

Original-Titel:

Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study

Autoren:

Yamashita et al.; Lancet Oncol. 2021; 22:132-141

Kommentar:

Prof. Dr. Hendrik Bläker, Leipzig, 27.7.2021

Die Mikrosatelliten-Instabilitäts (MSI)-Diagnostik beim kolorektalen Karzinom ist aus zwei Gründen relevant. Sie spielt seit langem eine zentrale Rolle bei der Identifizierung von Lynch Syndrom Patienten und hat darüber hinaus Bedeutung für die Entscheidung über Einsatz und Art einer adjuvanten Therapie in den Stadien II und IV.

MSI Kolonkarzinome sehen unter dem Mikroskop häufig etwas anders aus als das konventionelle Karzinom ohne MSI. Die besonderen, histologischen Charakteristika von MSI Karzinomen sind seit längerem bekannt und erlauben eine Vorhersage zur MSI Wahrscheinlichkeit. Die Frage nach einer MSI Prädiktion bei kolorektalen Karzinomen eignet sich damit auch als Fragestellung für die auf künstlicher Intelligenz (KI) basierenden Bildanalyse 2. Nun ist jüngst zum Thema eine Arbeit von Yamashita et al. in Lancet Oncology¹ mit einem anwendungsorientierten Ansatz erschienen, in dem die die Performance des für die Studie entwickelten MSINet KI-Systems mit der von Pathologen verglichen wurde.

Methodisch ist die Studie klar strukturiert und kurz erklärbar. Zusammengefasst wurde ein KI-basiertes MSI Prädiktions-Modell (MSINet) anhand von konventionell Hämatoxylin-Eosin (HE) gefärbten Schnittpräparaten von 100 kolorektalen Karzinomen aus den Archiven des Pathologischen Instituts der Stanford University entwickelt. Der MSI Status war bekannt, 50 Karzinome waren mit und 50 ohne MSI. Es erfolgte daraufhin die Evaluation der Leistungsfähigkeit des MSINet in der MSI Prädiktion zum einen anhand eines Stanford-internen Kollektiv von 15 Fällen (8 mit und 7 ohne MSI) zum anderen anhand 479 eingescannte Schnittpräparate kolorektaler Karzinome (77 mit und 402 ohne MSI) des öffentlich zugänglichen „The Cancer Genome Atlas“ (TCGA) Projekts. Eingescannte Schnitte von 40 Fällen dieses TCGA Datensätze (20 mit und 20 ohne MSI) wurden zum Vergleich 5 Pathologen zur Ansicht gegeben. Sie sollten einschätzen, ob eine MSI vorliegt.

Im Ergebnis zeigt sich, dass die Einschätzungen der 5 Pathologen im Durchschnitt der Auswertung durch MSINet unterlegen war. Signifikant war der Unterschied bei der Sensitivität und dem negativen prädiktiven Wert bei etwa gleicher Spezifität. Die Performance jedes einzelnen Pathologen war allerdings unterschiedlich, von gleich gut wie MSINet bis so schlecht wie wenn man eine Münze wirft. Dies deutet darauf hin, dass einige der beteiligten Pathologen oder Pathologinnen hinsichtlich der Fragestellung nicht besonders gut eingedacht sind. Für universitäres Pathologie Personal mit einem Schwerpunkt in der Gastrointestinalen Pathologie ist das etwas ungewöhnlich aber die Überlegenheit der KI in der MSI Prädiktion dürfte durchaus gegenüber der pathologischen Versorgung in der Fläche gültig sein. Die Autoren schlussfolgern aus den Ergebnissen, dass ein Einsatz von MSINet als Unterstützung für Pathologen bei einer im TCGA Datensatz nachgewiesenen Sensitivität von 76%, einer Spezifität von 66% und einem negativen prädiktiven Wert von 93.9% sinnvoll ist. Damit, so die Argumentation, könnte man den Umfang der Proben, die einer weitergehenden, immunhistologischen oder molekularen MSI Testung zugeführt werden um 62,8% reduzieren.

Wie sind diese Ergebnisse nun zu werten?

Als erstes muss man festhalten, dass eine MSI nur anhand einer immunhistologischen oder molekularen Testung mit der im klinischen Setting notwendigen Sicherheit diagnostiziert werden kann. Ich nenne diese Tests im Folgenden einmal „echte“ MSI-Tests. Weder die KI noch die traditionelle Histopathologie erkennt einen MSI Tumor am HE Schnitt mit der notwendigen, diagnostischen Sicherheit 2,3. Vielmehr werden morphologische Kriterien geprüft anhand deren Gewichtung eine dichotome Einstufung in eher MSI und eher nicht-MSI Karzinom erfolgt und somit eine Vorauswahl für die weitere, „echte“ MSI Testung getroffen wird. Aus den Zahlen der Studie kann man schlussfolgern, dass MSINet von den 479 Fällen des TCGA Datensatzes 296 als nicht-MSI und 183 als MSI einstuft. Unter den als 183 als MSI eingestuften Fällen finden sich 59 mit nachgewiesener MSI-H, unter den als nicht MSI-eingestuften finden sich 18 mit nachgewiesener MSI-H.

Das entspricht einer falsch-positiv Rate von 68% und einer falsch-negativ Rate von 6%. Am Ende würde das von Yamashita et al. vorgeschlagene Vorgehen zwar 62% der Untersuchungen einsparen, dabei aber 23% (18 von 77 in der Studie) aller MSI Tumoren als nicht-MSI einstufen und damit von der „echten“ MSI Diagnostik ausschließen.

Schon vor KI Zeiten gab es Ansätze klinisch-pathologische Scoring System wie z. B. MSPATH3, die etwas weniger spezifisch (50-60%) aber deutlich sensitiver (>90%) bei der MSI Prädiktion als KI basierte Verfahren sind. Obwohl diese Scoring Systeme besser als MSINet sind, haben sie sich in der Routine nicht durchgesetzt und das hat sowohl inhaltliche als auch methodische Gründe.

Wenn eine Pathologin oder ein Pathologe aus Lynch Syndrom diagnostischen oder allgemein therapeutischen Gründen wissen muss, ob ein kolorektales Karzinom MSI ist, dann wird eine immunhistologische und/oder molekulare MSI Analytik eingeleitet, die am Ende zu einem klaren Ergebnis führt. Das Vorgehen ist Routine. Wieso soll man einen Zwischenschritt wie MSINet einschieben, der mit einer erheblichen diagnostischen Unsicherheit behaftet ist? In Kenntnis der 6% falsch negativen Fälle in dieser Studie, müsste man trotzdem alle Fälle „echt“ testen. Denn eine Fehlerrate dieser Höhe ist im diagnostischen Alltag völlig inakzeptabel. Auch ist der negative prädiktive Wert (NPW) von 94% Augenwischerei. Denn bei einer bekannten 11% igen Prävalenz von MSI Karzinomen (Wert der größten Kohorte in Europa, DACHS5) ist der NPW der KI kaum verschieden von der Ratewahrscheinlichkeit von 89% für MSS.

Die Möglichkeiten der künstlichen Intelligenz sind faszinierend und mit Sicherheit hat die Methodik Einsatzbereiche in der Pathologie. Wenn man das automatisierte Screening von PAP Abstrichen im Rahmen der gynäkologischen Vorsorge als primitive Form der KI versteht, wird die KI in der Pathologie schon seit fast 20 Jahren eingesetzt. Gerade für Fragestellungen, bei denen es darum geht, in einer sehr großen Menge von Einsendungen die wenigen Fälle mit zytologisch definierbaren Zellen, die auf einen Tumor oder dessen Vorstufen hinweisen, zu identifizieren, wäre die KI eine überaus hilfreiche und Qualität sichernde Unterstützung für die Pathologen. Von diesen Fragestellungen gibt es in der Pathologie einige. Solch eine Diagnostik ist ermüdend und zeitintensiv, beides Faktoren, die die Leistungsfähigkeit von Menschen aber nicht von Maschinen beeinflussen.

Zur Studie möchte ich abschließend sagen, dass sie belegt, dass die KI ebenso wie der Mensch, ein MSI Karzinom erraten kann aber eben nicht erkennt. Es gibt MSI Tumoren, vorzugsweise von Lynch Syndrom Patienten 6,7 und Immunsupprimierten, die die typischen morphologischen Kriterien eines MSI Tumors nicht aufweisen und es gibt nicht-MSI Tumoren, die Stigmata von MSI Tumoren haben.

Eine Fehlerrate bei der MSI Prädiktion liegt in der Natur der Sache, daher kann man der KI keine Vorwürfe machen. Eher schon den Autoren der Studie dafür, dass sie nicht auf die offensichtlichen Schwachpunkte von MSINet eingehen.

1. Yamashita et al. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol.* 2021; 22:132-141
2. Kather JN et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med.* 2019;25:1054-1056
3. Halvarsson et al. Clinicopathologic Factors Identify Sporadic Mismatch Repair–Defective Colon Cancers. *Am J Clin Pathol* 2008;129 238–244
4. Jenkins et al. Pathology Features in Bethesda Guidelines Predict Colorectal Cancer Microsatellite Instability: A Population-Based Study. *Gastroenterology.* 2007;133: 48–56
5. Alwers E et al. External validation of molecular subtype classifications of colorectal cancer based on microsatellite instability, CIMP, BRAF and KRAS. *BMC Cancer* 2019;19:681
6. Jass J. HNPCC and sporadic MSI-H colorectal cancer: a review of the morphological similarities and differences. *Fam Cancer* 2004;3:93-100.
7. Yamada R et al. Differences in histological features and PD-L1 expression between sporadic microsatellite instability and Lynch-syndrome-associated disease in Japanese patients with colorectal cancer. *Int J Clin Oncol* 2018;23:504-513.