Consensus-Based Minimum Requirements for the Adoption of a Computational Pathology Algorithm for Tumor Budding in Colorectal Cancer: An Early Health Technology Assessment

Julie E.M. Swillens, PhD[1]; Iris D. Nagtegaal, MD, PhD[2]; Alessandro Lugli, MD, PhD[3]; Jeroen A.W.M. Van der Laak, MD, PhD[2,4]; and Marcia Tummers, PhD[1]

DOI https://doi.org/10.1200/CCI-26-00034

ABSTRACT

PURPOSE Tumor budding (TB) is an independent prognostic biomarker for colorectal cancer (CRC), yet its clinical adoption is limited by labor-intensive and subjective visual scoring. Computational pathology (CPath) algorithms using deep learning offer potential to improve efficiency and reproducibility. Using an international Delphi study, we established consensus requirements to guide validation and clinical adoption of CPath algorithms, aiming to enhance diagnostic accuracy and patient care.

METHODS A two-round international Delphi process was conducted, involving international experts, to reach consensus on predefined statements. In round 1, baseline characteristics were collected and participants were asked to rank statements representing the minimal requirements for the implementation of a CPath algorithm for TB assessment. After completion of this round, participants received a personalized feedback report summarizing interim results for items that lacked consensus. In round 2, participants re-evaluated their initial responses to statements without consensus, in the light of anonymized group feedback provided in their personalized report.

RESULTS Fifty-nine pathologists participated in round 1, with a 90% response rate in round 2. Consensus was reached for 21 of 29 (72%) minimal requirements. All reached consensus by agreement. Eight statements (28%) remained without consensus after round 2.

CONCLUSION Agreement was reached on the technical, organizational, ethical, and legal aspects, although some requirements were not agreed upon. This highlights the importance of evaluations that take specific contexts into account, as well as the need for continuous stakeholder engagement throughout the development and implementation phases.

ACCOMPANYING CONTENT

Appendix

Data Supplement

Accepted July 2, 2026
Published July 31, 2026

JCO Clin Cancer Inform
10:e2600034

INTRODUCTION

Histopathologic assessment of the presence of tumor budding (TB) yields a promising biomarker for colorectal cancer (CRC). TB is defined as isolated single cells or small clusters of up to four tumor cells located at the invasive tumor front.[1,2] These biomarkers possess independent prognostic information, in addition to TNM staging of CRC. TB is of particular importance in early CRC, as it is frequently encountered in population screening programs. The assessment is necessary for improved decision making on therapeutic strategies and for prevention of possible undertreatment and overtreatment.[3] Although TB assessment is recognized in the WHO classification and international clinical guidelines,[4-7] clinical adoption is hampered as human visual scoring remains labor intensive and subjective, resulting in a moderate interobserver agreement in hematoxylin and eosin slides.

TB scoring can be improved by supporting pathologists with computational pathology (CPath) algorithms, which are mostly based on deep learning (DL) technology.[8] In previous research, a DL model was developed to recognize TB in digitized CRC tissue sections. The DL model demonstrated high sensitivity for TB detection.[8-10]

Identifying the potential value of CPath algorithms in real-world clinical practice will greatly benefit from early health

CONTEXT

Key Objective

What do GI pathologists consider minimal requirements for the use of a computational pathology (CPath) algorithm for tumor budding (TB) assessment in daily clinical practice?

Knowledge Generated

An international panel of GI pathologists agreed upon a set of 21 minimal requirements for the implementation of a CPath algorithm for TB assessment in clinical practice. The additional set of requirements on which there was no agreement highlights key considerations in the implementation process.

Relevance *(P. Innominato)*

Although this Delphi process consensually identified a set of minimal requirements for the routine implementation of a potentially clinically relevant CPath algorithm, some aspects remained dependent on the working environment and artificial intelligence-proficiency of the stakeholders. This study provides a first appropriate consensus and warrants additional improvement cycles to further optimize this set for universal contextualization, both geographically and temporally.*

*Relevance section written by *JCO Clinical Cancer Informatics* Associate Editor Pasquale Innominato, MD, PhD.

technology assessment (early HTA). Early HTA can be defined as all methods used to inform decisions about subsequent development, research, and/or investment by explicitly evaluating the potential value of a conceptual or actual health technology.[11] In a previous study, potential barriers and facilitators for CPath adoption were explored among pathologists, showing that the quality of evidence supporting the clinical use of CPath algorithms and their compatibility with current pathology laboratory workflows are important for pathologists to consider before using such algorithm in daily practice.[12]

A CPath algorithm has the potential to improve efficiency and reproducibility in pathology. However, to facilitate the implementation of the algorithm, it is crucial to define minimal requirements for acceptance of this technology. Therefore, we conducted a Delphi study with an international panel of pathologists to establish consensus on the minimal requirements that must be met before implementation in clinical practice. These standards will support pathology laboratories in validating and implementing CPath solutions, ultimately enhancing diagnostics accuracy and patient care.

METHODS

Study Design

We conducted a qualitative study using a two-round Delphi method to systematically determine the minimum requirements for the adoption of a CPath algorithm for TB in CRC.[13] Minimum requirements were defined as the relevant conditions that must be satisfied for pathology laboratories to implement the CPath artificial intelligence (AI) TB to function correctly, safely, and effectively in daily clinical practice. A Delphi method is an iterative process used to achieve consensus among a panel of experts, particularly in areas where empirical evidence is limited, to guide future directions. Participants remain anonymous and respond to structured questions in multiple rounds. After each round, controlled feedback is given, allowing experts to refine their opinion. Consensus is reached through the analyses of group responses.[14] No formal, prespecified stopping rule was applied. Instead, adequacy of convergence was assessed pragmatically on the basis of the proportion of statements reaching consensus after two iterative rounds, with re-evaluation limited to the statements that did not reach consensus in round 1. No predefined quantitative convergence metrics were used to determine whether additional Delphi rounds were required.

Study Population and Recruitment

For this study, we recruited a panel of GI pathologists, who are the projected end-users of the CPath algorithm for TB in CRC, using a purposive sampling approach. Inclusion criteria required participants to (1) be (self-considered) subspecialized GI pathologists and (2) have sufficient English language proficiency. Potential participants were recruited by email from diverse working groups (Australian Division of the International Academy of Pathology, GI Pathology Division; Rodger C. Haggitt Gastrointestinal Pathology Society; ESP Working Group of Digestive Diseases; British Society of Gastroenterology; and Dutch Expert Group on Gastrointestinal Pathology). In addition, we invited GI pathologists from the TB consortium group through email, authors of publications on TB in the past 5 years who were identified through PubMed contacted through email, and additional participants through a public message on LinkedIn posted by one of the researchers (I.D.N.). After the initial email

requesting participation, two reminders were sent to nonresponders.

Data Collection

The Delphi study was conducted over two rounds. The first round took place between January 4 and March 8, 2024. The second round took place from July 24 to September 17. In round 1, we collected baseline characteristics of the participants, including sex, age, country of practice, years of professional work experience, and prior experience with digital image analysis algorithms. Then, we introduced a case study describing a hypothetical CPath algorithm for TB assessment. This case study was developed to ensure all participants approached the questionnaires from a consistent clinical and technological context. Next, participants were asked to rank statements representing the minimal requirements for the implementation of the CPath algorithm for TB assessment. These statements were derived from a model for assessing the value of AI (MAS-AI), as proposed by Fasterholdt et al.[15] This model provides a structured framework to evaluate the maturity, impact, and implementation readiness of AI solutions in clinical practice. The MAS-AI model includes a two-step approach across nine domains, along with five process factors to guide multidisciplinary assessment. We focused on the first step of MAS-AI, which involves the early evaluation of AI technology by using five domains: (1) health problem and current use of technology, (2) development of the AI model, (3) deployment of the AI model, (4) ethical aspects, and (5) legal aspects. Each MAS-AI domain was deconstructed into subcomponents reflecting key assessment criteria for the use of CPath in clinical practice. These subcomponents were informed by a literature review and were iteratively refined for clarity and contextual relevance.[16] The project team consisted of a pathologist, a CPath expert, a HTA expert, and an implementation science expert. Each team member independently rephrased and refined the statements in three internal iterative rounds, drawing on their own area of expertise. Following this process, 29 concrete and evaluable statements were formulated to define minimal standards for AI-based TB assessment. Before development of the Delphi study, the statements were reviewed by the project team for clarity and face validity and piloted with two independent pathologist colleagues. No formal external pilot testing or wording validation was conducted. Participants were asked to indicate their level of agreement with each statement regarding potential clinical implementation of the CPath algorithm for TB assessment, using a nine-point Likert scale ranging from 1-strongly disagree to 9-strongly agree, with only the end points labeled.

After completion of the first Delphi round, each participant received a personalized feedback form summarizing the interim results for items that lacked consensus. These forms included the median score for each statement, participants' individual scores, and the distribution of scores across the entire panel of participants (percentages per answer option). An illustrative example of a feedback form is provided in the Data Supplement (File S1).

All individuals who responded in the first round were invited to participate in the second Delphi round, using the same nine-point Likert scale. The second round comprised of two parts: (1) information on the hypothetical CPath algorithm for TB to provide contextual grounding and (2) the statements for which no consensus was reached in round 1. Participants were encouraged to re-evaluate their initial responses in the light of the anonymized group feedback presented in their personalized report and rate each statement using the same nine-point Likert scale. The questionnaires of round 1 and round 2 are provided in the Data Supplement (File S2).

Data Analysis

The data from the Delphi study were analyzed according to predefined consensus criteria, as commonly used in a Delphi methodology.[17] The analysis aimed to assess the level of agreement and disagreement among participants regarding the minimal standards for a CPath algorithm for TB assessment. Consensus was defined as ≥70% of participants scoring 7, 8, or 9 (agreement) or ≥70% scoring 1, 2, or 3 (disagreement). A consensus threshold of 70% cutoff agreement or disagreement was applied, which is commonly used in Delphi studies to balance inclusivity with meaningful expert agreement.[14,17] This threshold reflects a pragmatic compromise that allows for inclusion of broadly supported statements while avoiding excessive stringency that may impede consensus formation in heterogeneous expert panels. Adopting a more conservative threshold (eg, 80%) would have resulted in more conservative consensus outcomes and potentially excluded statements with substantive but not near-unanimous support. Conversely, a more liberal cutoff (eg, 60%) would have increased the risk of classifying statements with weaker, more polarized, or less stable agreement as consensus. Statements that did not reach consensus were carried forward to round 2 for further evaluation. In round 2, the same process was repeated. Statements that reached consensus were considered resolved and led to the identification of minimal requirements for AI-based TB assessment. In addition to consensus proportions, median scores and IQRs were calculated and reported for each statement and for each Delphi round. These descriptive statistics provide insight into the strength and dispersion of expert agreement beyond binary consensus outcomes. For statements that were re-evaluated in round 2, changes in median scores across rounds reflect the magnitude and direction of change in expert opinion between rounds.

RESULTS

In round 1, 59 pathologists participated, all of whom were invited to participate in round 2 (response rate round 2 = 90%). Of the participants, 31 were male (53%) and 28 were

female (47%). Most of the respondents were between 45 and 54 years of age (36%), followed by 35-44 years (34%) and 55-64 years (22%). Regarding professional experience, most of the participants had more than 10 years of practice in pathology (69%), whereas 17% had 6-10 years of experience and 14% had ≤5 years. Subspecialized pathologists represented the largest group (73%), while 25% were general pathologists and 2% were residents. Most of the respondents were affiliated with academic hospitals (70%), whereas 30% worked in nonacademic settings. Participants were geographically diverse, with the largest group residing in The Netherlands (29%) and the United States (27%). Other countries represented included the United Kingdom (14%); Australia (7%); Switzerland (5%); Germany, Canada, Ireland, and Nigeria (each 3%); and France, India, and Italy (each 2%).

Concerning digital pathology, 53% of the respondents reported prior experience with digital image analysis or AI algorithms in the research and development phase, while 47% had no such experience. Within diagnostic practice, 49% indicated experience with digital image analysis or AI, compared with 51% without such exposure.

In round 1, 17 statements (out of a total of 29; 59%) reached consensus as minimal requirements for the use of the TB CPath algorithm in daily clinical practice. After the first round, no consensus was reached on 12 statements. In the second round, an additional four statements reached consensus, resulting in a total of 21 (72%) minimal requirements. Of these last four, all were based on agreement. The consensus-based minimal requirements are presented in Figure 1. The results of all the requirements are presented in Data Supplement (File S3). Eight statements (28%) remained without consensus after round 2. These nonconsensus statements are reported separately in Figure 2 and are visually highlighted in Data Supplement (File S3) using a red box to provide transparency regarding areas where expert agreement was not achieved. Inspection of the response distributions for these statements revealed heterogeneous patterns of agreement, which are visualized in Data Supplement (File S4) using pie charts to illustrate the spread of expert opinions.

DISCUSSION

An international panel of GI pathologists reached consensus on a set of 21 minimal requirements for the implementation of a CPath algorithm for TB assessment in clinical practice. On examining the use of the CPath TB algorithm, there was no agreement on the requirement for this algorithm to fully replace the pathologist, but the result was close to rejection. This finding aligns with the current perspective on the use of AI in pathology and, more generally, in health care: There should always be an expert in the loop.[18] We recommend that pathology laboratories first establish a workflow that integrates the pathologist. However, no consensus was reached on the timing when to run the CPath TB algorithm: before, during, and/or after the diagnostic evaluation of the pathologist. This lack of consensus may reflect that the more critical question is not when to run the algorithm, but rather when to use the information it provides. Research by Fogliato et al[19] found that the timing of using the AI model does influence decisions. This study found that radiologists who conducted their own evaluation before using the AI model were less likely to agree with the AI output. Consequently, workflow integration must be established before implementation is undertaken, given its impact on clinical decision making. Moreover, our questionnaire provided only minimal information, for example, details on the risk categorization of the CPath TB algorithm according to the EU Medical Device Regulation were lacking.[20] Such risk categorization could help to determine at which time point in the workflow it is considered safe to run and use the CPath TB algorithm.[21]

Regarding the model development, GI pathologists agreed that they need additional information, such as the data used, the training approach, and the evaluation metrics. This aligns with a key facilitator of AI implementation, namely transparency in AI development.[22-24] However, no consensus was reached on the requirement whether information about the model architecture should be provided to enable reconstruction. This requirement showed a wide dispersion of scores and is highly technical, as pathologists may not fully understand its implications or possess the skills to perform such a task. While a substantial proportion of pathologists rated this requirement as highly relevant, others considered it of limited importance, suggesting divergent views on the necessity and feasibility of technical transparency for clinical use. A recent study on the competencies required for pathologists to use CPath in daily practice does not state this as an essential competency. However, it does mention that a basic understanding of the different steps in AI model development should be acquired, as well as awareness of the potential risks in the design and reporting of AI.[25]

All requirements regarding the deployment of the AI model were deemed necessary before implementing the CPath TB algorithm in daily clinical practice. To evaluate whether the CPath TB algorithm meets requirements such as compatibility, flexibility, and usability in a specific setting, pathology laboratories could use a structured validation process, as described in the review by van der Laak et al.[24]

Many participants acknowledged the relevance of the development of the algorithm using data from different ethnic and cultural groups, but the wide dispersion of scores suggests context-dependent prioritization, potentially influenced by differences in local patient populations, data availability, and regulatory expectations. Although equity and diversity were not prioritized as minimal requirements in this study, these aspects remain critical for the responsible adoption of AI in health care.[26] Algorithms trained on homogeneous datasets risk introducing bias, which can lead to

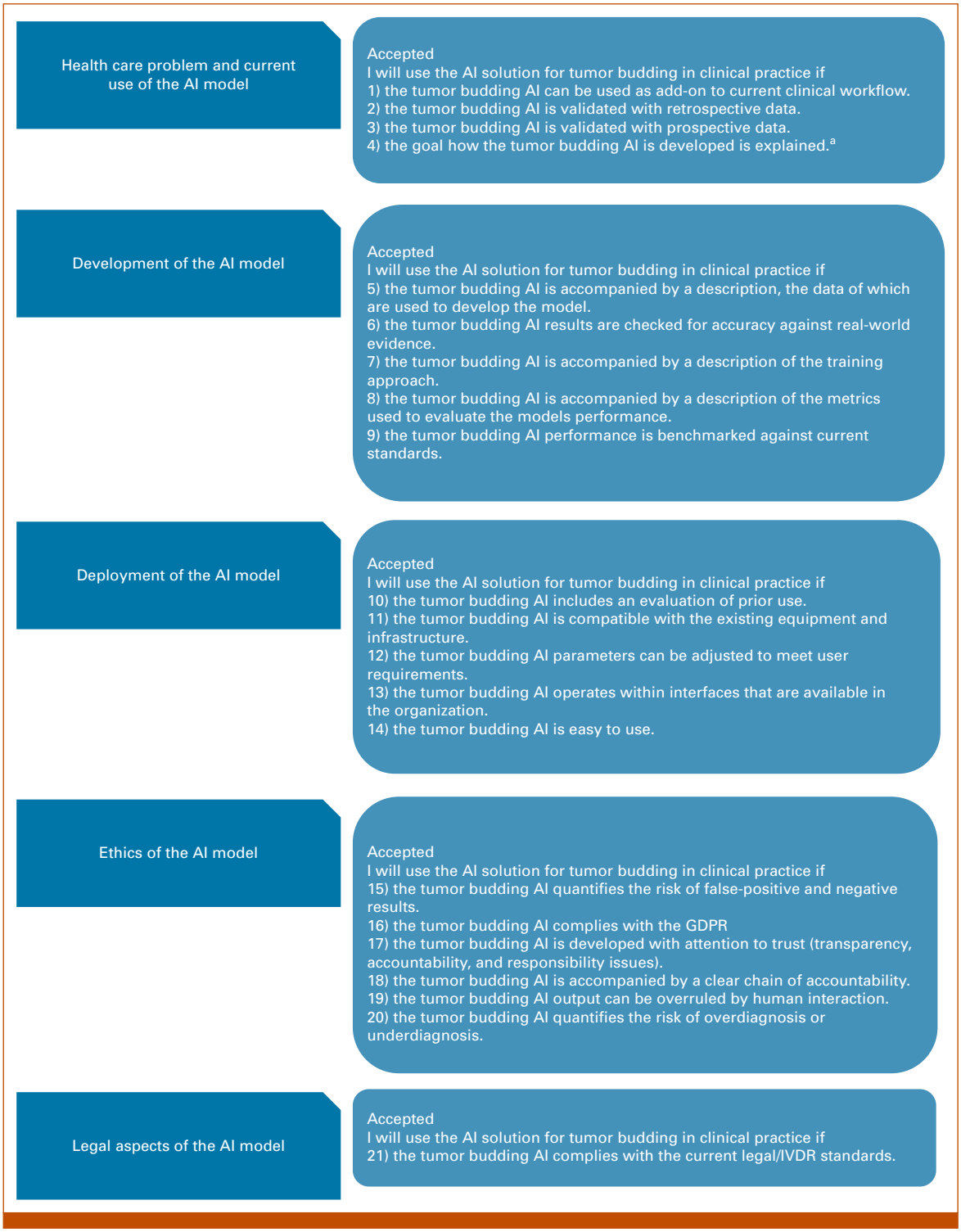

FIG 1. Consensus-based minimal requirements of TB algorithm implementation. [a]Model creation, exploratory study, feasibility study, noninferiority study, etc. AI, artificial intelligence; GDPR, General Data Protection Regulation; IVDR, in vitro diagnostic regulation; TB, tumor budding.

unequal diagnostic performance across patient populations. Therefore, as implementation settings become clearer, stakeholders should revisit these requirements and consider strategies to ensure inclusiveness, such as incorporating diverse datasets and monitoring algorithmic performance across demographic subgroups. Addressing these issues proactively can help build trust among clinicians and patients and align AI development with broader ethical and legal standards.

Together, the nonconsensus findings indicate that workflow, transparency, and equity-related requirements are

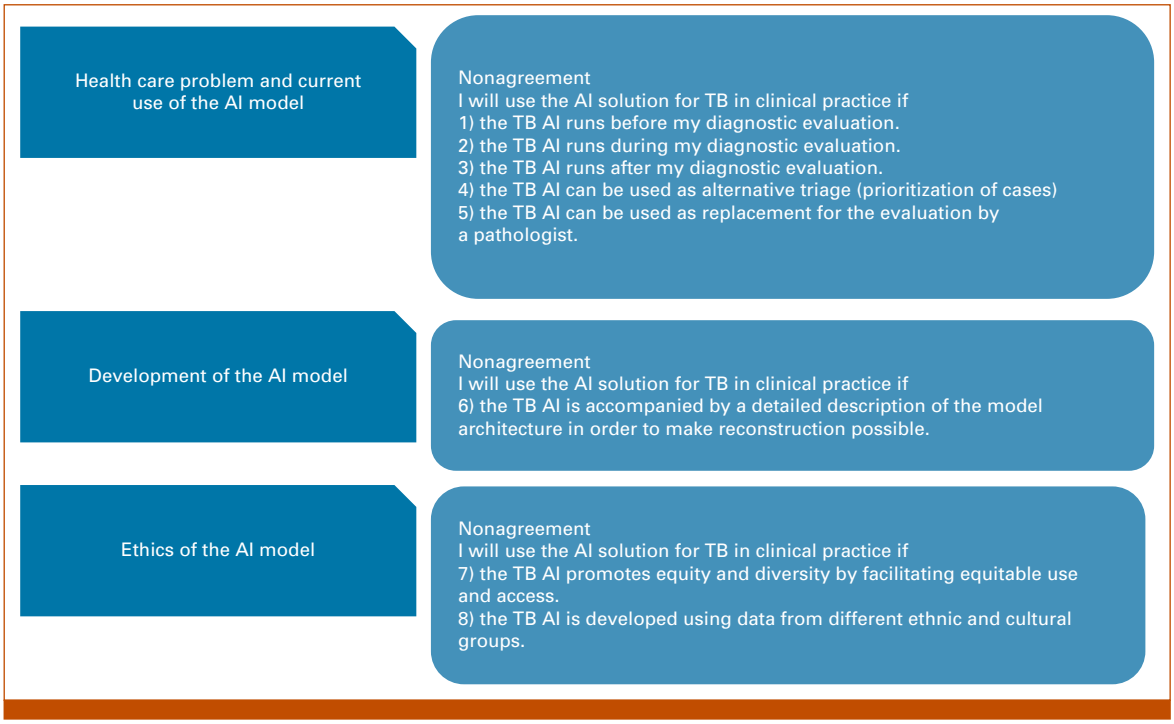

FIG 2. Nonconsensus statements of TB algorithm implementation. AI, artificial intelligence; TB, tumor budding.

recognized as important by many experts, but are weighed differently depending on practical and contextual considerations. From an early HTA perspective, this heterogeneity highlights areas where further empirical evidence and context-specific evaluation will be needed before uniform implementation expectations can be defined.

The consensus-based minimum requirements identified in this study do not constitute empirical evidence of analytical validity, clinical validity, or clinical utility. Rather, they reflect stakeholder-derived expectations that precede and inform these validation stages. From an early HTA perspective, these minimum requirements are intended to guide subsequent empirical validation and context-specific implementation decisions, rather than to replace formal assessments of analytical performance, clinical effectiveness, or clinical impact.

A strength of this study is that we included a broad international panel of pathologists in the early phase of development, to assess the minimal requirements for future implementation of the CPath TB algorithm into daily clinical practice. We used various recruitment strategies to include GI pathologists from a range of different countries. A limitation related to this panel is that most were from academics, while this algorithm is also relevant in nonacademic settings. Next to this, almost half of our panel was involved in AI development research in pathology. Both could have an influence on our results, for example, being more critical regarding the development of the AI model and wanting to have all the information present. The predominance of academic, subspecialized pathologists and those with AI experience may also result in a consensus that is more optimistic about the feasibility and potential benefits of integrating AI into pathology workflows. However, this perspective may not fully capture the barriers and concerns that pathologists in nonacademic laboratories might encounter, such as resource limitations, limited exposure to AI technologies, and different organizational priorities. No formal sensitivity or subgroup analyses were performed to explore differences in consensus by professional experience, practice setting (academic versus nonacademic), or prior exposure to AI. Despite these limitations consensus was reached for 21 out of 29 statements, indicating broad agreement on a core set of minimal requirements that must be met before the clinical implementation, even with this panel composition. Another limitation is that the explanation of the AI model in the survey may lack important information for the study participants to decide on the minimal requirements. However, some requirements may also be flexible, depending on the local context and preferences of pathology laboratories. Therefore, the minimal requirements that are still under discussion can be reviewed again by pathology laboratories, depending on their specific context.

In conclusion, this Delphi study identified stakeholder-derived minimal expectations for the potential implementation of a CPath algorithm for tumor budding in CRC. These expectations reflect expert consensus on prerequisites considered important in an early phase of technology development, rather than empirically validated

implementation standards. While consensus was reached on many technical, organizational, ethical, and legal aspects, several requirements remained context dependent or unresolved. This highlights the need for specific-context evaluation and ongoing stakeholder dialog as development and implementation progresses.

AFFILIATIONS

[1]Radboud University Medical Center, IQ Health Science Department, Nijmegen, the Netherlands
[2]Department of Pathology, Radboud University Medical Center, Nijmegen, the Netherlands
[3]Institute of Tissue Medicine and Pathology, University of Bern, Bern, Switzerland
[4]Center for Medical Image Science and Visualization, Linköping University, Linköping, Sweden

CORRESPONDING AUTHOR

Julie E.M. Swillens, PhD; e-mail: Julie.Swillens@radboudumc.nl.

SUPPORT

Supported by the Dutch Cancer Society (number 10602).

AUTHOR CONTRIBUTIONS

Conception and design: All authors
Financial support: Iris D. Nagtegaal, Jeroen A.W.M. Van der Laak
Collection and assembly of data: Julie E.M. Swillens, Marcia Tummers
Data analysis and interpretation: Julie E.M. Swillens, Marcia Tummers
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/cci/author-center.
Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

Julie E.M. Swillens
Employment: Radboudumc
Research Funding: Radboudumc (Inst)
Travel, Accommodations, Expenses: Radboudumc

Alessandro Lugli
Honoraria: Baxter (I)
Travel, Accommodations, Expenses: Nestle (I)

Jeroen A.W.M. Van der Laak
Stock and Other Ownership Interests: Aiosyn, Netherlands

No other potential conflicts of interest were reported.

ACKNOWLEDGMENT

The authors acknowledge the Delphi panelists who participated in this project. Panelists who agreed to be named are listed in the Appendix.

REFERENCES

1. Lugli A, Kirsch R, Ajioka Y, et al: Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol 30:1299-1311, 2017
2. Lugli A, Zlobec I, Berger MD, et al: Tumour budding in solid cancers. Nat Rev Clin Oncol 18:101-115, 2021
3. Brierley JD, Gospodarowicz MK, Wittekind C (eds): TNM Classification of Malignant Tumours. 8th ed. Oxford, UK: Wiley-Blackwell, 2017
4. Nagtegaal ID, Odze RD, Klimstra D, et al: The 2019 WHO classification of tumours of the digestive system. Histopathology 76:182-188, 2020
5. Loughrey MB, Webster F, Arends MJ, et al: Dataset for pathology reporting of colorectal cancer: Recommendations from the International Collaboration on Cancer Reporting (ICCR). Ann Surg 275: e549-e561, 2022
6. Rosty C, Webster F, Nagtegaal ID, et al: Pathology reporting of colorectal local excision specimens: Recommendations from the International Collaboration on Cancer Reporting (ICCR). Gastroenterology 161:382-387, 2021
7. Specialisten FM: CRC Guideline. 2025. https://richtlijnendatabase.nl/richtlijn/colorectaal_carcinoom_crc/startpagina_-_colorectaal_carcinoom.html
8. Bokhorst J-M, Ciompi F, Öztürk SK, et al: Fully automated tumor bud assessment in hematoxylin and eosin-stained whole slide images of colorectal cancer. Mod Pathol 36:100233, 2023
9. Öztürk SK, Bokhorst J-M, Baumann E, et al: Exploring intratumoral budding in colorectal cancer using computational pathology: A biopsy-based evaluation. Mod Pathol 38:100655, 2025
10. Haddad T, Bokhorst J, Berger M, et al: Combining immunoscore and tumor budding in colon cancer: An insightful prognostication based on the tumor-host interface. J Transl Med 22:1090, 2024
11. Grutters JP, Bouttell J, Abrishami P, et al: Defining early health technology assessment: Building consensus using Delphi technique. Int J Technol Assess Health Care 41:e34, 2025
12. Swillens JE, Nagtegaal ID, Engels S, et al: Pathologists' first opinions on barriers and facilitators of computational pathology adoption in oncological pathology: An international study. Oncogene 42:2816-2827, 2023
13. Fitch K, Bernstein SJ, Aguilar MD, et al: RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA, RAND Corporation, 2000
14. Hohmann E, Cote MP, Brand JC: Research pearls: Expert consensus based evidence using the Delphi method. Arthroscopy 34:3278-3282, 2018
15. Fasterholdt I, Kjølhede T, Naghavi-Behzad M, et al: Model for ASsessing the value of Artificial intelligence in medical imaging (MAS-AI). Int J Technol Assess Health Care 38:e74, 2022
16. Fasterholdt I, Naghavi-Behzad M, Rasmussen BS, et al: Value assessment of artificial intelligence in medical imaging: A scoping review. BMC Med Imaging 22:187, 2022
17. Diamond IR, Grant RC, Feldman BM, et al: Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 67:401-409, 2014
18. Jarrahi MH, Davoudi V, Haeri M: The key to an effective AI-powered digital pathology: Establishing a symbiotic workflow between pathologists and machine. J Pathol Inform 13:100156, 2022
19. Fogliato R, Chappidi S, Lungren M, et al: Who Goes First? Influences of human-AI Workflow on Decision Making in Clinical Imaging, 2022. pp 1362-1374
20. Union E: Guidance on Classification of Medical Devices. Medical Device Coordination Group. https://health.ec.europa.eu/system/files/2021-10/mdcg_2021-24_en_0.pdf
21. Daye D, Wiggins WF, Lungren MP, et al: Implementation of clinical artificial intelligence in radiology: Who decides and how? Radiology 305:555-563, 2022
22. Kelly CJ, Karthikesalingam A, Suleyman M, et al: Key challenges for delivering clinical impact with artificial intelligence. BMC Med 17:195-199, 2019
23. Fehr J, Citro B, Malpani R, et al: A trustworthy AI reality-check: The lack of transparency of artificial intelligence products in healthcare. Front Digit Health 6:1267290, 2024
24. Van der Laak J, Litjens G, Ciompi F: Deep learning in histopathology: The path to the clinic. Nat Med 27:775-784, 2021
25. Vos S, Hebeda K, Milota M, et al: Making pathologists ready for the new artificial intelligence era: Changes in required competencies. Mod Pathol 38:100657, 2025
26. Chauhan C, Gullapalli RR: Ethics of AI in pathology: Current paradigms and emerging issues, in Chauhan C, Cohen S (eds): Artificial Intelligence in Pathology (ed 2). Amsterdam, the Netherlands, Elsevier, pp 159-180, 2025

APPENDIX. DELPHI STUDY PARTICIPANTS

Only participants who provided their details in the second round of the Delphi study are listed.

1. Marnix Jansen, MD, PhD, MSc, Associate Professor, Department of Pathology, UCL Cancer Institute
2. Dr Newton A.C.S. Wong, Honorary Associate Professor in Pathology, University of Bristol
3. Professor Maurice Loughrey, Queen's University Belfast, United Kingdom
4. Chella van der Post, MD, PhD, Pathologist
5. Iván A. González, MD, Assistant Professor, Department of Pathology and Laboratory Medicine, Indiana University School of Medicine
6. Sonay Kuş Öztürk, Ms, Department of Pathology, Radboud University Medical Center
7. Paromita Roy, Senior Consultant, Tata Medical Center, Kolkata
8. Pr Magali SVRCEK, Sorbonne Université, AP-HP, Saint-Antoine Hospital, Department of Pathology
9. Michael Vieth, Prof. Dr Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Klinikum Bayreuth, Bayreuth, Germany
10. David Schaeffer, MD, Vancouver General Hospital, Vancouver, BC Canada
11. Professor Heike I Grabsch, Maastricht University (NL) and University of Leeds (United Kingdom)
12. Professor Dr Kirsten Mertz, Institute of Medical Genetics and Pathology, University Hospital Basel, Basel, Switzerland
13. Uchenna Ezenkwa, Dr, Department of Histopathology, Federal University of Health Sciences, Azare, Nigeria
14. Christophe Rosty, MD, PhD, Envoi Pathology, University of Queensland, Brisbane, QLD, Australia
15. Prof David Gibbons, Department of Pathology, St Vincent's University Hospital, Dublin, Ireland
16. Raul S. Gonzalez, MD, Emory University School of Medicine
17. Marilyne Lange, MD, PhD, Amsterdam UMC
18. Prof Anthony J. Gill, MD, FRCPA, University of Sydney, Sydney, NSW, Australia
19. Professor Adrian C. Bateman, Consultant Histopathologist, University Hospital Southampton NHS Foundation Trust
20. Professor Kieran Sheahan, St Vincent's University Hospital and UCD School of Medicine, Dublin
21. Adoke Kasimu Umar, Consultant Pathologist, Federal Medical Centre Birnin Kebbi, Nigeria
22. Valeria Barresi, University of Verona, Italy
23. A. Klooster, MD-PhD, Pathologie Friesland, Leeuwarden
24. Dr FC Meeuwsen, MD from Erasmus MC Rotterdam, the Netherlands
25. Prof. Petur Snaebjornsson, MD, PhD, Netherlands Cancer Institute, Department of Pathology, Amsterdam, the Netherlands, and University of Iceland, Faculty of Medicine, Reykjavik, Iceland.
26. Dr N Knijn, Pathology DNA, Arnhem
27. Gursah Kats-Ugurlu, MD, PhD, Department of Pathology, University Medical Center Groningen, University of Groningen
28. A.H.A.G. Ooms, MD, PhD, Pathology Department, Pathan BV, Rotterdam
29. Melissa W. Taggart, Professor, University of Texas M.D. Anderson Cancer Center
30. Noam Harpaz, MD, PhD, Professor of Pathology and Gastroenterology, Icahn School of Medicine at Mount Sinai, New York.
31. Stijn Crobach, Department of Pathology, LUMC Leiden
32. Adam L. Booth, Assistant Professor of Pathology and Immunology, Washington University School of Medicine
33. Joseph Misdraji, Professor of Pathology, Yale University, Yale New Haven Hospital
34. Jinru Shia, MD, Memorial Sloan Kettering Cancer Center, New York
35. Dr Nick West, University of Leeds
36. E.E.M. Peters, MD, PhD, Haaglanden Medical Center
37. Jennifer Do, MD, FCAP, FASCP, GI Pathologist; Advanced Pathology Solutions, Little Rock, AR
38. Michail Doukas, MD, PhD, Department of Pathology, Erasmus University Medical Center Rotterdam, the Netherlands
39. Dr Richard Standish