

8.2 Exkurs Stichprobenprüfungen

Methodische Überlegungen zu Stichprobenprüfungen am Beispiel der Koloskopie

Stichproben sind ein vom Gesetzgeber vorgesehenes und etabliertes Instrument zur Sicherung und Förderung der Qualität ärztlicher Leistungen. Dabei werden ärztliche Dokumentationen erbrachter Leistungen zufällig ausgewählt und nach definierten Qualitätskriterien gutachterlich bewertet. Solche Stichproben wurden mit der Überarbeitung der Qualitätssicherungsvereinbarung Koloskopie eingeführt. Sie sollen der Aufrechterhaltung der fachlichen Befähigung und Fertigkeiten von koloskopierenden Ärzten in der vertragsärztlichen Versorgung dienen.

Die Gestaltung eines Stichprobenverfahrens ist abhängig von dem konkreten medizinischen Versorgungsaspekt. So waren bei Erarbeitung des Stichprobenverfahrens folgende Rahmenbedingungen zu beachten:

- > Die Anzahl der Koloskopieleistungen (2006: über 1,2 Millionen) ist hoch. Sie werden von knapp 3.000 niedergelassenen Ärzten erbracht.
- > 17 Qualitätssicherungskommissionen können die Überprüfungen im Auftrag der Kassenärztlichen Vereinigungen anhand von Bild- oder Videodokumentationen durchführen.
- > Bei (wiederholt) festgestellten Mängeln besteht die Möglichkeit zur Wiederholung der Überprüfung oder sogar zum Widerruf der Genehmigung zur Durchführung von Koloskopien.

Ferner mussten zur Beurteilung einer Stichprobe angemessene Qualitätskriterien für eine fachgerechte Koloskopie festgelegt werden. Danach ist eine Koloskopie ordnungsgemäß erbracht, wenn die Koloskopie

- > ohne Mängel war, das heißt, die Sicht nicht durch Verschmutzung des Dickdarms oder mangelhafte Optik behindert war, und
- > vollständig unter Darstellung des gesamten Dickdarms einschließlich des Blinddarms durchgeführt wurde.

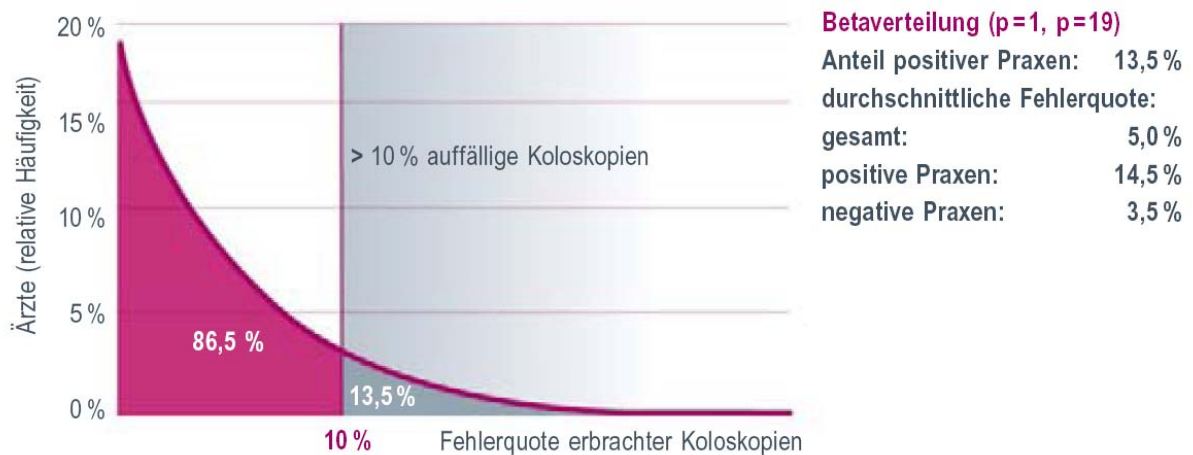
Kontrollierte Studien zeigen, dass eine unangemessene Reinigung des Dickdarms nicht nur einen negativen Einfluss auf Schwierigkeit und Dauer der Untersuchung hat, sondern vor allem die Entdeckungswahrscheinlichkeit von Polypen deutlich reduziert. Die Qualität der Koloskopie hinsichtlich ihrer Mängelfreiheit und Totalität kann dokumentiert werden, indem charakteristische anatomische Strukturen (Bauhin'sche Klappe, Zoekumtriangel, Appendixorifizium) bei Erreichen des Blinddarms fotografiert werden. Allerdings sind solche Fotografien für einen Gutachter in etwa zehn Prozent der Fälle nicht eindeutig zu bewerten (Rex 2000), selbst wenn der koloskopierende Arzt bei der Untersuchung den Blinddarm mit Sicherheit identifiziert hat, das heißt, die Koloskopie vollständig und korrekt durchgeführt wurde. Diese Unsicherheit bei der gutachterlichen Bewertung wird zurückgeführt auf anatomische Variationen zwischen Patienten mit der Folge, dass charakteristische Strukturen des Blinddarms bei einigen Patienten nicht klar dargestellt werden können.

Daraus ergibt sich folgende Definition von guter Qualität koloskopischer Leistungen:

Gute Qualität von einem Arzt durchgeführter Koloskopien liegt vor, wenn bei mindestens 90 Prozent der eingereichten Dokumentationen die vorgegebenen charakteristischen Strukturen des Blinddarms identifiziert werden können.

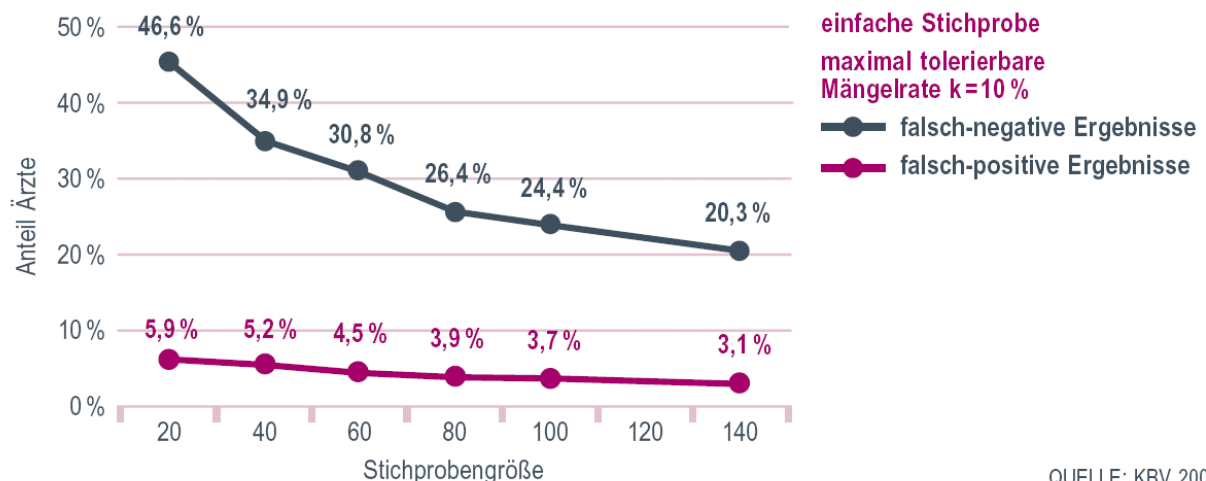
Aufgrund der Aufwändigkeit der Überprüfung und der hohen Anzahl von Leistungen wurde festgelegt, dass eine Stichprobe der erbrachten Leistungen bei allen koloskopierenden Ärzten gerechtfertigt ist. Allerdings muss das Stichprobenergebnis hinreichend genau sein, da übersehene Ärzte mit mangelhaften Koloskopien (= falsch-negative Ergebnisse) zur schlechteren Erkennung von Frühformen des Darmkrebses führen. Auch muss die mögliche Sanktionierung zu Unrecht einer schlechten Qualität bezichtigter Ärzte (falsch-positive Ergebnisse) im Rahmen eines rechtssicheren Verfahrens verhindert werden.

ABB. 1 ANNAHME: RELATIVE VERTEILUNG AUFFÄLLIGER KOLOSKOPIEN



QUELLE: KBV 2007

**ABB. 2 TESTGÜTE IN ABHÄNGIGKEIT VON DER STICHPROBENGROSSE
 FALSCH-POSITIV UND FALSCH-NEGATIV**

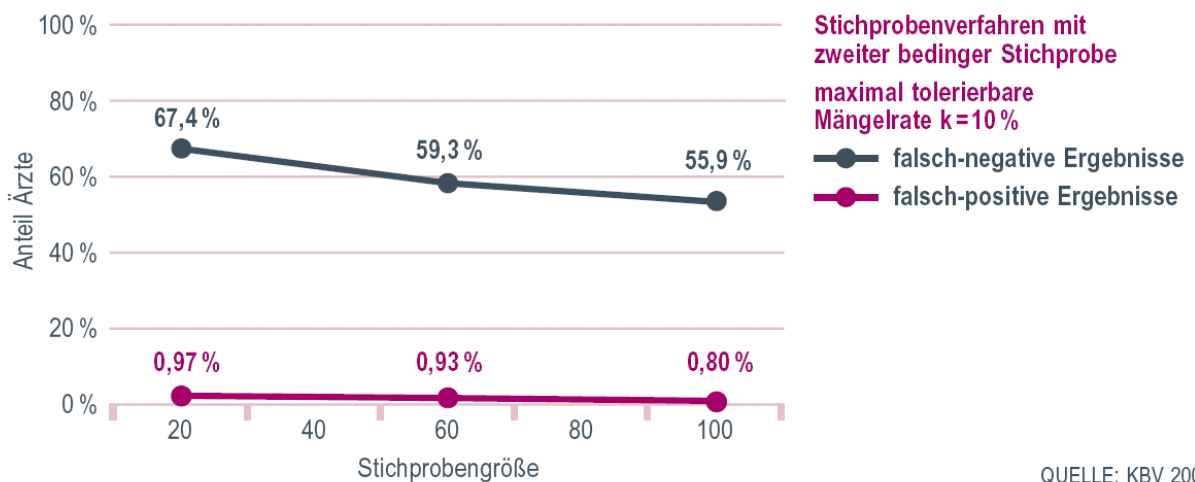


QUELLE: KBV 2007

ABB. 3 STICHPROBENVERFAHREN MIT ZWEITER BEDINGTER STICHPROBE



ABB. 4 TESTGÜTE IN ABHÄNGIGKEIT VON DER STICHPROBENGROSSE DER ZWEITEN BEDINGTEN STICHPROBE FALSCH-POSITIV UND FALSCH-NEGATIV



QUELLE: KBV 2007

Mit statistischer Unterstützung des Institutes für Statistik und Ökonometrie der Universität Hamburg (Prof. Karl Wegscheider) sind PD Dr. Rupert Pfandzelter und Andreas Walter aus der Abteilung Qualitätssicherung der KBV der Frage nachgegangen, wie Umfang und Verfahren der Stichprobenziehung zu wählen sind, um eine ausreichende Testgüte bei vertretbarem Aufwand zu gewährleisten. Hierzu wurden analog zu einem diagnostischen Test wesentliche Kenngrößen von unterschiedlichen Stichprobenverfahren berechnet, wie die falsch-positive und falsch-negative Rate. Die Berechnungen erfolgten analytisch beziehungsweise bei komplexen Stichprobenszenarien numerisch mittels Computersimulationen. Zur Berechnung der Testgüte vergleicht man die Testergebnisse mit der tatsächlichen Situation beziehungsweise dem sogenannten Goldstandard. Je höher die Übereinstimmung desto besser ist die Testgüte. Da im vorliegenden Fall die Verteilung tatsächlicher Mängelraten der koloskopierenden Ärzte nicht bekannt ist, ist eine Annahme darüber notwendig, wie sich die Mängelraten unter allen koloskopierenden Ärzten verteilen könnten. Als den Berechnungen zu Grunde liegende Verteilung wurde eine sogenannte Betaverteilung gewählt (Abb. 1). Danach würden 13,5 Prozent aller Ärzte mehr als zehn

Prozent mangelhafte Koloskopien erbringen bei einer durchschnittlichen Mängelrate aller Ärzte von fünf Prozent.

Die Berechnungen zeigen, dass eine einfache einstufige Stichprobenziehung zu kaum akzeptablen Werten für die falsch-positiven und falsch-negativen Raten führt, wenn der Aufwand für Gutachter und Arzt begrenzt bleiben soll. Bei einer Stichprobe von 20 Koloskopien je Arzt und einer noch akzeptablen Mängelrate von zehn Prozent liegt die falsch-positive und falsch-negative Rate bei 5,9 beziehungsweise 46,6 Prozent (Abb. 2). Selbst bei Verdoppelung der Stichprobe auf 40 ergeben sich noch inakzeptable Werte (5,2 beziehungsweise 34,9 Prozent). Um die aus Ärztesicht besonders bedeutsame falsch-positive Rate weiter zu reduzieren, wurde ein Stichprobenverfahren mit einer zweiten bedingten Stichprobe analysiert. Diejenigen Ärzte, die in der ersten Stichprobe eine höhere Mängelrate als zehn Prozent hatten, werden einer zweiten Stichprobe unterzogen (Abb. 3). Hier liegt die falsch-positive Rate lediglich bei 0,97 Prozent (Größe beider Stichproben: 20; Mängelrate: zehn Prozent; Abb. 4) bei einem zusätzlichen Prüfaufwand von 12,3 Prozent bezogen auf eine solitäre Prüfung. Eine weitere Vergrößerung der zweiten Stichprobe zeigt keine nennenswerte Verbesserung der Werte. Durch dieses konsekutive Stichprobenverfahren steigt jedoch die Rate der falsch-negativen Ergebnisse auf 67,4 Prozent an. Hierbei ist allerdings zu berücksichtigen, dass diese Stichprobenziehung jedes Jahr wiederholt wird. So beträgt nach fünf Jahren die gehäufte (kumulierte) Rate an falsch-negativen Ergebnissen noch 13,9 Prozent. Das bedeutet, dass nach fünf Jahren ungefähr jeder siebte Arzt, der jedes Jahr mehr als zehn Prozent mangelhafte Koloskopien erbracht hat, im Rahmen der Stichprobenprüfung nicht auffällig geworden wäre. Außerdem sinkt die Wahrscheinlichkeit, ungerechtfertigt nicht aufzufallen, mit zunehmender Mängelrate. Während die falsch-negative Rate bei allen Ärzten mit einer Mängelrate von über zehn Prozent bei 67,4 Prozent liegt, beträgt sie bei Ärzten, die über 20 Prozent mangelhafte Koloskopien erbringen, nur noch 23,9 Prozent (Abb. 5). Nach fünf Jahren würden von diesen Ärzten lediglich 0,07 Prozent unentdeckt bleiben.

Aufgrund dieser Ergebnisse haben sich die Partner der Bundesmantelverträge darauf geeinigt, die Überprüfung der erbrachten Koloskopien durch Anwendung des Stichprobenverfahrens mit zwei aufeinander folgenden bedingten Stichproben mit jeweils 20 zufällig ausgewählten Koloskopien durchzuführen. Angesichts der bei einem Stichprobenverfahren immer zu erwartenden – wenn auch in diesem Fall niedrigen – Rate falsch-positiver Ergebnisse ist im Falle des Nichtbestehens der beiden Stichprobenprüfungen eine Vollerhebung im folgenden Jahr vorgesehen (Abb. 6).

Allgemein lässt sich zu Stichproben in der Qualitätssicherung sagen, dass im Gegensatz zu Vollerhebungen Stichproben lediglich Wahrscheinlichkeitsaussagen erlauben. Neben der falsch-negativen Rate ist die falsch-positive Rate dann von besonderer Bedeutung, wenn eine bestimmte in der Stichprobe entdeckte Mängelrate grundsätzlich zu akzeptieren ist und wenn auf der Basis einer positiven Stichprobe Konsequenzen zu ergreifen sind. Auf die jeweilige Problemstellung angewandte Wahrscheinlichkeitsrechnungen gestatten es, einen für die erwarteten Ziele einer Stichprobenprüfung günstigen Arbeitspunkt hinsichtlich Umfang und Verfahren der Stichprobenziehung zu finden.

**ABB. 5 TESTGÜTE IN ABHÄNGIGKEIT VOM ANTEIL AUFFÄLLIGER KOLOSKOPIEN
 FALSCH-NEGATIV**

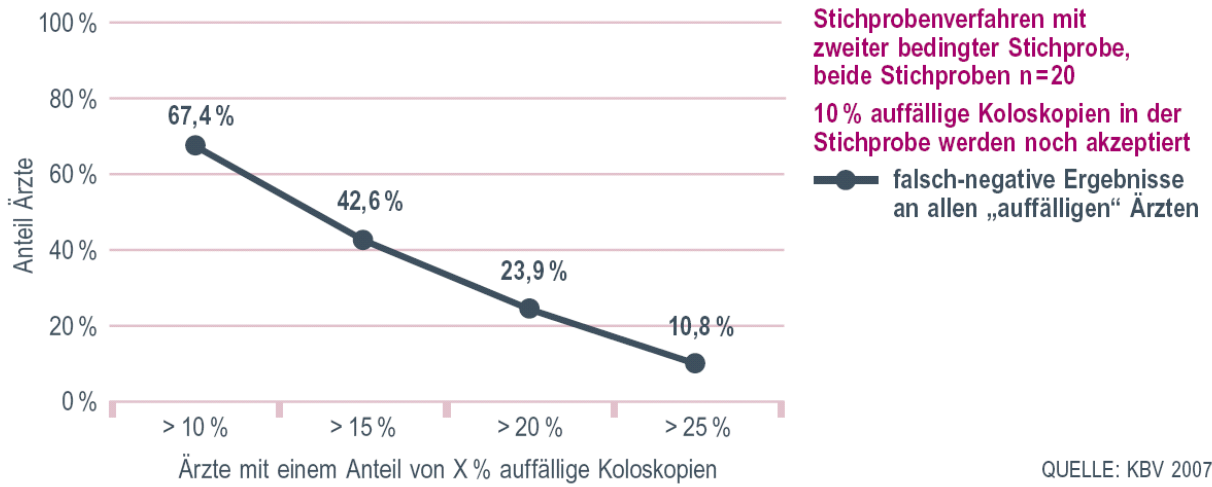


ABB. 6 PRÜFVERFAHREN IN DER KOLOSKOPIE-VEREINBARUNG

