

**Entwicklung einer Methodik zur Gewichtung patientenrelevanter
Endpunkte in Health Technology Assessments am Beispiel der
terminalen Niereninsuffizienz**

Dissertation

zur Erlangung des akademischen Grades

„Doctor of Public Health“

Vorgelegt von Inger Miriam Janßen

Juni 2017

Erstgutachter: Prof. Dr. Ansgar Gerhardus

Zweitgutachter: Prof. Dr. Wolfgang Greiner

Zusammenfassung

Hintergrund

Bei der Evaluation medizinischer Interventionen durch HTA (Health Technology Assessment)-Berichte werden Nutzen und Schaden anhand von Endpunkten beurteilt. Die Auswahl und Gewichtung der Endpunkte wird bislang häufig von den HTA-Erstellern oder medizinischen Experten vorgenommen. Die Einbeziehung von Patienten wird häufig gefordert, aber es gibt noch kein standardisiertes Vorgehen. In der vorliegenden Dissertation wird untersucht, ob und auf welche Art und Weise eine Einbeziehung von Patientenpräferenzen bereits erfolgt beziehungsweise erfolgen kann.

Methoden

In dem ersten Manuskript wurde untersucht, inwieweit eine Einbeziehung von Patientenpräferenzen im deutschen Gesundheitswesen bereits erfolgt und welche Entwicklungspotenziale vorhanden sind. Im zweiten Manuskript - eine deskriptive Übersichtsarbeit - wurde ein Überblick über mögliche Methoden zur Priorisierung von Endpunkten gegeben. Die empirische Datenerhebung wurde am Beispiel der Indikation terminale Niereninsuffizienz durchgeführt (Manuskript 3 und 4). In dem dritten Manuskript wurden relevante Behandlungsziele unter Einbezug von Patienten identifiziert, um darauf aufbauend 23 relevante Endpunkte zu definieren. Anschließend fand durch eine repräsentative Patientenstichprobe von Hämodialyse-Patienten eine Gewichtung der ermittelten Endpunkte mittels einer Rating Skala statt. In dem vierten Manuskript wurden die Ergebnisse der repräsentativen Patientenstichprobe mit Gewichtungen durch andere Gruppen (Nephrologen, HTA-Ersteller und Patienten einer Selbsthilfegruppe) verglichen.

Ergebnisse

Patientenpräferenzen werden auf verschiedenen Ebenen im deutschen Gesundheitswesen einbezogen: im Gemeinsamen Bundesausschuss (G-BA) durch die Repräsentation der Patienten im Beratungsprozess und beim Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) bei der Festlegung von Endpunkten. In der deskriptiven Übersichtsarbeit wurden drei Methodengruppen identifiziert. Studien, die Ranking oder Rating Methoden verwendet hatten, untersuchten im Median die meisten Endpunkte, gefolgt von Utility Assessment Methoden. Die geringste Anzahl untersuchter Endpunkte fand sich bei den Methoden der multikriteriellen Entscheidungsfindung. Der Vergleich von unterschiedlichen Teilnehmergruppen oder Methoden innerhalb einer Studie ergab häufig eine

unterschiedliche Hierarchisierung von Endpunkten. Insgesamt konnte keine überlegene Methode identifiziert werden. Bei der Anwendung einer Rating Skala in der Indikation terminale Niereninsuffizienz zeigte sich, dass die repräsentative Patientengruppe Endpunkte, die Behandlungsaspekte und Lebensqualität betreffen, als ebenso wichtig wie Mortalität und Morbidität bewerten oder sogar als wichtiger. Der Vergleich mit Gruppen, die bislang häufig für eine Priorisierung von Endpunkten in diesem Kontext eingesetzt werden (Patienten einer Selbsthilfegruppe, Ärzte und HTA-Ersteller), zeigte, dass insbesondere zwischen den HTA-Erstellern und der repräsentativen Patientengruppe erhebliche Unterschiede in den Rangfolgen der Endpunkte feststellbar sind.

Diskussion und Ausblick

Eine stärkere Einbeziehung von Patientenpräferenzen bei der Erstellung von HTA-Berichten kann durch die Beteiligung von Patienten bei der Priorisierung von Endpunkten erfolgen. Verschiedene Methoden sind hierzu vorhanden, die alle spezifische Vor- und Nachteile besitzen. Insbesondere die Anzahl der Endpunkte, die bewertet werden sollen, kann als ein Auswahlkriterium für die geeignete Methode fungieren. Mit einfacheren Rating oder Ranking Skalen können mehr Endpunkte gewichtet werden als mit komplexeren Methoden der multikriteriellen Entscheidungsfindung.

Den dargestellten Ergebnissen ist zu entnehmen, dass Patienten einer Selbsthilfegruppe als geeignete Proxys für eine repräsentative Patientenstichprobe eingesetzt werden können. Notwendig hierfür ist eine hinreichende Strukturgleichheit bezüglich wichtiger Baselinecharakteristika zwischen beiden Patientengruppen.

Die Forderung nach einer stärkeren Einbeziehung von Patientenpräferenzen bei der Auswahl und Gewichtung von Endpunkten impliziert nicht, dass alleinig die Präferenzen der Patienten bei der Evaluation von medizinischen Interventionen entscheidend sind. Vielmehr zeigen die Ergebnisse der vorliegenden Arbeit, dass jede Gruppe bei einer Bewertung einen wichtigen Beitrag leisten kann. Deswegen sollten alle Gruppen fester Bestandteil im Prozess der Auswahl und Gewichtung von Endpunkten in HTA-Berichten sein. Dieser Prozess sollte transparent gestaltet werden, damit nachvollziehbar ist, welche Präferenzen bei der Bewertung in welchem Maße eingeflossen sind.

Aus den Erkenntnissen dieser Arbeit resultiert noch kein standardisiertes Vorgehen für die Einbeziehung von Präferenzen von Patienten und anderen Gruppen bei der Auswahl und Gewichtung von Endpunkten, sie ist vielmehr als ein erster wichtiger Schritt zu betrachten. Für die Entwicklung eines standardisierten Vorgehens ist weitere Forschung notwendig. Die

in der vorliegenden Arbeit gefundene hohe Übereinstimmung zwischen einer repräsentativen Patientenstichprobe und Patienten einer Selbsthilfegruppe sollte in weiteren Indikationen überprüft werden. Auch die Übereinstimmung anderer Gruppen (Ärzte und HTA-Ersteller) in anderen Indikationen sollte weiter untersucht werden.

Falls diese Überprüfungen ähnliche Übereinstimmungsmuster zeigen wie die in der vorliegenden Arbeit, könnte in einem weiteren Schritt eine generalisierbare Vorgehensweise erarbeitet werden, die den Prozess der Einbeziehung aller Gruppen standardisiert.

Inhaltsverzeichnis

Zusammenfassung	i
<i>Hintergrund</i>	i
<i>Methoden</i>	i
<i>Ergebnisse</i>	i
<i>Diskussion und Ausblick</i>	ii
Liste der einbezogenen Publikationen	1
Hintergrund	2
Einleitung	2
Einführung in das Thema	3
Relevanz des Themas	4
Theoretischer Bezugsrahmen	5
<i>Patientenbeteiligung</i>	5
<i>Methoden zur Erhebung von Werten und Präferenzen</i>	7
Terminale Niereninsuffizienz	8
Ziele und Fragestellungen	10
Situation in Deutschland	12
Manuskript 1 – Status quo und mögliche Weiterentwicklungen	12
<i>Patientenpartizipation auf der Mikroebene</i>	12
<i>Patientenpartizipation auf der Makroebene</i>	13
Systematische Übersicht der Methoden	14
Manuskript 2 – Deskriptive Übersichtsarbeit möglicher Methoden	14
<i>Systematische Suche</i>	14
<i>Ein- und Ausschlusskriterien</i>	14
<i>Datenextraktion und Darstellung der Ergebnisse</i>	15
<i>Ergebnisse</i>	15
Empirische Datenerhebung	16
Manuskript 3 – Präferenzen Patienten	16
<i>Identifikation relevanter Behandlungsziele</i>	16
<i>Präferenzenerhebung</i>	16
<i>Datenanalyse</i>	17
<i>Ergebnisse</i>	17
Manuskript 4: Vergleich der Präferenzen verschiedener Gruppen	18

<i>Identifikation relevanter Gruppen</i>	19
<i>Fragebogen</i>	19
<i>Datenanalyse</i>	20
<i>Ergebnisse</i>	20
Diskussion	21
Methodisches Vorgehen und Ergebnisse	22
Methodische Überlegungen zur Präferenzmessung	26
Bezug zu anderer Forschung	28
<i>Festlegung relevanter Fragestellungen mit Einbezug von Patienten</i>	28
<i>Auswahl relevanter Endpunkte</i>	29
<i>Einbeziehung von Patientenpräferenzen bei der Gewichtung von Endpunkten</i>	31
Implikationen für die Forschung	31
Ausblick	32
Referenzen	34
Eidesstattliche Erklärung	39
Danksagung	40

Liste der einbezogenen Publikationen

1. Scheibler F, **Scheike IM**, Dintsios CM. Patientenpartizipation bei Festlegung und Gewichtung von Behandlungszielen – Status quo und Entwicklungspotenziale. Z Evid Fortbild Qual Gesundhwes 2008; 102(6): 373-377.
2. **Janssen IM**, Gerhardus A, Schroer-Gunther MA, Scheibler F. A descriptive review on methods to prioritize outcomes in a health care context. Health Expect 2015; 18: 1873–1893 (epub 2014 Aug 25).
3. **Janssen IM**, Gerhardus A, von Gersdorff GD, Baldamus CA, Schaller M, Barth C et al. Preferences of patients undergoing hemodialysis - results from a questionnaire-based study with 4,518 patients. Patient Prefer Adherence 2015; 9:847-855.
4. **Janssen IM**, Scheibler F, Gerhardus A. Importance of hemodialysis-related outcomes: comparison of ratings by a self-help group, clinicians, and health technology assessment authors with those by a large reference group of patients. Patient Prefer Adherence 2016; 10:2491-2500

Hintergrund

Einleitung

„Die Arzeneikunde umfasst drei Stücke: die Krankheit, den Kranken und den Arzt. Der Arzt sei ein Diener der Heilkunst; der Kranke soll zugleich mit ihm der Krankheit entgegenwirken. (Hippokrates)“

Die Medizin ist eine komplexe Wissenschaft und bedarf umfassender Kenntnisse über den menschlichen Körper, dessen Funktionsweise und über medizinische Interventionen. Bei allen Kenntnissen seitens der Mediziner sollte der Patient im Sinne Hippokrates‘ mit dem Mediziner gemeinsam seiner Krankheit entgegenwirken.[1,2] Seine Sicht, seine Bedürfnisse und seine Präferenzen nicht außer Acht zu lassen, sondern einzubeziehen, kann maßgeblich dazu beitragen, dass eine Behandlung erfolgreich ist und das Leben des Patienten insgesamt verbessert wird.

Die Einbeziehung von Patientenpräferenzen kann an vielen Stellen der medizinischen Versorgung von Patienten erfolgen, in der individuellen direkten Behandlungssituation oder in Situationen, bei denen es um kollektive Entscheidungen geht, wie zum Beispiel die Erstattung von medizinischen Leistungen, die ihrerseits einen Einfluss auf direkte Behandlungssituationen haben können.

Die vorliegende kumulative Dissertation befasst sich mit der Frage der Einbeziehung von Patientenpräferenzen im Bereich des Health Technology Assessments - einer Evaluation von medizinischen Interventionen auf Systemebene. Im Rahmen solcher Evaluationen werden Ergebnisse aus klinischen Studien (Evidenz) zusammengetragen, analysiert und ausgewertet, um die Vor- und Nachteile medizinischer Interventionen zu untersuchen. Diese Vor- und Nachteile werden durch Endpunkte wie zum Beispiel Mortalität oder gesundheitsbezogene Lebensqualität operationalisiert, um die Effekte der Intervention messbar zu machen. Die Beantwortung der Frage, ob und auf welche Art und Weise eine Einbeziehung von Patientenpräferenzen bei der Auswahl und Gewichtung dieser Endpunkte bereits erfolgt bzw. erfolgen kann, bildet den Kern der vorliegenden Untersuchung.

Diese kumulative Dissertation besteht aus vier Manuskripten, die zum einen den theoretischen Rahmen darstellen (Manuskript 1 und 2) und zum anderen die empirische Untersuchung abbilden (Manuskript 3 und 4). Im theoretischen Teil werden zunächst die Einsatzmöglichkeiten der Patienteneinbeziehung bei der Endpunktgewichtung geschildert sowie Methoden zusammengefasst, mit denen eine Gewichtung durchgeführt werden kann. Zudem

wird ein kurzer Überblick zur terminalen Niereninsuffizienz gegeben, die dem in dieser Dissertation eingeschlossenen Patientenkollektiv als Erkrankung zugrunde liegt. Im empirischen Teil dieser Dissertation wird in einem ersten Schritt ermittelt, wie eine repräsentative Gruppe von Patienten mit terminaler Niereninsuffizienz Endpunkte gewichtet. Daran anschließend wird in einem zweiten Schritt eruiert, welche Gewichtungen von Endpunkten andere Gruppen vornehmen, um diese anschließend mit der Gewichtung des Patientenkollektivs zu vergleichen. Abschließend werden mögliche Konsequenzen für die Evaluation von medizinischen Interventionen abgeleitet.

Einführung in das Thema

Im Gesundheitswesen gibt es seit einigen Jahren eine starke Besinnung auf Patientenzentrierung, die sich zunehmend auch durch den Wunsch einer stärkeren Einbeziehung von Patienten in klinischer Forschung und bei der Evaluation von medizinischen Interventionen bemerkbar macht.[3-6]

Das amerikanische „Institute of Medicine (IOM)“ definierte eine patientenzentrierte Gesundheitsversorgung wie folgt:

„Providing care that is respectful of and responsive to individual patient preferences, needs, and values, and ensuring that patient values guide all clinical decisions.“[7]

Es gibt viele Bereiche der Gesundheitsversorgung, in denen im Rahmen der Patientenzentrierung Entwicklungen stattfinden und Forschung durchgeführt wird, z. B. bei der Auswahl von wichtigen Forschungsschwerpunkten oder bei der Identifikation und Beseitigung von Barrieren, die eine Studienteilnahme von Patienten erschweren.[5,8,9] Ein Bereich von besonderem Interesse ist die Einbeziehung von Patienten bei der Auswahl von Endpunkten für klinische Studien und für Evaluationen von medizinischen Interventionen auf Basis solcher klinischen Studien im Rahmen eines HTA-Berichts.

Dieser Prozess der Evaluation medizinischer Interventionen mit Bezug zur gesundheitlichen Versorgung der Bevölkerung wird als Health Technology Assessment (HTA) bezeichnet.[10] In HTA-Berichten soll zu der jeweiligen Fragestellung die vorhandene Evidenz systematisch recherchiert, kritisch beurteilt und evaluiert werden. Während HTA-Berichte über eine Nutzenbewertung der zu prüfenden Intervention hinaus auch Informationen über den gesundheitsökonomischen Stellenwert sowie den sozialen, ethischen und legalen Rahmen liefern[11], ist ihr Grundstein die systematische Nutzenbewertung. Ohne diesen Grundstein verlieren die weiteren Schritte an Bedeutung.

Ziel aller klinischen Forschung – und damit auch im HTA-Bereich - sollte sein, die Gesundheit von Patienten zu verbessern. Hierbei ist eine wichtige Voraussetzung, dass die Effekte der medizinischen Interventionen anhand von prospektiv festgelegten Endpunkten untersucht werden, die für Patienten relevant sind [12,13], das heißt, bei denen ein Nutzen für die Betroffenen feststellbar ist.

Dabei haben medizinische Interventionen häufig Effekte auf mehrere patientenrelevante Endpunkte gleichzeitig (z. B. Mortalität, Morbidität, gesundheitsbezogene Lebensqualität). Deshalb sollten sowohl in den klinischen Studien als auch in den systematischen Bewertungen dieser Studien alle relevanten Endpunkte erfasst werden. Im Idealfall zeigt eine neue Intervention Verbesserungen in allen relevanten Endpunkten und ist somit einer Vergleichsintervention eindeutig überlegen. Geht jedoch mit einer Verbesserung des einen Endpunktes die Verschlechterung eines anderen Endpunktes einher, ist das Ergebnis der Nutzenbewertung nicht unmittelbar eindeutig. Die Komplexität der Bewertung kann mit der Zahl der relevanten und erhobenen Endpunkte zunehmen. Um einzuordnen, ob eine Intervention überlegen ist, müssen Entscheidungen über die relative Bedeutung der Endpunkte vorgenommen werden. Diese Entscheidungen werden bislang im Regelfall von medizinischen Experten vorgenommen. Eine patientenzentrierte Gesundheitsversorgung erfordert aber eine stärkere Einbeziehung der Patienten.

Relevanz des Themas

Dass eine medizinische Intervention in einigen Endpunkten Verbesserungen und in anderen Endpunkten Verschlechterungen verursacht, tritt häufig bei chronischen, insbesondere bei vielen onkologischen Erkrankungen auf. Bisher werden diese Priorisierungen entweder implizit oder explizit von medizinischen Experten vorgenommen. Patienten werden bei der Priorisierung bislang nicht regelhaft einbezogen. Eine einheitliche systematische Vorgehensweise bei der Priorisierung von Endpunkten existiert nicht.

Es haben sich bereits einige Studien mit der Frage beschäftigt, ob die Ergebnisse der Priorisierung durch Ärzte und Patienten konsistent sind. Es gibt bislang jedoch noch keine eindeutigen Empfehlungen für die Erstellung von Standards, die für die Durchführung von einer Endpunktgewichtung bei der Evaluation von medizinischen Interventionen verwendet werden können.[14-18]

Die Gewichtung der Endpunkte wird zunehmend gefordert. So wird beispielsweise im Handbuch der Cochrane Collaboration angeführt, dass die Endpunkte für die Beantwortung der Fragestellung in primäre und sekundäre eingeteilt werden sollen. Allerdings finden sich

keine Angaben im Handbuch, durch wen diese Einteilung vorgenommen werden soll. Ein systematisches Vorgehen wird nicht beschrieben, ebenso wenig die Einbeziehung von Patientenpräferenzen.[19]

Dass eine unterschiedliche Bewertung der einzelnen Endpunkte einen Einfluss auf das Ergebnis einer systematischen Übersicht haben kann, wird auch in den Methoden des Instituts für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) beschrieben:

„Sowohl Nutzen- als auch Schadenauswirkungen können eine unterschiedliche Wichtigkeit für die Betroffenen haben, die sich ggf. durch qualitative Erhebungen oder bereits bei der Beratung durch Betroffene, Patientenvertretungs- und/oder Verbraucherorganisationen im Zusammenhang mit der Definition patientenrelevanter Endpunkte abzeichnet... In einer solchen Situation kann es sinnvoll sein, eine Hierarchisierung von Endpunkten vorzunehmen. Allgemeine Nutzen- bzw. Schadenaussagen werden sich dann in erster Linie auf Belege bezüglich der höher gewichteten Zielgrößen stützen. Geplante Subgruppen- und Sensitivitätsanalysen werden dann im Wesentlichen für die höher gewichteten Zielgrößen durchgeführt, während solche Analysen für die verbleibenden Zielgrößen nicht routinemäßig durchlaufen werden.“ [20]

Auch das Grading of Recommendations Assessment, Development and Evaluation (GRADE)-System empfiehlt eine Gewichtung. Zur Bewertung von Evidenz aus systematischen Übersichten und Leitlinien sieht es in einem ersten Schritt eine Benennung von patientenrelevanten Endpunkten vor.[19] In einem zweiten Schritt werden diese Endpunkte auf einer Skala von 1 (unwichtig) bis 9 (sehr wichtig) gewichtet. Die Festlegung und Gewichtung der Endpunkte erfolgte hier durch die Ersteller, in der Regel Mediziner und Methodiker.

Zusammenfassend kann festgehalten werden, dass sich verschiedene Institutionen des Gesundheitssystems einig sind, dass ein Einbezug von Betroffenen bei der Gewichtung von Endpunkten als hilfreich erachtet wird. Gleichwohl mangelt es bislang an standardisierten und validen Verfahren.

Theoretischer Bezugsrahmen

Patientenbeteiligung

Die Konzeptualisierung der Patientenbeteiligung an Gesundheits- und Sozialforschung wurde in einer systematischen Übersichtsarbeit untersucht.[21] Ein explizites theoretisches Modell konnte nicht identifiziert werden, das die Patientenbeteiligung umfassend beschreiben und

erklären könnte. Allerdings konnten einzelne Elemente identifiziert werden, die für die Erstellung einer umfassenden Theorie von Bedeutung sind.

Patientenbeteiligung wird häufig als Kontinuum zwischen einer niedrigen und einer hohen Stufe der Partizipation beschrieben.[22,23] Hierbei werden häufig drei Stufen formuliert:[24] Die niedrigste Stufe bildet die Beratung durch Patienten; Meinungen und Einschätzungen der Patienten werden hierbei zur Information und Unterstützung der Forschung eingeholt. Die Patienten nehmen auf dieser Stufe lediglich eine beratende Funktion ein, häufig nur punktuell, und sind nicht an Entscheidungen beteiligt.[23] Auf der zweiten Stufe wird eine Zusammenarbeit mit den Patienten angestrebt, die durch eine gewisse Kontinuität der Beteiligung gekennzeichnet wird.[22] Die Patienten agieren als Partner im Forschungsprozess, werden an Entscheidungen beteiligt und tragen die Verantwortung für den Forschungsprozess mit.[23] Die dritte und höchste Stufe der Patientenbeteiligung ist die nutzergeleitete Forschung. Auf dieser Stufe liegt die Verantwortung des gesamten Forschungsprozess oder auch nur für einzelne Teilbereiche bei den Patienten, die dabei Unterstützung durch Forscher bekommen.

Basierend auf vorhergehenden Arbeiten wurde durch Bagley et al.[25] eine Aufreihung einzelner Schritte bei der Patientenbeteiligung erstellt, die mit der Definition relevanter Forschungsfragen beginnt und mit der Messung des Einflusses der Ergebnisse und der Information für zukünftige Forschung endet.

Eine systematische Übersichtsarbeit zur Patientenpartizipation in Forschungsprozessen untersuchte, welche Methoden häufig zur Einbeziehung von Patienten verwendet wurden. In den meisten Studien wurden Fokusgruppen, Interviews oder Surveys eingesetzt. Eine aktivere Teilnahme wurde durch die Beteiligung an Gremien oder an einem Beirat gewährleistet. Auch wurde in einigen Studien die regelmäßige Teilnahme an Besprechungen mit den Forschern als Möglichkeit beschrieben. In einigen wenigen Studien wurden Patienten bei operativen Aufgaben, wie der Datenerfassung oder Auswahl der Endpunkte, einbezogen.[5]

Die Einbeziehung von Patienten bei der Erstellung von systematischen Übersichten erfolgt mittlerweile in einigen Organisationen. Aber auch hier basiert die Einbeziehung häufig auf Konsultationen von „professionellen“ Patienten, z. B. Vertretern von Selbsthilfegruppen. Eine regelmäßige und intensivere Einbindung größerer Patientengruppen findet nur selten statt.[26,27]

Die Arbeit von Esmail et al. [28] untersuchte den angenommenen Einfluss einer Patientenbeteiligung auf die Forschung. Eine ausführliche Liste wurde erstellt, um die Aspekte abzubilden, auf die eine Patientenbeteiligung Einfluss haben könnte. Eine Patienten-

beteiligung könnte demnach Einfluss auf Forschungsthemen, Patientenrekrutierung, untersuchte Interventionen, Implementierung von Studienergebnissen und auf die Relevanz der Forschung haben. Zudem könnte eine Patientenbeteiligung die Übertragung und Akzeptanz der Studienergebnisse erhöhen und dazu führen, dass andere Endpunkte untersucht werden. Ein weiterer Einfluss wurde in der Veränderung der Kenntnisse und Einstellungen der Forscher antizipiert. Ebenso könnte die Beteiligung ein wachsendes Empowerment der Patienten erreichen. Auch ein Einfluss auf allgemeinere Aspekte wurde beschrieben; hierunter fallen Aspekte der Demokratie und Verantwortung (z. B. Transparenz, Rechtmäßigkeit, öffentliches Vertrauen) und moralische Aspekte (z. B. Gerechtigkeit, ethisch akzeptable Forschung).

Neben den denkbaren Vorteilen einer Patientenbeteiligung sollten ebenso kritische Aspekte und Hindernisse betrachtet werden. So kann eine intensivere Patientenbeteiligung Forschungsprozesse verlängern. Zudem müssen Patienten und auch Forscher Zeit für die Beteiligung aufbringen. Daher sollte die Einbeziehung auf sinnvolle Weise geschehen, damit beide Gruppen die eingebrachten Ressourcen als sinnvolle Maßnahme betrachten. Forscher befürchten häufig, dass durch eine Patientenbeteiligung Aspekte zu einer Studie hinzugefügt werden, die die Durchführbarkeit erschweren. Patienten könnten frustriert werden, wenn sie das Gefühl haben, dass sie eher pro forma eingebunden werden, ihre Anregungen aber nicht umgesetzt werden.[5]

Zusammenfassend zeigt sich allerdings sehr deutlich, dass eine Einbeziehung von Patienten in vielen Bereichen der medizinischen Forschung diskutiert wird. Standardisierte Prozesse, wie diese Einbeziehung tatsächlich umgesetzt werden kann, sind allerdings noch nicht vorhanden.

Methoden zur Erhebung von Werten und Präferenzen

Methoden, die für die Erhebung von Werten und Präferenzen verwendet werden, können konzeptuell in zwei Gruppen eingeteilt werden: 1) Methoden, die Einstellungen messen, und 2) Methoden, die Präferenzen messen.[29]

- 1) Bei der Messung von Einstellungen werden Rangfolgen- oder Rating-Methoden verwendet, um die Wichtigkeit eines diskreten Sets von Variablen, hier einem Set von Endpunkten, zu erfassen. Bei diesen Methoden werden die Endpunkte gemäß ihrer Priorität für die individuelle Person in eine Hierarchisierung gebracht. Die Stärke der Präferenz für einen Endpunkt im Vergleich zu einem anderen kann mit diesen Methoden nicht dargestellt werden. Bewertungen basieren vielmehr auf der Betrachtung der wichtigsten Endpunkte. Hierbei wird unterstellt, dass die Endpunkte,

die als unwichtiger eingestuft wurden, keinen oder nur einen sehr geringen Einfluss auf die Entscheidung der Individuen haben.

- 2) Bei der Präferenzmessung werden Methoden verwendet, die ihren Ursprung in der Ökonomie haben. Präferenzen werden als Nutzwerte definiert, die Individuen beim Konsum von Gesundheitsprodukten (z. B. Medikamenten) und Gesundheitsleistungen (z. B. Versorgungsaspekten) an einzelne Aspekte vergeben.[30] Präferenzmessungsmethoden vergeben einen spezifischen Nutzwert für jeden Endpunkt, sodass alle Endpunkte dann gemäß ihrer Größen sortiert werden können. Endpunkte mit einem höheren Nutzwert haben daher einen größeren Einfluss auf die Bewertung als Endpunkte mit einem geringeren Nutzwert. Mit diesen Methoden kann die Präferenz für jeden einzelnen Endpunkt detailliert dargestellt werden, auch im Verhältnis zueinander. Klassische Methoden zur Präferenzmessung stellen Methoden der multikriteriellen Entscheidungsfindung dar, wie die Conjoint Analyse (CA) und der Analytic Hierarchy Process (AHP). Eine weitere methodische Gruppe, die zu den Präferenzmessungsmethoden gezählt werden kann, sind Utility Assessment Methoden.

Welche dieser Methoden am besten für die empirische Datenerhebung dieser Dissertation geeignet ist, wurde im Rahmen einer deskriptiven Übersichtsarbeit untersucht (Manuskript 2). In dieser Übersichtsarbeit wurden alle Studien zusammengetragen, die Methoden der Präferenzmessung oder der Messung von Einstellungen verwendet hatten, um Endpunkte zu gewichten, und ihre Vor- und Nachteile untersucht. Die Bewertung von vielen Endpunkten, so wie in dieser Dissertation vorgesehen, kann am besten mit einer Methode zur Messung von Einstellungen durchgeführt werden, da bei den anderen Methoden (z. B. Conjoint-Analyse) die Komplexität zu groß würde. Daher wurde bei dieser Dissertation eine Ratingmethode verwendet.

Terminale Niereninsuffizienz

Die Frage der Einbeziehung von Patientenpräferenzen wurde im Rahmen der vorliegenden Arbeit anhand einer Gruppe von Patienten mit terminaler Niereninsuffizienz exploriert. Im Folgenden wird dieses Krankheitsbild in Kürze dargestellt.

Die terminale Niereninsuffizienz ist ein zunehmendes Gesundheitsproblem weltweit. Die Zunahme lässt sich unter anderem durch das Ansteigen relevanter Risikofaktoren wie Bluthochdruck, steigendes Lebensalter, Diabetes oder Adipositas erklären.[31-33] Um den Schweregrad der Erkrankung festzustellen, wird die glomeruläre Filtrationsrate (GFR) verwendet, die die Filtermöglichkeit der Nieren beschreibt. Der Schweregrad einer Nieren-

insuffizienz lässt sich in fünf Stadien einteilen. Während in Stadium I die Filtrationsrate noch normal oder sogar erhöht ist, sinkt sie mit zunehmendem Stadium. In Stadium V, das als terminale Niereninsuffizienz bezeichnet wird, ist die Nierenfunktion massiv eingeschränkt oder fällt vollkommen aus.

Eine Niereninsuffizienz des Stadiums V manifestiert sich in einer Vielzahl von Symptomen, unter anderem Hypertonie, Anämie, Azidose, kardialen Arrhythmien und Ödemen. Zusätzlich treten auch häufig Depressionen, Gewichtsverlust, Übelkeit, eine eingeschränkte körperliche Leistungsfähigkeit und periphere Nervenschäden auf.[34,35] In diesem Stadium kann eine Niereninsuffizienz nur noch mit einer Nierenersatztherapie behandelt werden. Diese Therapie besteht entweder aus einer Nierentransplantation oder einer Dialyse.[36] Während 25% aller Patienten in diesem Stadium eine Nierentransplantation erhalten, werden ungefähr 65% der Patienten mit einer Hämodialyse und 10% mit einer peritonealen Dialyse behandelt.[33]

Da im Rahmen dieser Dissertation ausschließlich Patienten mit terminaler Niereninsuffizienz eingeschlossen wurden, die mit der häufigsten Therapieform – der Hämodialyse – behandelt wurden, wird diese im Folgenden detaillierter beschrieben.

Die Hämodialyse, die in der Regel in einem Dialysezentrum durchgeführt wird, ist ein extrakorporales Verfahren, bei dem mithilfe eines venösen oder arteriellen Zugangs das Blut aus dem Körper durch eine Maschine geleitet wird, in der sich das Dialysat befindet. Durch unterschiedliche Konzentrationen von molekularen Substanzen wird mittels Osmose das Blut von Giftstoffen befreit. Eine Hämodialysebehandlung dauert in der Regel 4-5 Stunden und wird dreimal pro Woche durchgeführt. Patienten mit einer terminalen Niereninsuffizienz leiden häufig an Ödemen, daher ist ein zusätzliches Ziel der Dialysebehandlung die diuretische Funktion der Niere zu ersetzen. Da die Behandlung aber nicht kontinuierlich, sondern nur in der beschriebenen Frequenz durchgeführt wird, müssen Hämodialyse-Patienten darauf achten, nur eine geringe Flüssigkeitsmenge zu sich zu nehmen. Eine Alternative ist eine Nacht-Hämodialyse, die acht Stunden dauert, oder eine Heim-Hämodialyse, bei der die Patienten eine häufigere, meistens kürzere Dialysebehandlung durchführen können.[37-39]

Hämodialyse-Patienten haben nur eine relativ niedrige gesundheitsbezogene Lebensqualität, sind häufig unzufrieden mit ihrer Behandlung und fühlen sich durch die Behandlung in ihrem Alltag beeinträchtigt.[40-42]

Aspekte, wie eine gefühlte Unterstützung durch den behandelnden Arzt oder die Zufriedenheit mit der Behandlung, haben für Patienten nicht nur einen direkten Einfluss auf

ihr Wohlergehen. Sie haben darüber hinaus auch Einfluss auf die Behandlungstreue und die Kooperation der Patienten, was wiederum langfristige Auswirkungen auf die Lebenserwartung und Morbidität der Patienten hat.[43-46] Die Kenntnis über Präferenzen und Erfahrungen von Patienten ist daher essenziell für eine gute Behandlung.

Ziele und Fragestellungen

Ziel der vorliegenden Arbeit war es, eine Methodik zu entwickeln, anhand derer man die Priorisierung von patientenrelevanten Endpunkten im Rahmen eines HTAs vornehmen kann. Da die Auswahl von relevanten Endpunkten und die Gewichtung dieser Endpunkte nicht nur für die Evaluation medizinischer Interventionen relevant ist, sondern gleichermaßen auch für die Generierung von Evidenz in klinischen Studien und darüber hinaus insbesondere für die klinische Praxis (z. B. bei der individuellen Entscheidungsfindung), wurde die empirische Datenerhebung nicht auf eine HTA-Situation beschränkt.

Im Folgenden werden die spezifischen Fragestellungen, die in den einzelnen Manuskripten untersucht werden, kurz dargelegt:

Manuskript 1 (M1) untersucht die grundsätzliche Bedeutung einer Einbeziehung von Patientenpräferenzen bei der Gewichtung von patientenrelevanten Endpunkten. Auf Basis von vorhandenen Strukturen im deutschen Gesundheitssystem werden Konzepte auf Mikro- und Makroebene dargestellt, wie eine Patientenpartizipation bei der Festlegung und Gewichtung von Endpunkten erfolgen kann. Weiterhin werden mögliche Weiterentwicklungen vorgestellt und diskutiert.

In Manuskript 2 (M2) wird die Frage untersucht, welche Methoden bislang für die Gewichtung von Endpunkten in einem medizinischen Kontext eingesetzt wurden. Diese Frage wird anhand einer deskriptiven Übersichtsarbeit bearbeitet, in der auch die Vor- und Nachteile der identifizierten Methoden zur Gewichtung von Endpunkten dargestellt werden.

In Manuskript 3 (M3) wird empirisch untersucht, welche Endpunkte für eine repräsentative Stichprobe von Patienten mit terminaler Niereninsuffizienz relevant sind und welche Priorisierung die Patienten den ausgewählten Endpunkten geben. Subgruppenanalysen werden durchgeführt, um zu prüfen, ob die Gewichtungen stabil sind, oder ob sie sich z. B. bezüglich der Dauer der Dialysetherapie oder dem sozioökonomischen Status unterscheiden.

In Manuskript 4 (M4) wird empirisch der Frage nachgegangen, inwiefern sich die Gewichtung der relevanten Endpunkte durch die repräsentative Patientenstichprobe von

Gewichtungen unterscheidet, die von anderen Gruppen vorgenommen werden. Daten von Patienten, die Mitglieder einer Selbsthilfegruppe sind, von Nephrologen und von HTA-Erstellern werden hierfür erhoben und mit den Daten der repräsentativen Patientenstichprobe verglichen.

In der Gesamtschau ergeben sich somit folgende Fragestellungen:

- 1) Welche Bedeutung hat eine Patientenbeteiligung bei der Priorisierung von patientenrelevanten Endpunkten? (M1)
- 2) Welche Methoden zur Priorisierung von Endpunkten gibt es und welche Vor- und Nachteile haben die Priorisierungsmethoden? (M2)
- 3) Welche Endpunkte erscheinen Patienten mit terminaler Niereninsuffizienz relevant? (M3)
 - a. Welche Priorität haben diese Endpunkte?
 - b. Wie stabil ist diese Gewichtung? Verändert sie sich
 - i. bei unterschiedlichen Komorbiditäten?
 - ii. mit zunehmender Dialysedauer?
 - iii. bei unterschiedlichen sozioökonomischen bzw. soziodemographischen Merkmalen der Patienten?
- 4) Gibt es einen Unterschied zwischen einer repräsentativen Auswahl von Patienten mit terminaler Niereninsuffizienz und anderen Gruppen? (M4)
 - a. Angehörige einer Selbsthilfegruppe?
 - b. Ärzten mit Spezialisierung in der Nephrologie?
 - c. HTA-Experten?

Im Folgenden werden die einzelnen Manuskripte kurz dargestellt, in der Diskussion werden die Erkenntnisse aller vier Manuskripte gemeinsam berücksichtigt und besprochen.

Situation in Deutschland

Manuskript 1 – Status quo und mögliche Weiterentwicklungen

In diesem Manuskript wurde der Frage nachgegangen, inwieweit eine Einbeziehung von Patientenpräferenzen bei der Festlegung und Gewichtung von Behandlungszielen bereits erfolgt und welche Entwicklungspotenziale vorhanden sind.

Diese Betrachtung erfolgte auf Mikro- sowie auf Makroebene. Auf der Mikro-Ebene, die vor allem die direkte Interaktion zwischen einem Patienten und medizinischem Fachpersonal im Gesundheitswesen beschreibt, wurden exemplarisch Evidenzbasierte Patienteninformationen und die Partizipative Entscheidungsfindung (Shared decision making) als Konzepte näher analysiert.

Auf der Makroebene, auf der typischerweise Entscheidungen getroffen werden, die für weite Teile des Gesundheitswesens von Relevanz sind, wurden zwei Institutionen als Beispiel herangezogen und deren Konzepte zur Einbeziehung von Patientenpräferenzen analysiert: der Gemeinsame Bundesausschuss (G-BA) und das IQWiG.

Über die Beschreibung der genannten Konzepte hinaus wurden jeweils ein möglicher Ausblick und Anregungen für Weiterentwicklungen vorgestellt.

Patientenpartizipation auf der Mikroebene

Die Patientenpartizipation auf der Mikro-Ebene im Gesundheitswesen findet zum Beispiel in Evidenzbasierten Gesundheitsinformationen Berücksichtigung. Diese beruhen zum einen auf wissenschaftlich belegten Aussagen zu Erkrankungen und deren Behandlungsmöglichkeiten. Zum anderen berücksichtigen sie aber auch die Erfahrungen und Bedürfnisse betroffener Patienten. Gesundheitsinformationen müssen für die Allgemeinbevölkerung verständlich sein. Des Weiteren müssen sie relevant sein, wobei unter Relevanz in diesem Kontext verstanden wird, dass als Behandlungsziele auch solche dargestellt werden, die für Patienten bedeutsam sind. Hierunter fällt zum Beispiel neben der Lebenserwartung auch die gesundheitsbezogene Lebensqualität. Eine mögliche Weiterentwicklung des Konzepts der Evidenzbasierten Patienteninformationen wäre eine Evaluation der Wirkung dieser Informationen auf die Patienten. Denn trotz evidenzbasierter Inhalte können z. B. bestimmte Darstellungsformen, Begriffe oder suggestive Erfahrungsberichte unerwünschte Wirkungen verursachen (z. B. unnötige Verängstigung).

Ein weiteres Konzept auf Mikroebene stellt die Partizipative Entscheidungsfindung (Shared decision making) dar, dessen Kernelement die Einbeziehung von Patientenpräferenzen ist.

Dieses Konzept betrachtet die Patienten und Ärzte hinsichtlich der Auswahl einer Behandlung als gleichberechtigte Partner. Eine unterschiedliche Gewichtung der Behandlungsziele kann dazu führen, dass Ärzte und Patienten zu divergierenden Entscheidungen kommen. Eine transparente Darstellung dieser Behandlungsziele ist daher für eine gemeinsame Entscheidung notwendig. Eine mögliche Weiterentwicklung des bislang bestehenden Konzepts der partizipativen Entscheidungsfindung wäre eine Evaluation der Wirksamkeit dieser Maßnahmen unter Alltagsbedingungen.

Patientenpartizipation auf der Makroebene

Auf der Makroebene wurden die Vorgehensweisen zur Einbeziehung von Patientenpräferenzen von G-BA und IQWiG untersucht. Der G-BA ist das oberste Beschlussgremium der gemeinsamen Selbstverwaltung der Ärzte, Krankenhäuser und Krankenkassen in Deutschland. Eine der Hauptaufgaben des G-BA ist die Erstellung von Richtlinien, anhand derer der Leistungskatalog der gesetzlichen Krankenversicherung (GKV) festgelegt wird. Eine Patientenbeteiligung im G-BA ist dabei gesetzlich verankert. Gemäß der Verfahrensordnung des G-BA und der Patientenbeteiligungsverordnung haben Patientenvertreter ein Mitberatungsrecht ohne Stimmrecht.[47,48] Im Sinne einer Weiterentwicklung der Patientenpartizipation im G-BA könnte eine bessere Ausstattung der Institution Patientenvertretung es in Zukunft möglich machen, mit den Vertretern der stimmberechtigten Bänke auf gleicher Augenhöhe zu verhandeln.

Das IQWiG erstellt im Auftrag des G-BA systematische Übersichtsarbeiten zum Nutzen und Schaden von medizinischen Interventionen. Bislang ist ein wichtiger Aspekt der Patientenbeteiligung bei der Berichterstattung durch das IQWiG, dass Patienten zu Beginn der jeweiligen Projektarbeit bei der Identifikation relevanter Endpunkte eingebunden werden. Eine mögliche Erweiterung könnte die Gewichtung von Behandlungszielen betreffen. Die Gewichtung wird bislang häufig und eher implizit von medizinischen oder methodischen Experten vorgenommen, die an der Erstellung der systematischen Übersichtsarbeiten beteiligt sind. Eine Einbeziehung von Patientenpräferenzen bei der Hierarchisierung von patientenrelevanten Endpunkten wird in den allgemeinen Methoden des Instituts [49] seit der Version 4.0 (veröffentlicht im Jahr 2011) als optionale Möglichkeit angeführt. Bislang findet diese allerdings nicht regelhaft statt, bietet aber Weiterentwicklungspotenzial.

Systematische Übersicht der Methoden

Manuskript 2 Deskriptive Übersichtsarbeit möglicher Methoden

Um die Frage zu beantworten, welche Methoden zur Gewichtung von Endpunkten verwendet werden und welche Vor- und Nachteile diese Methoden haben, wurde eine deskriptive Übersichtsarbeit durchgeführt.

Systematische Suche

Im Februar 2014 wurde in den Literaturdatenbanken MEDLINE, EMBASE, PsycINFO und den Literaturdatenbanken der Cochrane Bibliothek anhand einer für jede Datenbank angepassten Suchstrategie nach relevanter Literatur gesucht.

Ein- und Ausschlusskriterien

Studien, die Präferenzen für Endpunkte in bestimmten Krankheitsbildern untersuchten, wurden eingeschlossen, die Art (Methode) der Präferenzmessung wurde nicht vorab eingeschränkt. Als Endpunkte wurden Mortalität, Morbidität, gesundheitsbezogene Lebensqualität aber auch weitere Aspekte erfasst, die im Zusammenhang mit dem Krankheitsbild und dessen Behandlung stehen konnten, wie z. B. Fahrtzeit zur Behandlung, Flexibilität bei der Dosierung, Art der Anwendung sowie der Einfluss auf das tägliche Leben. Abhängig von der untersuchten Fragestellung galten die Präferenzen sowohl von Patienten als auch von Familienmitgliedern, Betreuern und der Allgemeinbevölkerung (z. B. bei Screeningstudien) als adäquat und wurden berücksichtigt. Präferenzen von medizinischem Personal (Ärzten und Pflegern) wurden nur als zusätzliche Information berücksichtigt. Studien, die nur die Präferenzen dieser Berufsgruppen darstellten ohne direkt oder indirekt Betroffene einzubeziehen, wurden ausgeschlossen.

Es wurden nur Studien eingeschlossen, die nach dem Jahr 2000 veröffentlicht wurden, da eine Darstellung der aktuell verwendeten Methoden zur Präferenzmessung durchgeführt werden sollte. Studien, die lediglich methodische Aspekte der Präferenzmessung, aber keine Ergebnisse dargestellt hatten, wurden ausgeschlossen. Selbiges galt für Editorials, Leserbriefe oder Konferenzabstracts. Des Weiteren wurden Studien ausgeschlossen, in denen die Zielsetzung nicht zur Fragestellung der vorliegenden Arbeit passte. Dies betraf zum Beispiel Studien, die entweder Präferenzen für andere Fragestellungen erhoben (z. B. Priorisierung von Forschungsfragen) oder gar keine Ergebnisse für einzelne Endpunkte, sondern einen generellen Nutzwert für verschiedene Behandlungsmöglichkeiten darstellten.

Datenextraktion und Darstellung der Ergebnisse

Mithilfe eines standardisierten Datenextraktionsbogens wurden folgende Studiencharakteristika von allen eingeschlossenen Studien extrahiert: Zielsetzung, Krankheitsbild, Studienteilnehmer (z. B. Patienten, Allgemeinbevölkerung), Anzahl der Studienteilnehmer und Anzahl der Teilnehmer, bei denen die Präferenzerhebung vollständig vorlag, Präferenzerhebungsmethode und verwendete Endpunkte.

Die eingeschlossenen Studien wurden anhand der verwendeten Präferenzerhebungsmethode gruppiert. Anschließend wurden Vergleiche zwischen den verschiedenen Präferenzerhebungsmethoden anhand der extrahierten Studiencharakteristika vorgenommen. In einigen Studien wurden zudem die Präferenzen von mehreren Gruppen erhoben. Im Ergebnisteil wurden auch diese Vergleiche dargestellt.

Ergebnisse

Die systematische Suche lieferte 9740 Referenzen, von denen nach Dubletten-Entfernung und Sichtung der Titel und Abstracts 219 Referenzen im Volltext gesichtet wurden. Insgesamt wurden 124 Referenzen als relevante Studien eingestuft und in die deskriptive Übersichtsarbeit eingeschlossen.

Die Studien konnten vier Gruppen zugeordnet werden: 1) Methoden der multikriteriellen Entscheidungsfindung (z. B. Conjoint Analyse (CA) oder Analytic Hierarchy Process (AHP)) (n = 71), 2) Ranking oder Rating Methoden (n = 25), 3) Utility Assessment Methoden (z. B. Time-trade-off oder Willingness-to-pay) (n = 5) und 4) ein Vergleich verschiedener Methoden innerhalb einer Studie (n = 23).

Ein zentrales Ergebnis war, dass sich die Anzahl der erhobenen Endpunkte zwischen den einzelnen methodischen Gruppen unterschied. Die mediane Anzahl erhobener Endpunkte war in der Gruppe von Studien, die Ranking oder Rating Methoden verwendet hatten, mit zwölf Endpunkten am größten. Studien, die Methoden der multikriteriellen Entscheidungsfindung verwendet hatten, untersuchten im Median sechs Endpunkte. Dies galt ebenfalls für Studien, die mehrere Methoden miteinander verglichen, was vermutlich daran lag, dass hier häufig eine Methode der multikriteriellen Entscheidungsfindung mit einer anderen Methode verglichen wurde. Studien, die Utility Assessment Methoden verwendet hatten, untersuchten im Median zehn Endpunkte.

Studien, in denen Vergleiche verschiedener Präferenzerhebungsmethoden vorgenommen wurden, sowie Studien, in denen die Präferenzen von unterschiedlichen Teilnehmergruppen

miteinander verglichen wurden, kamen häufig zu unterschiedlichen Hierarchisierungen von Endpunkten.

Insgesamt konnte auf Basis der eingeschlossenen Studien keine Methode identifiziert werden, die sich am besten für den Einsatz in HTAs eignet. Alle Methoden haben Vor- und Nachteile. Keine zeigte eine starke Überlegenheit gegenüber einer anderen Methode. Die Wahl der Gruppe für eine Gewichtung von Endpunkten ist entscheidend, da die Hierarchisierung von Endpunkten zwischen verschiedenen Teilnehmergruppen zum Teil sehr unterschiedlich war.

Empirische Datenerhebung

Manuskript 3 – Präferenzen Patienten

Um zu untersuchen, welche Behandlungsziele für Patienten mit terminaler Niereninsuffizienz relevant sind und welche Gewichtung sie diesen Behandlungszielen geben würden, wurde eine empirische Studie durchgeführt. Diese Studie beinhaltete die Identifizierung relevanter Behandlungsziele und eine Präferenzenerhebung.

Identifikation relevanter Behandlungsziele

Für die Identifizierung relevanter Behandlungsziele wurde eine Gruppendiskussion mit Patienten, die an einer terminalen Niereninsuffizienz leiden und einer regionalen Selbsthilfegruppe (Sektion Niederrhein, Interessengemeinschaft Niere e.V.) angehören, durchgeführt. Diese Gruppendiskussion fand im Rahmen eines regulären Gruppentreffens statt, an dem sieben Patienten teilnahmen. Die Patienten wurden aufgefordert, für sie relevante Behandlungsziele während einer Dialysebehandlung zu benennen.

Präferenzenerhebung

Die Mitschrift der Gruppensitzung wurde verwendet, um aus den genannten Behandlungszielen 23 Endpunkte zu definieren. Die als Endpunkte formulierten Behandlungsziele wurden in einem Prätest an weiteren fünf Patienten, die nicht an der Gruppendiskussion teilgenommen hatten, auf Verständlichkeit und Vollständigkeit überprüft. Die vollständige Liste mit den 23 Endpunkten bildet die Grundlage für die Präferenzenerhebung. Alle aufgenommenen Endpunkte sind patientenrelevant, d. h. sie können unmittelbar durch den Patienten wahrgenommen werden. Rein klinische Endpunkte, wie Surrogatparameter, wurden von den Patienten nicht genannt und daher nicht aufgenommen. Da durch die deskriptive Übersichtsarbeit (Manuskript 2) aufgezeigt wurde, dass viele komplexe Präferenzenerhebungs-

methoden wie z. B. die Conjoint Analyse aus Praktikabilitätsgründen nur mit wenigen Endpunkten verwendet wurden, wurde eine Rating-Methode ausgewählt, um die Gewichtung der Endpunkte durchzuführen. Die Rating-Skala war in der Form einer diskreten visuellen Analogskala mit Werten von 1 bis 9 aufgebaut, wobei die beiden Extremwerte mit schriftlichen Ankern „überhaupt nicht wichtig“ und „sehr wichtig“ anstelle der Zahlen versehen wurden.

Die Befragung fand im 4. Quartal 2008 in den Zentren des KfH (Kuratorium für Dialyse und Nierentransplantation e.V.) statt, dem größten privaten Anbieter von Hämodialyse in Deutschland. In mehr als 200 Zentren werden schätzungsweise 18 000 Patienten versorgt. Als Teil des Qualitätssicherungsprogramms der KfH-Zentren wird dort jährlich eine Patientenbefragung durchgeführt (Qualität in der Nephrologie, QiN). Durch eine Kooperation mit den KfH-Zentren konnte die Präferenzhebung in die jährliche Patientenbefragung aufgenommen werden. Die Kooperation beinhaltete die Bereitstellung der beantworteten Fragebögen, inklusive der Gewichtung der 23 Endpunkte, sowie weitere Patientencharakteristika, die zum Teil durch die Verknüpfung mit einer elektronischen Patientenakte ergänzt wurden.

Datenanalyse

Die Gewichtung der Endpunkte wurde mithilfe von deskriptiver Statistik beschrieben (Mittel- und Medianwerte, Standardabweichungen). Anhand der Mittelwerte der einzelnen Endpunkte wurde eine Hierarchisierung vorgenommen. Weiter wurden die Präferenzbewertungen in drei Kategorien klassifiziert. Die Kategorie „sehr wichtig“ beinhaltet die Werte 7, 8, oder 9; die Kategorie „eher wichtig“ beinhaltet die Werte 4, 5, oder 6 und die Kategorie „nicht wichtig“ beinhaltet die Werte 1 (= überhaupt nicht wichtig), 2 oder 3. Subgruppenanalysen wurden mit der Mann-Whitney-U-Teststatistik für dichotome Variablen durchgeführt (Geschlecht, Komorbidität, Diabetes) oder mit dem Kruskal-Wallis-Test für kategoriale Variablen (Dialysedauer, Alter, Bildung, Beruf).

Ergebnisse

Insgesamt lagen die Daten von 4518 Patienten vollständig vor und konnten ausgewertet werden. Obwohl aus organisatorischen Gründen keine Angaben zu der Rücklaufquote gemacht werden konnten, konnte die Patientenstichprobe als repräsentative Stichprobe verwendet werden. Die Stichprobe bildete die Charakteristika der Hämodialyse-Patienten in Deutschland zufriedenstellend ab. [39] Die Patienten waren durchschnittlich 66,6 Jahre alt und 58% waren männlich. Die Mehrheit der Patienten war bereits im Ruhestand (76%), nur

10% waren noch berufstätig. Von den Patienten, die Angaben zu ihrer Bildung machten, hatte die Mehrzahl (51%) einen Hauptschulabschluss. Die Patienten wurden durchschnittlich seit 4,4 Jahren mit Hämodialyse behandelt und 40% der Patienten hatten Diabetes.

Die Mittelwerte der 23 Endpunkte lagen im Bereich zwischen 4,5 und 8,3 auf einer Skala von 0 bis 9. Die Unterschiede zwischen den Mittelwerten der einzelnen Endpunkte waren eher schmal und die beobachteten Konfidenzintervalle überlappten weitestgehend. Die drei am wichtigsten bewerteten Endpunkte waren *Sicherheit der Behandlung*, *gesundheitsbezogene Lebensqualität* und *Zufriedenheit mit der Betreuung im Zentrum*. Diese drei Endpunkte wurden auch von mehr als 90% der Patienten der Kategorie „sehr wichtig“ zugeordnet. Die *Möglichkeit, den Beruf weiter auszuführen* wurde von 40% der Patienten als „überhaupt nicht wichtig“ eingestuft. Die *Ansprechbarkeit vom Pflegepersonal* wurde von 90% der Patienten als „sehr wichtig“ eingestuft und nur von 4% der Patienten als „überhaupt nicht wichtig“. Ähnliche Prozentzahlen ergaben sich auch für die Endpunkte *Individuelle Informationen* und *Lebensverlängerung*.

Bei der Subgruppenanalyse ergaben sich die auffälligsten Effekte für die Charakteristika Alter, Beruf und Bildung. Der Endpunkt *Möglichkeit, den Beruf weiter auszuüben* war vor allem für Patienten sehr wichtig, die jünger waren, noch berufstätig waren und einen höheren Bildungsstand aufwiesen. *Wenig eigene Mitarbeit bei der Behandlung* war für ältere Patienten, Patienten im Ruhestand und für Patienten mit einem höheren Bildungsstand sehr wichtig. Eine maximale *Lebensverlängerung* war sehr wichtig für jüngere Patienten. Dies galt ebenso für die Endpunkte *gesundheitsbezogene Lebensqualität* und *körperliche Leistungsfähigkeit*. Die *Möglichkeit, das Pflegepersonal zu wählen*, war insbesondere für Patienten mit einer niedrigen Bildung sehr wichtig und eher nicht wichtig für Patienten mit einem hohen Bildungsstand.

Zusammenfassend waren neben den intuitiv bedeutsamen Endpunkten *Mortalität* und *Morbidität* viele weitere Endpunkte, wie z. B. die *Ansprechbarkeit von Pflegepersonal* für die Patienten wichtig, die derzeit nicht standardmäßig in der Forschung, in Evaluationsstudien oder in Qualitätsprogrammen Berücksichtigung finden.

Manuskript 4: Vergleich der Präferenzen verschiedener Gruppen

Bei der Durchführung von klinischen Studien und HTA-Berichten ist eine Ermittlung von Präferenzen unter Heranziehung großer Patientenstichproben, wie sie im 3. Artikel beschrieben wurde, nicht jederzeit möglich. Deshalb wird häufig auf Präferenzen von Gruppen zurückgegriffen, die mit geringerem Aufwand zu ermitteln sind. Inwiefern die

Präferenzen dieser Gruppen mit denen einer repräsentativen Patientenstichprobe vergleichbar sind, wurde im Rahmen des vierten Manuskripts evaluiert.

Identifikation relevanter Gruppen

Die Referenzgruppe bestand aus der repräsentativen Patientenstichprobe der QiN-Befragung, die im 3. Artikel ausführlich beschrieben wurde. Drei weitere Gruppen wurden herangezogen, um ihre Präferenzen für die Endpunkte mit den Präferenzen der QiN-Stichprobe zu vergleichen: Patienten einer Selbsthilfegruppe, Nephrologen und HTA-Ersteller.

Die Patienten der Selbsthilfegruppe (Sektion Niederrhein, Interessengemeinschaft Niere e.V.) wurden auf verschiedenen Wegen kontaktiert. Zum einen wurde der Fragebogen auf einer Mitgliederversammlung verteilt, des Weiteren wurde die Geschäftsstelle mit Fragebögen, in Papierform und als online-Version (abrufbar über einen Weblink), versorgt.

Die Nephrologen wurden per E-Mail kontaktiert, die den Link zu der online-Version des Fragebogens enthielt. Dazu wurde mithilfe einer google-Internet-Suche nach Nephrologen in den jeweils größten Städten aller deutschen Bundesländer gesucht. Insgesamt 80 Nephrologen, die eine Webseite und eine E-Mail-Adresse besitzen, wurden angeschrieben. Des Weiteren wurde durch einen kooperierenden Nephrologen der Link zu dem Fragebogen auf fachspezifischen Tagungen innerhalb Deutschlands bekannt gegeben. Zusätzlich wurden alle kontaktierten Nephrologen darauf hingewiesen, dass sie den Link zum Fragebogen gerne an interessierte Kollegen weiterleiten könnten.

Die HTA-Ersteller wurden aus der Sachverständigen-Datenbank des IQWiG identifiziert oder über persönliche Kontakte der Autorin. Insgesamt wurden 25 HTA-Ersteller per E-Mail angeschrieben, die – wie bei den Nephrologen – den Link zu der online-Version des Fragebogens enthielt und eine Aufforderung, diesen Link an interessierte Kollegen weiterzuleiten.

Fragebogen

Es wurde der gleiche Fragebogen wie bei der QiN-Befragung verwendet. Lediglich die Eingangsfrage unterschied sich bei den Nephrologen und HTA-Erstellern zu der Version, die die Patienten erhielten. Die Experten wurden aufgefordert, die Endpunkte zu bewerten, so wie sie denken, dass diese für die Patienten von Bedeutung sind. Patienten wurden nach ihrer persönlichen Einschätzung gefragt.

Datenanalyse

Die Gewichtung der Endpunkte wurde ebenfalls mithilfe von deskriptiver Statistik beschrieben (Mittel- und Medianwerte, Standardabweichungen). Anhand der Mittelwerte der einzelnen Endpunkte wurde eine Hierarchisierung vorgenommen. Die Ergebnisse der Befragung der repräsentativen Patientenstichprobe wurden als Referenzwerte verwendet, mit denen die drei übrigen Gruppen auf unterschiedliche Weise verglichen wurden.

In der ersten Analyse wurden die sieben Endpunkte jeder Gruppe verglichen, die die höchste Präferenz bekamen. In HTA-Berichten wird nur eine begrenzte Anzahl Endpunkte betrachtet. Die Cochrane Collaboration verwendet sieben Endpunkte, daher wurde diese Anzahl für die Auswertung verwendet.[19]

In einer zweiten Analyse wurde die Abweichung der Gewichtung aller 23 Endpunkte jeder Gruppe mit der Referenzgruppe verglichen. Des Weiteren wurden die sieben wichtigsten Endpunkte der drei Vergleichsgruppen (Patienten der Selbsthilfe, Nephrologen und HTA-Ersteller) untereinander verglichen, ohne Berücksichtigung der Referenzgruppe.

Ergebnisse

Die Patienten der Selbsthilfe waren jünger, bekamen schon längere Zeit eine Dialyse und hatten weniger häufig Diabetes im Vergleich zu der repräsentativen Patientenstichprobe. Die HTA-Ersteller hatten im Durchschnitt 6 Jahre Berufserfahrung, die Nephrologen behandelten im Durchschnitt seit 20 Jahren Patienten mit terminaler Niereninsuffizienz und waren im Durchschnitt für 40-150 Patienten zuständig.

Drei Endpunkte waren bei allen vier Gruppen unter den sieben Endpunkten mit den jeweils höchsten Präferenzen: *Sicherheit der Behandlung*, *gesundheitsbezogene Lebensqualität* und *emotionaler Gemütszustand*. Die Patienten der Selbsthilfegruppe hatten insgesamt fünf Endpunkte mit der Referenzgruppe gemeinsam bei der Einschätzung der wichtigsten sieben Endpunkte. Die Nephrologen und die HTA-Ersteller hatten jeweils vier Endpunkte mit der Referenzgruppe gemeinsam.

Beim Vergleich aller Endpunkte zwischen den Gruppen hatten die Patienten der Selbsthilfe 18 Endpunkte nur mit sehr kleinen oder kleinen Abweichungen im Vergleich zur Referenzgruppe bewertet. Die Nephrologen bewerteten 17 Endpunkte mit sehr kleinen oder kleinen Abweichungen und die HTA-Ersteller 13 Endpunkte.

Zusammenfassend unterscheiden sich die Präferenzen der untersuchten Gruppen in unterschiedlichem Ausmaß von den Präferenzen einer großen Referenzgruppe an Patienten

mit terminaler Niereninsuffizienz. Die Präferenzen aller Gruppen sollte die Grundlage für einen transparenten Ansatz bilden, um eine valide Liste der wichtigsten Endpunkte zu entwickeln.

Diskussion

Bislang existiert kein einheitliches Vorgehen bei der Priorisierung von Endpunkten bei der Evaluation von medizinischen Interventionen durch HTA-Berichte. Es fehlen Standards dafür, welche Methoden eingesetzt werden sollen und welche Personengruppen die Priorisierung durchführen sollen.

HTA-Organisationen thematisieren die Priorisierung von Endpunkten bereits häufig bei ihrem methodischen Vorgehen.[19,50,51] Dennoch konnte sich noch kein Standard durchsetzen bzw. wurde noch nicht ausreichend definiert. Ein wesentlicher Punkt, der noch wenig thematisiert wird, ist die Auswahl der Personengruppe, die an dem Prozess beteiligt werden soll. Häufig wird die Einbindung von Patienten in diesem Kontext als optionale Möglichkeit aufgeführt. Eine regelhafte Beteiligung ist allerdings bislang noch nicht eingeführt worden.[50,51]

Mit der vorliegenden Dissertation wurde am Beispiel der terminalen Niereninsuffizienz eine Methodik entwickelt, mit der patientenrelevante Endpunkte in HTA-Berichten gewichtet werden können. Zusätzlich zu der Identifikation einer geeigneten Methode wurde der Frage nachgegangen, ob Patienten in diesem Prozess einbezogen werden können. Für diese Dissertation wurden vier Fragestellungen definiert, die diesen Prozess strukturieren:

- 1) Welche Bedeutung hat eine Patientenbeteiligung bei der Priorisierung von patientenrelevanten Endpunkten?
- 2) Welche Methoden zur Priorisierung von Endpunkten gibt es und welche Vor- und Nachteile haben diese Priorisierungsmethoden?
- 3) Welche Endpunkte erscheinen Patienten mit terminaler Niereninsuffizienz relevant und welche Priorität haben diese Endpunkte?
- 4) Gibt es einen Unterschied zwischen einer repräsentativen Patientengruppe und anderen Gruppen?

Im Folgenden werden das methodische Vorgehen und die erarbeiteten Ergebnisse für jede Fragestellung diskutiert.

Methodisches Vorgehen und Ergebnisse

Zur Beantwortung der ersten Frage nach der Bedeutung einer Priorisierung von patientenrelevanten Endpunkten wurde eine theoretische Aufarbeitung der Patientenpartizipation im deutschen Gesundheitswesen erstellt (Manuskript 1). Vorhandene Strukturen und Konzepte auf der Mikro- und Makroebene in Bezug auf die Festlegung und Gewichtung patientenrelevanter Endpunkte wurden dargestellt und anschließend mögliche Weiterentwicklungen diskutiert.

Im Manuskript 1 wurde die Gewichtung von patientenrelevanten Endpunkten bei der Nutzenbewertung als eine mögliche Weiterentwicklung bei der Evaluation von medizinischen Interventionen durch das IQWiG beschrieben. Zu dem Zeitpunkt der Manuskripterstellung und Entwicklung des Themas dieser Dissertation war dieses Vorgehen noch nicht in den Methoden des IQWiG integriert. In den allgemeinen Methoden des Instituts wurde 2011 zum ersten Mal in der Version 4.0 die Gewichtung von Endpunkten thematisiert.[49] Es wurden zwei Methoden, der Analytic Hierarchy Process (AHP) und die Conjoint Analyse, beispielhaft aufgeführt, die hierzu verwendet werden können. Auch in der aktuellen (vorläufigen) Version 5.0 der allgemeinen Methoden des IQWiG werden diese beiden Verfahren weiterhin als geeignet beschrieben.[50] Allerdings wird vom Institut darauf hingewiesen, dass eine routinemäßige Anwendung dieser Methoden nicht vorgesehen ist. Als Begründung werden derzeit noch ungelöste methodische Probleme beider Verfahren aufgeführt.[52] Eine regelhafte Einbeziehung von Patienten bei der Hierarchisierung von Endpunkten wird – neben der Beschreibung zur Problematik, eine geeignete Methodik hierfür zu definieren – nicht weiter thematisiert oder gefordert.

Zur Beantwortung der zweiten Fragestellung bezüglich möglicher Methoden für eine Gewichtung von Endpunkten wurde eine systematische Suche durchgeführt und eine Übersichtsarbeit erstellt (Manuskript 2).

Die Ergebnisse zeigen, dass sich für die Priorisierung von Endpunkten grundsätzlich eine relativ große Anzahl an Methoden eignet.[29] Die in den Studien angewendeten Methoden wurden drei Gruppen zugeordnet:

- 1) Rating oder Ranking Skalen

Es wurden einfache Rating Skalen sowie Ranking Skalen verwendet.

- 2) Methoden der multikriteriellen Entscheidungsfindung

Die am häufigsten verwendeten Methoden waren die Conjoint Analyse Methoden und der Analytic Hierarchy Process.

3) Utility Assessment Methoden

Es wurden die „Standard Gamble“, „Time-Trade-off“ und „Willingness-to-pay“ Methoden verwendet, teilweise mit leichten Abänderungen der Konzepte, so wurde auch ein „Sleep-Trade-off“ oder ein „Willingness-to-travel“ Instrument eingesetzt.

Mit Rating oder Ranking Methoden kann aufgrund ihrer geringen Komplexität eine große Anzahl Endpunkte bewertet werden. Das größte Problem dieser Methoden ist die Variabilität der Priorisierungen zwischen den Teilnehmern und ein häufig vorhandener Deckeneffekt. Diese Methoden ermöglichen es nicht, einen relativen Präferenzwert zwischen den einzelnen Endpunkten zu generieren. [29]

Methoden der multikriteriellen Entscheidungsfindung zeichnen insbesondere die Möglichkeit aus, für mehrere Endpunkte den relativen Präferenzwert zu erheben. Dies folgt einer ökonomischen Überlegung, dass Konsumenten bei der Entscheidung für ein Konsumprodukt mehrere Attribute gegeneinander abwägen, um sich für das für sie passendste zu entscheiden. Die Abwägung zwischen einzelnen Endpunkten sichtbar zu machen, ist eine Stärke dieser Methoden. Die Komplexität ist allerdings auch die Schwäche dieser Methoden.[53] Obwohl häufig nur wenige Endpunkte bewertet wurden, hatten viele Studienteilnehmer Probleme bei der vollständigen Durchführung der Priorisierung, was sich in einer unvollständig beantworteten Rückmeldung äußerte. Ein weiteres Problem dieser Methoden ist, dass Teilnehmer sich bei zunehmender Komplexität – also Abwägung zwischen einer großen Anzahl von Endpunkten – häufig einer Vereinfachungsregel bedienen. Dies bedeutet konkret, dass sich die Teilnehmer auf eine sehr kleine Anzahl Endpunkte konzentrierten und die übrigen Endpunkte in der Priorisierung vernachlässigten.[54]

Die unterschiedlichen Utility Assessment Methoden können ähnlich wie die Methoden der multikriteriellen Entscheidungsfindung ebenfalls einen relativen Präferenzwert für mehrere Attribute erheben. Der hypothetische Charakter dieser Methoden ist allerdings ebenfalls häufig ein Nachteil dieser Methoden. [54]

Bislang hat sich noch keine der angewendeten Methoden endgültig durchgesetzt, um als Standard für die Priorisierung von Endpunkten eingesetzt zu werden. Dies ist vermutlich darauf zurückzuführen, dass jede Methode – wie beschrieben – sowohl Vor- als auch Nachteile aufweist.

Um die dritte Fragestellung zu beantworten, wurde eine empirische Datenerhebung durchgeführt. Hierzu wurde in einem ersten Schritt eine Gruppendiskussion mit Patienten mit

terminaler Niereninsuffizienz während eines regulären Selbsthilfegruppentreffens durchgeführt, um relevante Endpunkte zu erheben.

Insgesamt wurden 23 Endpunkte identifiziert; neben solchen Endpunkten, die häufig bei der Evaluation von medizinischen Interventionen verwendet werden wie *Mortalität*, *gesundheitsbezogene Lebensqualität* verschiedene Morbiditätsendpunkte (zum Beispiel: *Krankenhausaufenthalte*, *körperliche Funktionsfähigkeit*) oder Nebenwirkung einer Behandlung (zum Beispiel *Blutdruckabfälle und Übelkeit während der Behandlung*), wurden auch Endpunkte aufgeführt, die den interventions- und erkrankungsbezogenen Aufwand und die Patientenzufriedenheit darstellen. Hierunter fallen zum Beispiel Endpunkte wie: *Möglichkeit den Beruf weiter auszuüben*, *Fahrzeit zur Dialyse*, *Individuell angepasste Dialysedauer*, *Zufriedenheit mit der Betreuung im Zentrum* oder *Ansprechbarkeit von Pflegekräften oder Ärzten während der Behandlung*.

In einem zweiten Schritt wurden diese Endpunkte in einem Fragebogen aufgenommen, der an Hämodialyse-Patienten verteilt wurde. Aus dem Methodenvergleich (Manuskript 2) wurde deutlich, dass eine Rating oder Ranking Skala für die Gewichtung einer großen Anzahl Endpunkte geeignet ist. Aufgrund der großen Anzahl zu gewichtender Endpunkte wurde deshalb eine Rating-Skala für die Fragebogenerhebung herangezogen.

Der Nachteil dieser Methode ist allerdings das Auftreten eines Deckeneffekts. In der vorliegenden Arbeit vergaben die Patienten häufig hohe Bewertungen für viele der 23 Endpunkte. Die Mittelwerte der einzelnen Endpunkte lagen dicht beieinander mit überlappenden Konfidenzintervallen. Dieser Effekt muss bei der Interpretation der Rangfolgen berücksichtigt werden. Allerdings zeigt sich hierdurch auch, dass viele der Endpunkte für einen Großteil der Befragten relevant sind. Die drei Endpunkte mit dem höchsten Mittelwert waren *Sicherheit der Behandlung*, *Verbesserung der gesundheitsbezogenen Lebensqualität* und *Zufriedenheit mit der Betreuung im Zentrum*. Die Gewichtung für einige Endpunkte unterschied sich in einigen Subgruppen, insgesamt waren aber die meisten Endpunkte in allen Subgruppen stabil.

Eine Rücklaufquote war im Rahmen dieser Datenerhebung aus organisatorischen Gründen nicht zu ermitteln, sodass keine Angaben möglich sind, wie viele Patienten die Rating Methode problematisch fanden oder diese nicht vollständig abschließen konnten. Grundsätzlich wird eine Rating Skala aber von Patienten als leicht verständlich aufgenommen.[54]

Der Anteil älterer Patienten, die sich bereits im Ruhestand befanden, war in der repräsentativen Patientengruppe sehr groß, die Interpretation der Priorisierung sollte diese

Tatsache berücksichtigen. So wurde der Endpunkt *Möglichkeit den Beruf weiter auszuüben* von der Subgruppe der jüngeren Patienten als sehr wichtig eingestuft, von älteren Patienten als eher unwichtig.

Für die Beantwortung der vierten Fragestellung, ob sich die Priorisierung durch eine repräsentative Patientengruppe von der Priorisierung durch andere Gruppen unterscheidet, wurden drei weitere Personengruppen aufgefordert, die Priorisierung durchzuführen. Neben Patienten einer Selbsthilfegruppe wurden Ersteller von HTA-Berichten und Ärzte (hier im Speziellen Nephrologen) befragt, da diese Gruppen häufig bei der Priorisierung von Endpunkten im Kontext einer HTA-Berichterstellung herangezogen werden.

Die Gewichtung der Endpunkte unterschied sich geringfügig zwischen der repräsentativen Auswahl von Patienten und den Patienten, die einer Selbsthilfegruppe angehörten (Manuskript 4). Insgesamt wurden 18 der 23 Endpunkte mit einem sehr kleinen oder kleinen Unterschied in der Rangfolge eingestuft. Die Gewichtungen zwischen Ärzten und Patienten unterschied sich ebenfalls eher gering (17 von 23 Endpunkten).

Die Gewichtungen unterschieden sich deutlicher zwischen HTA-Experten und Patienten. HTA-Experten hatten nur bei 13 von 23 Endpunkten einen sehr kleinen oder kleinen Unterschied in der Rangfolge im Vergleich zu der repräsentativen Patientengruppe. Sowohl die Nephrologen als auch die HTA-Ersteller stimmten in vier Endpunkten mit der Patientengruppe überein, wenn nur die wichtigsten sieben Endpunkte betrachtet wurden.

Obwohl die Bewertung der Patienten der Selbsthilfegruppe mehr Übereinstimmungen mit der Referenzgruppe zeigte als die Bewertungen der Ärzte oder HTA-Experten, lagen auch hier Unterschiede vor. Diese Unterschiede können durch eine Stichprobenverzerrung hervorgerufen werden, die durch eine nicht randomisierte Gruppenzusammenstellung verursacht werden kann.[55-57] In der vorliegenden Untersuchung waren die Patienten der Selbsthilfegruppe z. B. deutlich jünger als die Patienten der QiN-Befragung. Eine explorative Betrachtung jüngerer QiN-Patienten im Vergleich zu den Selbsthilfepatienten erhöhte die Übereinstimmung allerdings nicht merklich. Andere Faktoren, die wesentlich schwieriger zu benennen sind, könnten somit einen Einfluss auf die Bewertung der Patienten der Selbsthilfe gehabt haben.

Methodische Überlegungen zur Präferenzmessung

Ein wichtiges, bislang ungelöstes Problem bei der Verwendung von Methoden zur Priorisierung von Endpunkten ist der fehlende Goldstandard für die Messung von Präferenzen. Unabhängig von der verwendeten Methode – ob Rating Skala so wie in der vorliegenden Dissertation, oder aber auch einer komplexeren multikriteriellen Methode – gibt es bislang keine Möglichkeit, abschließend zu beweisen, dass die festgestellte Priorisierung tatsächlich valide ist. Ein Goldstandard für die Präferenzen einzelner Endpunkte, mit dem man die Messmethoden vergleichen könnte, existiert nicht. Mögliche Kriterien für eine Validitätsprüfung in diesem Kontext sind: Plausibilität, interne Validität, prädiktive Validität und Konvergenzvalidität.

In der vorliegenden deskriptiven Übersicht wurden die Angaben zur Validitätsprüfung der eingeschlossenen Studien zusammengetragen. Allerdings wurde der Versuch, Validitätskriterien anzuwenden, um die Ergebnisse zu überprüfen, nur von einer sehr geringen Anzahl der Studien beschrieben.[54] Häufig wurde das Kriterium der Konvergenzvalidität angewendet und die Priorisierungsergebnisse unterschiedlicher Verfahren miteinander verglichen.

Allerdings liefert die Überprüfung der Konvergenz nur bei identischen Ergebnissen der verglichenen Methoden eine mögliche Aussage zur Validität, wobei auch in dieser Situation nicht von einem Beweis ausgegangen werden kann. Denn auch, wenn zwei Methoden konkordante Ergebnisse zeigen, können diese Ergebnisse sich diskordant zu den wahren Priorisierungsergebnissen verhalten, die mit einem hypothetischen Goldstandard festgehalten werden könnten. Diskordante Ergebnisse zwischen zwei Methoden erschweren ebenfalls eine Interpretation darüber, welche Methode valider ist oder ob das wahre Priorisierungsergebnis gemessen mit einem hypothetischen Goldstandard von den Ergebnissen beider Methoden abweicht.

Im Rahmen einer Plausibilitätsprüfung wurden in einigen Studien vorab erstellte Hypothesen über die Wichtigkeit von Endpunkten mit den erhaltenen Ergebnissen der Präferenzmessung verglichen. In anderen Studien wurde die test-retest-Korrelation untersucht – die Reproduzierbarkeit eines Ergebnisses bei einer Wiederholung des Testes -, um Aussagen zur Konstruktvalidität zu machen. Allerdings ist auch eine hohe Korrelation ein Garant dafür, dass die gemessenen Ergebnisse die tatsächlichen Präferenzen der Teilnehmer valide wiedergeben.[54] Insbesondere im Bereich der multikriteriellen Methoden wird diese Problematik thematisiert und darauf hingewiesen, dass eine Interpretation der Ergebnisse

durch die fehlende Validitätsprüfung schwierig ist.[58-61] Als einen ersten Schritt wurde eine Leitlinie erstellt, die eine Standardisierung der Durchführung von Präferenzhebungen regeln soll.[62]

Unterschiedliche Präferenzen zwischen zwei Gruppen, zum Beispiel Patienten und Ärzten, können darauf hinweisen, dass die Präferenzen aufgrund von unterschiedlichen Sichtweisen auf die Behandlungssituation tatsächlich unterschiedlich sind. Allerdings kann nicht ausgeschlossen werden, dass die verwendeten Methoden die Präferenzen in beiden Gruppen unterschiedlich valide erheben können. Da es, aufgrund eines fehlenden Goldstandards, bislang keine Möglichkeit gibt, die Validität der Ergebnisse abschließend zu bewerten, ist es essenziell, unterschiedliche Präferenzen einzelner Gruppen transparent darzustellen, da nur durch diese Transparenz eine ausgewogene Diskussion über unterschiedliche Präferenzen im Behandlungsprozess erfolgen kann.[54,63]

Ein weiterer wichtiger Aspekt neben der Auswahl der Gruppen, deren Präferenzen erhoben werden sollen, ist die Übertragbarkeit. In der vorliegenden Dissertation wurden Präferenzen von Hämodialyse-Patienten mit terminaler Niereninsuffizienz erhoben. Obwohl die repräsentative Stichprobe aus über 4500 Patienten bestand, können die gefundenen Ergebnisse nicht auf andere Patientengruppen, wie zum Beispiel Patienten mit terminaler Niereninsuffizienz, die mit einer Peritonealdialyse behandelt werden oder eine Nierentransplantation hatten, übertragen werden. Eine Übertragung der Ergebnisse auf Patienten mit anderen chronischen Erkrankungen erscheint noch schwieriger und sollte daher in weiteren Studien in anderen Indikationen überprüft werden.

Die vorliegende Dissertation hat an einer konkreten Behandlungssituation die Präferenzen erhoben und damit aufgezeigt, dass eine Priorisierung von Endpunkten mithilfe einer Rating Skala durchführbar ist. Ob diese Methode auch bei anderen Erkrankungen und anderen Behandlungen ebenfalls geeignet ist, muss in weiteren Studien überprüft werden.

Weitere Untersuchungen sind ebenfalls notwendig, um eine Aussage darüber treffen zu können, ob auch bei anderen Erkrankungen die Priorisierung von Endpunkten durch Patienten aus einer Selbsthilfegruppe mit der Priorisierung einer repräsentativen Patientengruppe zu einem großen Teil übereinstimmt, oder ob es sich dort anders verhält. Auch die in dieser Arbeit identifizierten Diskrepanzen zwischen den Präferenzen der Experten und der Patienten können erst einmal nur für die hier untersuchte Indikation gelten. Ob sich ähnliche Muster auch in anderen Indikationen zeigen, muss weiter untersucht werden.

Bezug zu anderer Forschung

Die Einbeziehung von Patienten rückt in vielen Forschungsbereichen immer stärker in den Fokus der Aufmerksamkeit.[9,64,65] Viele verschiedene Ansätze sind vorhanden, die sich entlang eines konsekutiven Forschungsstrangs bewegen, angefangen bei der Planung von Studien, der Durchführung, der Evaluation bis hin zu der Implementierung.[66-69] Im Folgenden soll für einige dieser Schritte beispielhaft gezeigt werden, wie eine Einbeziehung von Patienten erfolgen kann und welche Schritte momentan unternommen werden.

Festlegung relevanter Fragestellungen mit Einbezug von Patienten

Da Patienten diejenigen sind, die sich in der Behandlungssituation befinden, in der medizinische Interventionen angewendet werden, ist die Einbeziehung von Patienten bei der Festlegung relevanter Fragestellungen nicht nur aus moralischen Gründen wichtig. Patienten können auch durch ihren speziellen Blickwinkel auf die Situation Fragestellungen und Themen benennen, die Forschern nicht unbedingt bewusst sein müssen.[9] Häufig werden Fragestellungen untersucht, die für Forscher interessant oder für Interessensgruppen wie der pharmazeutischen Industrie lukrativ sind, aber nicht den tatsächlichen Prioritäten der Patienten entsprechen.[9,70,71] Die Festlegung von wichtigen Forschungsthemen und Fragestellungen wird zum Beispiel von der James Lind Alliance als Kooperation zwischen Ärzten und Patienten durchgeführt. In Konsultationsgruppen werden die Präferenzen von beiden Gruppen gemeinsam erarbeitet und fließen in Empfehlungen ein, welche Fragestellungen bei einer speziellen Erkrankung die höchste Priorität haben.[70] Die Konsultationsgruppen bestehen aus Patienten, Angehörigen und Ärzten der entsprechenden Fachgruppe. Die Teilnehmer der Konsultationsgruppen führen Befragungen bei einer großen Gruppe Patienten, Angehörigen und Ärzten durch, um unbeantwortete Fragen oder Unsicherheiten bei der Behandlung zu identifizieren, die erforscht werden sollten. Die gesammelten Fragen werden dann durch die Konsultationsgruppe gesichtet und anhand einer Evidenzrecherche wird überprüft, ob die Fragen tatsächlich noch nicht ausreichend erforscht wurden. Die verbleibenden Fragestellungen werden durch die Mitglieder der Konsultationsgruppe priorisiert, wobei eine Liste der zehn dringendsten Fragestellungen erstellt wird, die der Forschungsgemeinschaft durch eine Publikation zur Verfügung gestellt wird.[70,71] Wie wichtig diese Initiative ist, wurde von Crowe et al. in einer Analyse dargestellt, bei der untersucht wurde, welchen Kategorien die durch die Konsultationsgruppen identifizierten Fragestellungen der Jahre 2003-2012 zugeordnet werden konnten. Die Kategorien waren 1) Arzneimittel, 2) Nichtmedikamentöse Verfahren (hier: Radiotherapie,

Operationen, Medizinprodukte und Diagnostik) und 3) sonstige Verfahren (z. B. Psycho-, Physio-, Bewegungs-, Ernährungs- oder alternative Therapien). In den Kategorien 1 und 2 konnten jeweils 20% der priorisierten Fragestellungen eingeordnet werden, 60% der Fragestellungen der Kategorie 3. Zusätzlich untersuchten Crowe et al., wie viele Studien in diesem Zeitraum registriert wurden und wie diese Studien den drei Kategorien zugeordnet werden konnten, aufgeteilt in Industrie-unabhängige Studien und Industrie-finanzierte Studien. Bei der Gegenüberstellung wird sehr schnell deutlich, dass es eine große Unverhältnismäßigkeit gibt: 85% der Industrie-finanzierten und 40% der Industrie-unabhängigen Studien befassten sich mit Fragestellungen zu Arzneimitteln (Kategorie 1), 10% der Industrie-finanzierten und 30% der Industrie-unabhängigen Studien befassten sich mit Fragestellungen der Kategorie 2. Nur 3% der Industrie-finanzierten und 30% der Industrie-unabhängigen Studien befassten sich mit Fragestellungen zu sonstigen Therapieverfahren. Hier hatten die Konsultationsgruppen aber erheblichen Forschungsbedarf identifiziert. [70] Im Sinne einer verantwortlichen Nutzung von Ressourcen, Zeit und Aufwand der Patienten und Ärzte, sollten sinnvolle Fragestellungen untersucht werden, die auch mit der Hilfe von Patienten identifiziert werden können. [72]

Auswahl relevanter Endpunkte

In einer systematischen Übersicht von Gargon et al. [73] wurde untersucht, in welchen Indikationen es bereits Bestrebungen gibt, ein sogenanntes „Core Outcome Set“, also eine Festlegung von Endpunkten, die in jeder Studie zu dieser Indikation berücksichtigt werden sollten, zu entwickeln, oder ob solche Sets sogar schon bestehen. Sie fanden heraus, dass, obwohl es diese Bestrebungen in vielen unterschiedlichen Indikationen gibt, nur in ca. 16% der Fälle Patienten oder Betroffene zu der Auswahl dieser Endpunkte hinzugezogen wurden. Eine stärkere Einbeziehung wird daher auch von den Autoren als wichtige Handlungsempfehlung formuliert. Die „Core Outcome Measures in Effectiveness Trials (COMET)“ Initiative hat sich zum Ziel gesetzt, die Entwicklung eines Kernsets relevanter Endpunkte für medizinische Indikationen zu fördern. Auch bei dieser Initiative liegt der Fokus auf der Einbeziehung der Präferenzen verschiedener Gruppen. So werden neben klinischen Experten auch Patienten oder die Allgemeinbevölkerung in dem Entwicklungsprozess einbezogen.[74]

In der Indikation Niereninsuffizienz wird dieser Prozess durch die SONG-HD Initiative (Standardized Outcomes in Nephrology – Haemodialysis) durchgeführt, bei der für die

Auswahl der relevanten Endpunkte Patienten und andere Gruppen, wie medizinische Experten, Pflegepersonal oder Angehörige, einbezogen werden.[75]

Im Rahmen dieser Initiative beschreibt die Arbeit von Evangelidis et al [76] ein Delphi-Verfahren, das als erster Schritt für die Entwicklung eines Kernsets relevanter Endpunkte fungierte. Bei diesem Verfahren wurden die Präferenzen von 1181 Teilnehmern (17% Patienten oder Angehörige, 83% medizinische Experten) in einem mehrstufigen Delphi-Verfahren zu relevanten Endpunkten bei der Hämodialyse erhoben. Die Auswahl der Endpunkte basierte auf einer systematischen Übersicht zu bereits verwendeten Endpunkten in randomisierten Studien, Interviews mit verschiedenen Gruppen und einer nominalen Gruppentechnik, die mit Hämodialyse-Patienten und ihren Angehörigen durchgeführt wurden.[75,77] Insgesamt wurden 34 Endpunkte in der ersten Runde des Delphi-Verfahrens von den Teilnehmern bewertet. Diese Bewertung wurde mit einer neunstufigen Rating Skala durchgeführt, die wie bei der GRADE-Gruppe in drei Kategorien unterteilt wurde („Weniger wichtige Endpunkte“, „wichtig, aber nicht kritisch“ und „besonders wichtig“).

Nur solche Endpunkte, die sowohl von den Patienten als auch von den medizinischen Experten als besonders wichtig eingestuft wurden, wurden in der zweiten Runde nochmals bewertet und ein weiteres Mal in einer dritten Runde durchgeführt. Das Ergebnis stellt somit eine Schnittmenge zwischen den Präferenzen der Patienten und der medizinischen Experten dar. Die Endpunkte, die von beiden Gruppen benannt wurden, waren: „Komplikationen beim Gefäßzugang“, „Dialyseeffektivität“, „Fatigue“, „Kardiovaskuläre Erkrankungen“ und „Mortalität“.

Die Autoren des Delphi-Verfahrens weisen allerdings ausdrücklich darauf hin, dass viele Endpunkte, die eher die gesundheitsbezogene Lebensqualität oder das tägliche Leben mit einer Hämodialysebehandlung betreffen, von den Patienten als besonders wichtig eingestuft wurden, diese Einschätzung aber nicht von den medizinischen Experten geteilt wurde und diese Endpunkte somit nicht in das Kernset aufgenommen wurden. Unter Berücksichtigung dieser zusätzlichen Information erscheint die Endpunkt-Gewichtung der Patienten bei dem Delphi-Verfahren ähnliche Ergebnisse zu liefern wie in der vorliegenden empirischen Datenerhebung. Allerdings erschwert das beschriebene Vorgehen aufgrund des Schnittmengenansatzes eine Darstellung von Diskrepanzen bei der Einschätzung von wichtigen Endpunkten der unterschiedlichen Gruppen.

Einbeziehung von Patientenpräferenzen bei der Gewichtung von Endpunkten

Bei der Erstellung von systematischen Übersichten ist die Gewichtung von Endpunkten insbesondere dann wichtig, wenn ein sogenannter Zielkonflikt vorliegt, also eine medizinische Intervention Vorteile in einigen Endpunkten zeigt, allerdings auch Nachteile in anderen. In so einer Situation kann eine Gewichtung von Endpunkten helfen, die Ergebnisse besser zu interpretieren. Wie bereits zuvor dargestellt, beschreibt das IQWiG in seinen Methoden, dass eine Gewichtung von Endpunkten eine Option ist, die mithilfe von zum Beispiel einer Conjoint Analyse oder einer AHP durchgeführt werden kann. Allerdings verweist das Institut in seinen Methoden auch auf bestehende methodische Probleme, wie z. B. die Validität der zur Verfügung stehenden Methoden, die einen regelhaften Einsatz derzeit nicht möglich machen.[50]

Implikationen für die Forschung

Die Wahl der relevanten Endpunkte hat einen Einfluss auf die Beurteilung der Wirksamkeit in klinischen Studien und dadurch auch mittelbar bei der Beurteilung von medizinischen Interventionen in HTA-Berichten.[78] Die Einbeziehung von Patientenpräferenzen, aber auch die transparente Darstellung der Präferenzen von anderen Interessengruppen, wie Ärzten oder HTA-Experten, kann den bislang häufig impliziten Prozess der Gewichtung von Endpunkten bei der Interpretation von Ergebnissen nachvollziehbarer machen. Obwohl die Einbeziehung von Patientenpräferenzen stärker in den Fokus gerückt ist, fehlt häufig noch ein standardisiertes Verfahren.[79]

Es bleibt bislang offen, wie Patientenpräferenzen bei der Gewichtung von Endpunkten in einer validen und praktikablen Weise erhoben werden können, um in HTA-Berichten verwendet zu werden.[80] Die vorliegende Dissertation gibt Hinweise darauf, dass die Präferenzen von Patienten aus einer Selbsthilfegruppe eine annehmbare Annäherung an Präferenzen einer repräsentativen Patientenstichprobe sein können. Ob sich diese Hinweise auch in anderen Indikationen als der terminalen Niereninsuffizienz zeigen, muss in weiteren Studien geklärt werden.

Der Umgang mit divergierenden Präferenzen unterschiedlicher Gruppen, wie zum Beispiel Ärzten, HTA-Experten und Patienten, ist bislang auch noch nicht abschließend in Handlungsempfehlungen eingegangen. Hier liegen weitere relevante Forschungsfragen, deren Klärung notwendig ist, um ein standardisiertes Vorgehen erarbeiten zu können.

Ausblick

Die Erkenntnisse dieser Dissertation können sowohl für die Planung klinischer Studien und die Evaluation von medizinischen Interventionen als auch für die Qualitätsüberprüfung oder Patientenberatung verwendet werden. Die Schlussfolgerung dieser Dissertation bezieht sich letztendlich allerdings auf die Anwendbarkeit der Methodik im Rahmen eines HTA-Berichts und für diese Situation werden auch die abschließenden Empfehlungen formuliert.

Bei der Evaluation medizinischer Interventionen werden in der Regel Endpunkte herangezogen, um Nutzen und Schaden zu bewerten. Die Frage, mittels welcher Endpunkte die Vor- und Nachteile operationalisiert werden beziehungsweise werden sollten, ist Gegenstand vieler Diskussionen. Im Idealfall werden in den Prozess der Auswahl und Gewichtung von Endpunkten die Präferenzen repräsentativer Patientenstichproben einbezogen. Dies ist aufgrund des damit verbundenen Aufwands häufig nicht realisierbar, weshalb mitunter auf Präferenzen von anderen Gruppen zurückgegriffen wird.

Den dargestellten Ergebnissen ist zu entnehmen, dass diese Herangehensweise ein probates Mittel sein kann und Patienten einer Selbsthilfegruppe als geeignete Proxys für eine repräsentative Patientenstichprobe eingesetzt werden können. Notwendig hierfür ist eine hinreichende Strukturgleichheit bezüglich wichtiger Baselinecharakteristika zwischen beiden Patientengruppen.

Die Forderung nach einer stärkeren Einbeziehung von Patientenpräferenzen bei der Auswahl und Gewichtung von Endpunkten impliziert nicht, dass alleinig die Präferenzen der Patienten bei der Evaluation von medizinischen Interventionen entscheidend sind. Vielmehr zeigen die Ergebnisse der vorliegenden Arbeit, dass jede Interessensgruppe bei einer Bewertung einen wichtigen Beitrag leisten kann. So haben Patienten einen Überblick über die Auswirkungen einer Intervention auf ihren gesamten Lebensbereich, sie können daher eine Vielzahl Endpunkte benennen und als wichtig einstufen, die anderen Gruppen nicht bewusst sind. Obwohl in der vorliegenden Dissertation nur patientenrelevante Endpunkte untersucht wurden und Ärzte und HTA-Ersteller gegebenenfalls weitere Endpunkte hinzugefügt hätten, wurden die vorhandenen Endpunkte von den Expertengruppen basierend auf ihrer eigenen Expertise eingeschätzt. Hierbei zeigte sich, dass diesen beiden Gruppen andere Endpunkte wichtig waren als den Patienten. Denn Ärzte können klinisch relevante Endpunkte benennen, deren Bedeutung den Patienten nicht zwangsläufig ersichtlich sein müssen. HTA-Ersteller sind aufgrund ihrer methodischen Expertise in der Lage einzuschätzen, welche Endpunkte in

klinischen Studien erhoben werden können, um diese im Rahmen einer Evaluation im Rahmen eines HTA-Berichts zu verwenden.

Deswegen sollten alle drei Interessengruppen fester Bestandteil im Prozess der Auswahl und Gewichtung von Endpunkten in HTA-Berichten sein. Der Prozess der Auswahl und Gewichtung von Endpunkten in HTA-Berichten unter Einbeziehung der Präferenzen von Patienten, Ärzten und HTA-Erstellern sollte transparent gestaltet werden, damit nachvollziehbar ist, welche Präferenzen bei der Bewertung in welchem Maße eingeflossen sind.

Aus den Erkenntnissen dieser Arbeit resultiert noch kein standardisiertes Vorgehen für die Einbeziehung von Präferenzen von Patienten und anderen Gruppen bei der Auswahl und Gewichtung von Endpunkten, sie ist vielmehr als ein erster wichtiger Schritt zu betrachten. Für die Entwicklung eines standardisierten Vorgehens ist weitere Forschung notwendig. Die in der vorliegenden Arbeit gefundene hohe Übereinstimmung zwischen einer repräsentativen Patientenstichprobe und Patienten einer Selbsthilfegruppe sollte in weiteren Indikationen überprüft werden. Auch die Übereinstimmung anderer Gruppen (Ärzte und HTA-Ersteller) in anderen Indikationen sollte weiter untersucht werden.

Falls diese Überprüfungen ähnliche Übereinstimmungsmuster zeigen wie die in der vorliegenden Arbeit, könnte in einem weiteren Schritt eine generalisierbare Vorgehensweise erarbeitet werden, die den Prozess der Einbeziehung aller drei Gruppen standardisiert.

Referenzen

1. Bürgerliches Gesetzbuch (BGB), §630c (Gesetz zur Verbesserung der Rechte von Patientinnen und Patienten) [online]. [Zugriff: 29.05.2017]. URL: https://www.gesetze-im-internet.de/bgb/_630c.html.
2. Department of Health. NHS Constitution for England [online]. 2015. [Zugriff: 26.05.2017]. URL: <https://www.gov.uk/government/publications/the-nhs-constitution-for-england>.
3. Facey K, Boivin A, Gracia J, Hansen HP, Lo Scalzo A, Mossman J et al. Patients' perspectives in health technology assessment: a route to robust evidence and fair deliberation. *Int J Technol Assess Health Care* 2010; 26(3): 334-340.
4. Facey KM, Hansen HP. Patient-focused HTAs. *Int J Technol Assess Health Care* 2011; 27(4): 273-274.
5. Domecq JP, Prutsky G, Elraiyah T, Wang Z, Nabhan M, Shippee N et al. Patient engagement in research: a systematic review. *BMC Health Serv Res* 2014; 14: 89.
6. Muhlbacher AC, Juhnke C. [Involving patients, the insured and the general public in healthcare decision making]. *Z Evid Fortbild Qual Gesundheitswes* 2016; 110-111: 36-44.
7. Institute of Medicine (U.S.). Committee on Quality of Health Care in America. Crossing the quality chasm : a new health system for the 21st century. Washington, D.C.: National Academy Press; 2001.
8. Ryan M, Kinghorn P, Entwistle VA, Francis JJ. Valuing patients' experiences of healthcare processes: towards broader applications of existing methods. *Soc Sci Med* 2014; 106: 194-203.
9. Crowe S, Giles C. Making patient relevant clinical research a reality. *BMJ* 2016; 355: i6627.
10. Office of Technology Assessment. Strategies for Medical Technology Assessment. Washington DC: Office of Technology Assessment; 1982.
11. European network for Health Technology Assessment. Common Questions. What is Health Technology Assessment (HTA) [online]. [Zugriff: 15.03.2017]. URL: <http://www.eunetha.eu/about-us/faq#t287n73>.
12. DasMahapatra P, Raja P, Gilbert J, Wicks P. Clinical trials from the patient perspective: survey in an online patient community. *BMC Health Serv Res* 2017; 17(1): 166.
13. Williamson P, Clarke M. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative: Its Role in Improving Cochrane Reviews. *Cochrane Database Syst Rev* 2012; (5): ED000041.
14. Bederman SS, Mahomed NN, Kreder HJ, McIsaac WJ, Coyte PC, Wright JG. In the eye of the beholder: preferences of patients, family physicians, and surgeons for lumbar spinal surgery. *Spine (Phila Pa 1976)* 2010; 35(1): 108-115.
15. Chancellor J, Martin M, Liedgens H, Baker MG, Muller-Schwefe GH. Stated preferences of physicians and chronic pain sufferers in the use of classic strong opioids. *Value Health* 2012; 15(1): 106-117.
16. de Bekker-Grob EW, Essink-Bot ML, Meerding WJ, Koes BW, Steyerberg EW. Preferences of GPs and patients for preventive osteoporosis drug treatment: a discrete-choice experiment. *Pharmacoeconomics* 2009; 27(3): 211-219.
17. Faggioli G, Scalone L, Mantovani LG, Borghetti F, Stella A, group Ps. Preferences of patients, their family caregivers and vascular surgeons in the choice of abdominal aortic aneurysms treatment options: the PREFER study. *Eur J Vasc Endovasc Surg* 2011; 42(1): 26-34.
18. Thrumurthy SG, Morris JJ, Mughal MM, Ward JB. Discrete-choice preference comparison between patients and doctors for the surgical management of oesophagogastric cancer. *Br J Surg* 2011; 98(8): 1124-1131; discussion 1132.
19. In: Higgins J, Green S (Ed). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. 2011. URL: www.cochrane-handbook.org.
20. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.2. Köln: IQWiG; 2015. URL: https://www.iqwig.de/download/IQWiG_Methoden_Version_4-2.pdf.
21. Brett J, Staniszewska S, Mockford C, Seers K, Herron-Marx S, Bayliss H. The PIRICOM study: a systematic review of the conceptualisation, measurement, impact and outcomes of patients and

- public involvement in health and social care research. University of Warwick [online] 2010. URL: <http://www.ukcrc.org/wp-content/uploads/2014/03/Piricom+Review+Final+2010.pdf>.
22. Robinson A. Patient and public involvement: in theory and in practice. *J Laryngol Otol* 2014; 1-8.
 23. Boote J, Telford R, Cooper C. Consumer involvement in health research: a review and research agenda. *Health Policy* 2002; 61(2): 213-236.
 24. Oliver SR, Rees RW, Clarke-Jones L, Milne R, Oakley AR, Gabbay J et al. A multidimensional conceptual framework for analysing public involvement in health services research. *Health Expect* 2008; 11(1): 72-84.
 25. Bagley HJ, Short H, Harman NL, Hickey HR, Gamble CL, Woolfall K et al. A patient and public involvement (PPI) toolkit for meaningful and flexible involvement in clinical trials – a work in progress. *Research Involvement and Engagement* 2016; 2(1): 15.
 26. Kreis J, Puhan MA, Schunemann HJ, Dickersin K. Consumer involvement in systematic reviews of comparative effectiveness research. *Health Expect* 2013; 16(4): 323-337.
 27. Gagnon MP, Lepage-Savary D, Gagnon J, St-Pierre M, Simard C, Rhainds M et al. Introducing patient perspective in health technology assessment at the local level. *BMC health services research* 2009; 9: 54.
 28. Esmail L, Moore E, Rein A. Evaluating patient and stakeholder engagement in research: moving from theory to practice. *J Comp Eff Res* 2015; 4(2): 133-145.
 29. Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen ER, Russell EM et al. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technol Assess* 2001; 5(5): 1-186.
 30. Phillips KA, Johnson FR, Maddala T. Measuring what people value: a comparison of "attitude" and "preference" surveys. *Health Services Research* 2002; 37(6): 1659-1679.
 31. Levey AS, Atkins R, Coresh J, Cohen EP, Colins AJ, Eckardt K-U et al. Chronic kidney disease as a global public health problem: Approaches and initiatives – a position statement from Kidney Disease Improving Global Outcomes Kidney International 2007; 72: 247–259.
 32. National Kidney and Urologic Diseases Information Clearinghouse. *Kidney Disease Statistics for the United States*. Bethesda, USA: U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2012. URL: http://www.niddk.nih.gov/health-information/health-statistics/Documents/KU_Diseases_Stats_508.pdf.
 33. Abbasi MA, Chertow GM, Hall YN. End-stage renal disease. *Clin Evid (Online)* 2010; 2010.
 34. Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW et al. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med* 2003; 139(2): 137-147.
 35. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. *N Engl J Med* 2004; 351(13): 1296-1305.
 36. Eckardt KU, Coresh J, Devuyst O, Johnson RJ, Kottgen A, Levey AS et al. Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet* 2013; 382(9887): 158-169.
 37. Wong B, Collister D, Muneer M, Storie D, Courtney M, Lloyd A et al. In-Center Nocturnal Hemodialysis Versus Conventional Hemodialysis: A Systematic Review of the Evidence. *Am J Kidney Dis* 2017.
 38. Locatelli F. Dose of dialysis, convection and haemodialysis patients outcome--what the HEMO study doesn't tell us: the European viewpoint. *Nephrol Dial Transplant* 2003; 18(6): 1061-1065.
 39. Frei U, Schober-Halstenberg HJ. Nierenersatztherapie in Deutschland - Bericht über Dialysebehandlung und Nierentransplantation in Deutschland 2006/2007. *Quasi Niere*; 2008. URL: http://www.bundesverband-niere.de/fileadmin/user_upload/Quasi-Niere-Bericht_2006-2007.pdf.
 40. Palmer SC, de Berardis G, Craig JC, Tong A, Tonelli M, Pellegrini F et al. Patient satisfaction with in-centre haemodialysis care: an international survey. *BMJ Open* 2014; 4(5): e005020.
 41. Juergensen E, Wuertth D, Finkelstein SH, Juergensen PH, Bekui A, Finkelstein FO. Hemodialysis and peritoneal dialysis: patients' assessment of their satisfaction with therapy and the impact of the therapy on their lives. *Clin J Am Soc Nephrol* 2006; 1(6): 1191-1196.

42. Wasserfallen JB, Moinat M, Halabi G, Saudan P, Perneger T, Feldman HI et al. Satisfaction of patients on chronic haemodialysis and peritoneal dialysis. *Swiss Med Wkly* 2006; 136(13-14): 210-217.
43. Krespi R, Bone M, Ahmad R, Worthington B, Salmon P. Haemodialysis patients' beliefs about renal failure and its treatment. *Patient Educ Couns* 2004; 53(2): 189-196.
44. Tong A, Sainsbury P, Chadban S, Walker RG, Harris DC, Carter SM et al. Patients' experiences and perspectives of living with CKD. *Am J Kidney Dis* 2009; 53(4): 689-700.
45. National Kidney Foundation. KDOQI Clinical Practice Guidelines and Clinical Practice Recommendations for 2006 Updates: Hemodialysis Adequacy, Peritoneal Dialysis Adequacy and Vascular Access. *Am J Kidney Dis* 2006; 48: S1-S322 (suppl 321).
46. Kovac JA, Patel SS, Peterson RA, Kimmel PL. Patient satisfaction with care and behavioral compliance in end-stage renal disease patients treated with hemodialysis. *Am J Kidney Dis* 2002; 39(6): 1236-1244.
47. Verfahrensordnung des Gemeinsamen Bundesausschusses vom 20. September 2005 (Ed). zuletzt geändert am 18. April 2006, veröffentlicht im Bundesanzeiger 2006, S. 4876, in Kraft getreten am 7. Juli 2006. URL: http://www.g-ba.de/downloads/62-492-83/VerfO_2006-04-18.pdfS
48. Verordnung zur Beteiligung von Patientinnen und Patienten in der Gesetzlichen Krankenversicherung (Patientenbeteiligungsverordnung - PatBeteiligungsV). Vom 19. Dezember 2003. [online]. URL: <https://www.gesetze-im-internet.de/patbeteiligungsV/BJNR275300003.html>.
49. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Version 4.0 [online]. 23.09.2011 [Zugriff: 11.03.2013]. URL: http://www.iqwig.de/download/IQWiG_Methoden_Version_4_0.pdf.
50. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Allgemeine Methoden: Entwurf für Version 5.0 [online]. 07.12.2016 [Zugriff: 28.03.2017]. URL: https://www.iqwig.de/download/IQWiG_Entwurf_Methoden_Version_4-0.pdf.
51. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. London: NICE; 2013. URL: <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781>.
52. Neidhardt K WT, Schmid A.,. Die Gewichtung multipler patientenrelevanter Endpunkte: ein methodischer Vergleich von Conjoint Analyse und Analytic Hierarchy Process unter Berücksichtigung des Effizienzgrenzenkonzepts des IQWiG; Diskussionspapier [online]. *Wirtschaftswissenschaftliche Diskussionspapiere*, Band 02-12, [online] 2012. URL: http://www.fwi.uni-bayreuth.de/de/download/WP_02-12.pdf.
53. Muhlbacher A, Johnson FR. Choice Experiments to Quantify Preferences for Health and Healthcare: State of the Practice. *Appl Health Econ Health Policy* 2016; 14(3): 253-266.
54. Janssen IM, Gerhardus A, Schroer-Gunther MA, Scheibler F. A descriptive review on methods to prioritize outcomes in a health care context. *Health Expect* 2015; 18: 1873-1893
55. Kalsbeek W, Heiss G. Building bridges between populations and samples in epidemiological studies. *Annu Rev Public Health* 2000; 21: 147-169.
56. Tyrer S, Heyman B. Sampling in epidemiological research: issues, hazards and pitfalls. *BJPsych Bull* 2016; 40(2): 57-60.
57. Neumark DE, Stommel M, Given CW, Given BA. Research design and subject characteristics predicting nonparticipation in a panel survey of older families with cancer. *Nurs Res* 2001; 50(6): 363-368.
58. Wallenius J, Dyer JS, Fishburn PC, Steuer RE, Zionts S, Deb K. Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science* 2008; 54(7): 1336-1349.
59. Triantaphyllou E. *Multi-Criteria Decision Making Methods: A Comparative Study*. Kluwer Academic Publishers 2000. (Applied Optimization Series).

60. Scholl A, Mathey L, Helm R, Steiner M. Solving multiattribute design problems with analytic hierarchy process and conjoint analysis: An empirical comparison. *European Journal of Operational Research* 2005; 164(3): 760-777.
61. Mulye R. An empirical comparison of three variants of the AHP and two variants of conjoint analysis *Journal of Behavioral Decision Making* 1998; 11(4): 263-280.
62. Marsh K, M IJ, Thokala P, Baltussen R, Boysen M, Kalo Z et al. Multiple Criteria Decision Analysis for Health Care Decision Making-Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force. *Value Health* 2016; 19(2): 125-137.
63. Muhlbacher AC, Juhnke C. Patient preferences versus physicians' judgement: does it make a difference in healthcare decision making? *Appl Health Econ Health Policy* 2013; 11(3): 163-180.
64. Gagnon MP, Desmartis M, Gagnon J, St-Pierre M, Rhainds M, Coulombe M et al. Framework for User Involvement in Health Technology Assessment at the Local Level: Views of Health Managers, User Representatives, and Clinicians. *Int J Technol Assess Health Care* 2015; 31(1-2): 68-77.
65. Finney Rutten LJ, Morris MA, Schrader LM, Manemann SM, Pathak J, Dimler R et al. Approaching patient engagement in research: what do patients with cardiovascular disease think? *Patient Prefer Adherence* 2015; 9: 1061-1064.
66. Frank L, Forsythe L, Ellis L, Schrandt S, Sheridan S, Gerson J et al. Conceptual and practical foundations of patient engagement in research at the patient-centered outcomes research institute. *Qual Life Res* 2015; 24(5): 1033-1041.
67. Raeymaekers P with contributions from Crowe S, Cowan K, Broerse J, and Hertz-Pannier L. Mind the gap! Multi-stakeholder dialogue for priority setting in health research. A publication of the King Baudouin Foundation. Belgium: 2016. URL: <http://en.calameo.com/read/0017742952f4a0659203a?authid=vEWVdbAPaMvd>.
68. Harris J, Croot L, Thompson J, Springett J. How stakeholder participation can contribute to systematic reviews of complex interventions. *J Epidemiol Community Health* 2016; 70(2): 207-214.
69. Methodology Committee of the Patient-Centered Outcomes Research I. Methodological standards and patient-centeredness in comparative effectiveness research: the PCORI perspective. *JAMA* 2012; 307(15): 1636-1640.
70. Crowe S, Fenton M, Hall M, Cowan K, Chalmers I. Patients', clinicians' and the research communities' priorities for treatment research: there is an important mismatch. *Research Involvement and Engagement* 2015; 1(1): 2.
71. Whiting C. The James Lind Alliance - involving patients and their health professionals jointly in setting priorities for research. *ENT and Audiology News* 2016; 24(6): 56-57.
72. Chalmers I, Atkinson P, Fenton M, Firkins L, Crowe S, Cowan K. Tackling treatment uncertainties together: the evolution of the James Lind Initiative, 2003-2013. *J R Soc Med* 2013; 106(12): 482-491.
73. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M et al. Choosing important health outcomes for comparative effectiveness research: a systematic review. *PLoS One* 2014; 9(6): e99111.
74. Gorst SL, Gargon E, Clarke M, Smith V, Williamson PR. Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and Identification of Gaps. *PLoS One* 2016; 11(12): e0168403.
75. Tong A, Manns B, Hemmelgarn B, Wheeler DC, Evangelidis N, Tugwell P et al. Establishing Core Outcome Domains in Hemodialysis: Report of the Standardized Outcomes in Nephrology-Hemodialysis (SONG-HD) Consensus Workshop. *Am J Kidney Dis* 2016.
76. Evangelidis N, Tong A, Manns B, Hemmelgarn B, Wheeler DC, Tugwell P et al. Developing a Set of Core Outcomes for Trials in Hemodialysis: An International Delphi Survey. *Am J Kidney Dis* 2017.
77. Urquhart-Secord R, Craig JC, Hemmelgarn B, Tam-Tham H, Manns B, Howell M et al. Patient and Caregiver Priorities for Outcomes in Hemodialysis: An International Nominal Group Technique Study. *Am J Kidney Dis* 2016; 68(3): 444-454.
78. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012; 13: 132.
79. Wahlster P, Goetghebeur M, Kriza C, Niederlander C, Kolominsky-Rabas P, National Leading-Edge Cluster Medical Technologies 'Medical Valley EMN. Balancing costs and benefits at different stages of

medical innovation: a systematic review of Multi-criteria decision analysis (MCDA). *BMC Health Serv Res* 2015; 15: 262.

80. Angelis A, Kanavos P. Value-Based Assessment of New Medical Technologies: Towards a Robust Methodological Framework for the Application of Multiple Criteria Decision Analysis in the Context of Health Technology Assessment. *Pharmacoeconomics* 2016; 34(5): 435-446.

Eidesstattliche Erklärung

Ich erkläre, dass ich die vorgelegte Arbeit ohne fremde Hilfe verfasst, keine anderen Quellen und Hilfsmittel als die angegebenen benutzt sowie Textstellen, die anderen Werken wörtlich oder dem Sinne nach entnommen sind, unter Angabe der Quelle kenntlich gemacht habe.

Ich versichere, dass die Arbeit nicht anderweitig in dieser oder ähnlicher Form als Dissertation eingereicht wurde und ich bisher keine weiteren Promotionsversuche unternommen habe.

Bedburg-Hau, den 09.06.2017

Inger Miriam Janßen

Danksagung

An einer wissenschaftlichen Arbeit sind meist mehrere Personen beteiligt. Deshalb möchte ich hier allen, die einen fachlichen Beitrag zum Gelingen dieser Arbeit beigetragen haben, recht herzlich danken:

Mein ganz besonderer Dank gilt meinem Doktorvater, Herrn Professor Dr. Ansgar Gerhardus für seine wissenschaftliche Betreuung und Unterstützung über diesen langen Zeitraum. Ebenso gilt mein besonderer Dank Herrn Professor Dr. Oliver Razum für die Übernahme der Zweitbetreuung und die damit verbundene fachliche Unterstützung.

Ganz besonders möchte ich mich bei Dr. Fülöp Scheibler für die wissenschaftliche und motivierende Unterstützung während der gesamten Promotion bedanken.

Für den methodisch-wissenschaftlichen Austausch und motivierenden Zuspruch möchte ich mich recht herzlich bei Evelyn, Anne und Linda bedanken. Mein besonderer Dank gilt Milly für die Unterstützung bei der Übersichtsarbeit. Des Weiteren bedanke ich mich bei Natalie und Consuela für Ihre Unterstützung bei den englischsprachigen Artikeln.

Ich bedanke mich bei den Mitarbeitern der KfH-Zentren und des QiN-Programms für Ihre Unterstützung bei der Befragung.

Mein ausdrücklicher Dank gilt allen Teilnehmern der Befragung, insbesondere den Patienten der KfH-Zentren und den Mitgliedern der Selbsthilfegruppe I. G. Niere NRW e.V.