

Original-Titel

Efficacy of artificial intelligence in reducing miss rates of GI adenomas, polyps, and sessile serrated lesions: a meta-analysis of randomized controlled trials

Autoren

X.-F. Jin, H.-Y. Ma, J.-W. Shi und J.-T. Cai, *Gastrointestinal Endoscopy* 99:5, pp. 667-675.e1, Mai 2024, doi: 10.1016/j.gie.2024.01.004

Kommentar

Dr. med. Rüdiger Schmitz, Hamburg, 24.6.24

It's deja vu all over again¹.

In der aktuellen Ausgabe von GIE präsentieren Jin et al. eine Metaanalyse des Einsatzes von KI bei der Vorsorgekoloskopie². Anders als die Vielzahl der bekannten Metaanalysen von RCTs betrachten die Autoren hierbei Adenoma Miss Rate (AMR) statt Adenom Detection Rate (ADR). Zusammengefasst – die Erkenntnisse bleiben die bekannten: 1. Endoskopische KI kann Polypen erkennen. 2. Endoskopische KI führt im RCT-Setting zu mehr entdeckten bzw. hier nun weniger verpassten Polypen. 3. Endoskopische KI erhöht die Zahl entdeckter kleiner Läsionen (diminutiver und sessiler Adenome).

Was lohnt am Betrachten dieser Studie? Dass die Metaanalyse von sieben RCTs, von denen sechs positiv sind und eines siebenten mit 5% Gewicht eingeht, wiederum positiv ist, kann nicht überraschen. Dass KI technisch in der Lage ist, Polypen zu erkennen, ist bekannt. Dass KI dies in RCTs als Näherung an die Realität beweisen kann ebenfalls. Dasselbe in Metaanalysen von RCTs und auch in Metaanalysen von Metaanalysen³.

Sind es die relativ deutlichen Effekte, was diese Studie interessant macht – sei es nach Hedges' g oder nach der Zahl der für sich bereits signifikanten Studien (als etwas unsauberes Surrogat)? Immerhin sind sechs Studien bereits einzeln signifikant und der gepoolte Effekt zumindest moderat stark (Hedges' g -0.3 bis -0.4)? Nicht direkt: Die Rate verpasster Adenoma (AMR) wird notwendigerweise in einem Tandem-Design erhoben, d.h. in einem Setting von zwei aufeinander folgenden Untersuchungen. Dies steht im Gegensatz zum Design des klassischen „ADR mit oder ohne Intervention X“-RCT, welcher zwei Gruppen parallel je der Untersuchung mit oder ohne X zuweist. Da ein Tandem-Design notwendigerweise diverse Varianten eröffnet, z.B. der Untersuchungsabfolge oder hinsichtlich der Variation der Untersucher, und AMR weniger einheitlich definiert wird als die ADR, ist das Pooling von AMR-Tandem-Studien methodisch grundsätzlich problematischer ist das als von vergleichsweise homogenen Parallel-ADR-Studien⁴. Auch ist das Tandem-Setting weiter weg vom realen Anwendungsfall und gilt als anfälliger für Biases (inkl. Hawthorne)⁴. Dass die Effekte der analysierten AMR-Studien verglichen mit der Studienlage per ADR stark sind, vollzieht nach, was in verschiedenen Settings bereits untersucht wurde: „more positive results with tandem than with parallel studies“⁴ – nun also auch bei der KI.

Es bleibt nach dieser Studie also dabei: KI kann Polypen erkennen – die im Bild sind. Und KI führt zu mehr entdeckten Polypen – im Studiensetting. Das ist nicht nichts und eine Leistung für sich und ein gewaltiger technischer Erfolg. Auch ist die vorliegende Studie solide gemacht und liefert sicherlich interessante Zahlen. Gleichwohl: Die Polypen Detektion per KI bleibt im Studiensetting gut belegt, muss sich aber weiterhin in der Praxis bewähren und in prospektiven Studien ihren Benefit zeigen. Und die (Vorsorge-)Koloskopie als Gesamtprozess von Einbestellung über Vorbereitung, Untersuchung, ggf. Resektion und Nachverfolgung bleibt mehr als die Detektion von Polypen, die der Untersucher sauber im Bild dargestellt hat. Die dicken Bretter bleiben ungebohrt, auch für die KI.

Dieses mit AMR statt ADR wiederholte (Meta-)Deja vu demonstriert Bekanntes – über die Fähigkeiten von KI und den Charakter von Tandemstudien; es bringt uns, frei nach dem eingangs zitierten Baseballer Yogi Berra, dem home run der Koloskopie-KI noch nicht näher.

Referenzen

1 https://en.wikipedia.org/wiki/Yogi_Berra#%22Yogi-isms%22, abgerufen am 21.06.2024

2 X.-F. Jin, H.-Y. Ma, J.-W. Shi, und J.-T. Cai, „Efficacy of artificial intelligence in reducing miss rates of GI adenomas, polyps, and sessile serrated lesions: a meta-analysis of randomized controlled trials“, *Gastrointestinal Endoscopy* 99:5, pp. 667-675.e1, Mai 2024, doi: 10.1016/j.gie.2024.01.004.

3 K. Zimmermann-Fraedrich und T. Rösch, „Artificial intelligence and the push for small adenomas: all we need?, *Endoscopy* 55:04, pp. 320–323, Apr. 2023, doi: 10.1055/a-2038-7078.

4 K. Zimmermann-Fraedrich u. a., „Designs of colonoscopic adenoma detection trials: more positive results with tandem than with parallel studies - an analysis of studies on imaging techniques and mechanical devices“, *Gut*, S. gutjnl-2020-320984, Mai 2020, doi: 10.1136/gutjnl-2020-320984.

Conflict of Interest

Der Autor Rüdiger Schmitz ist Gründer und Gesellschafter von casuu.health und tätig im Bereich der KI-gestützten Weiterbildung von nicht-muttersprachlichem medizinischen Personal.