraft.
- Der Zusammenhang zwischen Surrogat und klinisch relevantem Endpunkt muss biologisch und pathophysiologisch (d. h. bzgl. der Funktion des Organismus unter Krankheit) plausibel sein.

Surrogatparameter werden nicht nur in Studien über die Wirksamkeit von Arzneimitteln eingesetzt. Auch Studien über die Effekte von medizinischen und chirurgischen Prozeduren berichten häufig über Ergebnisparameter, die nicht das ultimative, patientennahe Ziel der Intervention wiedergeben. Bei nicht-klinischen Interventionen (z. B. Patientenschulungen, bevölkerungsbezogenen Inter-

ventionen etc.) werden auch häufig Zielgrößen gemessen, die einen intermediären Charakter haben.

Die Verwendung von Surrogatendpunkten ist jedoch nicht unproblematisch. In der Vergangenheit gab es viele Situationen, in denen das Vertrauen auf Surrogatendpunkte trotz starker Korrelation mit dem klinischen Endpunkt irreführend war bzw. fatale Folgen hatte. Diese Problematik ist mittlerweile mehr als 30 Jahre bekannt. Prominentes und mittlerweile klassisches Beispiel stellen die Klasse I-Antiarrhythmika dar. Einige Produkte, die auf der Basis von Surrogatendpunkten zugelassen wurden, mußten später vom Markt genommen werden, da sich in Studien mit Mortalitäts- bzw. Morbiditätsendpunkten, die Nutzen-Risiko-Abwägung umkehrte. Aber ebenso kann das Festhalten an Surrogatendpunkten auch dazu führen, dass therapeutische Chancen nicht genutzt werden. So gelten beispielsweise Betablocker über lange Zeit bei Patienten mit Herzinsuffizienz als kontraindiziert, da sie zu einer Verringerung der Herzfrequenz führen, was unter pathophysiologischen Überlegungen zu negativen Effekten bei diesen Patienten führen sollte.

3. Forschungsfragen

1. Welche Kriterien muss ein Surrogatparameter erfüllen, um als valider Endpunkt angesehen werden zu können?
2. Welche Methoden werden zur Validierung von Surrogatendpunkten in der Literatur diskutiert?
3. Welche methodischen Vorgaben machen internationale Agenturen aus den Bereichen Health Technology Assessment (HTA) oder Arzneimittelnutzenbewertung hinsichtlich des Einsatzes von Surrogatendpunkten?
4. Welchen Stellenwert haben Surrogatendpunkte in HTA-Berichten internationaler HTA-Agenturen und in Berichten, die in Deutschland produziert werden?

4. Methodik

Entsprechend der Fragestellungen dieses Berichts, werden drei verschiedene methodische Herangehensweisen gewählt.

Um die erste und die zweite Frage bzgl. der methodischen Diskussionen um die Anwendung von Surrogatendpunkten zu beantworten, wird ein Review der einschlägigen methodischen Literatur durchgeführt. Folgende Suchbegriffe werden verwendet:

SURROGATE END POINT; SURROGATE END POINTS; SURROGATE ENDPOINT; SURROGATE ENDPOINTS; ENDPOINT, SURROGATE; ENDPOINTS, SURROGATE; ENDPOINT, SURROGATE; END POINTS, SURROGATE; BIOLOGICAL MARKER; BIOLOGICAL MARKERS; VALIDATION; STATISTICS; BIOMETRY; DECISION SUPPORT TECHNIQUES; ENDPOINT DETERMINATION; CAUSALITY

Die methodische Literatur wird in Form eines narrativen Reviews zusammengefasst. Sie besteht aus einem Teil, in dem die Anforderungen dargestellt werden, die ein Surrogat erfüllen muss, um als valide zu gelten. Im zweiten Teil wird ein Überblick der in der Literatur diskutierten

statistischen Methoden für die Surrogatendpunkt-Validierung gegeben.

Für die Beantwortung der dritten Fragestellung bzgl. des Umgangs von internationalen Agenturen mit dem Thema Surrogatendpunkte werden die Methodenpapiere dieser Institutionen untersucht. Dabei werden alle Mitglieder der INAHTA berücksichtigt, sowie eine Gruppe von Agenturen, die in der Arzneimittelnutzenbewertung im Rahmen von Kostenerstattung- und Preisentscheidungen involviert ist.

Für die Beantwortung der vierten Frage bzgl. der Anwendung von Surrogatendpunkten in HTA-Berichten werden die in abgeschlossenen und veröffentlichten internationalen sowie deutschen HTA-Berichten berücksichtigten Outcomeparameter analysiert. Dabei wird eine Stichprobe aus der HTA-Berichtsdatenbank der INAHTA untersucht. Für Deutschland werden alle Berichte mit nicht-methodischen Inhalten berücksichtigt, die in der HTA-Datenbank der Deutschen Agentur für HTA (DAHTA) gesammelt sind.

5. Ergebnisse

5.1 Literaturrecherche

Insgesamt werden mit der Recherchestrategie N=1.109 Treffer erzielt. Von diesen werden im ersten Selektionsschritt n=2 Dubletten entfernt, und n=1.007 Literaturstellen ausgeschlossen, da sie für das Ziel des Berichts nach Durchsicht von Titel und Abstract als nicht-relevant eingestuft werden. Insgesamt werden n=100 Literaturstellen ausgewählt und für eine weitere Überprüfung im Volltext bestellt. Eine weitere Studie wird aus den Literaturverzeichnissen der Publikationen ausgewählt und ebenfalls im Volltext bestellt. Nach der Selektion werden aus der elektronischen Literaturrecherche 25 Studien für das Review zu Kriterien zur Beurteilung von Surrogatendpunkten und Methoden zur Validierung von Surrogatendpunkten eingeschlossen.

Die in der Literatur diskutierten Kriterien, die ein Surrogatendpunkt erfüllen muss, damit dieser als valider Endpunkt eingestuft werden kann, lassen sich wie folgt zusammenfassen:

- **Biologische Plausibilität:** Es liegt Evidenz aus tierexperimentellen oder epidemiologischen Studien vor, die auf einen kausalen Zusammenhang zwischen dem Surrogatparameter und dem klinischen Endpunkt hinweisen. Der Surrogatparameter liegt demnach auf dem pathophysiologischen Kausalpfad zum klinisch relevanten Endpunkt.
- **Ausmaß der Beziehung zwischen Surrogatendpunkt und klinischem Endpunkt:** Epidemiologische Studien zeigen wiederholt, dass durch eine Veränderung des Surrogatendpunkts eine qualitativ und quantitativ gleichgerichtete Veränderung des klinisch relevanten Endpunkts erreicht werden kann. Das Surrogat sollte den Effekt der Intervention vollständig ausbilden.
- **Wirkungsnachweis im Rahmen von RCT:** Es liegen Ergebnisse aus RCT vor, die zeigen, dass durch eine

therapiebedingte Veränderung des Surrogatendpunkts eine gleichgerichtete Veränderung des klinisch relevanten Endpunkts erreicht werden kann. Der Surrogatendpunkt muss den Effekt der Intervention vollständig abbilden. Dabei ist zu beachten, dass es nicht nur bei pharmakologisch unterschiedlichen Wirkstoffen, sondern auch bei Stoffen der gleichen Wirkstoffgruppe zu abweichenden Wirkmechanismen kommen kann und daher die Übertragbarkeit der Studienergebnisse von einer Substanz auf eine andere zusätzlich geprüft werden muss.

In der Literatur werden verschiedene statistische Methoden zur Validierung von Surrogatendpunkten beschrieben, die im Bericht ausführlich dargestellt werden. Insgesamt kann festgehalten werden, dass es derzeit keinen Goldstandard für die Validierung von Surrogatendpunkten gibt. Aufgrund der Fehleranfälligkeit, die mit der Verallgemeinerung von Ergebnissen aus Einzelstudien verbunden ist, sind allerdings Ansätze auf Basis mehrerer Studien, wie dies z. B. bei Metaanalysen der Fall ist, zu bevorzugen.

5.2 Analyse von Methodenpapieren aus HTA-Agenturen

Es werden 23 Methodenpapiere von 14 INAHTA-Mitgliedern (elf Länder) identifiziert, deren Aussagen bzgl. der in der Bewertung zu verwendenden Ergebnisparameter im Allgemeinen und zur Anwendung von Surrogatendpunkten im Besonderen zusammengefasst werden. Darüber hinaus werden weitere elf von Arzneimittelbewertungsagenturen herausgegebene methodologische Leitlinien identifiziert.

Insgesamt geben 13 der 23 analysierten Methodenpapiere der INAHTA-Mitglieder und sieben der elf analysierten Leitlinien von Arzneimittelbewertungsinstitutionen Empfehlungen zur Auswahl der in der Bewertung der Technologie zu verwendenden Ergebnisparameter. Dabei sind sich alle Institutionen darüber einig, dass in HTA-Berichten und Nutzenbewertungen patientenrelevanter Ergebnisparameter vorrangig berücksichtigt werden müssen. Alle Agenturen betonen in ihren Methodenpapieren, dass grundsätzlich die Verwendung von harten klinischen Endpunkten gegenüber Surrogatendpunkten zu bevorzugen ist, wobei die Mehrzahl der Institutionen beschreibt, dass bei der Erfüllung bestimmter Voraussetzungen die Verwendung von Surrogatendpunkten als Ausnahme akzeptabel sein kann. Dazu ist es im Wesentlichen erforderlich, dass die oben beschriebene Validität des verwendeten Endpunkts nachgewiesen werden kann. Dies soll i. d. R. durch die Vorlage zusätzlicher Evidenz erfolgen, die den kausalen Zusammenhang zwischen Surrogat und klinischem Endpunkt belegt. In keinem der hier berücksichtigten Methodenpapiere wird eine Liste von anerkannten Surrogatparametern formuliert.

5.3 Analyse von HTA-Berichten

Insgesamt werden 140 internationale und 131 HTA-Berichte von deutschen Institutionen untersucht.

Die internationalen Berichte decken verschiedene Technologien ab, wobei das Assessment von medizinischen und chirurgischen Prozeduren in der Mehrheit der Berichte Thema ist. In weniger als der Hälfte der internationalen HTA-Berichte werden die im Assessment zu extrahierenden Ergebnisparameter bei der Formulierung der Fragestellung bzw. im Methodenteil explizit und detailliert festgelegt. Insgesamt wird in 87 (62%) der internationalen HTA-Berichte die Verwendung von Surrogatparametern geplant bzw. Surrogatparameter aus den für das Assessment vorliegenden Studien extrahiert und berichtet. Fast alle Berichte betrachten einen klinisch relevanten, patientennahen Endpunkt. Nur in fünf Berichten werden ausschließlich Surrogatendpunkte betrachtet. In diesen Berichten werden diagnostische Technologien bewertet; hier sind Surrogate Testcharakteristika.

Nicht in allen deutschen HTA-Berichten werden die im Assessment zu extrahierenden Ergebnisparameter bei der Formulierung der Fragestellung bzw. im Methodenteil explizit und detailliert festgelegt. In ca. einem Drittel der analysierten deutschen HTA-Berichte wird lediglich erwähnt, dass der Bericht sich beispielsweise mit dem „Nutzen und Risiken“, „Wirksamkeit und Sicherheit“ beschäftigt, ohne genauer darzulegen ob und welche Ergebnisparameter prospektiv festgelegt werden. Insgesamt wird in 74 (56%) der deutschen HTA-Berichte die Verwendung von Surrogatparametern geplant bzw. Surrogatparameter aus den für das Assessment vorliegenden Studien extrahiert und berichtet. Fast alle Berichte betrachten einen klinisch relevanten, patientennahen Endpunkt. Nur in sechs der 74 Berichte, die Surrogatparameter verwenden, werden keine weiteren Parameter berücksichtigt. Das heißt, dass von den 131 deutschen HTA-Berichten nur 4,5% ausschließlich auf Surrogatendpunkten beruhen. Da in diesen Berichten diagnostische Technologien bewertet werden, gehört die ausschließliche Verwendung von Surrogaten zu den Testcharakteristika.

6. Diskussion

6.1 Diskussion der Methodik

Die Beantwortung der diesen Bericht leitenden Fragestellungen erfordert verschiedene methodische Ansätze. Den ersten zwei Fragen wird im Rahmen eines Literaturreviews nachgegangen. Es werden im Literaturreview drei aktuelle Übersichtsarbeiten identifiziert, die neben den durch die Recherche gefundenen Arbeiten auch weitere relevante Artikel zusammenfassen. Dadurch ist ein umfassender Überblick zu den Anforderungen an Surrogatparameter und zu Validierungsmethoden möglich.

Die Beantwortung der dritten Frage beruht auf der Analyse von methodischen Leitlinien internationaler Institutionen, die in HTA involviert sind. Damit stellt diese Arbeit einen repräsentativen Überblick dar, welchen Stellenwert inter-

ationale HTA-Organisationen Surrogatendpunkten einräumen.

Für die Beantwortung der vierten Frage werden deutsche HTA-Berichte sowie eine Stichprobe von HTA-Berichten internationaler HTA-Agenturen analysiert. Der Ansatz ermöglicht eine Abschätzung des Stellenwerts von Surrogatendpunkten in HTA-Berichten auf der Basis einer quantitativen Analyse.

6.2 Diskussion der Fragestellungen

Damit ein Surrogatendpunkt als valider Endpunkt für die Bewertung des Nutzens einer Technologie berücksichtigt werden kann, müsste dieser mehrere Kriterien erfüllen. Aus diesem Grund reichen die Ergebnisse statistischer Validierungsverfahren allein nicht aus, um die Eignung eines Surrogatendpunkts zu beurteilen. Die Einbeziehung von Informationen über biologische und klinische Faktoren ist unverzichtbar. Die Validierung eines Surrogatendpunkts für die Bewertung des Nutzens ist in diesem Sinn technologiespezifisch. Es hängt vom Wirkmechanismus der Technologie ab, ob das Surrogat deren Effekte abbilden kann oder nicht. Dies ist unabhängig davon, ob es ein starker und konsistenter Zusammenhang zwischen Veränderungen im Surrogat sowie Veränderungen in den klinisch relevanten Endpunkten statistisch nachgewiesen worden ist.

Zusammenfassend zeigen die Empfehlungen der hier berücksichtigten Institutionen dahingehend eine kritische Einstellung zur Verwendung von Surrogatendpunkten in HTA. Klinisch relevante Endpunkte wie Mortalität oder Morbidität haben Vorrang, weil sie eine valide Beurteilung der Wirksamkeit und der Sicherheit ermöglichen, die die Grundlage jeglicher Nutzenbewertung darstellt. Die von den HTA-Institutionen gestellten Anforderungen und Überlegungen zur Akzeptanz eines Surrogatendpunkts im Rahmen von HTA-Berichten, sind im Einklang mit denen, die in der theoretischen Literatur diskutiert werden. Aus unserer quantitativen Analyse lässt sich feststellen, dass Surrogatendpunkten ein sehr geringer Stellenwert in HTA-Berichten eingeräumt wird. Die Ergebnisse der Analyse der deutschen HTA-Berichte sind annähernd identisch zu denen aus der Stichprobe internationaler HTA-Berichte. Es liegt kein Hinweis dafür vor, dass in

Deutschland ein strengerer Standard als in den Ländern des internationalen Umfeldes beim Umgang mit Surrogatendpunkten praktiziert wird.

7. Schlussfolgerung

Die Validierung eines Surrogatendpunkts erfordert ausführliche Forschung, einschließlich der Durchführung von RCT mit klinisch relevanten Endpunkten. Die Validität eines Surrogatendpunkts ist eher technologie- als krankheitsspezifisch, so dass die Ergebnisse der Validierung eines Surrogatendpunkts nicht auf andere Technologien übertragen werden können (auch nicht bei angeblich ähnlichem Wirkmechanismus). Um ein Höchstmaß an Sicherheit zu erreichen, muss die Validität eines Surrogatendpunktes bei jeder Technologie bzw. jedem Wirkstoff einzeln geprüft werden.

Nach wie vor ist der Einsatz von Surrogatendpunkten bei der Nutzenbewertung sehr kritisch zu betrachten.

Korrespondenzadresse:

Marcial Velasco Garrido

Technische Universität Berlin, Strasse des 17. Juni 135, 10623 Berlin, Tel.: 030 31428420

marcial.velasco@tu-berlin.de

Bitte zitieren als

Mangiapane S, Velasco Garrido M. Surrogatendpunkte als Parameter der Nutzenbewertung. *GMS Health Technol Assess.* 2009;5:Doc12.

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/hta/2009-5/hta000074.shtml>

Der vollständige HTA-Bericht steht zum kostenlosen Download zur Verfügung unter:

http://portal.dimdi.de/de/hta/hta_berichte/hta250_bericht_de.pdf

Copyright

©2009 Mangiapane et al. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.