

Immunologische Stuhlbluttests zur Darmkrebsfrüherkennung

Schnelltests sind keine vertretbare Option für ein qualitätsgesichertes Programm

Ulrike Haug, Nikolaus Becker

Das Krebsfrüherkennungs- und -registergesetz (KFRG) sieht die Einführung eines organisierten, qualitätsgesicherten Darmkrebs-Screenings vor. Dabei werden fäkale immunologische Tests auf Hämoglobin (FITs) zum Einsatz kommen. Dieser Beitrag erläutert, warum im Rahmen eines qualitätsgesicherten Screenings keine Schnelltests verwendet werden sollten.

Unterschiedliche immunologische Stuhlbluttests

Immunologische Stuhlbluttests | Im Vergleich zu Guajak-basierten Stuhlbluttests (z.B. Hämocult®) bieten FITs u.a. Vorteile hinsichtlich diagnostischer Wertigkeit und Teilnahmeraten [1, 2]. Eine grundlegende Frage ist jedoch, welche Art von FIT zum Einsatz kommt:

- ▶ qualitative Verfahren in Form von Schnelltests (auch Point-of-Care-Tests genannt) oder
- ▶ quantitative Tests.

Quantitative Verfahren | Bei quantitativen FITs werden die Stuhlproben im Labor unter standardisierten Bedingungen i.d.R. automatisiert ausgewertet; der Hämoglobingehalt wird quantitativ bestimmt. Liegt er oberhalb des Schwellenwertes, der für das Programm vorgegeben ist, wird das Ergebnis als positiv gewertet.

Schnelltests | Die qualitativen FITs hingegen erfordern keine Laboranalytik. Sie sind so konzipiert, dass sie in Arztpraxen bzw. grundsätzlich auch von Laien ausgewertet werden können. Die Auswertung basiert auf einem immunochromatographischen Verfahren. Dabei trägt man die gelöste Stuhlprobe auf eine dazugehörige Testkassette auf; das Testergebnis wird visuell als positiv oder negativ klassifiziert.

Anders als bei den quantitativen FITs ist damit der Schwellenwert bei qualitativen FITs im Testsystem verankert und durch den Hersteller festgelegt, d. h. er ist nicht unmittelbar durch das Programm kontrollierbar.

Problematik der Schwellenwerte

Kontrolle des Schwellenwerts entscheidend | Die Kontrolle des Schwellenwertes durch das Programm vs. den Hersteller ist damit ein wesentlicher Unterschied zwischen qualitativen und quantitativen FITs. Er ist für die Qualitätssicherung von zentraler Bedeutung. Es erscheint im Rahmen ei-

nes qualitätsgesicherten Programms nicht vertretbar, dass Teilnehmer A einen Test mit einer niedrigen Spezifität (d. h. einer hohen Rate falsch-positiver Ergebnisse) und einer hohen Sensitivität erhält, Teilnehmer B aber einen Test mit einer hohen Spezifität und einer niedrigen Sensitivität. Dies wäre jedoch die Konsequenz, wenn der Schwellenwert variiert.

Möglichkeiten zur Kontrolle | Bezüglich der Kontrolle des Schwellenwertes ist es wichtig, die Phase vor der Programm-Zulassung des Tests von der Phase zu unterscheiden, in der das Screening-Programm bereits im Gange ist. Vor der Zulassung sind Studien erforderlich, bei denen eine hinreichend große Anzahl an Personen den jeweiligen Stuhlbluttest durchführt und sich (unabhängig vom Testergebnis) einer Koloskopie unterzieht. Daraus lassen sich Sensitivität und Spezifität ermitteln, d. h. es kann festgestellt werden, ob der Schwellenwert des Tests zum Zeitpunkt der Zulassung adäquat ist.

- ▶ Nach der Zulassung kann bei quantitativen FITs durch das Programm dauerhaft sichergestellt werden, dass dieser Schwellenwert beibehalten wird.
- ▶ Bei qualitativen FITs hingegen stellt sich die Frage, ob der Schwellenwert im laufenden Programm tatsächlich beibehalten wird, zumal es zu produktionsbedingten Verschiebungen oder Chargen-Schwankungen kommen könnte.

Schwankungen im Schwellenwert | Die Bedenken bezüglich eines inadäquaten Schwellenwertes bei qualitativen FITs sind nicht unberechtigt. Insbesondere die Festsetzung und dauerhafte Einhaltung eines Schwellenwertes, der klinisch zu einer definierten Rate falsch-positiver und falsch-negativer Ergebnisse führt, ist bei qualitativen FITs problematisch. Dies haben mehrere Studien gezeigt [3–9].

- ▶ Beispielsweise wurden in einer Studie aus Deutschland sechs qualitative FITs untersucht. Davon hatten nur zwei einen Schwellenwert, der zu einer Falsch-positiv-Rate unter 10% führte, während die anderen FITs eine Falsch-positiv-Rate von 20–40% aufwiesen [3].

- ▶ In einer Studie von Levy et al. [6] sollte bei 1026 Teilnehmern ein qualitativer FIT durchgeführt werden. Der Test musste im Laufe der Studie mehrfach gewechselt werden – meist weil es Qualitätsprobleme mit dem jeweiligen qualitativen FIT gab.
- ▶ Für zwei qualitative FITs, die laut Hersteller beide auf einen Schwellenwert von 3 µg/g Stuhl eingestellt waren, zeigten sich im direkten Vergleich Spezifitäten von 64% bzw. 93% [9].

Qualitätssicherung

Kontrolle nach Zulassung | Um die Qualität der FITs nach der Zulassung zu sichern, wurde eine Datenerfassung vorgeschlagen, die es ermöglicht, verschiedene klinische Prozessparameter zu ermitteln. Diese Prozessparameter erlauben es jedoch nicht, Schwankungen im Schwellenwert bei qualitativen FITs zuverlässig zu erkennen, zumal sie nicht eindeutig interpretierbar sind.

Prozessparameter | Im laufenden Programm führen Screening-Teilnehmer zunächst einen Stuhltest durch. Nur bei positivem Ergebnis erfolgt eine Koloskopie. Dokumentiert man die Testergebnisse und die Befunde der Abklärungs-Koloskopien, lassen sich für die jeweiligen Stuhltests folgende Prozessparameter berechnen:

- ▶ die Positivitätsrate (PR): Anzahl positiver Testergebnisse geteilt durch die Anzahl der Getesteten
- ▶ die Detektions- oder Erkennungsrate (DR): Anzahl der richtig positiven Testergebnisse geteilt durch die Anzahl der Getesteten
- ▶ der positiv-prädiktive Wert (PPV): Anzahl der richtig positiven Testergebnisse geteilt durch die Anzahl aller positiven Testergebnisse

Begrenzt aussagefähig | Aus einer Veränderung dieser Parameter kann man allerdings nicht unmittelbar auf einen veränderten Schwellenwert schließen: Sie werden nicht nur vom Schwellenwert, sondern auch von der Prävalenz der Zielläsionen (kolorektale Karzinome und Adenome) bei den Getesteten beeinflusst. Diese Prävalenz wiederum ist u. a. abhängig vom

- ▶ Alter und Geschlecht der Getesteten
- ▶ ebenso wie von deren Screening-Vorgeschichte (inkl. Koloskopie-Inanspruchnahme).

In einer niederländischen Kohorte beispielsweise variierte die PR von 5–12%, die DR von 1–5% und der PPV von 22–42% in Abhängigkeit von der Screening-Vorgeschichte der Teilnehmer [10].

Selbst wenn Informationen zu Alter, Geschlecht und Screening-Vorgeschichte vorliegen (insbesondere letzteres ist eine Herausforderung), ist die Schätzung der Prävalenz der Zielläsionen im jeweiligen Kollektiv der Getesteten mit weiteren Unsicherheiten und Annahmen verbunden.

Insofern wäre es angreifbar und nicht stichhaltig, diese Parameter heranzuziehen, um auf einen inadäquaten Schwellenwert hinzuweisen. Diese Problematik betrifft auch eine kürzlich vorgeschlagene Methode zur Qualitätssicherung von FITs [11]. Eine weitere Einschränkung der Nutzung dieser Prozessparameter zur Qualitätssicherung ist die Erkennung von Chargen-Schwankungen. Da es zur präzisen Schätzung der Parameter größerer Fallzahlen bedarf, könnten Chargen-Schwankungen nicht zuverlässig von Zufallsschwankungen unterschieden werden.

Konsequenzen | Folglich gibt es bei qualitativen FITs keine Möglichkeit sicherzustellen, dass deren Sensitivität und Spezifität dauerhaft den Vorgaben des Programms entspricht. Dies ist nicht nur problematisch im Hinblick auf die Qualitätssicherung, sondern auch hinsichtlich der informierten Entscheidung im Sinne des KFRG.

Beispielsweise kann den zu screenenden Personen keine verlässliche Auskunft erteilt werden, mit welcher Wahrscheinlichkeit bei Nichterkrankten ein falsch-positives Testergebnis zu erwarten ist.

Zu bedenken sind auch die Unwägbarkeiten bzgl. der erforderlichen Ressourcen und Koloskopie-Kapazitäten, falls Testverfahren eingesetzt werden, deren Spezifität nicht den Planungen des Programms entspricht. Zudem gibt es bei qualitativen FITs Probleme bei der Qualitätssicherung, die sich aus der Variabilität (z.B. verzögerte Ablesezeit) und Subjektivität der Testauswertung ergeben.

Bewertung der Gegenargumente

Zweifelhafte Argumente | Bei Diskussionen zur Frage, ob quantitative oder qualitative FITs zur Darmkrebsfrüherkennung eingesetzt werden sollen, werden diese zentralen Aspekte in der Regel nicht thematisiert. Stattdessen werden Argumente angeführt, die rational schwer nachvollziehbar sind oder denen keine Evidenz zugrunde liegt.

Arztpraxis versus Labor | Als Vorteil qualitativer FITs wurde angeführt, dass die Testauswertung und Ergebnismitteilung direkt in der Arztpraxis erfolgen kann. Bei quantitativen FITs wird die Stuhlprobe zunächst an ein Labor geschickt und dort ausgewertet. In Anbetracht der effizienten Auswertungsverfahren liegt aber auch bei quantitativen FITs das Testergebnis in kurzer Zeit vor und kann an die Arztpraxis übermittelt werden.

Grundsätzlich ist jedoch zu hinterfragen, ob es für ein Screening-Programm ökonomisch vertretbar ist, dass alle Teilnehmer zur Ergebnismitteilung in die Arztpraxis kommen.

Dabei ist zu bedenken, dass bei ca. 90% der Screening-Teilnehmer ein negatives Testergebnis zu erwarten ist. Effizienter scheint ein Vorgehen wie beim Zervixkarzinom-Screening, bei dem Patienten nur dann wieder einbestellt werden, wenn das Testergebnis positiv ist.

Datenschutz | Auch datenschutzrechtliche Bedenken wurden geäußert, wenn der Teilnehmer die Stuhlprobe an ein Labor schickt und das Ergebnis dann vom Labor an eine Arztpraxis übermittelt wird. Dies sollte jedoch – analog zum Vorgehen in anderen Ländern – durch ein unterschriebenes Formular zu lösen sein, das der Teilnehmer zusammen mit der Stuhlprobe einstellt. Daraus geht hervor, welchem Arzt das Ergebnis mitgeteilt werden soll bzw. ob es der Teilnehmer selbst unmittelbar erfahren soll. Ansonsten gäbe es noch die Möglichkeit, die Probe in einer Arztpraxis abzugeben, die sie dann ans Labor schickt. Dann könnten zwar Datenschutzregelungen wie bei sonstiger Laboranalytik angewandt werden, aber wie erwähnt hätte dieses Verfahren andere Nachteile.

Postversand | Des Weiteren wurde zu bedenken gegeben, dass der Postversand von Stuhlproben, der bei quantitativen FITs i. d. R. erforderlich ist, problematisch sein könnte und außerdem zu einer nicht vertretbaren Kostensteigerung führen würde. Seitens der Deutschen Post gibt es jedoch spezifische Vorgaben für die Verpackung und den Versand bei Vorsorgeuntersuchungen basierend auf Stuhlbluttests [12]. Erforderlich sind:

- ▶ eine auslaufsichere Innen- und Zwischenverpackung
- ▶ aufsaugendes Material
- ▶ eine reißfeste Außenverpackung
- ▶ Versand als Groß- oder Maxibrief zu erfolgen.

Unter Berücksichtigung dieser Vorgaben ist zu erwarten, dass Verpackungs- und Versandkosten insgesamt unter 2,45 € liegen.

Kosten | Was den Kostenvergleich zwischen qualitativen und quantitativen FITs betrifft, sollte nicht nur die Logistik, sondern auch die Analytik betrachtet werden. Dahingehend ist nicht auszuschließen, dass qualitative FITs bei großflächigem Einsatz mit höheren Kosten einhergehen als quantitative FITs.

Rolle des Hausarztes | Auch wurde angemerkt, dass ein Test, der im Labor ausgewertet wird, der bedeutenden Stellung der niedergelassenen Ärzte im Bereich Prävention nicht gerecht würde. Die zusätzliche Beteiligung einer großen Anzahl an Laborärzten könne das Programm zu komplex und kostenintensiv machen.

Zunächst sei festzuhalten, dass die Sicherstellung einer qualitätsgesicherten Analytik oberste Priorität haben sollte.

Es steht aber außer Frage, dass niedergelassene Ärzte eine wichtige Rolle spielen – insbesondere sollten sie Personen mit einem positiven Testergebnis zur Abklärungskoloskopie motivieren.

Labor-Analytik | Ähnlich wie in anderen europäischen Ländern empfiehlt sich die Einrichtung einer begrenzten Anzahl an Screening-Labors, die auf die Testauswertung spezialisiert sind. Daraus ergeben sich zahlreiche Vorteile im Vergleich zur dezentralen Analytik, z. B. hinsichtlich (Kosten-) Effizienz und Datenmanagement.

- ▶ Für eine Labor-Analytik spricht außerdem, dass dem Teilnehmer ein weiterer Gang zum Arzt erspart wird, d. h. die Logistik ist einfacher. Dies wirkt der Benachteiligung von Personen entgegen, die in strukturschwachen Regionen leben, und vermeidet weitere Selektionseffekte.
- ▶ Auch ist davon auszugehen, dass das direkte Einsenden der Proben an ein Labor die Zeit von der Probensammlung bis zur Probenanalyse verkürzt und das angestrebte, maximale Zeitfenster von 3 Tagen besser eingehalten werden kann.

Untersuchungen in England zeigten, dass bei einer Außentemperatur von 35 °C der Hämoglobin-Gehalt von gelagerten Stuhlproben innerhalb von 2–3 Tagen um 50% abnimmt [13]. Insofern ist insbesondere in den heißen Sommermonaten eine Überschreitung des Zeitfensters zu vermeiden.

Anzahl erforderlicher Arztbesuche | Nicht nachvollziehbar sind Argumente, dass quantitative FITs mit mehr Arztbesuchen einhergehen als qualitative. Dies würde eher für letztere zutreffen, wenn Teilnehmer zur Probenübergabe und Testauswertung eine Arztpraxis aufsuchen. Auch Argumente, dass qualitative FITs mit höheren Teilnahmeraten einhergehen, entbehren jeglicher Grundlage. Mit dem postalischen Versand von quantitativen FITs wurden in den Niederlanden Teilnahmeraten von 60% erzielt [1, 2]. Es ist nicht geklärt, ob mit einer Testauswertung in der Arztpraxis – wie es für qualitative FITs vorgesehen ist – ähnliche Teilnahmeraten erzielt werden. Das Vorgehen könnte auch mit einem größeren Selektionseffekt und niedrigeren Teilnahmeraten einhergehen.

Konsequenz für Klinik und Praxis

- ▶ Qualitative FITs bringen keine erkennbaren Vorteile.
- ▶ Vielmehr haben sie im Vergleich zu quantitativen FITs deutliche Nachteile hinsichtlich der Qualitätssicherung.
- ▶ Insofern erscheint der Einsatz qualitativer FITs im Rahmen eines qualitätsgesicherten Screenings nicht vertretbar.
- ▶ Es wäre ein Sonderweg, der von den europäischen Leitlinien abweicht [14] und der Darmkrebsfrüherkennung in Deutschland nachhaltig schaden könnte.



Prof. Dr. sc. hum. Ulrike Haug

ist Leiterin der Abteilung Klinische Epidemiologie, Leibniz-Institut für Präventionsforschung und Epidemiologie, BIPS GmbH Bremen und Professorin an der Universität Bremen. haug@bips.uni-bremen.de



Prof. Dr. rer. nat. Nikolaus Becker

ist Leiter der Abteilung NCT-Krebsregister, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg. becker@dkfz-heidelberg.de

Vollständiges Literaturverzeichnis unter <http://dx.doi.org/10.1055/s-0042-104048>

Interessenkonflikt

Die Autoren geben an, dass kein Interessenkonflikt besteht.

DOI 10.1055/s-0042-104048
Dtsch Med Wochenschr
2016; 141: 729–731
© Georg Thieme Verlag KG · Stuttgart · New York · ISSN 0012-0472

Literatur

- 1 van Rossum LG, van Rijn AF, Laheij RJ et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. *Gastroenterology* 2008; 135: 82–90
- 2 Hol L, van Leerdam ME, van Ballegooijen M et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. *Gut* 2010; 59: 62–68
- 3 Hundt S, Haug U, Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. *Ann Intern Med* 2009; 150: 162–169
- 4 Tannous B, Lee-Lewandrowski E, Sharples C et al. Comparison of conventional guaiac to four immunochemical methods for fecal occult blood testing: implications for clinical practice in hospital and outpatient settings. *Clin Chim Acta* 2009; 400: 120–122
- 5 Huang Y, Ge W, London V et al. Diagnostic inconsistency of faecal immunochemical tests for haemoglobin in population screening of colorectal cancer. *Clin Chem Lab Med* 2013; 51: 2173–2180
- 6 Levy BT, Bay C, Xu Y et al. Test characteristics of faecal immunochemical tests (FIT) compared with optical colonoscopy. *J Med Screen* 2014; 21: 133–143
- 7 Huang Y, Li Q, Ge W et al. Predictive power of quantitative and qualitative fecal immunochemical tests for hemoglobin in population screening for colorectal neoplasm. *Eur J Cancer Prev* 2014; 23: 27–34
- 8 Park MJ, Choi KS, Lee YK et al. A comparison of qualitative and quantitative fecal immunochemical tests in the Korean national colorectal cancer screening program. *Scand J Gastroenterol* 2012; 47: 461–466
- 9 Tao S, Brenner H. Well adjusted qualitative immunochemical faecal occult blood tests could be a promising alternative for inexpensive, high-quality colorectal cancer screening. *Eur J Cancer Prev* 2013; 22: 305–310
- 10 Kapidzic A, Grobbee EJ, Hol L et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. *Am J Gastroenterol.* 2014; 109: 1257–1264
- 11 Brenner H, Hoffmeister M, Stock C. Neue Tests zur Darmkrebsfrüherkennung. Qualitätssicherung in der Routine ist notwendig und gut realisierbar. *Dtsch Arztebl* 2014; 19: 842–843
- 12 Deutsche Post. Regelungen für die Beförderung von gefährlichen Stoffen und Gegenständen. https://www.deutschepost.de/content/dam/mlm.nf/dpag/images/b/brief_postkarte_national/broschuere_befoerderung_von_gefaehrlichen_stoffen_und_gegenstaenden_teil_1a.pdf (letzter Zugriff 25.10.2015)
- 13 World Endoscopy Organization. Evaluation of quantitative faecal immunochemical tests for Haemoglobin. http://www.worldendo.org/assets/downloads/pdf/activities/fit_reports/gmec_fit_evaluation_report.pdf (letzter Zugriff 06.04.2016)
- 14 European Colorectal Cancer Screening Guidelines Working Group. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. *Endoscopy* 2013; 45: 51–59