
**Stellungnahme des DNVF e.V. zum Vorbericht:
Sektorenübergreifende Qualitätssicherung nach § 137a SGB V "Entwicklung einer
Patientenbefragung im Rahmen des Qualitätssicherungsverfahrens Arthroskopie
am Kniegelenk"**

04.05.2015

Rahmenbedingungen und Zielsetzung

Das AQUA-Institut wurde im März 2013 beauftragt Instrumente zur Abbildung der Patientenperspektive zu erarbeiten. Konkret ging es um die Entwicklung eines Instruments, das Ergebnis- und Prozessqualität bei Arthroskopie am Kniegelenk valide erheben kann.

Das Ergebnis dieser Arbeiten liegt in Form eines Vorberichts (177 Seiten), eines Anhangs zum Vorbericht (81 Seiten) und eines Indikatorensets (49 Seiten) vor.

Als Abgabetermin wurde der 10. März 2015 festgelegt. Eingeladene Institutionen wurden mit einer Frist zum 05.05.2015 aufgefordert, am Stellungnahmeverfahren teilzunehmen.

Das Deutsche Netzwerk Versorgungsforschung (DNVF) e.V. nimmt an diesem Stellungnahmeverfahren teil und fasst seine Kommentare hier zusammen.

Genereller Eindruck

Kernstück des Arbeitsergebnisses des AQUA-Instituts ist ein Fragebogen, der insgesamt 12 Seiten umfasst und Aspekte der Ergebnisqualität (basierend auf dem KOOS-Fragebogen) und der Prozessqualität (selbstentwickelte Fragen auf der Vorlage von Beispielen aus der Literatur) sowie biographische Angaben umfasst. Dieser Fragebogen soll flächendeckend zur Qualitätssicherung eingesetzt werden. Die Patientenbefragung soll über eine zentrale Stelle erfolgen und die entsprechende Methodik und Logistik ist in Ausarbeitung.

Besonders positiv hervorzuheben ist, dass durch diese Initiative endlich patientenbasierte Endpunkte (Patient-Reported Outcomes, PRO) in den Mittelpunkt einer flächendeckenden Qualitätssicherung für ein wichtiges und vielfach kontrovers diskutiertes Gesundheitsthema gerückt werden. Der Fokus auf Patientenorientierung wurde konsequent umgesetzt. Das Messinstrument erfasst Lebensqualität (im Sinne von Ergebnisqualität) sowie Information, Aufklärung und shared decision making (im Sinne von Prozessqualität). Patienten wurden stark in die Entwicklung des Instruments eingebunden.

Insgesamt bestechen das systematische und sorgfältige Vorgehen, sowie die transparente Beschreibung von Prozessen und Ergebnissen. Beispielhaft sei genannt, dass in der Entwicklung mehrere qualitative Verfahren kombiniert werden (Fokusinterviews, Einbindung von Experten, Recherche zu bereits vorliegenden Instrumenten) und dass jeder Schritt ausführlich und oft unter Verweis auf die Literatur begründet wird. Dieses sorgfältige Vorgehen ist auch im Vergleich zur typischen Entwicklung von PROs für Forschungszwecke überdurchschnittlich. Die sorgfältige Vorgehensweise ist aber durchaus gerechtfertigt, weil der Einsatz des Instruments konkrete Folgen für die Gesundheitsversorgung und die bewerteten Institutionen haben kann.

Das entwickelte Instrument weist positive Eigenschaften auf, insbesondere eine theoretische und empirische Fundierung, die Passung zum Messzweck „Qualitätssicherung“ und gute oder zumindest hinreichende psychometrische Messeigenschaften. Zur Überprüfung wurden Methoden der klassischen Testtheorie eingesetzt.

Kritikpunkte

Methodik

Momentan ist in der *scientific community* ein Trend zu beobachten, die Fülle der möglichen klinischen und patientenberichteten Outcomes zu systematisieren und in Konsensusprozessen in Form von Core Outcome Sets (COS) zu verabschieden und zu bündeln. Derartige Entwicklungen sind bei einer Reihe von Krankheitsbildern beobachtbar (Hauterkrankungen, Rheuma, onkologische Erkrankungen, Tinnitus) und die Initiativen bzw. Arbeitsgruppen verbergen sich hinter den Abkürzungen wie CONSORT-PRO, COMET, COSMIN, ICHOM, OMERACT oder TINNET. Es wäre wünschenswert gewesen, wenn der vorliegende Bericht zu diesem Trend (kritisch) Stellung genommen hätte, um so den AQUA-Ansatz besser einordnen zu können.

Auswahl des Bogens

Der KOOS-Fragebogen (Knee injury and Osteoarthritis Outcome Score) wurde als Instrument ausgewählt und Gründe hierfür wurden genannt. Was zu kurz kommt, ist der Vergleich zu den übrigen sechs Bögen, die ebenfalls in deutscher Sprache vorliegen. Die Vorteile des KOOS hätten hier besser herausgearbeitet werden können. Auch wird nicht angegeben, wer oder welches Gremium letztlich die Entscheidung getroffen hat.

Verständlichkeit

Die Prüfung der Verständlichkeit des Instruments ist nicht ausreichend erfolgt.

Es wurden kognitive Pretests durchgeführt, die jedoch „überwiegend die Schritte des Probing und General Probing“ angewandt haben. Die beiden anderen Methoden „Paraphrasing“ und „Think Aloud“ kommen nur zur Anwendung, wenn gezielt überprüft werden soll, ob Patienten bestimmte Begriffe, Fragen oder Antwortvorgaben anders interpretieren, als vom Fragebogenentwickler beabsichtigt. „Sofern der Patient bei der Begründung unsicher wirkte, wurde er aufgefordert, die Frage mit eigenen Worten zu formulieren (Paraphrasing).“ (Vorbericht, S. 52).

Die Methode des Paraphrasing hätte durchgehend und konsequent angewendet werden sollen. Die Tatsache, dass ein Patient kein Verständnisproblem wahrnimmt oder zum Ausdruck bringt, bedeutet nicht, dass er das Item tatsächlich richtig verstanden hat. Vielleicht hat er/sie etwas anderes verstanden als gemeint war.

Zudem sind die kognitiven Interviews mit einer kleinen Stichprobe ($n = 7$) durchgeführt worden. Nur sehr große Verständnisprobleme werden dadurch detektiert.

Bei einer Prävalenz eines Verständnisproblems von 5% ist eine Stichprobe von $n = 32$ Probanden nötig, um es mit einer Power von 80% aufzudecken. (Perneger TV, Courvoisier DS, Hudelson PM, Gayet-Ageron: A. Sample size for pre-tests of questionnaires. Qual Life Res. 2015 Jan;24(1):147-5).

Veränderungsmessung

Zwar wird ausführlich auf Basis der Literatur diskutiert, ob eine retrospektive Erhebung ohne Pretest (wie nun vorgesehen) valide ist; hier ist natürlich die Gefahr des Recall Bias gegeben. Dieses Problem wurde laut Bericht „sehr kontrovers diskutiert“, aber in „den Fokusgruppen hingegen wurde überwiegend angegeben, dass keine Schwierigkeiten bei der Erinnerung bestehen“ (Vorbericht, S. 73).

Es ist als wissenschaftliches Argument nicht stichhaltig, wenn Probanden annehmen oder subjektiv den Eindruck haben, sich zuverlässig erinnern zu können. Auch kognitiv unbeeinträchtigte Menschen können die eigene Erinnerungsfähigkeit in vielen Fällen nicht zuverlässig einschätzen – auch sehr deutliche Erinnerungen können falsch sein, wie die psychologische Gedächtnisfor-

schung zeigt. Daher sollte die Erinnerbarkeit der erfragten Inhalte (z.B. Schmerzen vor dem Eingriff) nicht angenommen, sondern empirisch getestet werden anhand eines Pre-Post-Vergleichs. Und auch dieser kann lediglich nur einen Hinweis auf die Validität retrospektiver Angaben geben, da Pre-Post-Vergleiche wiederum durch Rekalibrierungs-Response-Shift verzerrt sein können (Blome C, Augustin M.: Measuring change in quality of life: bias in prospective and retrospective evaluation. Value Health. 2015 Jan;18(1):110-5.).

Länge des Fragebogens und Reihenfolgeeffekte

Der Fragebogen umfasst 12 Seiten und 103 Items. Das dürfte an der Grenze dessen sein, was Patienten in einer derartigen Umfrage zu beantworten willens sind. Aus der Umfrageforschung ist bekannt, dass Reihenfolgeeffekte sowie Ermüdung das Antwortverhalten stark beeinflussen.

Es gilt daher das Prinzip, Wichtiges stets zuerst abzufragen. Entgegen dieser Grundregel wird die zentrale Messgröße „aktuelle Kniebeschwerden“ erst ganz am Ende des Fragebogens erhoben. Notwendigerweise wird dieser entscheidende Parameter der Ergebnisqualität durch vorab abgefragte Angaben zu Kniebeschwerden vor der Operation und Aspekte der Prozessqualität überlagert.

Interpretationskriterien

Die Interpretation von PRO-Messergebnissen ist ein in Forschung und Praxis notorisch vernachlässigtes Problem. Auch in diesem Bericht kommt dieses Thema viel zu kurz. Lediglich eine Beispielabbildung zu den Mittelwertsverteilungen ist auf S. 152 zu sehen.

Anwender brauchen aber eindeutige Interpretationskriterien und müssen auf einen Blick erkennen können, ob ein Behandlungsergebnis „gut“ oder „schlecht“ ist. Daher wäre es sinnvoll, durch eine bundesweite repräsentative Umfrage in der Normalbevölkerung Referenzwerte zum KOOS-Fragebogen zu generieren und diese Werte zur Interpretation der Qualitätssicherungsergebnisse zu verwenden.

Diese normbasierte Methode der Ergebnisinterpretation würde es zudem ermöglichen, das Befragungsinstrument deutlich zu verkürzen und den ersten Teil des Fragebogens („Kniebeschwerden vor der Operation“) ganz wegzulassen.

Ein derartiger Schritt bringt mehrere Vorteile mit sich:

1. Das Befragungsinstrument ist insgesamt kürzer und dies steigert die Rücklaufquote,
2. Das methodische Problem der quasi-indirekten/retrospektiven Veränderungsmessung wird umgangen,
3. Das Befragungsinstrument kann gleich zu Beginn mit dem zentralen Parameter der Ergebnisqualität (nämlich aktuelle Kniebeschwerden nach der Operation) beginnen.

Last, but not least sei darauf hingewiesen, dass mittlerweile der IQWiG-Bericht zur Kniearthroskopie vorliegt (<https://www.iqwig.de/de/presse/pressemitteilungen/pressemitteilungen/arthroskopie-des-kniegelenks-bei-arthrose-kein-nutzen-erkennbar.6108.html>). Dieses wichtige Dokument muss der AQUA-Vorbericht unbedingt aufgreifen.

Autoren: Prof. Dr. M. Koller, Dr. Christine Blome, PD Dr. M. Klinkhammer-Schalke und Prof. Dr. M. Geraedts

Kontakt: DNVF e.V.
Geschäftsstelle c/o IMVR
Eupener Str. 129, 50933 Köln

Dr. Gisela Nellesen-Martens
Tel. 0221-47897115
E-Mail: dnvf@uk-koeln.de