

Knie-Scores als Quelle von vorgetäuschem Nutzen

Luzi Dubs

Dr. med., Facharzt für Orthopädische Chirurgie und Traumatologie

Bei der Bemessung eines Gesundheitszustandes fokussiert sich die Orthopädie vornehmlich auf Erhebungsdaten der Lebensqualität. Der Patient interessiert sich letztlich nur für seine Fähigkeiten. Er will wissen, was er kann oder allenfalls nicht mehr kann. Speziell am Kniegelenk eignet sich zur Bestimmung der sportlichen und beruflichen körperlichen Leistungsfähigkeit seit Jahrzehnten der validierte Tegner-Score [1]*, eine 10-Punkte-Skala mit Fähigkeits-Abstufungen vom gerade noch Gehfähigen (1 Punkt) bis zum Leistungssportler in Mannschaftskontakt-Sportarten (10 Punkte). Mit diesem pragmatisch einfachen Fähigkeitsassessment lassen sich auch Aussagen machen, wenn es darum geht, über Fähigkeiten vor 20 bis 40 Jahren nachträglich zu berichten. Neben dem Recall-Bias und dem Gefälligkeits-Bias lässt sich damit vor allem der gefährliche Score-Bias [2, 3] ausschliessen.

Der Score-Bias

Ein Misch-Score wie z.B. der häufig eingesetzte Lysholm-Score [4] besteht in der Regel aus 100 Punkten. Der Hauptanteil der Punkte verteilt sich auf die sogenannten Surrogatparameter. Das sind Messwerte, die den Patienten letztlich nicht interessieren. Zu diesen gehören auch Punkte für die messbare Gelenkstabilität. Der Hauptdenkfehler liegt darin, dass jemand, dessen Knie z.B. durch vorderen Kreuzbandersatz operativ behandelt und stabilisiert worden ist und bei guter Durchführung letztlich auch in den Stabilitätstests besser abschneidet, 10 Punkte bekommt, derweil der Nichtoperierte, dessen Knie instabiler erscheint, was

aber nicht von klinischer Bedeutung sein muss, diese Chance des Punktegewinns gar nicht bekommt. Es handelt sich offensichtlich um eine Ausgangslage mit systematischem Potential des Selbstbetruges, indem selbstredend davon ausgegangen wird, dass durch eine Veränderung im Sinne der Besserung eines Surrogatmesswertes ein Patientennutzen entstehen soll.

Ein gesondertes Problem eines Scores kann die Zuteilung von bis zu 30 Punkten für das Fehlen von Schmerzen sein. Wer sich schon und somit Fähigkeiten einschränkt, um schmerzfrei zu sein, verliert vielleicht 10 Fähigkeitspunkte, gewinnt dafür aber 30 Punkte wegen Schmerzfreiheit, was den Gesamtscore erhöht und die Aussage verfälschen kann. Die Resultatgewichtung mit Schmerzpunkten kann als irrelevant angesehen werden, denn die Schmerzwahrnehmung drückt sich automatisch in einem Fähigkeitsverlust aus. Wenn jemand starke Schmerzen hat, kann er seine Situation günstig beeinflussen, indem er auf gewisse Fähigkeiten verzichten muss.

Seit einigen Jahren gibt es komplexere Fragebögen mit Zuordnung von Scorepunkten (KOOS, IKDC 2000, WOMAC, SF36 usw.). Sie sind komplex in viele Untergruppen unterteilt. Sie sind nur in prospektiven Studien einsetzbar, da der Patient diesen Fragebogen nicht mehr zuverlässig ausfüllen kann, was die Beschreibung seines früheren Zustandes betrifft. Auch in diesen Messinstrumenten ist die Schmerzmessung enthalten.

Problematisch kann die Auswertung einer Studie werden, wenn über Operationserfolge berichtet wird, die den Score-Wert vor der Operation zum Ausgangswert nehmen und nicht denjenigen vor dem Unfall. Dies hat schon oft dazu geführt, dass beim Riss des vorderen Kreuzbandes die Frühoperation empfohlen wurde, da der Punktegewinn über zwei Jahre der Verlaufsbeobachtung hinweg besser ausfiel als bei der zeitlich verzögerten Operation. Die Frühoperierten profitieren von den natürlichen Heilungsfortschritten, die sich unabhängig bzw. trotz der operativen Belastung ohnehin einstellen und nicht der Operation zugeschrieben werden können. Problematisch sind derartige Studienauswertungen auch deshalb, weil die Spätoperierten meistens einer negativen Selektion entsprechen.

* Die Literatur findet sich unter www.saez.ch → Aktuelle Ausgabe oder → Archiv → 2015 → 45.

Les scores du genou à l'origine de bénéfices fictifs

Au niveau international, on mesure et interprète généralement les résultats d'une reconstruction du ligament croisé antérieur du genou à l'aide de scores. Parmi ces scores, on trouve des scores validés à large échelle, qui avantagent systématiquement le patient opéré. En effet, ceux-ci contiennent des critères de substitution comme la mesure de la stabilité, qui ne sont pas pertinents pour le patient mais uniquement lors de comparaisons avec des personnes non opérées. Ce biais doit donc être reconnu et éliminé.

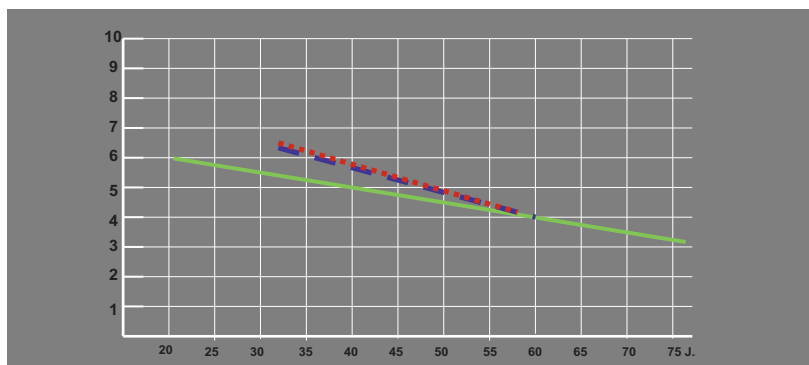


Abbildung 1: 14 Matched Pairs (12m, 2f) mit VKB-Rekonstruktion und Follow-up von 25,7 Jahren (Tegner von 6,57 zu 4,28; - - - - ohne VKB-Rekonstruktion und Follow-up von 27,5 Jahren (Tegner von 6,35 zu 4); — Normkurve Tegner Score (1–10).

Langzeitstudie möglichst ohne Bias

Dieser Frage ist der Autor in einer eigenen Studie mit je 24 operativ und konservativ behandelten Kniegelenken und einer Beobachtungsdauer von je 25 Jahren im Jahr 2014 nachgegangen [5]. Im Sinne einer Vereinfachung auf das Patientenrelevante wurden alle Mischscores mit einem möglichem Score-Bias weggelassen und es wurde nur das Fähigkeitsassessment des Tegner-Scores eingesetzt. Die Resultate zeigten in den beiden Gruppen keine relevanten Unterschiede. Natürlich interessierten bei den Nachkontrollen separat die Beweglichkeit, die Stabilität und das Röntgenbild. Die Stabilitätsmessung zeigte erwartungsgemäss bei den Operierten bessere Werte, was aber, wie sich zeigen wird, nicht von Bedeutung sein muss. Um die Vergleichbarkeit der Gruppen nochmals zu optimieren, wurde in einer darin enthaltenen Matched-Pair-Analyse auf das vergleichbare Geschlecht, das Alter und den Tegner-Score zum Zeitpunkt des Unfalls bei je 14 Patienten mit einer Beobachtungsdauer von 25 Jahren geachtet. Der primäre Endpunkt war also lediglich der Vergleich des Fähigkeitsverlustes der damals Operierten gegenüber den Kreuzbandverletzten im Spontanverlauf ohne Operation am Kreuzband, wobei konsequent der Tegner-Score vor dem Unfall und derjenige zum Zeitpunkt der Nachuntersuchung verwendet wurden. Die Kohorte der Nichtoperierten entspricht eher einer negativen Selektion, da praktisch alle wegen einer medialen und/oder lateralen Meniskusschädigung, teils korbhenkelartig, zugewiesen wurden. Sie hatten alle eine gesicherte frühere Kreuzbandruptur, wurden deswegen aber nie operiert.

Altersabhängiger Fähigkeitsverlust als Standard

Um den Stellenwert des normalen altersabhängigen Fähigkeitsverlustes, analog zur MARA-Kurve [6], zu

illustrieren, konnte eine Bevölkerungsuntersuchung aus England von Yarashi [7] als Grundlage eingesetzt werden, die zur Feststellung gelangt, dass pro 20 Jahre ziemlich linear 1 Tegnerpunkt an Fähigkeiten verloren geht. Diese «Messlatte» wird sich künftig als sehr wertvoll erweisen, um einzelne Resultatmitteilungen altersabhängig einzuordnen. Das Resultat bei dieser eher kleinen Patientenzahl von je 14 Personen hat aber letztlich doch deutlich gezeigt, dass zwischen den Operierten und Nicht-Operierten im Langzeitverlauf hinsichtlich eines Fähigkeitsverlustes kein Unterschied festgestellt werden kann (Abb. 1). Der Fähigkeitsverlust in beiden Gruppen von etwas mehr als zwei Tegnerpunkten nach 25 Jahren erscheint eigentlich nicht besorgniserregend und unterscheidet sich vom physiologischen Fähigkeitsverlust nur um rund einen Punkt. Hätte man einen Mischscore gemäss Lysholm eingesetzt, hätte die Gruppe der Operierten nur schon wegen der besseren messbaren Stabilität ein besseres Resultat zeigen müssen, was Anlass gegeben hätte, die Operation als notwendig zu propagieren.

Folgerungen

Alle Mischscores erscheinen nicht mehr zeitgemäss, um die patientenrelevanten Fragen der Fähigkeitsveränderungen und somit den Patientennutzen abzubilden. Sie sollten in klinischen Studien systematisch weggelassen werden. Es sind primäre Endpunkte in den Vordergrund zu stellen, die patientenrelevant sind, d.h. Messungen von Fähigkeitsveränderungen. Der Score-Bias muss ausgerottet werden. Seit Jahren bekämpfen sich die Exponenten der Knieexpertengruppe von swiss orthopaedics und das Swiss Medical Board (SMB) zum Thema, ob die «Rekonstruktion» des vorderen Kreuzbandes Sinn macht. Der Autor steht als einer der damals beigezogenen Spartenvertreter des SMB und als langjähriges, eng verbundenes Mitglied der Orthopädischen Fachgesellschaft zwischen Hammer und Amboss. In der jüngsten Stellungnahme beruft sich swiss orthopaedics auf neuere Studien in hochkarätigen amerikanischen Zeitschriften, wo der Nutzen und sogar auch die Kosten bei operativer Behandlung in einem vorteilhaften Licht erscheinen sollen. Bei genauer Betrachtung fällt jedoch auf, dass die Nutzenbewertung nicht in obigem Sinne erfolgte und der Score-Bias auch in hochkarätigen Zeitschriften virulent ist. Es wäre somit zu fordern, dass alle, die letztlich am Tisch debattieren, eine parkett-sichere Expertise in methodischen Fragen mitbringen, um über die spezifische Sachlage patientengerecht befinden zu können.

Korrespondenz:
Dr. med. Luzi Dubs
Merkurstrasse 12
CH-8400 Winterthur
Tel. 052 213 22 23
dubs.luzi[at]bluewin.ch

Literatur

- 1 Tegner Y, Lysholm J. Rating systems in the evaluation of the knee ligament injuries. *Clin Orthop*. 198, 1985;43–9.
- 2 Hefti F. Numerische Scores zur Evaluation von Behandlungsergebnissen am Bewegungsapparat. *Schweiz Ärztezeitung*. 1996;77(12):492–6.
- 3 Dubs L. Der Score Bias. *Schweiz Ärztezeitung*. 2000;81(6):292–5.
- 4 Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale. *Am J Sports Med*. 1982;10:150–4.
- 5 Dubs L. Long term results in two cohorts comparing operative and non operative treatment of anterior cruciate ligament rupture of the knee based on ability assessment by the Tegner score. Vortrag SGOT-Kongress Basel am 25.6.2015.
- 6 Dubs L. Erlaubt ist, was gefällt – wie lange noch? *Schweiz Ärztezeitung*. 1996;77(12):485–8.
- 7 Yarashi T. et al. Knee outcome scores – do we get people to normal? *JBJS Br*. 2011; Vol 93B no. Supp I 14.