

# Die arthroskopische Gelenksäuberung bei Kniearthrose

## Teil 3: Was versteht man unter Diagnostik nach der Sherlock-Holmes-Methode?

In dieser Ausgabe soll gezeigt werden, wie man aufgrund von Anamnese und klinischer Untersuchung auch ohne bildgebende Zusatzdiagnostik eine Entscheidungsschwelle zur arthroskopischen Operation überschreiten kann, wenn in der diagnostischen Abklärung nach der Sherlock-Holmes-Methode vorgegangen wird.

LUZI DUBS

Sherlock Holmes ist bekannt dafür, dass er in seiner Verbrecherjagd stets verschiedene Indizien sorgfältig gesammelt und für den Informationsgewinn zusammengetragen hat, um letztlich die Entscheidungsschwelle zu überschreiten und die Täterschaft zu identifizieren. Seine Stärke war der



Der Nutzen einer arthroskopischen Gelenksäuberung wird bis heute kontrovers beurteilt. Liegt der Gewinn womöglich eher beim Operateur als beim Patienten? In dieser Arbeit werden Mythen, Fakten und vielleicht etwas Neuland aus der Sicht der evidenzbasierten Medizin in 5 Folgen vorgestellt:

- Teil 1: Einleitung und Literatur: Spreu oder Weizen?
- Teil 2: Ist die Kniearthrose eine Folge der Meniskusoperation?
- Teil 3: Was versteht man unter Diagnostik nach der Sherlock-Holmes-Methode?
- Teil 4: Wie viel Information bringen Anamnese und klinische Tests?
- Teil 5: Röntgen oder MRI? Beides oder keines?

gewissenhafte Umgang mit Wahrscheinlichkeiten. Es entzieht sich unserer Kenntnis, ob Sherlock Holmes beziehungsweise sein Autor Sir Arthur Conan Doyle seinen Ko-Patrioten, den Priester Thomas Bayes (1702–1761) von Tunbridge Wells, gekannt hat. Diesem ist jedenfalls das Bayes-Theorem, der rationale, mathematische Zugang zur Indizientchnik, zu verdanken, welches darauf beruht, dass der Wert eines Tests von der Wahrscheinlichkeit der gesuchten Krankheit abhängt (1).

### Die Vierfeldertafel

Um die Sherlock-Holmes-Methode im klinischen Alltag anwenden zu können, benötigen wir zuerst einige Grundkenntnisse, einerseits über die Krankheitswahrscheinlichkeiten (Vor-, Nachtestwahrscheinlichkeit, Informationsgewinn, Entscheidungsschwelle), andererseits solche über die Testeigenschaften (Sensitivität, Spezifität, Likelihood-Ratio). Wenn die Aussagekraft eines (dichotomen) Tests, wie etwa Kniegelenkerguss Ja oder Nein, mit dem Vorliegen oder Fehlen eines Meniskusschadens berechnet werden soll, bedienen wir uns mit Vorliebe der Vierfeldertafel, welche Auskunft sowohl über die Krankheitswahrscheinlichkeiten als auch über die Testeigenschaften geben kann. Die vier Felder geben die vier möglichen Zusammenhänge an (Abbildung 1).

		Krankheit	
		+	-
Test	+	richtig-positiv <b>a</b>	falsch-positiv <b>b</b>
	-	falsch-negativ <b>c</b>	richtig-negativ <b>d</b>

a = Der Test ist positiv, die gesuchte Krankheit ist vorhanden. Zu Recht hat der positive Test die Krankheit erkannt (richtigpositiv).  
 b = Der Test ist positiv, die gesuchte Krankheit ist jedoch nicht vorhanden. Der Test hat Fehlalarm geschlagen (falschpositiv).  
 c = Der Test ist negativ, die gesuchte Krankheit ist jedoch vorhanden. Dem Test ist es nicht gelungen, die Krankheit zu erkennen (falschnegativ).  
 d = Der Test ist negativ, die gesuchte Krankheit ist nicht vorhanden. Zu Recht hat der negative Test die Krankheit ausgeschlossen (richtignegativ).

Abbildung 1: Die Vierfeldertafel

Daraus lassen sich alle weiteren Informationen über die Krankheitswahrscheinlichkeiten und die Testeigenschaften ableiten.

## Fazit .....

- ❖ Eine diagnostische Abklärung besteht aus verschiedenen Tests. Anamnesefragen und klinische Untersuchungsbefunden, wobei jeder Test mit seiner Sensitivität und Spezifität einen eigenen Beitrag zum Informationsgewinn auf der Suche nach einer Krankheit liefert.
- ❖ Ausgehend von einer Krankheitswahrscheinlichkeit vor dem Test (Vortestwahrscheinlichkeit) erfolgt eine Umwandlung in eine Krankheitswahrscheinlichkeit nach dem Test (Nachtestwahrscheinlichkeit), die Differenz ist der numerische Informationsgewinn.
- ❖ Die Berechnung der Likelihood-Ratio und die Anwendung des Fagan-Nomogramms erleichtern die Erfassung des Informationsgewinns und ermöglichen die praktische Anwendung einer Diagnostik in der Testserie.

### Krankheitswahrscheinlichkeiten

1. Die Vortestwahrscheinlichkeit ist der Anteil der Personen mit der gesuchten Krankheit unter den untersuchten Personen  $(a+c/a+b+c+d)$ . Wenn es sich um grosse Populationen handelt, kann auch der Begriff der Prävalenz eingesetzt werden.
2. Die positive Nachtestwahrscheinlichkeit (oder der positiv prädiktive Wert) ist der Anteil der Erkrankten unter den Testpositiven  $(a/a+b)$ . Dieser wichtige Wert gibt an, wie gross die Wahrscheinlichkeit der gesuchten Krankheit nach dem positiven Test geworden ist.
3. Analog wäre die negative Nachtestwahrscheinlichkeit (oder der negativ prädiktive Wert) der Anteil der Gesunden unter den Testnegativen  $(d/c+d)$ .
4. Der numerische Informationsgewinn durch den Test ist die Differenz zwischen Nach- und Vortestwahrscheinlichkeit. Wenn ein Test sehr schwach ist, kann auch ein Informationsverlust resultieren (Summe von Sensitivität und Spezifität  $< 100$ ). Damit ist angedeutet, dass über die Aussagekraft eines Tests vordergründig die Testeigenschaften Sensitivität und Spezifität entscheiden. Aus den beiden Werten lässt sich dann das positive oder negative Wahrscheinlichkeitsverhältnis (Likelihood-Ratio) errechnen.
5. Die Sensitivität erfasst die Anzahl der Testpositiven unter denjenigen mit der gesuchten Krankheit  $(a/a+c)$ .
6. Die Spezifität eines Tests erfasst die Anzahl der Testnegativen unter denjenigen, welche die gesuchte Krankheit nicht aufweisen  $(d/b+d)$ .

### Die Likelihood-Ratio

Die «Einschlusskraft» mit einem positiven Test kann mit der positiven Likelihood-Ratio (LR pos) bestimmt werden. Die Richtigpositiv-Rate (Sensitivität) wird geteilt durch die Falschpositiv-Rate (100-Spezifität). Damit weiss man, wie viel Mal häufiger das Testresultat richtigpositiv als falschpositiv ist. Man weiss, wenn der Test positiv ist, wie viel Mal häufiger er bei den Kranken gegenüber den Gesunden vorkommt. Wenn diese Verhältniszahl über 10 liegt, darf man

von einem guten Test sprechen, die gesuchte Krankheit zu erkennen.

Analog errechnet sich die «Ausschlusskraft» eines Tests. Man bestimmt die Falschnegativ-Rate (100-Sensitivität) und teilt diese durch die Richtignegativ-Rate (Spezifität). Man weiss, wenn der Test negativ ist, wie viel Mal seltener er bei den Kranken gegenüber den Gesunden vorkommt. Werte tiefer als 0,1 zeugen von einem guten Test.

### Das Fagan-Nomogramm

Der grosse Vorteil der Likelihood-Ratio (LR) liegt in der Möglichkeit, den Zahlenwert mit der Vortestwahrscheinlichkeit zu multiplizieren, um direkt den Wert der Nachtestwahrscheinlichkeit zu errechnen. Diese Berechnung wird etwas erschwert durch die Tatsache, dass sie in einer Odds-Einheit erfolgen muss, worauf hier nicht näher eingegangen werden soll. Glücklicherweise verfügen wir über eine Nomogramm-Tabelle nach Fagan (Abbildung 2), welche uns die Arbeit erleichtert. Wenn zwei der drei Werte aus Vor-, Nachtestwahrscheinlichkeit und Likelihood-Ratio bekannt sind, kann eine entsprechende Verbindungslinie zwischen den bekannten Werten gezogen und dadurch der dritte Wert ermittelt werden.

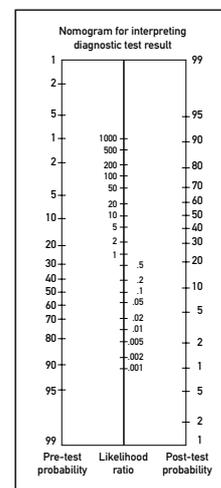


Abbildung 2: Das Fagan-Nomogramm

Ein weiterer, grosser Vorteil der LR ist die Möglichkeit, während einer Testserie die LR der einzelnen Tests zu multiplizieren, was mit den Angaben der Sensitivität und Spezifität allein nicht machbar ist. Die Gesamt-LR kann uns in der diagnostischen Kette erheblich weiterbringen. Wichtig ist dabei, dass die einzelnen Tests voneinander unabhängig sind.

Es ist nicht sinnvoll, den Test «Knieschmerz beim Fernsehen» mit dem Test «Knieschmerz beim Zeitunglesen» zu kombinieren.

Leider ist in der medizinischen Nutzenforschung das Thema der Testeigenschaften von einzelnen Anamnesefragen und klinischen Untersuchungsbefunden zur Pflege der Sherlock-Holmes-Methode bisher von den Entscheidungsträgern in der Medizin und in der Politik stiefmütterlich angegangen und kaum gefördert worden. Dies mag einerseits damit zusammenhängen, dass weiterhin der Irrglaube herrscht, nur eine maximale Bildgebung habe vor einem Richter Bestand, und andererseits mit den ökonomischen Anreizen, die Hilfsmethoden eher einzusetzen, als sie zu unterlassen.

Der Internist Daniel Pewsner versuchte zusammen mit Kollegen und mit Unterstützung der Schweizerischen Akademie der Medizinischen Wissenschaften bereits vor zehn Jahren, analog zur Cochrane Collaboration eine Art Bayes Collaboration ins Leben zu rufen, eine Datenbank über Krankheitswahrscheinlichkeiten und Testeigenschaften. Seine in einer sehr lehrreichen Publikation (1) mit dem Titel «Der Intuition auf der Spur? – Das Bayes'sche Theorem und die Diagnostik in der Grundversorgung» geäusserten Gedanken haben trotz anhaltend hoher Aktualität und klinischer Bedeutung bedauerlicherweise bis heute noch keine weitere Resonanz erfahren.

## Sherlock-Holmes-Methode am klinischen Beispiel

Ausgehend von der Erfahrung im Praxisalltag, dass das degenerativ veränderte Kniegelenk in der Regel nicht nur eine chondrale, sondern auch eine meniskale Schädigung aufweist und die jeweilige Symptomatik zugeordnet werden muss, befassen wir uns im Folgenden mit der spannenden Anwendung der Sherlock-Holmes-Methode, welche einen sinnvollen Ablauf der Diagnostik aufgrund der Anamnese und der klinischen Befunde ermöglicht. Im Vordergrund stehen die Bemühungen, wegen der besseren arthroskopischen Behandlungsoptionen speziell den Stellenwert der meniskalen Schädigungen zu erkennen. Zweifellos wäre es bequemer, nach der «Radarschirm-Methode» vorzugehen und jeden Patienten mit Knieschmerzen in die MR-Abklärung zu schicken. Es gilt jedoch in Erinnerung zu behalten, dass die MR-Bilder die klinische Relevanz nicht schlüssig abbilden können, dass – wie bei jedem Test – falschpositive und falschnegative Befunde entstehen und letztlich aus ökonomischer Sicht eine kritische Indikationsstellung für das MRI ein Gebot der Stunde ist. Die Sherlock-Holmes-Methode scheint verlässlich genug, unnötige und somit teure Bilddiagnostik in der grossen Mehrzahl der Fälle zu vermeiden (2).

## Informationsgewinn bei Kniegelenkerguss

Am Beispiel einer Schwellung des Kniegelenks im Sinne eines Gelenkergusses als Bestandteil der klinischen Diagnosestellung bei einem schmerzhaften Kniegelenk auf der Suche nach einer Meniskusschädigung sollen die Begriffe praktisch durchexerziert werden.

Der Test soll dichotom sein, entweder ist ein Erguss vorhanden, oder es ist kein Erguss vorhanden. Ein Gelenkerguss kann auch aus anderen Gründen als bei einer Meniskusschädigung vorliegen, zum Beispiel bei Knorpelschädigung, freien Gelenkkörpern, rheumatoider Arthritis oder Borreliose. Auf der anderen Seite gibt es Meniskusschädigungen ohne Gelenkerguss. Welchen Informationsgewinn gibt uns nun das Vorliegen eines Ergusses hinsichtlich der Diagnose einer Meniskusschädigung? Das folgende Zahlenbeispiel in der Vierfeldertafel ist fiktiv unter der Modellrechnung von 500 Knieschmerzpatienten im Umfeld einer Hausarztstunde gedacht (Abbildung 3).

Die Vortestwahrscheinlichkeit ist die Wahrscheinlichkeit einer Meniskusschädigung bei einem Patienten mit Knieschmerzen beim Eintritt ins Sprechzimmer, bevor er untersucht wird. Gehen wir einmal davon aus, dass dieser Wert

		Meniskusschaden		
		+	-	
Gelenkerguss	+	50 a	100 b	150
	-	50 c	300 d	350
		100	400	500

Abbildung 3: Die Vierfeldertafel – fiktive Modellrechnung bei 500 Knieschmerzpatienten

20 Prozent (100 von 500) beträgt. Die anderen 400 verteilen sich auf andere Krankheitsbilder wie fortgeschrittene Arthrose, entzündliche Gelenkerkrankungen, beginnende Osteonekrose, ins Knie ausstrahlende Coxarthrose oder spondylogene Schmerzen, muskuläre Knieschmerzen und so weiter. Ein Kniegelenkerguss tritt in unserem Beispiel bei 150 von 500 Personen mit Knieschmerzen (30%) auf. Unter diesen 150 Testpositiven haben 50 einen Meniskusschaden, somit beträgt die gesuchte Nachtestwahrscheinlichkeit beziehungsweise der positiv prädiktive Wert 33 Prozent. Vor der Untersuchung betrug die Wahrscheinlichkeit eines Meniskusschadens 20 Prozent, der Informationsgewinn als Differenz der beiden Krankheitswahrscheinlichkeiten beträgt bei Vorliegen eines Ergusses somit 13 Prozent. Dieser bescheidene Wert erklärt sich durch die eher mässigen Testeigenschaften. Es sind immer noch viele Falschnegative und Falschpositive vorhanden. Die Sensitivität beträgt 50 Prozent, die Spezifität 75 Prozent. Die LR pos (Sensitivität/100-Spezifität) errechnet sich somit aus  $50/25 = 2$ . Mit diesem Test allein kann man also noch keine Entscheidungsschwelle zur Arthroskopie überschreiten. Im Fagan-Nomogramm lässt sich der Informationsgewinn von 13 Prozent bei der LR pos von 2 sehr schön ablesen (Abbildung 4).

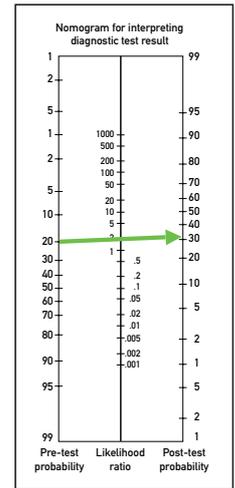


Abbildung 4: Fagan-Nomogramm (Gelenkerguss)

## Die Testserie

Auch bei einer Knieabklärung haben jede anamnestische Frage und jeder Untersuchungsbefund, welche im diagnostischen Prozess zur Anwendung kommen, eine eigene Sensitivität und Spezifität. In der Testserie lassen sich die LR der einzelnen Tests multiplizieren. Dadurch kann der Informationsgewinn erheblich gesteigert werden, um letztlich einmal die Entscheidungsschwelle zu überschreiten. Das Ziel müsste sein, nach Erhebung der Anamnese, Durchführung der klinischen Untersuchung und eventuell nach konventionellem Röntgenbild eine Nachtestwahrscheinlichkeit von 90 Prozent oder mehr zu erreichen, um die Indikation zur Arthroskopie zu stellen, bevor man sich der Frage der MRI-Indikation stellt. In der nächsten Folge werden wir uns mit dieser Zielsetzung auseinandersetzen. ❖

Dr. med. Luzi Dubs

Facharzt für Orthopädische Chirurgie FMH

Wissenschaftlicher Beirat Ars Medici

Merkurstrasse 12, 8400 Winterthur

E-Mail: dubs.luzi@bluewin.ch

Interessenkonflikte: keine

Literatur:

1. Pewsner D. et al.: Der Intuition auf der Spur? – Das Bayes'sche Theorem und die Diagnostik in der Grundversorgung. Praxis (Schweiz Med Forum) Nr. 3, 17. Januar 2001, 41–52.
2. Dubs L.: Die Entscheidung zur arthroskopischen Meniskusoperation mit der «Sherlock-Holmes-Methode». SAEZ 2003; 84: Nr. 12, 541–544.