

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<u>http://bmjopen.bmj.com</u>).

If you have any questions on BMJ Open's open peer review process please email <u>info.bmjopen@bmj.com</u>

**BMJ** Open

# **BMJ Open**

#### The CONFIDENT-trial protocol: a pragmatic template for clinical implementation of artificial intelligence assistance in pathology

Journal:	BMJ Open
Manuscript ID	bmjopen-2022-067437
Article Type:	Protocol
Date Submitted by the Author:	17-Aug-2022
Complete List of Authors:	Flach, Rachel; University Medical Centre Utrecht, Department of Pathology Stathonikos, Nikolas; University Medical Centre Utrecht, Department of Pathology Nguyen, Tri; University Medical Centre Utrecht, Department of Pathology Ter Hoeve, Natalie; University Medical Centre Utrecht, Department of Pathology van Diest, P.J.; University Medical Centre Utrecht, Department of Pathology van Dooijeweert, Carmen; University Medical Centre Utrecht, Department of Pathology
Keywords:	Prostate disease < UROLOGY, Breast tumours < ONCOLOGY, PATHOLOGY



# The CONFIDENT-trial protocol: a pragmatic template for clinical implementation of artificial intelligence assistance in pathology

#### Authors

R.N. Flach (<u>r.n.flach-2@umcutrecht.nl</u>), N. Stathonikos (<u>nstatho2@umcutrecht.nl</u>), T.Q. Nguyen (<u>t.q.nguyen@umcutrecht.nl</u>), N.D. ter Hoeve (<u>N.terhoeve@umcutrecht.nl</u>), P.J. van Diest (<u>P.j.vandiest@umcutrecht.nl</u>), C. van Dooijeweert (<u>c.vandooijeweert3@umcutrecht.nl</u>)

### Affiliation

Department of Pathology, University Medical Centre Utrecht, Utrecht, The Netherlands

#### **Trial registration**

Applicable nor suitable.

#### **Protocol version**

Version 1.

## Funding

No funding was received for this research.

## **Roles and responsibilities**

As no funding was received, and as this is a single centre clinical trial, full authority over all study activities (design; collection, management, analysis and interpretation of data; writing of the report; and the decision to submit the report for publication) lies with the researchers in the UMC Utrecht.

ellen

## **Corresponding author**

Carmen van Dooijeweert

Internal mail no H04.312

T: +31 88 75 565 65

c.vandooijeweert3@umcutrecht.nl

## Abstract

#### Introduction

Artificial Intelligence (AI) has been on the rise in the field of pathology. Despite promising results in retrospective studies, and several CE-IVD certified algorithms on the market, prospective clinical implementation studies of AI have yet to be performed. In this trial, we will explore the benefits of an AI-assisted pathology workflow, while maintaining diagnostic safety standards.

#### Methods and analysis

This is a SPIRIT-AI compliant single-centre, controlled clinical trial, in a fully digital academic pathology laboratory. We will prospectively include prostate cancer patients who undergo prostate needle biopsies (CONFIDENT-P) and breast cancer patients who undergo a sentinel lymph node (SN) procedure (CONFIDENT-B) in the UMC Utrecht. For both the CONFIDENT-B and CONFIDENT-P trials, the specific pathology specimens will be pseudo-randomized to be assessed by a pathologist with- or without AI-assistance in a pragmatic (bi-)weekly sequential design. In the intervention group, pathologists will assess whole slide images (WSI) of the standard haematoxylin-eosin (HE)-stained sections assisted by the output of the algorithm. In the control group, pathologists will assess HE WSI according to the current clinical workflow. If no tumour cells are identified or when the pathologist is in doubt, staining by immunohistochemistry (IHC) will respectively may be performed. At least 80 patients in the CONFIDENT-P and 180 patients in the CONFIDENT-B trial will need to be enrolled to detect superiority, allocated as 1:1. Primary endpoint for both trials is the number of saved resources on IHC for detecting tumour cells, since this will clarify tangible cost savings that will build the business case for AI.

#### Ethics and dissemination

The ethics committee (MREC NedMec) waived the need of official ethical approval, as participants are not subjected to procedures and as they are not required to follow rules. Results of both trials (CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed journals.

## **Strengths and limitations**

- This is the first clinical trial to examine the added value of artificial intelligence in the daily pathology workflow.
- By maintaining the current diagnostic safety standards patients are not at risk of an inferior diagnosis during the trial.
- This is a pragmatic template for prospective AI-trials for object-identifying algorithms in pathology.
- A limitation is that this is a single-centre study, which may hamper generalizability.
- Due to the existing clinical workflow, randomization of patients and (double-)blinding of the participating pathologists and researchers is not possible.

#### **Keywords:**

Artificial intelligence, prospective trial, prostate cancer, breast cancer, digital pathology

Word count: 2864 words

#### Introduction

#### Background and rationale

Since the introduction of digital pathology, the number of studies on artificial intelligence (AI) within the field of pathology has increased exponentially.<sup>1,2</sup> Algorithms have been created for tumour detection, tumour grading, recognizing tumour subtypes, evaluating biomarkers and more.<sup>1,3</sup> Worldwide, a shortage of pathologists exists, while their workload is only increasing.<sup>2,4</sup> Therefore, AI has great potential to alleviate pathologists' workload.<sup>2</sup> At the same time, AI has great potential to improve diagnostics by improving accuracy, reproducibility and speed.<sup>2</sup> In fact, several algorithms have shown to be comparable, or even superior to pathologists (under time-constraint).<sup>2,5–10</sup> However, artificial and human intelligence are not mutually exclusive, but they complement each other, a concept which is known as "augmented intelligence", where AI enhances, rather than replaces human intelligence.<sup>11</sup> In the (very) early AI-adoption phase, and presumably also in later adoption phases, pathologist-supervision remains of key-importance.

This is particularly relevant as, despite the promising results of retrospective studies and the availability of CE-IVD approved algorithms, prospective validation and clinical implementation of AI are currently lacking. For example, six years after the successful CAMELYON-16 Grand Challenge<sup>6</sup>, the top algorithms have yet to be implemented in daily clinical practice, showing that the time between development of an AI model and clinical implementation is considerable. Likewise, numerous promising prostate cancer grading algorithms have been developed, yet implementation studies have yet to be performed.<sup>12–14</sup> In addition, nine AI pathology devices received CE-IVD approval in 2021.<sup>15</sup>

#### **Trial rationale**

As a pathology laboratory with a fully digital workflow for over seven years, we are eager to explore the full potential of working digitally by adding the benefit of AI in daily pathology practice. We decided to start with the object localisation task of tumour-detection, where an objective reference standard is in place in the routine clinical workflow (i.e. pathologist-supervision and/or immunohistochemistry (IHC) staining).

We developed the CONFIDENT-trial template, in which AI tumour detection algorithms can be safely implemented in prospective clinical trials, while ensuring that patients are not at risk of receiving an inferior diagnosis, since IHC is always performed when no tumour cells are visible, but also when pathologists need more confirmation about the diagnosis.

#### CONFIDENT-B and CONFIDENT-P

Our trials aim to prospectively investigate the added value of an AI-assisted pathology workflow in the identification of prostate cancer (PCa) in prostate needle biopsies (CONFIDENT-P) and the identification of sentinel lymph node (SN) metastases in breast cancer (BCa) patients (CONFIDENT-B). As both prostate cancer and breast cancer are the most common (non-skin) malignancies in men and women, respectively, implementation of AI-assistance may have a great impact on diagnostic processes.<sup>16</sup> However, it is important to emphasize that this trial serves as a template for other pragmatic AI-intervention trials for object-localisation tasks as well.

We obtained CE-IVD-approved algorithms for detection and grading of prostate cancer in prostate needle biopsies and an algorithm for detecting lymph node metastases in BCa patients. In

both cases the task of the pathologist is both labour-intensive and expensive, due to the performed IHC stains in case no tumour cells are morphologically observed. However, IHC is expensive and these costs sometimes even exceed reimbursement for the entire specimen (e.g. in case of multiple blocks of multiple SNs). This raises the question whether AI may be of added value to morphologically detect cancer cells without the need for IHC-use. Thereby the number of performed IHC stains may be reduced, which may lead to tangible costs savings that will help to build the business case for AI, while potentially decreasing the workload of pathologists as well.<sup>2</sup>

#### Study objective

The primary objective is to explore whether an AI-assisted workflow reduces the number of spent resources on IHC, while maintaining diagnostic safety standards in both PCa patients who underwent prostate needle biopsies (CONFIDENT-P), and BCa patients who underwent an SN procedure.

Secondary objectives are to investigate whether time management improves in an Alassisted workflow and to analyse how many IHC staining may have been safely omitted after Alimplementation.

### Methods and analysis

#### **Trial design**

The study protocol is structured following the SPIRIT-AI (Standard Protocol Items: recommendations for Interventional Trials - Artificial Intelligence) statement 2020<sup>17</sup>. This study is a single-centre, parallel-group controlled trial, assessing superiority. The allocation ratio is 1:1. Eligible patients will be assigned to arm 1 (control group) or arm 2 (AI-assisted workflow), based on a bi-weekly time-schedule. Eligibility criteria are summarised in Figure 1. The CONFIDENT trials will be carried out in 2022-2023.

#### Study setting

The trial will take place in the daily practice of a single academic hospital (University Medical Centre Utrecht, the Netherlands), with a fully digital pathology enabled clinical set-up, where all slides are digitised using ultrafast whole slide image (WSI) using Hamamatsu S360 scanners and reviewed using the Sectra pathology Picture Archiving and Communication System (PACS). Although the UMC Utrecht is an academic hospital, primary routine pathology diagnostics is performed for non-academic hospitals as well (i.e., the Alexander Monro Breast Cancer Hospital, Bilthoven, the Netherlands).

#### Study population

For PCa, WSI of all males who undergo a prostate needle biopsy in the UMC Utrecht will be included. For BCa, WSI of all females or males with BCa as primary malignancy (i.e. invasive breast cancer) who undergo an SN procedure in the Alexander Monro Breast Cancer Hospital or the UMC Utrecht will be included. Patients will be excluded, if they were redirected to the UMC Utrecht for a second opinion.

#### Assessment of specimen

During the study period, all WSI will be assessed by the same group of pathologists; i.e. two expert urological pathologists for the PCa biopsies, and three expert breast pathologists for the lymph node assessment from BCa patients.

For both the CONFIDENT-B and CONFIDENT-P trial, the specific pathology specimens will be assigned to be assessed by a pathologist with or without AI-assistance in a pragmatic (bi-)weekly sequential design. This is deemed feasible as case-mix variation and time trends are highly unlikely to occur within the envisioned inclusion-period of approximately six to nine months. Furthermore, both

 specialized breast- and urological pathologists within the UMC Utrecht work according to weekly schedules. Therefore, using AI every other week or every other two weeks, as opposed to switching by day, ensures that all pathologists are equally distributed between groups. Lastly, it would be impractical to switch from AI-assistance to no AI-assistance on a case-to-case basis. For obvious reasons, allocation concealment and blinding of pathologists is not applicable. In addition, as there is no room for interpretation, researchers who perform the data-analysis will also not be blinded.

#### Control and intervention

All eligible specimens will be assigned to either the control group or the intervention group. In the control group, pathologists will assess H&E stained WSI of patients digitally, according to the current clinical workflow. For PCa biopsies, IHC is routinely performed on all cases. For BCa lymph nodes, if no metastases or tumour are present, IHC staining will be performed. Additional IHC staining will also be performed by additional request of the pathologist in case of doubt.

In the intervention group, pathologists will assess the H&E-specimens digitally with the outcome of the algorithm provided in their first assessment of the specimen. For PCa, they will use the CE-IVD certified Paige Prostate Suite algorithm, and for BCa, pathologists will use the CE-IVD certified Metastasis Detection App by Visiopharm. These algorithms will be integrated within the Sectra PACS where the output of the algorithms will be graphically displayed. AI analysis of the WSI will be performed right after scanning to avoid delays in the clinical workflow. If the AI-assisted pathologist does not detect metastases or tumours on the H&E slide, routine additional IHC staining will be performed by P503S/p63/CK HMW for PCa and CAM5.2 for BCa, to ensure no metastases or tumours are missed. Pathologists can also request an additional IHC if they feel they need this to make an adequate diagnosis (Figure 2).

#### **Outcome measures**

Primary outcome for the CONFIDENT-P trial is the added value of AI-assistance in the detection of PCa and in the detection of tumour volume in prostate needle biopsies in daily pathology practice. The primary outcome for the CONFIDENT-B trial is the added value of AI-assistance in the detection of BCa SN-metastases. The outcome measures for both trials will be the number of spent resources, i.e. the number of IHC-stains performed in both groups.

Secondary outcome measures will be sensitivity and specificity of the AI-assisted pathologist, time spent on WSI analysis, the number of IHC stains that may have been omitted after AI-implementation and a pathologists' evaluation by a questionnaire om the AI-assisted work process.

#### Input data

Input data for the algorithm will be WSI of haematoxylin and eosin (H&E) stained slides scanned at 40x of either prostate needle biopsies, and WSI of H&E stained slides of BCa lymph nodes. As per routine in our daily clinical practice, WSI will be quality controlled after scanning for colour, focus quality and completeness of the scan. When necessary, the specimens will be rescanned.

#### Sample size

#### **CONFIDENT-P**

We performed power calculations using a two-sample proportion superiority test, using expected percentages of IHC staining in both study arms. We assume that the pathologists in the control arm can detect 50% of the tumours without using IHC. We expect AI-assisted pathologists to detect 80% of the tumours, without using IHC. These percentages were conservatively derived from the validity

study by Raciti (74% for pathologists without AI and 90% for pathologists with AI respectively)<sup>18</sup>, by expert pathologist opinion, and taking into account that pathologists under time constraint of daily practice do not detect tumours as well as pathologists without time constraint during retrospective studies<sup>19</sup>. We assume that this effect will be larger for the biopsies assessed without AI than with AI, as AI is assumed to make tumour detection easier.

A sample size of 60 (30 per arm) would give a power of approximately 80%, using a one-sided 5% significance level. However, uncertainties remain regarding the sample size parameters. We therefore inflated our sample size to 80 (40 per arm), in order to ensure study power and allowing us to detect smaller effect sizes.

For detection of tumour volume percentage, we performed a power calculation based on the assumption that AI should be able to replace at least 20% of the IHC stains, in order to be cost-effective. IHC is currently used in 100% of all prostate needle biopsies. Using a power of 80% and a one-sided significance level of 5%, this leads to 27 patients per arm.

#### CONFIDENT-B

Sample size calculations for the CONFIDENT-B trial are based on the assumption that the AIalgorithm can detect all metastases for which currently IHC is used, which are mainly micrometastases and isolated tumour cells (ITC). Approximately 15% of the SN-specimens in the UMC Utrecht contain a micrometastasis or ITC.

A sample size of 166 patients (83 per arm) with a one-sided 5% significance level therefore results in a power of 80%. Again, as there are uncertainties on the assumptions on what amount of the metastases will be detect by AI, we decided to be conservative and include 180 patients (90 per arm).

Overall, we are only interested in one-sided outcomes, as it is not possible that more IHC will be performed in the AI-assisted arm. IHC is performed to detect metastases, when they are macroscopically undetectable, rather than to confirm them when they are macroscopically visible. As AI would show only more metastases than the pathologist could macroscopically detect, this means that only a reduction of IHC is possible.

## Statistical methods

For baseline comparisons between both arms, the appropriate measures (parametric or nonparametric) for categorical (Chi-squared test/Fisher's exact) and continuous variables (T-test, Mann-Whitney U test) will be used. For the analysis of the primary outcome measure, we will compare the proportion of IHC-use in both arms, and calculate adjusted relative risks, using a log-binomial model.<sup>20–22</sup>

Missing data for baseline characteristics and for the primary outcome are not to be expected, as they are obligatory items in the structured pathology reports.

We will determine sensitivity and specificity of the conclusion of the AI-assisted pathologists without the use of IHC. Subsequently, we will focus on the cases with metastases that the AI-assisted pathologist misses, categorize them (i.e. macro-metastases, micro-metastases, ITC) and determine their clinical relevance (i.e. clinical consequences if these metastases are being missed). Data-analysis will be performed in R Statistical Software<sup>23</sup>, with a significance level set at p<0.05.

#### Data collection and management

All data (baseline and primary outcome measurements) will be retrieved from the structured pathology reports and will be managed and stored in Castor EDC<sup>24</sup>. For the secondary outcome measure of time spent by the pathologist on a slide, data will be collected on an interval basis for

practical reasons, as timing every assessment for months (by stopwatch) was not deemed feasible. For the secondary outcome measure of AI-assisted work process for pathologists, a questionnaire will be distributed to the participating pathologists. The final secondary assessment measurement, the number of IHC stains that may have been omitted after AI-implementation, will be determined by the researchers based on the data from the structured pathology reports (combination of IHC and AI-assisted diagnoses of the pathologist).

#### **Ethical approval**

Research within these trials is not subject to the (Dutch) Medical Research Involving Human Subjects Act (WMO), as participants are not subjected to procedures and as they are not required to follow rules. Therefore, the ethics committee (MREC NedMec) waived the need of ethical approval and informed consent.

#### **Risk of harm**

Patients are not at risk of any harm for an inferior diagnosis (i.e. missed tumour cells), as in both arms, IHC-staining will be performed when no tumour cells are visible, according to current clinical standards. As a rule in augmented intelligence, all cases will be evaluated by a pathologist, which further minimized the risk of a false diagnosis based on the AI algorithm. Taking all of the above into account, a data monitoring committee (DMC) is not required and adverse events are not to be expected. In theory, the algorithm could be more of a disturbance than a help to pathologists (for example, when it frequently reports false positive or negative results, which have to be corrected by the pathologist). However, the algorithms used are IVDR-approved, and thus have undergone extensive review for their intended purpose. Nonetheless, the experience and ease of use of pathologists working with the algorithm will be one of the secondary outcome measures.

#### Informed consent and data access

Informed consent was waived by the local quality coordinator (QC) and data protection officer (DPO) for the following reasons. First, in both arms patients receive standard care, while maintaining diagnostic safety standards (pathologists' supervision, IHC in all negative cases). Second, patients are not subjected to any procedures. Third, all data will be anonymized to the researchers by the pathologist who assessed the slide.

The collected (anonymous) research data will be stored in Castor EDC to ensure data security. Data will be kept for a period of 15 years. Data access in Castor will be restricted to two researchers (RF, CvD). Pathologists have access to the electronic patient files for the purpose of patient care. The researchers are not permitted access to these files. At no point will the data (both in Castor EDC and patient files) be accessed by the companies providing the algorithms (i.e. Visiopharm and Paige).

#### Patient and Public Involvement

None

#### Discussion

The promising retrospective results of AI-assisted pathology have not yet resulted in prospective clinical implementation studies. This may be due to a lack of digital transition in the majority of pathology laboratories, but it may also be partly due to the lack of a good implementation model. Fortunately, however, new guidelines for AI-trials have recently been proposed by the SPIRIT-AI and CONSORT-AI-steering groups, as well as roadmaps to routine use of AI in clinical practice.<sup>17,25,26</sup> Yet, to date, no pathology AI-trials have been published in PubMed or Web of Science or, to the best of

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

our knowledge, otherwise made public.

As a pathology laboratory with a fully digital workflow, we developed a clinical trial template for tumour detection models, as a first step to implement AI in daily pathology practice. We will start with an object localisation task (i.e. tumour cells) as a reference standard is in place in the routine clinical workflow. For classification tasks like tumour grading, a clinical trial design is more challenging, as no reference is in place in daily pathology practice and inter-laboratory and interpathologist variation is notorious.<sup>27–31</sup> Nevertheless, in future trials, implementing AI-assistance in the grading process might also reduce this variation. For now, results of the CONFIDENT-trials will provide the first assessment of the potential added value of AI in daily pathology practice. This evaluation will substantially contribute to a potential paradigm shift in tumour detection in pathology. The pragmatic template of the CONFIDENT trials may serve as example for other prospective AI implementation trials in diagnostic pathology.

#### Declaration of interests

PJvD is a member of the Scientific Advisory Board of Paige and Sectra. All other authors do not report conflict of interest.

#### Funding

No funding was obtained at the moment of writing this paper.

#### **Dissemination policy**

Results of both trials (CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed journals. Authorship will be acknowledged to all those that substantially contributed in the CONFIDENT-trials. Data will be available upon reasonable request.

#### Author statement

PJvD conceived of the study. RNF and CvD initiated the study design and NS, TQN and NDtH helped with implementation. RNF and CvD provided statistical expertise in clinical trial design and RNF and CvD are conducting the primary statistical analysis. All authors contributed to refinement of the study protocol and approved the final manuscript.

#### **Figure legends**

Figure 1. Flowchart with patient selection; SN = sentinel node; AI = artificial intelligence.

Figure 2. Study flow chart. H&E = hematoxylin & eosin; IHC = immunohistochemistry

## References

- 1. Jiang, Y., Yang, M., Wang, S., Li, X. & Sun, Y. Emerging role of deep learning-based artificial intelligence in tumour pathology. *Cancer Commun.* **40**, 154–166 (2020).
- 2. van der Laak, J., Litjens, G. & Ciompi, F. Deep learning in histopathology: the path to the clinic. *Nat. Med.* **27**, 775–784 (2021).
- 3. Steiner, D. F., Chen, P. H. C. & Mermel, C. H. Closing the translation gap: AI applications in digital pathology. *Biochim. Biophys. Acta Rev. Cancer* **1875**, 188452 (2021).
- 4. Robboy, S. J. *et al.* Pathologist workforce in the United States: I. Development of a predictive model to examine factors influencing supply. *Arch. Pathol. Lab. Med.* **137**, 1723–1732 (2013).
- 5. Bulten, W. *et al.* Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. *Lancet. Oncol.* **21**, 233–241 (2020).
- Ehteshami Bejnordi, B. *et al.* Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **318**, 2199–2210 (2017).
- 7. Steiner, D. F. *et al.* Evaluation of the Use of Combined Artificial Intelligence and Pathologist Assessment to Review and Grade Prostate Biopsies. *JAMA Netw. open* **3**, e2023267 (2020).
- 8. Ström, P. *et al.* Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet. Oncol.* **21**, 222–232 (2020).
- 9. Raciti, P. *et al.* Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. *Mod. Pathol.* **33**, 2058–2066 (2020).
- 10. Hekler, A. *et al.* Superior skin cancer classification by the combination of human and artificial intelligence. *Eur. J. Cancer* **120**, 114–121 (2019).
- 11. Harrison, J. H. *et al.* Introduction to Artificial Intelligence and Machine Learning for Pathology. *Arch. Pathol. Lab. Med.* **145**, 1228–1254 (2021).
- 12. Ström, P. *et al.* Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. *Lancet Oncol.* **21**, 222–232 (2020).
- 13. Steiner, D. F. *et al.* Evaluation of the Use of Combined Artificial Intelligence and Pathologist Assessment to Review and Grade Prostate Biopsies. *JAMA Netw. Open* **3**, 1–14 (2020).
- 14. Bulten, W. *et al.* Artificial intelligence assistance significantly improves Gleason grading of prostate biopsies by pathologists. *Mod. Pathol.* **34**, 660–671 (2021).
- 15. Muehlematter, U. J., Daniore, P. & Vokinger, K. N. Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *Lancet Digit. Heal.* **3**, e195–e203 (2021).
- 16. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **71**, 209–249 (2021).
- 17. Cruz Rivera, S. *et al.* Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *The Lancet Digital Health* vol. 2 e549–e560 at https://doi.org/10.1016/S2589-7500(20)30219-3 (2020).
- 18. Raciti, P. et al. Novel artificial intelligence system increases the detection of prostate cancer in

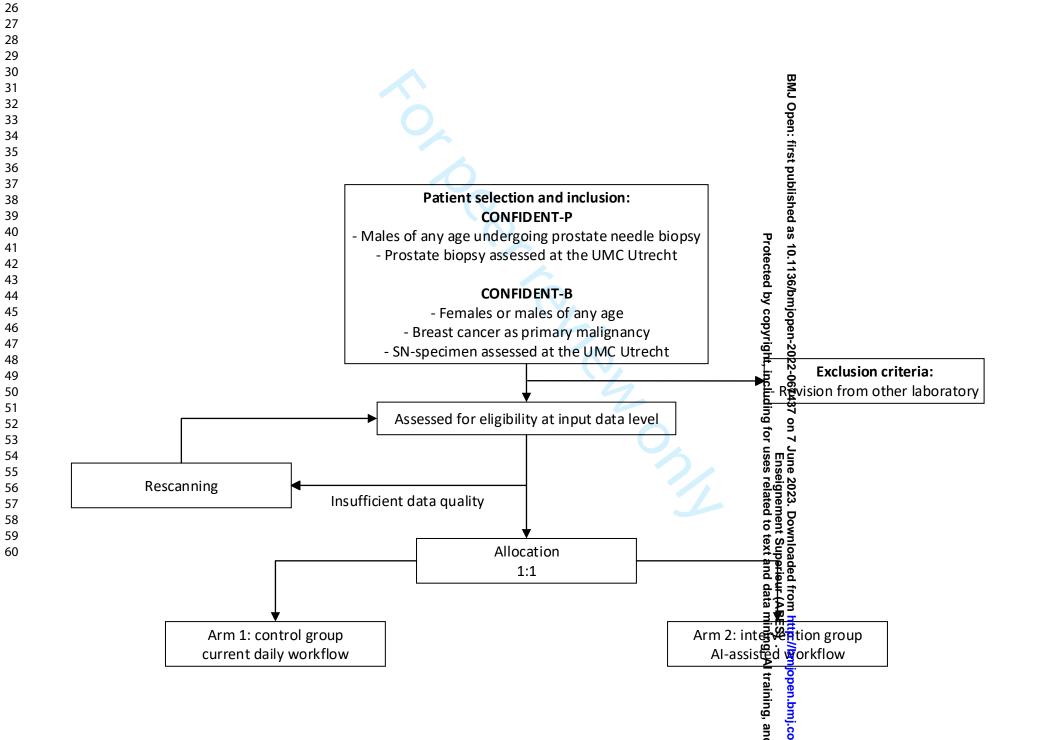
Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

2	
3	
4	
5	
6	
7	
, 0	
0	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
י∡ ר<	
∠∠ วว	
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 32 4 25 26 27 28 29 30 31 32 33 34 35 36 37 38	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
20	
3/	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
50 51	
52	
53	
54	
55	
56	
57	
58	
59	
60	

1

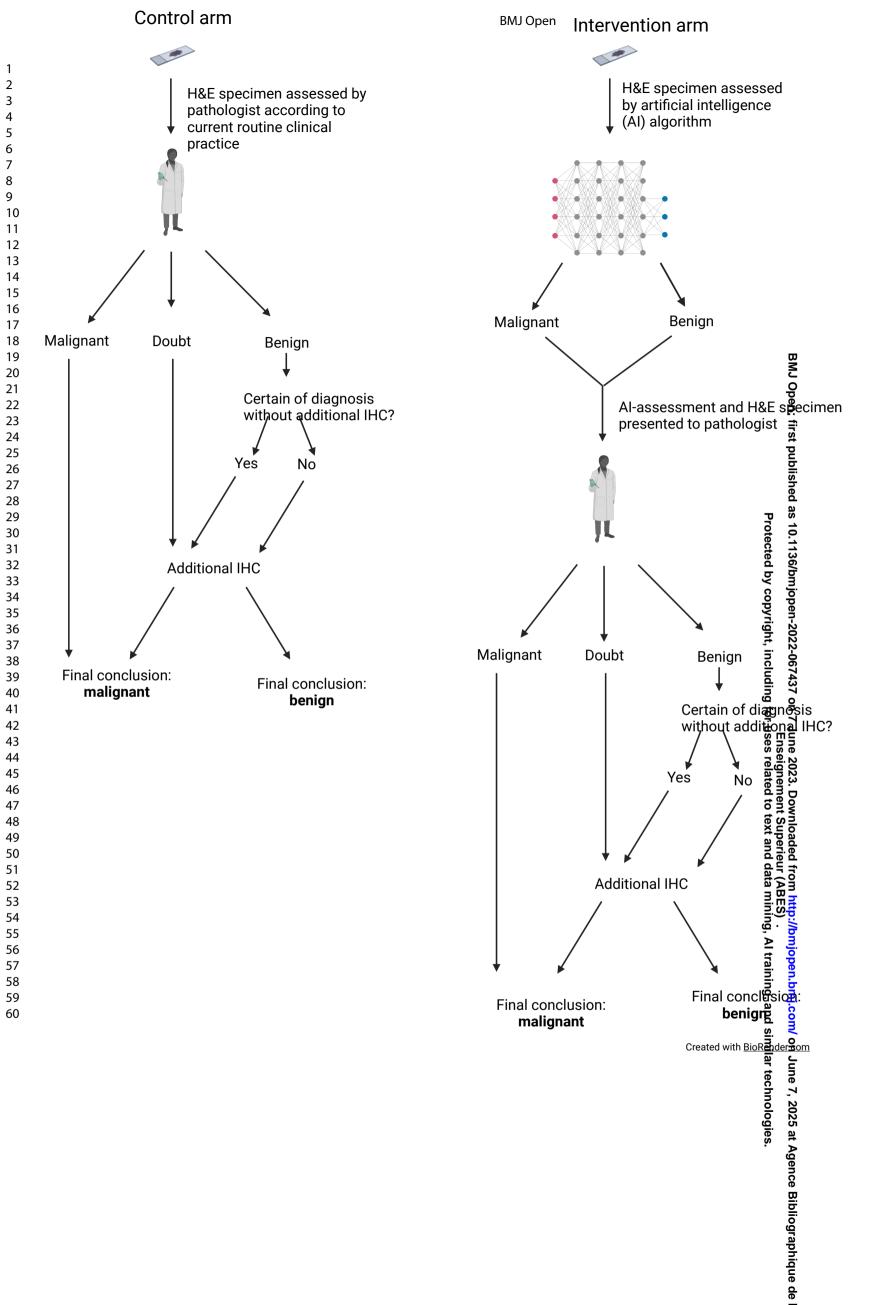
whole slide images of core needle biopsies. *Mod. Pathol.* **33**, 2058–2066 (2020).

- 19. Bejnordi, B. E. *et al.* Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA J. Am. Med. Assoc.* **318**, 2199–2210 (2017).
- 20. McNutt, L. A., Wu, C., Xue, X. & Hafner, J. P. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *Am. J. Epidemiol.* **157**, 940–943 (2003).
- 21. Lumley, T., Kronmal, R. & Ma, S. Relative risk regression in medical research: models, contrasts, estimators, and algorithms. . *UW Biostat. Work. Pap. Ser. 293 [http://www.bepress. com/uwbiostat/paper293].* (2006).
- 22. Zou, G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. *Am. J. Epidemiol.* **159**, 702–706 (2004).
- 23. R Core Team. R: A Language and Environment for Statistical Computing. at https://www.r-project.org/ (2020).
- 24. Castor Electronic Data Capture. at https://castoredc.com (2019).
- 25. Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J. & Denniston, A. K. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Lancet. Digit. Heal.* **2**, e537–e548 (2020).
- 26. Colling, R. *et al.* Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *J. Pathol. J Pathol* **249**, 143–150 (2019).
- 27. Kuijpers, C. C. H. J. *et al.* Interlaboratory variability in the histologic grading of colorectal adenocarcinomas in a nationwide cohort. *Am. J. Surg. Pathol.* **40**, 1100–1108 (2016).
- 28. Kuijpers, C. C. H. J. *et al.* Interlaboratory variability in the grading of dysplasia in a nationwide cohort of colorectal adenomas. *Histopathology* **69**, 187–197 (2016).
- Flach, R. N. *et al.* Significant Inter- and Intralaboratory Variation in Gleason Grading of Prostate Cancer: A Nationwide Study of 35,258 Patients in The Netherlands. *Cancers (Basel)*.
   13, (2021).
- 30. van Dooijeweert, C. *et al.* Significant inter- and intra-laboratory variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901 patients in the Netherlands. *Breast Cancer Res. Treat.* **174**, 479–488 (2019).
- 31. van Dooijeweert, C. *et al.* Significant inter- and intra-laboratory variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the Netherlands. *Int. J. Cancer* **146**, 769–780 (2020).



om/ on June 7, 2025 at Agence Bibliographique de l d similar technologies.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



# Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ. 2013;346:e7586

			Page
		Reporting Item	Number
Administrative information		°Z	1
Title	<u>#1</u>	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	1
Trial registration	<u>#2a</u>	Trial identifier and registry name. If not yet registered, name of intended registry	na
Trial registration: dat set	ta <u>#2b</u>	All items from the World Health Organization Trial Registration Data Set	na 1
Protocol version	<u>#3</u>	Date and version identifier	1
Funding	<u>#4</u>	Sources and types of financial, material, and other support	1
Roles and responsibilities: contributorship	<u>#5a</u>	Names, affiliations, and roles of protocol contributors	1
	For peer revi	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

#### BMJ Open

na

na

na

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de I Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

3

3

4

4

4-5

1 2 3 4 5 6	Roles and responsibilities: sponsor contact information	<u>#5b</u>	Name and contact information for the trial sponsor
7 8 9 10 11 12 13 14 15 16	Roles and responsibilities: sponsor and funder	<u>#5c</u>	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities
17 18 19 20 21 22 23 24 25	Roles and responsibilities: committees	<u>#5d</u>	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)
26 27 28 29 30 31 32 33	Introduction		
	Background and rationale	<u>#6a</u>	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention
34 35 36 37	Background and rationale: choice of	<u>#6b</u>	Explanation for choice of comparators
38 39	comparators		
40 41	Objectives	<u>#7</u>	Specific objectives or hypotheses
42 43 44 45 46 47 48	Trial design	<u>#8</u>	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)
49 50	Methods:		
51 52	Participants,		
53 54	interventions, and		
55	outcomes		
56 57 58	Study setting	<u>#9</u>	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will
59 60	Fc	or peer revie	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Page 15 of 19			BMJ Open	
1 2 3			be collected. Reference to where list of study sites can be obtained	
4 5 6 7 8 9	Eligibility criteria	<u>#10</u>	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	5
10 11 12 13 14 15	Interventions: description	<u>#11a</u>	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	0 O
16 17 18 19 20 21	Interventions: modifications	<u>#11b</u>	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	na n
22 23 24 25 26 27	Interventions: adherance	<u>#11c</u>	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	na uses related
28 29 30 31	Interventions: concomitant care	<u>#11d</u>	Relevant concomitant care and interventions that are permitted or prohibited during the trial	na
32 33 34 35 36 37 38 39 40 41 42	Outcomes	<u>#12</u>	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	6, 8 6, 8 and similarly, At training, and si
43 44 45 46 47 48	Participant timeline	<u>#13</u>	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	6 Similar recimionogies.
49 50 51 52 53 54 55	Sample size	<u>#14</u>	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	9 7
56 57 58 59 60	Recruitment Fo	<u>#15</u> r peer revie	Strategies for achieving adequate participant enrolment to reach target sample size ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	5

**BMJ** Open

Page	16 of '	19
------	---------	----

1 2 3	
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 22 23 24 25 26 27 28 29 30 31 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 37 37 37 37 37 37 37 37 37	
7 8 0	
9 10 11	
12 13 14	
15 16 17	
18 19 20	
21 22 23	
23 24 25	
20 27 28	
29 30 31	
32 33 34	
35 36 37	
38 39 40	
41 42 43	
44 45 46	
47 48 49	
50 51	
53 54	
55 56 57	
58 59	

60

Methods: Assignment of interventions (for controlled trials)		
Allocation: sequence generation	<u>#16a</u>	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions
Allocation concealment mechanism	<u>#16b</u>	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned
Allocation: implementation	<u>#16c</u>	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions
Blinding (masking)	<u>#17a</u>	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how
Blinding (masking): emergency unblinding	<u>#17b</u>	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial
Methods: Data collection, management, and analysis		
Data collection plan	<u>#18a</u>	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known.
For	ooor rovic	w only http://hmionon.hmi.com/sita/about/quidalinas.yhtml

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Page 1	7 of 19		BMJ Open		
1 2 3			Reference to where data collection forms can be found, if not in the protocol		вил оре
4 5 6 7 8 9	Data collection plan: retention	<u>#18b</u>	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	na	P
10 11 12 13 14 15 16 17 18	Data management	<u>#19</u>	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	8-9	rotected by copy
19 20 21 22 23 24 25	Statistics: outcomes	<u>#20a</u>	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	7-8	n-2022-06/43/ on / June Ense /right, including for uses
26 27 28	Statistics: additional analyses	<u>#20b</u>	Methods for any additional analyses (eg, subgroup and adjusted analyses)	7-8	eignement related to
29 30 31 32 33 34 35	Statistics: analysis population and missing data	<u>#20c</u>	Definition of analysis population relating to protocol non- adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	7-8	nloaded from http:/ Superieur (ABES) text and data minin
36 37	Methods: Monitoring				g, Al ti
<ul> <li>38</li> <li>39</li> <li>40</li> <li>41</li> <li>42</li> <li>43</li> <li>44</li> <li>45</li> <li>46</li> <li>47</li> <li>48</li> <li>49</li> </ul>	Data monitoring: formal committee	<u>#21a</u>	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	8	ES) . ES) . Ining, Al training, and similar technologies
50 51 52 53 54 55	Data monitoring: interim analysis	<u>#21b</u>	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	na	/, zuzs at Agence Bibliographique de hnologies.
56 57 58 59 60	<b>Harms</b> For	<u>#22</u> peer revie	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	8	onique de i

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de I Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

2 3 4	
2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 37 38 37 38 37 38 37 38 39 30 31 32 33 34 35 36 37 38 38 39 30 31 32 32 33 34 35 35 35 35 35 35 35 35 35 35	
9	
10 11	
12 13	
14 15	i
16	
17 18	
19 20	
21 22	
23	
24 25	
26 27	
28	
29 30	
31 32	
33 34	,
35	
36 37	
38 39	
40 41	
42	
43 44	
45 46	I
47	
48 49	
50 51	
52 53	
54	
55 56	
57 58	
59 60	1
00	

		and other unintended effects of trial interventions or trial conduct	
Auditing	<u>#23</u>	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	na
Ethics and dissemination			
Research ethics approval	<u>#24</u>	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	9
Protocol amendments	<u>#25</u>	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	na
Consent or assent	<u>#26a</u>	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	na
Consent or assent: ancillary studies	<u>#26b</u>	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	8
Confidentiality	<u>#27</u>	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	8 9
Declaration of interests	<u>#28</u>	Financial and other competing interests for principal investigators for the overall trial and each study site	9
Data access	<u>#29</u>	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	8-9
Ancillary and post trial care	<u>#30</u>	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	na
Dissemination policy: trial results	#31a peer revie	Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	9

Page	19	of	19	
· ~		•••		

#### **BMJ Open**

Page 19 of 19			BMJ Open	
1 2 3 4			public, and other relevant groups (eg, via publication, reporting in results databases, or other data sharing arrangements), including any publication restrictions	
	Dissemination policy: authorship	<u>#31b</u>	Authorship eligibility guidelines and any intended use of professional writers	9
	Dissemination policy: reproducible research	<u>#31c</u>	Plans, if any, for granting public access to the full protocol, participant-level dataset, and statistical code	9
	Appendices			
	Informed consent materials	<u>#32</u>	Model consent form and other related documentation given to participants and authorised surrogates	8
	Biological specimens	<u>#33</u>	Plans for collection, laboratory evaluation, and storage of biological specimens for genetic or molecular analysis in the current trial and for future use in ancillary studies, if applicable	na
	Commons Attribution Lie	cense C	aboration paper is distributed under the terms of the Creative IC-BY-NC. This checklist was completed on 12. August 2022 usi tool made by the EQUATOR Network in collaboration with	ing
	For	peer revie	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

**BMJ** Open

# **BMJ Open**

#### The CONFIDENT-trial protocol: a pragmatic template for clinical implementation of artificial intelligence assistance in pathology

Journal:	BMJ Open
Manuscript ID	bmjopen-2022-067437.R1
Article Type:	Protocol
Date Submitted by the Author:	22-Mar-2023
Complete List of Authors:	Flach, Rachel; University Medical Centre Utrecht, Department of Pathology Stathonikos, Nikolas; University Medical Centre Utrecht, Department of Pathology Nguyen, Tri; University Medical Centre Utrecht, Department of Pathology Ter Hoeve, Natalie; University Medical Centre Utrecht, Department of Pathology van Diest, P.J.; University Medical Centre Utrecht, Department of Pathology van Dooijeweert, Carmen; University Medical Centre Utrecht, Department of Pathology
<b>Primary Subject Heading</b> :	Pathology
Secondary Subject Heading:	Urology, Diagnostics, Oncology
Keywords:	Prostate disease < UROLOGY, Breast tumours < ONCOLOGY, PATHOLOGY

### SCHOLARONE<sup>™</sup> Manuscripts

## The CONFIDENT-trial protocol: a

# pragmatic template for clinical

## implementation of artificial intelligence assistance in pathology

- Authors
- R.N. Flach (r.n.flach-2@umcutrecht.nl), N. Stathonikos (nstatho2@umcutrecht.nl), T.Q. Nguyen
- (t.q.nguyen@umcutrecht.nl), N.D. ter Hoeve (N.terhoeve@umcutrecht.nl), P.J. van Diest
- (P.j.vandiest@umcutrecht.nl), C. van Dooijeweert (c.vandooijeweert-3@umcutrecht.nl)

#### Affiliation

Department of Pathology, University Medical Centre Utrecht, Utrecht, The Netherlands

#### Trial registration

- Not applicable

#### **Protocol version**

Version 1.

#### Funding

No funding was received for this research at the moment of writing this paper. In the meanwhile, the Hanarth Foundation has provided funding to support this study. 

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

#### **Roles and responsibilities**

- As no funding was received, and as this is a single centre clinical trial, full authority over all study activities (design; collection, management, analysis and interpretation of data; writing of the report;
- and the decision to submit the report for publication) lies with the researchers in the UMC Utrecht.

#### **Corresponding author**

- Carmen van Dooijeweert
- Internal mail no H04.312
- T: +31 88 75 565 65

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

**BMJ** Open

1 <u>c.vandooijeweert3@umcutrecht.nl</u>

to occurrence on the occurrence of the occurrenc

### Abstract

## 3 Introduction

Artificial Intelligence (AI) has been on the rise in the field of pathology. Despite promising results in
retrospective studies, and several CE-IVD certified algorithms on the market, prospective clinical
implementation studies of AI have yet to be performed, to the best of our knowledge. In this trial, we
will explore the benefits of an AI-assisted pathology workflow, while maintaining diagnostic safety
standards.

#### 15 9 17 10 Me

#### 10 Methods and analysis

This is a SPIRIT-AI compliant single-centre, controlled clinical trial, in a fully digital academic pathology laboratory. We will prospectively include prostate cancer patients who undergo prostate needle biopsies (CONFIDENT-P) and breast cancer patients who undergo a sentinel lymph node (SN) procedure (CONFIDENT-B) in the UMC Utrecht. For both the CONFIDENT-B and CONFIDENT-P trials, the specific pathology specimens will be pseudo-randomized to be assessed by a pathologist with- or without AI-assistance in a pragmatic (bi-)weekly sequential design. In the intervention group, pathologists will assess whole slide images (WSI) of the standard haematoxylin-eosin (HE)-stained sections assisted by the output of the algorithm. In the control group, pathologists will assess HE WSI according to the current clinical workflow. If no tumour cells are identified or when the pathologist is in doubt, immunohistochemistry (IHC) staining will be performed. At least 80 patients in the CONFIDENT-P and 180 patients in the CONFIDENT-B trial will need to be enrolled to detect superiority, allocated as 1:1. Primary endpoint for both trials is the number of saved resources of IHC staining procedures for detecting tumour cells, since this will clarify tangible cost savings that will support the business case for AI. 

# 26 Ethics and dissemination 39

The ethics committee (MREC NedMec) waived the need of official ethical approval, since participants
are not subjected to procedures nor are they required to follow rules. Results of both trials
(CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed journals.

## 31 Strengths and limitations

- This is the first clinical trial to examine the added value of artificial intelligence in the daily
   pathology workflow.
   By maintaining the current diagnostic safety standards patients are not at risk of an inferio
  - By maintaining the current diagnostic safety standards patients are not at risk of an inferior diagnosis during the trial.
- This is a pragmatic template for prospective AI-trials for object-identifying algorithms in
   pathology.
  - A limitation is that this is a single-centre study, which may hamper generalizability.
- 55 39
   56 39
   57 40
   56 Due to the existing clinical workflow, randomization of patients and (double-)blinding of the participating pathologists and researchers is not possible.

## 1 Keywords:

1 2 3

4

5 6

7 8 9

10 11

- 2 Artificial intelligence, prospective trial, prostate cancer, breast cancer, digital pathology
- 3 Word count: 2864 words

## 4 Introduction

### 5 Background and rationale

12 6 Since the introduction of digital pathology, the number of studies on artificial intelligence (AI) within 13 14 7 the field of pathology has increased exponentially.[1,2] Algorithms have been created for tumour 15 8 detection, tumour grading, tumour subtyping, evaluating biomarkers and more.[1,3] Due to 16 9 demographic trends, the needs for healthcare are increasing globally which combined with a lack of 17 10 specialists, increases the current workload .[2,4] Therefore, AI has great potential to alleviate 18 pathologists' workload[2] and improve diagnostics by improving accuracy, reproducibility and 19 11 20 12 speed.[2] In fact, several algorithms have shown to be comparable, or even superior to pathologists 21 13 (under time-constraint).[2,5–10] 22

Artificial and human intelligence are not mutually exclusive, they complement each other, a concept which is known as "augmented intelligence", where Al can enhance, rether then replace human

- which is known as "augmented intelligence", where AI can enhance, rather than replace human
   intelligence.[11] In the (very) early AI-adoption phase, and presumably also in later phases,
- intelligence.[11] In the (very) early Al-adoption phase, and presumably also in later phases,
   pathologist supervision remains of key importance. This is particularly relevant as, despite the
- pathologist supervision remains of key importance. This is particularly relevant as, despite the
   promising results of retrospective studies and the availability of CE-IVD approved algorithms,
- <sup>29</sup> 19 prospective validation and clinical implementation of AI is currently lacking. For example, six years
- after the successful CAMELYON-16 Grand Challenge[6], the top algorithms have yet to be
- implemented in daily clinical practice, showing that the time between development of an AI model
- and clinical implementation is considerable. Likewise, numerous promising prostate cancer grading
- algorithms have been developed, yet implementation studies have yet to be performed,[12–14]
- whereas nine AI pathology devices received CE-IVD approval in 2021.[15]
   whereas nine AI pathology devices received CE-IVD approval in 2021.[15]

## <sup>37</sup><sub>38</sub> 25 Trial rationale

As a pathology laboratory with a fully digital workflow for over seven years, we are eager to explore the full potential of working digitally by adding the benefit of AI in daily pathology practice. We decided to start with the object localisation task of tumour-detection, where an objective reference standard is in place in the routine clinical workflow (i.e. pathologist-supervision and/or

44 30 immunohistochemistry (IHC) staining).

4531We developed the CONFIDENT-trial template, in which AI tumour detection algorithms can4632be safely implemented in prospective clinical trials, while ensuring that patients are not at risk of4733receiving an inferior diagnosis, since IHC is always performed when no tumour cells are visible, but4934also when pathologists need more confirmation about the diagnosis.

#### 50 51 35 CONFIDENT-B and CONFIDENT-P

52 36 Our trials aim to prospectively investigate the added value of an AI-assisted pathology workflow in 53 37 the identification of prostate cancer (PCa) in prostate needle biopsies (CONFIDENT-P) and the 54 55 38 identification of sentinel lymph node (SN) metastases in breast cancer (BCa) patients (CONFIDENT-B). 56 39 As both prostate cancer and breast cancer are the most common (non-skin) malignancies in men and 57 40 women, respectively, implementation of AI-assistance may have a great impact on diagnostic 58 41 processes.[16] However, it is important to emphasize that this trial serves as a template for other 59 42 pragmatic AI-intervention trials for object-localisation tasks as well. 60

- We obtained CE-IVD-approved algorithms for detection and grading of prostate cancer in prostate needle biopsies and an algorithm for detecting lymph node metastases in BCa patients. In both cases the task of the pathologist is both labour-intensive and expensive, due to the performed IHC stains in case no tumour cells are morphologically observed. However, IHC is expensive and these costs sometimes even exceed reimbursement for the entire specimen (e.g. in case of multiple blocks of multiple SNs). This raises the question whether AI may be of added value to morphologically detect cancer cells without the need for IHC-use. Thereby the number of performed IHC stains may be reduced, which may lead to tangible costs savings that will help to build the business case for AI, while potentially decreasing the workload of pathologists as well.[2]
- <sup>15</sup> 10 Study objective

The primary objective is to explore whether an AI-assisted workflow reduces the number of spent resources on IHC, while maintaining diagnostic safety standards in both PCa patients who underwent prostate needle biopsies (CONFIDENT-P), and BCa patients who underwent an SN procedure. Secondary objectives are to investigate whether time management improves in an Al-assisted workflow and to analyse how many IHC staining may have been safely omitted after AI-implementation. 

### 17 Methods and analysis

#### 18 Trial design

The study protocol is structured following the SPIRIT-AI (Standard Protocol Items: recommendations for Interventional Trials - Artificial Intelligence) statement 2020[17]. This study is a single-centre, parallel-group controlled trial, assessing superiority. The allocation ratio is 1:1. Eligible patients will be assigned to arm 1 (control group) or arm 2 (AI-assisted workflow), based on a bi-weekly time-schedule. Eligibility criteria are summarised in Figure 1. The CONFIDENT trials will be carried out in 2022-2023. 

## <sup>37</sup> 25 Study setting

The trial will take place in the daily practice of a single academic hospital (University Medical Centre Utrecht, the Netherlands), with a fully digital pathology enabled clinical set-up, where all slides are digitised using ultrafast whole slide image (WSI) using Hamamatsu S360 scanners and reviewed using the Sectra pathology Picture Archiving and Communication System (PACS). Although the UMC Utrecht is an academic hospital, primary routine pathology diagnostics is performed for non-academic hospitals as well (i.e., the Alexander Monro Breast Cancer Hospital, Bilthoven, the Netherlands). 

#### 49 33 Study population

For PCa, WSI of all males who undergo a prostate needle biopsy in the UMC Utrecht will be included.
 For BCa, WSI of all females or males with BCa as primary malignancy (i.e. invasive breast cancer) who
 undergo an SN procedure in the Alexander Monro Breast Cancer Hospital or the UMC Utrecht will be
 included. Patients will be excluded, if they were redirected to the UMC Utrecht for a second opinion.

#### 56 38 Assessment of specimen

During the study period, all WSI will be assessed by the same group of pathologists; i.e. two expert
urological pathologists for the PCa biopsies, and three expert breast pathologists for the lymph node
assessment from BCa patients.

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

For both the CONFIDENT-B and CONFIDENT-P trial, the specific pathology specimens will be assigned to be assessed by a pathologist with or without AI-assistance in a pragmatic (bi-)weekly sequential design. This is considered feasible as the change in the case mix and time trends are unlikely to occur within the inclusion-period of about six to nine months. Furthermore, both specialized breast- and urological pathologists within the UMC Utrecht work according to weekly schedules. Therefore, using AI every other week or every other two weeks, as opposed to switching by day, ensures that all pathologists are equally distributed between groups. Lastly, it would be impractical to switch from AI-assistance in the intervention group to no AI-assistance in the control group on a case-to-case basis. For obvious reasons, allocation concealment and blinding of pathologists and researchers is not applicable.

#### 12 Control and intervention

All eligible specimens will be assigned to either the control group or the intervention group. In the control group, pathologists will assess H&E stained WSI of patients digitally, according to the current clinical workflow. For PCa biopsies, IHC is routinely performed on all cases. For BCa lymph nodes, if no metastases or tumour are present, IHC staining will be performed. Additional IHC staining will also be performed by additional request of the pathologist in case of doubt. 

In the intervention group, pathologists will assess the H&E-specimens digitally with the outcome of the algorithm provided in their first assessment of the specimen. For PCa, they will use the CE-IVD certified Paige Prostate Suite algorithms for tumor detection and tumor volume percentage calculations, which reaches sensitivity and specificity of 99% and 93% respectively and which are based on a weakly-supervised deep learning algorithm as described by Campanella et al.[18,19]For BCa, pathologists will use the CE-IVD certified Metastasis Detection App by Visiopharm, a deep-learning algorithm for lymph node metastases of BCa and colon carcinoma with a combined sensitivity and specificity of 98,7 and 99.6% respectively[20]. These algorithms will be integrated within the Sectra PACS where the output of the algorithms will be graphically displayed. Al analysis of the WSI will be performed right after scanning to avoid delays in the clinical workflow. If the AI-assisted pathologist does not detect metastases or tumours on the H&E slide, routine additional IHC staining will be performed by P503S/p63/CK HMW for PCa and CAM5.2 for BCa, to ensure no metastases or tumours are missed. Pathologists can also request an additional IHC if they feel they need this to make an adequate diagnosis (Figure 2). 

# 4344 32 Outcome measures

Primary outcome for the CONFIDENT-P trial is the added value of AI-assistance in the detection of PCa and in the detection of tumour volume in prostate needle biopsies in daily pathology practice. The primary outcome for the CONFIDENT-B trial is the added value of AI-assistance in the detection of BCa SN-metastases. The outcome measures for both trials will be the number of spent resources, i.e. the number of IHC-stains performed in both groups. 

Secondary outcome measures will be sensitivity and specificity of the AI-assisted pathologist, time spent on WSI analysis, the number of IHC stains that may have been omitted after AI-implementation and a pathologists' evaluation by a questionnaire on the Al-assisted work process. Sensitivity and specificity analyses of the algorithm itself have already been well documented, and is therefore outside the scope of the paper, as we focus on the combination of pathologist and AI to explore cost savings. 

#### Input data

Input data for the algorithm will be WSI of haematoxylin and eosin (H&E) stained slides scanned at

40x of either prostate needle biopsies, and WSI of H&E stained slides of BCa lymph nodes. As per

routine in our daily clinical practice, WSI will be quality controlled after scanning for colour, focus

quality and completeness of the scan. When necessary, the specimens are rescanned.

#### Sample size

#### **CONFIDENT-P**

We performed power calculations using a two-sample proportion superiority test, using expected percentages of IHC staining in both study arms. We assume that the pathologists in the control arm can detect 50% of the tumours without using IHC. We expect AI-assisted pathologists to detect 80% of the tumours, without using IHC. These percentages were conservatively derived from the validity study by Raciti (74% for pathologists without AI and 90% for pathologists with AI respectively)[21], by expert pathologist opinion, and taking into account that pathologists under time constraint of daily practice do not detect tumours as well as pathologists without time constraint during retrospective studies[22]. We assume that this effect will be larger for the biopsies assessed without AI than with AI, as AI is assumed to make tumour detection easier.

A sample size of 60 (30 per arm) would give a power of approximately 80%, using a one-sided 5% significance level. However, uncertainties remain regarding the sample size parameters. We therefore inflated our sample size to 80 (40 per arm), in order to ensure study power and allowing us to detect smaller effect sizes.

For detection of tumour volume percentage, we performed a power calculation based on the assumption that AI should be able to replace at least 20% of the IHC stains, in order to be cost-effective. IHC is currently used in 100% of all prostate needle biopsies. Using a power of 80% and a one-sided significance level of 5%, this leads to 27 patients per arm. 

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

#### **CONFIDENT-B**

Sample size calculations for the CONFIDENT-B trial are based on the assumption that the Al-

algorithm can detect all metastases for which currently IHC is used, which are mainly micro-

metastases and isolated tumour cells (ITC). Approximately 15% of the SN-specimens in the UMC Utrecht contain a micrometastasis or ITC.

A sample size of 166 patients (83 per arm) with a one-sided 5% significance level therefore results in a power of 80%. Again, as there are uncertainties on the assumptions on what amount of the metastases will be detect by AI, we decided to be conservative and include 180 patients (90 per arm).

Overall, we are only interested in one-sided outcomes, as it is not possible that more IHC will be performed in the AI-assisted arm. IHC is performed to detect metastases, when they are macroscopically undetectable, rather than to confirm them when they are macroscopically visible. As Al would show only more metastases than the pathologist could macroscopically detect, this means that only a reduction of IHC is possible.

#### Statistical methods

For baseline comparisons between both arms, the appropriate measures (parametric or non-parametric) for categorical (Chi-squared test/Fisher's exact) and continuous variables (T-test, Mann-

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

Whitney U test) will be used. For the analysis of the primary outcome measure, we will compare the proportion of IHC-use in both arms, and calculate adjusted relative risks, using a log-binomial model.[23-25] Missing data for baseline characteristics and for the primary outcome are not to be expected, as they are obligatory items in the structured pathology reports. We will determine sensitivity and specificity of the conclusion of the AI-assisted pathologists without the use of IHC. Subsequently, we will focus on the cases with metastases that the AI-assisted pathologist misses, categorize them (i.e. macro-metastases, micro-metastases, ITC) and determine their clinical relevance (i.e. clinical consequences if these metastases are being missed). Data-analysis will be performed in R Statistical Software[26], with a significance level set at p<0.05. Data collection and management All data (baseline and primary outcome measurements) will be retrieved from the structured pathology reports and will be managed and stored in Castor EDC[27]. For the secondary outcome measure of time spent by the pathologist on a slide, data will be collected on an interval basis for practical reasons, as timing every assessment for months (by stopwatch) was not deemed feasible. For the secondary outcome measure of Al-assisted work process for pathologists, a questionnaire will be distributed to the participating pathologists. The final secondary assessment measurement, the number of IHC stains that may have been omitted after AI-implementation, will be determined by the researchers based on the data from the structured pathology reports (combination of IHC and Al-assisted diagnoses of the pathologist). Ethical approval Research within these trials is not subject to the (Dutch) Medical Research Involving Human Subjects Act (WMO), as participants are not subjected to procedures and as they are not required to follow rules. Therefore, the ethics committee (MREC NedMec) waived the need of ethical approval and informed consent. Risk of harm Patients are not at risk of any harm for an inferior diagnosis (i.e. missed tumour cells), as in both arms, IHC-staining will be performed when no tumour cells are visible, according to current clinical standards. As a rule in augmented intelligence, all cases will be evaluated by a pathologist, which further minimized the risk of a false diagnosis based on the AI algorithm. Taking all of the above into account, a data monitoring committee (DMC) is not required and adverse events are not to be expected. In theory, the algorithm could be more of a disturbance than a help to pathologists (for example, when it frequently reports false positive or negative results, which have to be corrected by the pathologist). However, the algorithms used are IVDR-approved, and thus have undergone extensive review for their intended purpose. Nonetheless, the experience and ease of use of pathologists working with the algorithm will be one of the secondary outcome measures. Informed consent and data access Informed consent was waived by the local quality coordinator (QC) and data protection officer (DPO) for the following reasons. First, in both arms patients receive standard care, while maintaining diagnostic safety standards (pathologists' supervision, IHC in all negative cases). Second, patients are not subjected to any procedures. Third, all patient data will be anonymized to the researchers by the

**BMJ** Open

2		
3	1	The collected (anonymous) research data will be stored in Castor EDC to ensure data
4 5	2	security. Data will be kept for a period of 15 years. Data access in Castor will be restricted to two
6	3	researchers (RF, CvD). Pathologists have access to the electronic patient files for the purpose of
7	4	patient care. The researchers are not permitted access to these files. At no point will the data (both
8	5	in Castor EDC and patient files) be accessed by the companies providing the algorithms (i.e.
9	6	Visiopharm and Paige).
10 11		
12	7	Patient and Public Involvement
13	8	None
14		
15 16	9	Discussion
17	10	The promising retrospective results of AI-assisted pathology have not yet resulted in prospective
18	11	clinical implementation studies. This may be due to a lack of digital transition in the majority of
19 20	12	pathology laboratories, but it may also be partly due to the lack of a good implementation model.
20 21	13	Fortunately, however, new guidelines for AI-trials have recently been proposed by the SPIRIT-AI and
22	14	CONSORT-AI-steering groups, as well as roadmaps to routine use of AI in clinical practice.[28–30] Yet,
23	15	to date, no pathology AI-trials have been published in PubMed or Web of Science or, to the best of
24	16	our knowledge, otherwise made public.
25 26	17	As a pathology laboratory with a fully digital workflow, we developed a clinical trial template
20	18	for tumour detection models, as a first step to implement AI in daily pathology practice. We will start
28	19	with an object localisation task (i.e. tumour cells) as a reference standard is in place in the routine
29	20	clinical workflow. For classification tasks like tumour grading, a clinical trial design is more
30 31	21	challenging, as no reference is in place in daily pathology practice and inter-laboratory and inter-
32	22	pathologist variation is notorious.[31–35] Nevertheless, in future trials, implementing Al-assistance in
33	23	the grading process might also reduce this variation. For now, results of the CONFIDENT-trials will
34	24	provide the first assessment of the potential added value of AI in daily pathology practice. This
35 36	25	evaluation will substantially contribute to a potential paradigm shift in tumour detection in
37	26	pathology. The pragmatic template of the CONFIDENT trials may serve as example for other
38	27	prospective AI implementation trials in diagnostic pathology.
39		
40	28	Declaration of interests
41 42	29	PJvD is a member of the Scientific Advisory Board of Paige and Sectra.
43	30	All other authors do not report conflict of interest.
44		
45 46	31	Funding
40 47	32	No funding was obtained at the moment of writing this paper. In the meanwhile, the Hanarth
48	33	Foundation has provided funding to support this study.
49		
50 51	34	Dissemination policy
52	35	Results of both trials (CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed
53	36	journals. Authorship will be acknowledged to all those that substantially contributed in the
54	37	CONFIDENT-trials. Data will be available upon reasonable request.
55 56		
57	38	Author statement
58	39	PJvD conceived of the study. RNF and CvD initiated the study design and NS, TQN and
59		
60	40	NDtH helped with implementation. RNF and CvD provided statistical expertise in clinical
		9
		5

- 1 trial design and RNF and CvD are conducting the primary statistical analysis. All authors
  - 2 contributed to refinement of the study protocol and approved the final manuscript.

#### 3 Figure legends

- 4 Figure 1. Flowchart with patient selection; SN = sentinel node; AI = artificial intelligence.
- 5 Figure 2. Study flow chart. H&E = hematoxylin & eosin; IHC = immunohistochemistry

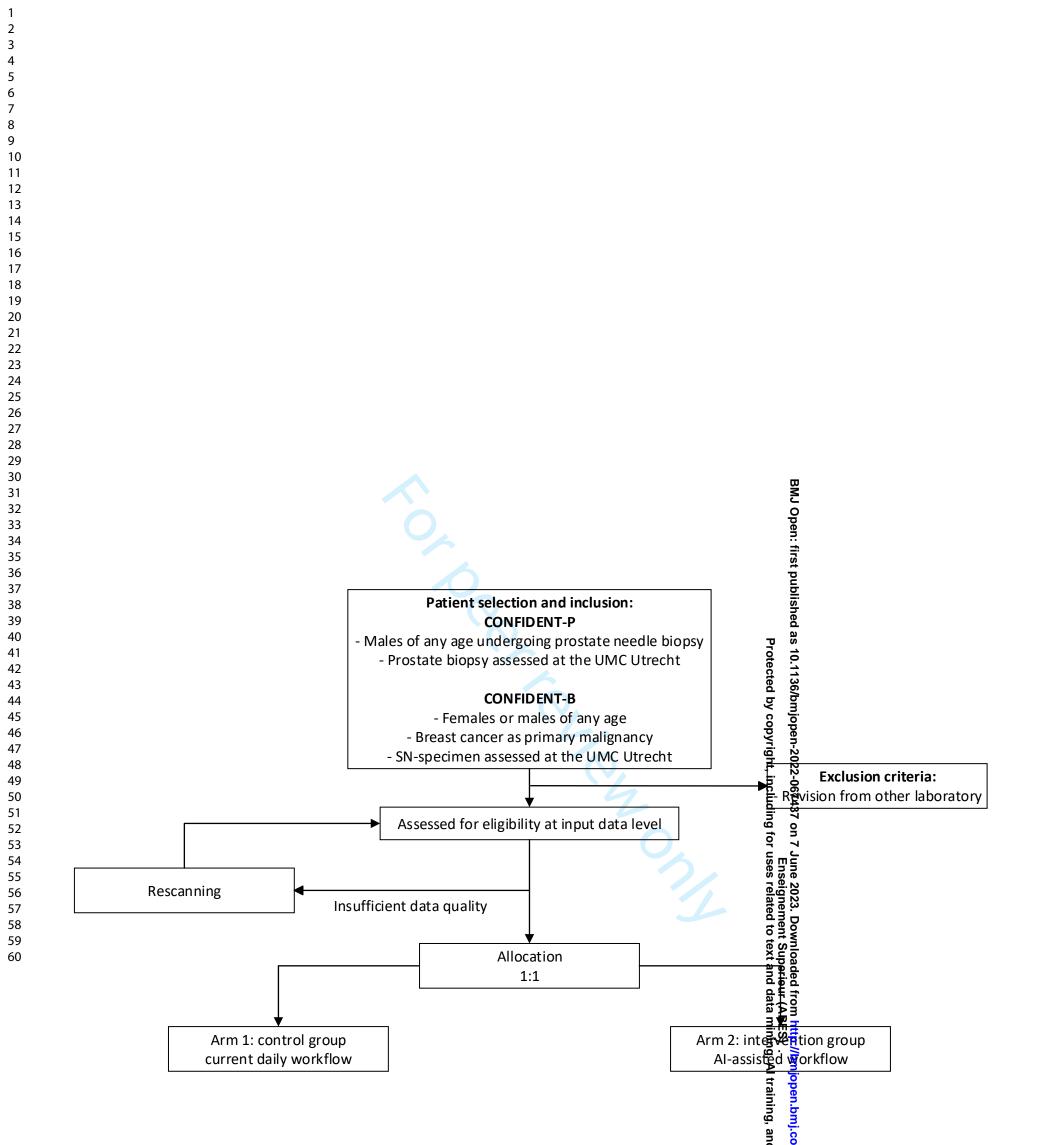
for peer teries only

1 2			
3 4	1	Ref	erences
4 5	2		
6 7 8	3 4	1	Jiang Y, Yang M, Wang S, <i>et al</i> . Emerging role of deep learning-based artificial intelligence in tumor pathology. <i>Cancer Commun</i> 2020; <b>40</b> :154–66. doi:https://doi.org/10.1002/cac2.12012
9 10 11	5 6	2	van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. <i>Nat</i> <i>Med</i> 2021; <b>27</b> :775–84. doi:10.1038/s41591-021-01343-4
12 13 14 15 16	7 8 9	3	Steiner DF, Chen PHC, Mermel CH. Closing the translation gap: AI applications in digital pathology. <i>Biochim Biophys Acta Rev Cancer</i> 2021; <b>1875</b> :188452. doi:10.1016/j.bbcan.2020.188452
17 18 19 20	10 11 12	4	Robboy SJ, Weintraub S, Horvath AE, <i>et al.</i> Pathologist workforce in the United States: I. Development of a predictive model to examine factors influencing supply. <i>Arch Pathol Lab</i> <i>Med</i> 2013; <b>137</b> :1723–32. doi:10.5858/arpa.2013-0200-OA
21 22 23 24	13 14 15	5	Bulten W, Pinckaers H, van Boven H, <i>et al</i> . Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study. <i>Lancet Oncol</i> 2020; <b>21</b> :233–41. doi:10.1016/S1470-2045(19)30739-9
25 26 27 28 29	16 17 18	6	Ehteshami Bejnordi B, Veta M, Johannes van Diest P, <i>et al.</i> Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA 2017; <b>318</b> :2199–210. doi:10.1001/jama.2017.14585
30 31 32 33	19 20 21	7	Steiner DF, Nagpal K, Sayres R, <i>et al.</i> Evaluation of the Use of Combined Artificial Intelligence and Pathologist Assessment to Review and Grade Prostate Biopsies. <i>JAMA Netw Open</i> 2020; <b>3</b> :e2023267. doi:10.1001/jamanetworkopen.2020.23267
34 35 36 37	22 23 24	8	Ström P, Kartasalo K, Olsson H, <i>et al</i> . Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. <i>Lancet Oncol</i> 2020; <b>21</b> :222–32. doi:10.1016/S1470-2045(19)30738-7
38 39 40 41	25 26 27	9	Raciti P, Sue J, Ceballos R, <i>et al.</i> Novel artificial intelligence system increases the detection of prostate cancer in whole slide images of core needle biopsies. <i>Modern Pathology</i> 2020; <b>33</b> :2058–66. doi:10.1038/s41379-020-0551-y
42 43 44 45 46	28 29 30	10	Hekler A, Utikal JS, Enk AH, <i>et al.</i> Superior skin cancer classification by the combination of human and artificial intelligence. <i>Eur J Cancer</i> 2019; <b>120</b> :114–21. doi:10.1016/j.ejca.2019.07.019
47 48 49 50	31 32 33	11	Harrison JH, Gilbertson JR, Hanna MG, <i>et al.</i> Introduction to Artificial Intelligence and Machine Learning for Pathology. <i>Arch Pathol Lab Med</i> 2021; <b>145</b> :1228–54. doi:10.5858/arpa.2020- 0541-cp
51 52 53 54	34 35 36	12	Ström P, Kartasalo K, Olsson H, <i>et al</i> . Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study. <i>Lancet Oncol</i> 2020; <b>21</b> :222–32. doi:10.1016/S1470-2045(19)30738-7
55 56 57 58 59 60	37 38 39	13	Steiner DF, Nagpal K, Sayres R, <i>et al.</i> Evaluation of the Use of Combined Artificial Intelligence and Pathologist Assessment to Review and Grade Prostate Biopsies. <i>JAMA Netw Open</i> 2020; <b>3</b> :1–14. doi:10.1001/jamanetworkopen.2020.23267

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

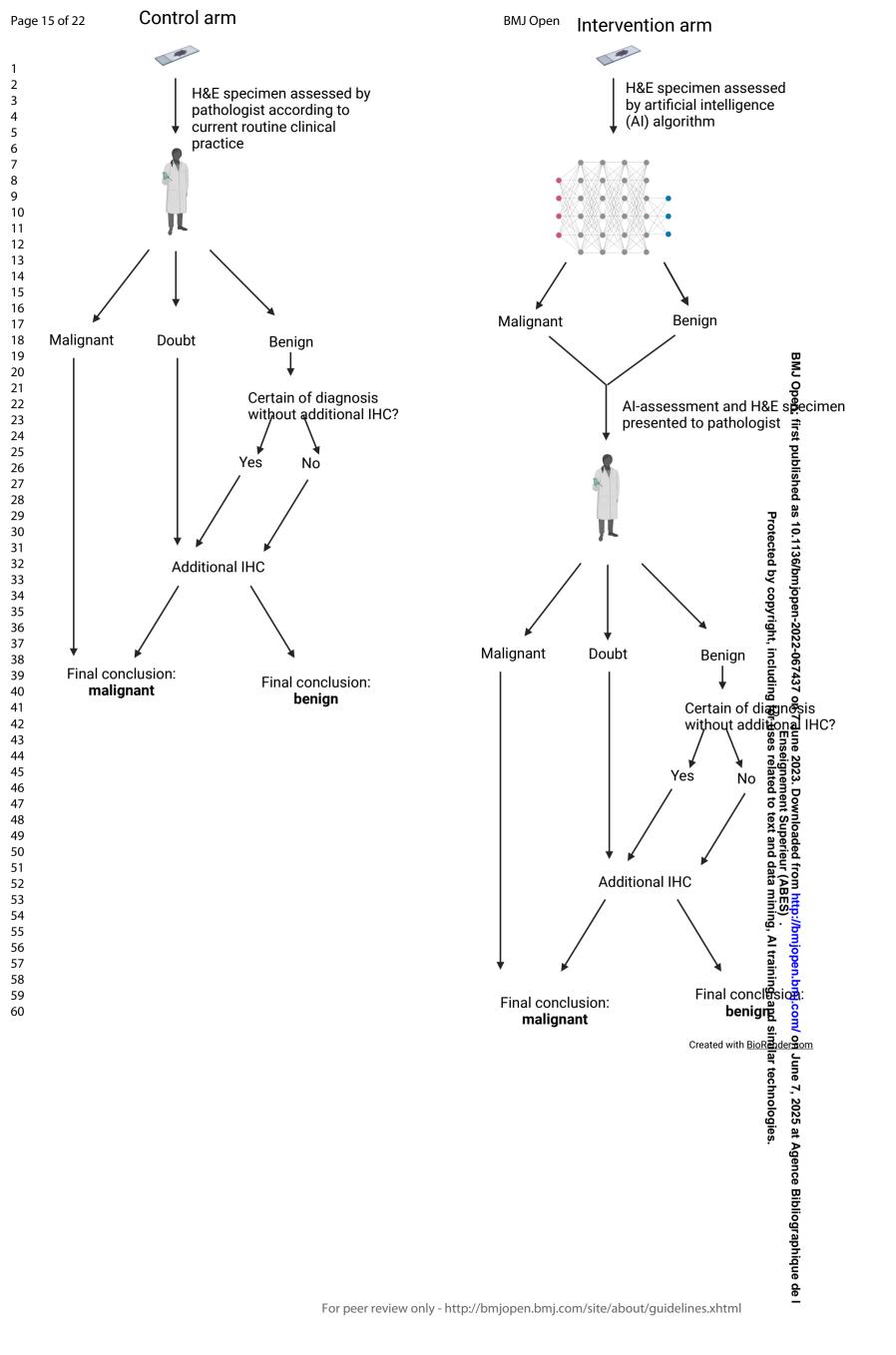
1 2			
3	1	14	Bulten W, Balkenhol M, Belinga JJA, et al. Artificial intelligence assistance significantly
4	2		improves Gleason grading of prostate biopsies by pathologists. Modern Pathology
5 6	3		2021; <b>34</b> :660–71. doi:10.1038/s41379-020-0640-y
7	4	15	Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine
8 9	5	10	learning-based medical devices in the USA and Europe (2015–20): a comparative analysis.
9 10	6		Lancet Digit Health 2021; <b>3</b> :e195–203. doi:10.1016/S2589-7500(20)30292-2
11	-	10	Suns II. Farlan I. Garad Bl., et al. Clabel Concern Statistics 2020; CLOBOCAN Estimates of
12 13	7 8	16	Sung H, Ferlay J, Siegel RL, <i>et al.</i> Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. <i>CA Cancer J Clin</i>
14	9		2021; <b>71</b> :209–49. doi:10.3322/caac.21660
15			
16 17	10	17	Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions
18	11		involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020;2:e549–60.
19	12		doi:10.1016/S2589-7500(20)30219-3
20	13	18	Campanella G, Hanna MG, Geneslaw L, et al. Clinical-grade computational pathology using
21 22	14		weakly supervised deep learning on whole slide images. Nat Med 2019; <b>25</b> :1301–9.
23	15		doi:10.1038/s41591-019-0508-1
24 25	16	19	da Silva LM, Pereira EM, Salles PGO, et al. Independent real-world application of a clinical-
25 26	17		grade automated prostate cancer detection system. <i>Journal of Pathology</i> 2021; <b>254</b> :147–58.
27	18		doi:10.1002/path.5662
28	10	20	Visionharm https://visionharm.com/ann.conter/ann/matastasis.datastion.ci/
29 30	19	20	Visiopharm. https://visiopharm.com/app-center/app/metastasis-detection-ai/.
31	20	21	Raciti P, Sue J, Ceballos R, et al. Novel artificial intelligence system increases the detection of
32	21		prostate cancer in whole slide images of core needle biopsies. Modern Pathology
33 34	22		2020; <b>33</b> :2058–66. doi:10.1038/s41379-020-0551-y
35	23	22	Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for
36	24		detection of lymph node metastases in women with breast cancer. JAMA - Journal of the
37 38	25		American Medical Association 2017; <b>318</b> :2199–210. doi:10.1001/jama.2017.14585
39	26	23	McNutt LA, Wu C, Xue X, et al. Estimating the relative risk in cohort studies and clinical trials
40	27	20	of common outcomes. Am J Epidemiol 2003; <b>157</b> :940–3. doi:10.1093/aje/kwg074
41 42	20	24	
43	28	24	Lumley T, Kronmal R, Ma S. Relative risk regression in medical research: models, contrasts,
44	29 30		estimators, and algorithms UW Biostatistics Working Paper Series 293 [http://www bepress com/uwbiostat/paper293] 2006.
45 46	50		
47	31	25	Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J
48	32		<i>Epidemiol</i> 2004; <b>159</b> :702–6. doi:10.1093/aje/kwh090
49 50	33	26	R Core Team. R: A Language and Environment for Statistical Computing. 2020.
51			
52	34	27	Castor Electronic Data Capture. 2019.
53 54	35	28	Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for
55	36		interventions involving artificial intelligence: the CONSORT-AI extension. Lancet Digit Health
56	37		2020; <b>2</b> :e537–48. doi:10.1016/S2589-7500(20)30218-1
57 58	38	29	Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions
59	39	-	involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020; <b>2</b> :e549–60.
60	40		doi:10.1016/S2589-7500(20)30219-3

1			
2		_	
3 4	1	30	Colling R, Pitman H, Oien K, et al. Artificial intelligence in digital pathology: a roadmap to
5	2		routine use in clinical practice. <i>Journal of Pathology J Pathol</i> 2019; <b>249</b> :143–50.
6	3		doi:10.1002/path.5310
7	4	31	Kuijpers CCHJ, Sluijter CE, von der Thüsen JH, et al. Interlaboratory variability in the histologic
8	5	51	grading of colorectal adenocarcinomas in a nationwide cohort. American Journal of Surgical
9 10	6		Pathology 2016;40:1100–8. doi:10.1111/his.12923
10	U		<i>1 uthology</i> 2010, <b>40</b> .1100 0. uol.10.1111/lli3.12325
12	7	32	Kuijpers CCHJ, Sluijter CE, von der Thüsen JH, et al. Interlaboratory variability in the grading of
13	8		dysplasia in a nationwide cohort of colorectal adenomas. <i>Histopathology</i> 2016;69:187–97.
14	9		doi:10.1111/his.12923
15 16	10	33	Flach BNI Willomso D DNA Suchmann BBNA at al Significant Inter and Intralaboratory
17		55	Flach RN, Willemse P-PM, Suelmann BBM, <i>et al.</i> Significant Inter- and Intralaboratory
18	11		Variation in Gleason Grading of Prostate Cancer: A Nationwide Study of 35,258 Patients in
19	12		The Netherlands. <i>Cancers (Basel)</i> 2021; <b>13</b> . doi:10.3390/cancers13215378
20	13	34	van Dooijeweert C, van Diest PJ, Willems SM, et al. Significant inter- and intra-laboratory
21 22	14		variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901
22	15		patients in the Netherlands. Breast Cancer Res Treat 2019;174:479–88. doi:10.1007/s10549-
24	16		018-05082-y
25			
26	17	35	van Dooijeweert C, van Diest PJ, Willems SM, et al. Significant inter- and intra-laboratory
27	18		variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the
28 29	19		Netherlands. Int J Cancer 2020; <b>146</b> :769–80. doi:10.1002/ijc.32330
30	20		
31	20		
32			
33			
34 35			
36			
37			
38			
39			
40 41			
42			
43			
44			
45 46			
46 47			
48			
49			
50			
51			
52 53			
53 54			
55			
56			
57			
58 59			
17			



om/ on June 7, 2025 at Agence Bibliographique de l d similar technologies.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



# Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ. 2013;346:e7586

30								
31 32	Reporting Item							
<ul> <li>33</li> <li>34</li> <li>35</li> <li>36</li> <li>37</li> <li>38</li> <li>39</li> <li>40</li> <li>41</li> <li>42</li> <li>43</li> <li>44</li> <li>45</li> <li>46</li> <li>47</li> </ul>	Administrative information			Number no ora ara mining, Artraining, 1				
	Title	<u>#1</u>	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	Al training				
	Trial registration	<u>#2a</u>	Trial identifier and registry name. If not yet registered, name of intended registry	na ang				
	Trial registration: data set	<u>#2b</u>	All items from the World Health Organization Trial Registration Data Set	na na 1				
48 49 50	Protocol version	<u>#3</u>	Date and version identifier	gles. 1 s.				
50 51 52 53 54 55 56 57 58 59	Funding	<u>#4</u>	Sources and types of financial, material, and other support	1				
	Roles and responsibilities: contributorship	<u>#5a</u>	Names, affiliations, and roles of protocol contributors	1				
60	For	peer revi	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml					

### BMJ Open

1 2 3 4 5 6	Roles and responsibilities: sponsor contact information	<u>#5b</u>	Name and contact information for the trial sponsor	na			
7 8 9 10 11 12 13 14 15 16	Roles and responsibilities: sponsor and funder	<u>#5c</u>	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	na			
17 18 19 20 21 22 23 24 25	Roles and responsibilities: committees	<u>#5d</u>	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	na			
26 27	Introduction						
28 29 30 31 32 33 34	Background and rationale	<u>#6a</u>	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	3			
35 36 37 38	Background and # rationale: choice of comparators		Explanation for choice of comparators	3			
	Objectives	<u>#7</u>	Specific objectives or hypotheses	4			
	Trial design	<u>#8</u>	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	4			
	Methods:						
	Participants,						
	interventions, and outcomes						
55 56 57 58 59	Study setting	<u>#9</u>	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will	4-5			
60	For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml						

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de I Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

1 2			be collected. Reference to where list of study sites can be obtained	
3 4 5 6 7 8 9	Eligibility criteria	<u>#10</u>	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	5
10 11 12 13 14	Interventions: description	<u>#11a</u>	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	6
15 16 17 18 19 20 21	Interventions: modifications	<u>#11b</u>	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	na
22 23 24 25 26 27	Interventions: adherance	<u>#11c</u>	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	na
28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55	Interventions: concomitant care	<u>#11d</u>	Relevant concomitant care and interventions that are permitted or prohibited during the trial	na
	Outcomes	<u>#12</u>	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	6, 8
	Participant timeline	<u>#13</u>	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	6
	Sample size	<u>#14</u>	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	7
56 57 58 59 60	Recruitment	<u>#15</u> peer revie	Strategies for achieving adequate participant enrolment to reach target sample size ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	5

BMJ Open

1 2 3 4 5 6	Methods: Assignment of interventions (for controlled trials)			
7 8 9 10 11 12 13 14 15 16 17 18	Allocation: sequence generation	<u>#16a</u>	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	5
19 20 21 22 23 24 25 26 27 28 29 30	Allocation concealment mechanism	<u>#16b</u>	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	na
	Allocation: implementation	<u>#16c</u>	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	na
31 32 33 34 35	Blinding (masking)	<u>#17a</u>	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	na
36 37 38 39 40	Blinding (masking): emergency unblinding	<u>#17b</u>	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	na
41 42 43 44 45 46 47	Methods: Data collection, management, and analysis			
44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59	Data collection plan	<u>#18a</u>	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known.	8
60	For	peer revie	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

		BMJ Open					
1 2 3			Reference to where data collection forms can be found, if not in the protocol	BMJ Ope			
4 5 6 7 8 9	Data collection plan: retention	<u>#18b</u>	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	BMJ Open: first published a na			
10 11 12 13 14 15 16 17 18	Data management	<u>#19</u>	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	as 10.1136/bmjopen-2022-067437 on 7 June Ens Protected by copyright, including for uses 99 8 <sup>-9</sup> 7 <sup>-8</sup> 7 <sup>-</sup>			
19 20 21 22 23 24 25	Statistics: outcomes	<u>#20a</u>	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	2-067437 on 7 June Ensi , including for uses 7-8 7			
26 27 28	Statistics: additional analyses	<u>#20b</u>	Methods for any additional analyses (eg, subgroup and adjusted analyses)	2023. Dow related to 7-8 7			
29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55	Statistics: analysis population and missing data	<u>#20c</u>	Definition of analysis population relating to protocol non- adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	Downloaded from http:/ nent Superieur (ABES) t o text and data mining 7-8 7-8			
	Methods: Monitoring						
	Data monitoring: formal committee	<u>#21a</u>	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	ttp://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de l :S) . ining, Al training, and similar technologies. ග ක			
	Data monitoring: interim analysis	<u>#21b</u>	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	, na Bibliograp			
56 57 58 59 60	Harms For	<u>#22</u> peer revie	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	8 8			

Page 2	1 of 22	BMJ Open			
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15			and other unintended effects of trial interventions or trial conduct		
	Auditing	<u>#23</u>	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	na	
	Ethics and dissemination			Protec	
	Research ethics approval	<u>#24</u>	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	9 9	
16 17 18 19 20 21 22 23 24	Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	Enseignement Superieur (AB) Protected by copyright, including for uses related to text and data m 9 a a 2 8	
25 26 27 28 29 30 31 32 33 34	Consent or assent	<u>#26a</u>	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	nseignement Su na na	
	Consent or assent: ancillary studies	<u>#26b</u>	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	uperieur (ABES (t and data min 8	
35 36 37 38 39 40 41	Confidentiality	<u>#27</u>	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	ES) . nining, Al training, and similar technologies ⊗	
42 43 44 45	Declaration of interests	<u>#28</u>	Financial and other competing interests for principal investigators for the overall trial and each study site	9 9	
46 47 48 49 50 51 52 53 54 55 56	Data access	<u>#29</u>	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	chnologies. 8-9	
	Ancillary and post trial care	<u>#30</u>	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	na	
50 57 58 59 60	Dissemination policy: trial results		Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the w only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	9	

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES).

her relevant groups (eg, via publication, sults databases, or other data sharing ), including any publication restrictions gibility guidelines and any intended use of 9 rriters or granting public access to the full 9 cipant-level dataset, and statistical code t form and other related documentation ipants and authorised surrogates ction, laboratory evaluation, and storage of 10 cimens for genetic or molecular analysis in al and for future use in ancillary studies, if t is distributed under the terms of the Creative checklist was completed on 12. August 2022 using he EQUATOR Network in collaboration with	er relevant groups (eg, via publication, sults databases, or other data sharing , including any publication restrictions	
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).		9
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).	0 01	
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).		v copyright, 8
data mining,	imens for genetic or molecular analysis in	including for uses
	checklist was completed on 12. August 2022 using	ur (ABES) . data mining,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

**BMJ** Open

Page 22 of 22

liographique de l

**BMJ** Open

# **BMJ Open**

### The CONFIDENT-trial protocol: a pragmatic template for clinical implementation of artificial intelligence assistance in pathology

Journal:	BMJ Open
Manuscript ID	bmjopen-2022-067437.R2
Article Type:	Protocol
Date Submitted by the Author:	28-Apr-2023
Complete List of Authors:	Flach, Rachel; University Medical Centre Utrecht, Department of Pathology Stathonikos, Nikolas; University Medical Centre Utrecht, Department of Pathology Nguyen, Tri; University Medical Centre Utrecht, Department of Pathology Ter Hoeve, Natalie; University Medical Centre Utrecht, Department of Pathology van Diest, P.J.; University Medical Centre Utrecht, Department of Pathology van Dooijeweert, Carmen; University Medical Centre Utrecht, Department of Pathology
<b>Primary Subject Heading</b> :	Pathology
Secondary Subject Heading:	Urology, Diagnostics, Oncology
Keywords:	Prostate disease < UROLOGY, Breast tumours < ONCOLOGY, PATHOLOGY

### SCHOLARONE<sup>™</sup> Manuscripts

### The CONFIDENT-trial protocol: a

# pragmatic template for clinical

## implementation of artificial intelligence assistance in pathology

- Authors
- R.N. Flach (r.n.flach-2@umcutrecht.nl), N. Stathonikos (nstatho2@umcutrecht.nl), T.Q. Nguyen
- (t.q.nguyen@umcutrecht.nl), N.D. ter Hoeve (N.terhoeve@umcutrecht.nl), P.J. van Diest
- (P.j.vandiest@umcutrecht.nl), C. van Dooijeweert (c.vandooijeweert-3@umcutrecht.nl)

#### Affiliation

Department of Pathology, University Medical Centre Utrecht, Utrecht, The Netherlands

#### Trial registration

- Not applicable

#### **Protocol version**

Version 1.

#### Funding

No funding was received for this research at the moment of writing this paper. In the meanwhile, the Hanarth Foundation has provided funding to support this study. 

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

#### **Roles and responsibilities**

- As no funding was received, and as this is a single centre clinical trial, full authority over all study activities (design; collection, management, analysis and interpretation of data; writing of the report;
- and the decision to submit the report for publication) lies with the researchers in the UMC Utrecht.

#### **Corresponding author**

- Carmen van Dooijeweert
- Internal mail no H04.312
- T: +31 88 75 565 65

Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

**BMJ** Open

1 <u>c.vandooijeweert3@umcutrecht.nl</u>

to occurrence on the occurrence of the occurrenc

### Abstract

#### Introduction

Artificial Intelligence (AI) has been on the rise in the field of pathology. Despite promising results in retrospective studies, and several CE-IVD certified algorithms on the market, prospective clinical implementation studies of AI have yet to be performed, to the best of our knowledge. In this trial, we will explore the benefits of an AI-assisted pathology workflow, while maintaining diagnostic safety standards.

#### 

#### Methods and analysis

This is a SPIRIT-AI compliant single-centre, controlled clinical trial, in a fully digital academic pathology laboratory. We will prospectively include prostate cancer patients who undergo prostate needle biopsies (CONFIDENT-P) and breast cancer patients who undergo a sentinel lymph node (SN) procedure (CONFIDENT-B) in the UMC Utrecht. For both the CONFIDENT-B and CONFIDENT-P trials, the specific pathology specimens will be pseudo-randomized to be assessed by a pathologist with- or without AI-assistance in a pragmatic (bi-)weekly sequential design. In the intervention group, pathologists will assess whole slide images (WSI) of the standard haematoxylin-eosin (HE)-stained sections assisted by the output of the algorithm. In the control group, pathologists will assess HE WSI according to the current clinical workflow. If no tumour cells are identified or when the pathologist is in doubt, immunohistochemistry (IHC) staining will be performed. At least 80 patients in the CONFIDENT-P and 180 patients in the CONFIDENT-B trial will need to be enrolled to detect superiority, allocated as 1:1. Primary endpoint for both trials is the number of saved resources of IHC staining procedures for detecting tumour cells, since this will clarify tangible cost savings that will support the business case for AI. 

#### Ethics and dissemination

The ethics committee (MREC NedMec) waived the need of official ethical approval, since participants are not subjected to procedures nor are they required to follow rules. Results of both trials (CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed journals.

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

#### Strengths and limitations

- This is the first clinical trial to examine the added value of artificial intelligence in the daily pathology workflow.
  - By maintaining the current diagnostic safety standards patients are not at risk of an inferior diagnosis during the trial.
- This is a pragmatic template for prospective AI-trials for object-identifying algorithms in pathology.
  - A limitation is that this is a single-centre study, which may hamper generalizability. -
- Due to the existing clinical workflow, randomization of patients and (double-)blinding of the participating pathologists and researchers is not possible.

### 1 Keywords:

1 2 3

4

5 6

7 8 9

10 11

- 2 Artificial intelligence, prospective trial, prostate cancer, breast cancer, digital pathology
- 3 Word count: 2864 words

### 4 Introduction

### 5 Background and rationale

12 6 Since the introduction of digital pathology, the number of studies on artificial intelligence (AI) within 13 14 7 the field of pathology has increased exponentially.[1,2] Algorithms have been created for tumour 15 8 detection, tumour grading, tumour subtyping, evaluating biomarkers and more.[1,3] Due to 16 9 demographic trends, the needs for healthcare are increasing globally which combined with a lack of 17 10 specialists, increases the current workload .[2,4] Therefore, AI has great potential to alleviate 18 pathologists' workload[2] and improve diagnostics by improving accuracy, reproducibility and 19 11 20 12 speed.[2] In fact, several algorithms have shown to be comparable, or even superior to pathologists 21 13 (under time-constraint).[2,5–10] 22

Artificial and human intelligence are not mutually exclusive, they complement each other, a concept which is known as "augmented intelligence", where Al can enhance, rether then replace human

- which is known as "augmented intelligence", where AI can enhance, rather than replace human
   intelligence.[11] In the (very) early AI-adoption phase, and presumably also in later phases,
- intelligence.[11] In the (very) early Al-adoption phase, and presumably also in later phases,
   pathologist supervision remains of key importance. This is particularly relevant as, despite the
- pathologist supervision remains of key importance. This is particularly relevant as, despite the
   promising results of retrospective studies and the availability of CE-IVD approved algorithms,
- <sup>29</sup> 19 prospective validation and clinical implementation of AI is currently lacking. For example, six years
- after the successful CAMELYON-16 Grand Challenge[6], the top algorithms have yet to be
- implemented in daily clinical practice, showing that the time between development of an AI model
- and clinical implementation is considerable. Likewise, numerous promising prostate cancer grading
- algorithms have been developed, yet implementation studies have yet to be performed,[12–14]
- whereas nine AI pathology devices received CE-IVD approval in 2021.[15]
   whereas nine AI pathology devices received CE-IVD approval in 2021.[15]

## <sup>37</sup><sub>38</sub> 25 Trial rationale

As a pathology laboratory with a fully digital workflow for over seven years, we are eager to explore the full potential of working digitally by adding the benefit of AI in daily pathology practice. We decided to start with the object localisation task of tumour-detection, where an objective reference standard is in place in the routine clinical workflow (i.e. pathologist-supervision and/or

44 30 immunohistochemistry (IHC) staining).

We developed the CONFIDENT-trial template, in which AI tumour detection algorithms can
be safely implemented in prospective clinical trials, while ensuring that patients are not at risk of
receiving an inferior diagnosis, since IHC is always performed when no tumour cells are visible, but
also when pathologists need more confirmation about the diagnosis.

### 50 51 35 CONFIDENT-B and CONFIDENT-P

52 36 Our trials aim to prospectively investigate the added value of an AI-assisted pathology workflow in 53 37 the identification of prostate cancer (PCa) in prostate needle biopsies (CONFIDENT-P) and the 54 55 38 identification of sentinel lymph node (SN) metastases in breast cancer (BCa) patients (CONFIDENT-B). 56 39 As both prostate cancer and breast cancer are the most common (non-skin) malignancies in men and 57 40 women, respectively, implementation of AI-assistance may have a great impact on diagnostic 58 41 processes.[16] However, it is important to emphasize that this trial serves as a template for other 59 42 pragmatic AI-intervention trials for object-localisation tasks as well. 60

- We obtained CE-IVD-approved algorithms for detection and grading of prostate cancer in prostate needle biopsies and an algorithm for detecting lymph node metastases in BCa patients. In both cases the task of the pathologist is both labour-intensive and expensive, due to the performed IHC stains in case no tumour cells are morphologically observed. However, IHC is expensive and these costs sometimes even exceed reimbursement for the entire specimen (e.g. in case of multiple blocks of multiple SNs). This raises the question whether AI may be of added value to morphologically detect cancer cells without the need for IHC-use. Thereby the number of performed IHC stains may be reduced, which may lead to tangible costs savings that will help to build the business case for AI, while potentially decreasing the workload of pathologists as well.[2]
- <sup>15</sup> 10 Study objective

The primary objective is to explore whether an AI-assisted workflow reduces the number of spent resources on IHC, while maintaining diagnostic safety standards in both PCa patients who underwent prostate needle biopsies (CONFIDENT-P), and BCa patients who underwent an SN procedure. Secondary objectives are to investigate whether time management improves in an Al-assisted workflow and to analyse how many IHC staining may have been safely omitted after AI-implementation. 

### 17 Methods and analysis

### 18 Trial design

The study protocol is structured following the SPIRIT-AI (Standard Protocol Items: recommendations for Interventional Trials - Artificial Intelligence) statement 2020[17]. This study is a single-centre, parallel-group controlled trial, assessing superiority. The allocation ratio is 1:1. Eligible patients will be assigned to arm 1 (control group) or arm 2 (AI-assisted workflow), based on a bi-weekly time-schedule. Eligibility criteria are summarised in Figure 1. The CONFIDENT trials will be carried out in 2022-2023. 

### <sup>37</sup> 25 Study setting

The trial will take place in the daily practice of a single academic hospital (University Medical Centre Utrecht, the Netherlands), with a fully digital pathology enabled clinical set-up, where all slides are digitised using ultrafast whole slide image (WSI) using Hamamatsu S360 scanners and reviewed using the Sectra pathology Picture Archiving and Communication System (PACS). Although the UMC Utrecht is an academic hospital, primary routine pathology diagnostics is performed for non-academic hospitals as well (i.e., the Alexander Monro Breast Cancer Hospital, Bilthoven, the Netherlands). 

#### 49 33 Study population

For PCa, WSI of all males who undergo a prostate needle biopsy in the UMC Utrecht will be included.
 For BCa, WSI of all females or males with BCa as primary malignancy (i.e. invasive breast cancer) who
 undergo an SN procedure in the Alexander Monro Breast Cancer Hospital or the UMC Utrecht will be
 included. Patients will be excluded, if they were redirected to the UMC Utrecht for a second opinion.

#### 56 38 Assessment of specimen

During the study period, all WSI will be assessed by the same group of pathologists; i.e. two expert
urological pathologists for the PCa biopsies, and three expert breast pathologists for the lymph node
assessment from BCa patients.

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

For both the CONFIDENT-B and CONFIDENT-P trial, the specific pathology specimens will be assigned to be assessed by a pathologist with or without AI-assistance in a pragmatic (bi-)weekly sequential design. This is considered feasible as the change in the case mix and time trends are unlikely to occur within the inclusion-period of about six to nine months. Furthermore, both specialized breast- and urological pathologists within the UMC Utrecht work according to weekly schedules. Therefore, using AI every other week or every other two weeks, as opposed to switching by day, ensures that all pathologists are equally distributed between groups. Lastly, it would be impractical to switch from AI-assistance in the intervention group to no AI-assistance in the control group on a case-to-case basis. For obvious reasons, allocation concealment and blinding of pathologists and researchers is not applicable.

### 12 Control and intervention

All eligible specimens will be assigned to either the control group or the intervention group. In the control group, pathologists will assess H&E stained WSI of patients digitally, according to the current clinical workflow. For PCa biopsies, IHC is routinely performed on all cases. For BCa lymph nodes, if no metastases or tumour are present, IHC staining will be performed. Additional IHC staining will also be performed by additional request of the pathologist in case of doubt. 

In the intervention group, pathologists will assess the H&E-specimens digitally with the outcome of the algorithm provided in their first assessment of the specimen. For PCa, they will use the CE-IVD certified Paige Prostate Suite algorithms for tumor detection and tumor volume percentage calculations, which reaches sensitivity and specificity of 99% and 93% respectively and which are based on a weakly-supervised deep learning algorithm as described by Campanella et al.[18,19]For BCa, pathologists will use the CE-IVD certified Metastasis Detection App by Visiopharm, a deep-learning algorithm for lymph node metastases of BCa and colon carcinoma with a combined sensitivity and specificity of 98,7 and 99.6% respectively[20]. These algorithms will be integrated within the Sectra PACS where the output of the algorithms will be graphically displayed. Al analysis of the WSI will be performed right after scanning to avoid delays in the clinical workflow. If the AI-assisted pathologist does not detect metastases or tumours on the H&E slide, routine additional IHC staining will be performed by P503S/p63/CK HMW for PCa and CAM5.2 for BCa, to ensure no metastases or tumours are missed. Pathologists can also request an additional IHC if they feel they need this to make an adequate diagnosis (Figure 2). 

# 4344 32 Outcome measures

Primary outcome for the CONFIDENT-P trial is the added value of AI-assistance in the detection of PCa and in the detection of tumour volume in prostate needle biopsies in daily pathology practice. The primary outcome for the CONFIDENT-B trial is the added value of AI-assistance in the detection of BCa SN-metastases. The outcome measures for both trials will be the number of spent resources, i.e. the number of IHC-stains performed in both groups. 

Secondary outcome measures will be sensitivity and specificity of the AI-assisted pathologist, time spent on WSI analysis, the number of IHC stains that may have been omitted after AI-implementation and a pathologists' evaluation by a questionnaire on the Al-assisted work process. Sensitivity and specificity analyses of the algorithm itself have already been well documented, and is therefore outside the scope of the paper, as we focus on the combination of pathologist and AI to explore cost savings. 

#### Input data

Input data for the algorithm will be WSI of haematoxylin and eosin (H&E) stained slides scanned at

40x of either prostate needle biopsies, and WSI of H&E stained slides of BCa lymph nodes. As per

routine in our daily clinical practice, WSI will be quality controlled after scanning for colour, focus

quality and completeness of the scan. When necessary, the specimens are rescanned.

# 

#### Sample size

#### **CONFIDENT-P**

We performed power calculations using a two-sample proportion superiority test, using expected percentages of IHC staining in both study arms. We assume that the pathologists in the control arm can detect 50% of the tumours without using IHC. We expect AI-assisted pathologists to detect 80% of the tumours, without using IHC. These percentages were conservatively derived from the validity study by Raciti (74% for pathologists without AI and 90% for pathologists with AI respectively)[21], by expert pathologist opinion, and taking into account that pathologists under time constraint of daily practice do not detect tumours as well as pathologists without time constraint during retrospective studies[22]. We assume that this effect will be larger for the biopsies assessed without AI than with AI, as AI is assumed to make tumour detection easier.

A sample size of 60 (30 per arm) would give a power of approximately 80%, using a one-sided 5% significance level. However, uncertainties remain regarding the sample size parameters. We therefore inflated our sample size to 80 (40 per arm), in order to ensure study power and allowing us to detect smaller effect sizes.

For detection of tumour volume percentage, we performed a power calculation based on the assumption that AI should be able to replace at least 20% of the IHC stains, in order to be cost-effective. IHC is currently used in 100% of all prostate needle biopsies. Using a power of 80% and a one-sided significance level of 5%, this leads to 27 patients per arm. 

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

#### **CONFIDENT-B**

Sample size calculations for the CONFIDENT-B trial are based on the assumption that the Al-

algorithm can detect all metastases for which currently IHC is used, which are mainly micro-

metastases and isolated tumour cells (ITC). Approximately 15% of the SN-specimens in the UMC Utrecht contain a micrometastasis or ITC.

A sample size of 166 patients (83 per arm) with a one-sided 5% significance level therefore results in a power of 80%. Again, as there are uncertainties on the assumptions on what amount of the metastases will be detect by AI, we decided to be conservative and include 180 patients (90 per arm).

Overall, we are only interested in one-sided outcomes, as it is not possible that more IHC will be performed in the AI-assisted arm. IHC is performed to detect metastases, when they are

macroscopically undetectable, rather than to confirm them when they are macroscopically visible. As

Al would show only more metastases than the pathologist could macroscopically detect, this means that only a reduction of IHC is possible. 

Sample sizes were calculated uwing the power.prop.test command in R version 4.2.2[23].

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

#### Statistical methods For baseline comparisons between both arms, the appropriate measures (parametric or non-parametric) for categorical (Chi-squared test/Fisher's exact) and continuous variables (T-test, Mann-Whitney U test) will be used. For the analysis of the primary outcome measure, we will compare the proportion of IHC-use in both arms, and calculate adjusted relative risks, using a log-binomial model.[24-26] Missing data for baseline characteristics and for the primary outcome are not to be expected, as they are obligatory items in the structured pathology reports. We will determine sensitivity and specificity of the conclusion of the AI-assisted pathologists without the use of IHC. Subsequently, we will focus on the cases with metastases that the AI-assisted pathologist misses, categorize them (i.e. macro-metastases, micro-metastases, ITC) and determine their clinical relevance (i.e. clinical consequences if these metastases are being missed). Data-analysis will be performed in R Statistical Software[23], with a significance level set at p<0.05. Data collection and management All data (baseline and primary outcome measurements) will be retrieved from the structured pathology reports and will be managed and stored in Castor EDC[27]. For the secondary outcome measure of time spent by the pathologist on a slide, data will be collected on an interval basis for practical reasons, as timing every assessment for months (by stopwatch) was not deemed feasible. For the secondary outcome measure of AI-assisted work process for pathologists, a questionnaire will be distributed to the participating pathologists. The final secondary assessment measurement, the number of IHC stains that may have been omitted after AI-implementation, will be determined by the researchers based on the data from the structured pathology reports (combination of IHC and Al-assisted diagnoses of the pathologist). Ethical approval Research within these trials is not subject to the (Dutch) Medical Research Involving Human Subjects Act (WMO), as participants are not subjected to procedures and as they are not required to follow rules. Therefore, the ethics committee (MREC NedMec) waived the need of ethical approval and informed consent. Risk of harm Patients are not at risk of any harm for an inferior diagnosis (i.e. missed tumour cells), as in both arms, IHC-staining will be performed when no tumour cells are visible, according to current clinical standards. As a rule in augmented intelligence, all cases will be evaluated by a pathologist, which further minimized the risk of a false diagnosis based on the AI algorithm. Taking all of the above into account, a data monitoring committee (DMC) is not required and adverse events are not to be expected. In theory, the algorithm could be more of a disturbance than a help to pathologists (for example, when it frequently reports false positive or negative results, which have to be corrected by the pathologist). However, the algorithms used are IVDR-approved, and thus have undergone extensive review for their intended purpose. Nonetheless, the experience and ease of use of pathologists working with the algorithm will be one of the secondary outcome measures. Informed consent and data access For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Informed consent was waived by the local quality coordinator (QC) and data protection officer (DPO)
for the following reasons. First, in both arms patients receive standard care, while maintaining
diagnostic safety standards (pathologists' supervision, IHC in all negative cases). Second, patients are
not subjected to any procedures. Third, all patient data will be anonymized to the researchers by the
pathologist who assessed the slide.
The collected (anonymous) research data will be stored in Castor EDC to ensure data
security. Data will be kept for a period of 15 years. Data access in Castor will be restricted to two

researchers (RF, CvD). Pathologists have access to the electronic patient files for the purpose of
patient care. The researchers are not permitted access to these files. At no point will the data (both
in Castor EDC and patient files) be accessed by the companies providing the algorithms (i.e.

- 15 11 Visiopharm and Paige).16
  - 12 Patient and Public Involvement
  - 13 None

### <sup>21</sup> 14 **Discussion**

The promising retrospective results of AI-assisted pathology have not yet resulted in prospective clinical implementation studies. This may be due to a lack of digital transition in the majority of pathology laboratories, but it may also be partly due to the lack of a good implementation model. Fortunately, however, new guidelines for AI-trials have recently been proposed by the SPIRIT-AI and CONSORT-AI-steering groups, as well as roadmaps to routine use of AI in clinical practice.[28–30] Yet, to date, no pathology AI-trials have been published in PubMed or Web of Science or, to the best of our knowledge, otherwise made public.

As a pathology laboratory with a fully digital workflow, we developed a clinical trial template for tumour detection models, as a first step to implement AI in daily pathology practice. We will start with an object localisation task (i.e. tumour cells) as a reference standard is in place in the routine clinical workflow. For classification tasks like tumour grading, a clinical trial design is more challenging, as no reference is in place in daily pathology practice and inter-laboratory and inter-pathologist variation is notorious.[31–35] Nevertheless, in future trials, implementing Al-assistance in the grading process might also reduce this variation. For now, results of the CONFIDENT-trials will provide the first assessment of the potential added value of AI in daily pathology practice. This evaluation will substantially contribute to a potential paradigm shift in tumour detection in pathology. The pragmatic template of the CONFIDENT trials may serve as example for other prospective AI implementation trials in diagnostic pathology. 

### 46 33 Declaration of interests

- 47
  48
  48
  49
  49
  44
  44
  45
  46
  47
  47
  48
  47
  48
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  49
  <
- 49 35 All other authors do not report conflict of interest.

#### 51 36 Funding

37 No funding was obtained at the moment of writing this paper. In the meanwhile, the Hanarth
 38 Foundation has provided funding to support this study.

# 56 39 Dissemination policy

40 Results of both trials (CONFIDENT-B and CONFIDENT-P) will be published in scientific peer-reviewed

- 59 41 journals. Authorship will be acknowledged to all those that substantially contributed in the
- 60 42 CONFIDENT-trials. Data will be available upon reasonable request.

Enseignement Superieur (ABES) Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies

#### 

### 1 Author statement

PJvD conceived of the study. RNF and CvD initiated the study design and NS, TQN and
NDtH helped with implementation. RNF and CvD provided statistical expertise in clinical
trial design and RNF and CvD are conducting the primary statistical analysis. All authors
contributed to refinement of the study protocol and approved the final manuscript.

### 6 Figure legends

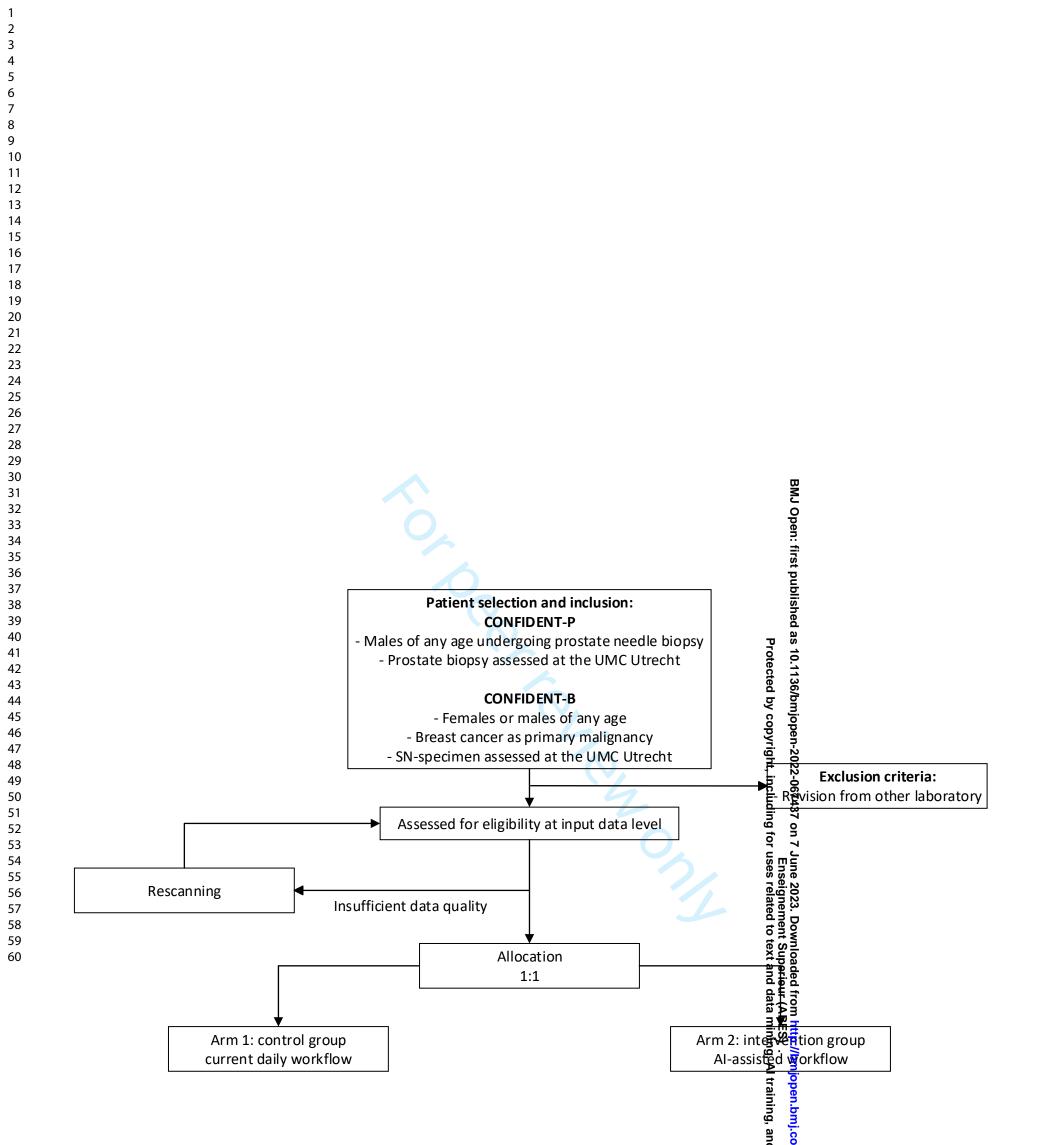
7 Figure 1. Flowchart with patient selection; SN = sentinel node; AI = artificial intelligence.

8 Figure 2. Study flow chart. H&E = hematoxylin & eosin; IHC = immunohistochemistry

2			
3	1	Ref	erences
4 5	2		
6 7 8	3 4	1	Jiang Y, Yang M, Wang S, <i>et al</i> . Emerging role of deep learning-based artificial intelligence in tumor pathology. <i>Cancer Commun</i> 2020; <b>40</b> :154–66. doi:https://doi.org/10.1002/cac2.12012
9 10	5 6	2	van der Laak J, Litjens G, Ciompi F. Deep learning in histopathology: the path to the clinic. <i>Nat</i> <i>Med</i> 2021; <b>27</b> :775–84. doi:10.1038/s41591-021-01343-4
11 12	0		Weu 2021, <b>21</b> .775–64. 001.10.1056/541591-021-01545-4
13	7	3	Steiner DF, Chen PHC, Mermel CH. Closing the translation gap: AI applications in digital
14 15 16	8 9		pathology. <i>Biochim Biophys Acta Rev Cancer</i> 2021; <b>1875</b> :188452. doi:10.1016/j.bbcan.2020.188452
17	10	4	Robboy SJ, Weintraub S, Horvath AE, et al. Pathologist workforce in the United States: I.
18 19	11		Development of a predictive model to examine factors influencing supply. Arch Pathol Lab
20	12		Med 2013; <b>137</b> :1723–32. doi:10.5858/arpa.2013-0200-OA
21 22	13	5	Bulten W, Pinckaers H, van Boven H, et al. Automated deep-learning system for Gleason
23	14		grading of prostate cancer using biopsies: a diagnostic study. <i>Lancet Oncol</i> 2020; <b>21</b> :233–41.
24	15		doi:10.1016/S1470-2045(19)30739-9
25 26	16	6	Ehteshami Bejnordi B, Veta M, Johannes van Diest P, <i>et al.</i> Diagnostic Assessment of Deep
27	17		Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer.
28 29	18		JAMA 2017; <b>318</b> :2199–210. doi:10.1001/jama.2017.14585
30	19	7	Steiner DF, Nagpal K, Sayres R, et al. Evaluation of the Use of Combined Artificial Intelligence
31	20		and Pathologist Assessment to Review and Grade Prostate Biopsies. JAMA Netw Open
32 33	21		2020; <b>3</b> :e2023267. doi:10.1001/jamanetworkopen.2020.23267
34	22	8	Ström P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of
35 36	23		prostate cancer in biopsies: a population-based, diagnostic study. Lancet Oncol 2020; <b>21</b> :222–
37	24		32. doi:10.1016/S1470-2045(19)30738-7
38 39	25	9	Raciti P, Sue J, Ceballos R, et al. Novel artificial intelligence system increases the detection of
40	26		prostate cancer in whole slide images of core needle biopsies. Modern Pathology
41	27		2020; <b>33</b> :2058–66. doi:10.1038/s41379-020-0551-y
42 43	28	10	Hekler A, Utikal JS, Enk AH, et al. Superior skin cancer classification by the combination of
44	29		human and artificial intelligence. <i>Eur J Cancer</i> 2019; <b>120</b> :114–21.
45	30		doi:10.1016/j.ejca.2019.07.019
46 47	71	11	
47 48	31 32	11	Harrison JH, Gilbertson JR, Hanna MG, <i>et al</i> . Introduction to Artificial Intelligence and Machine Learning for Pathology. <i>Arch Pathol Lab Med</i> 2021; <b>145</b> :1228–54. doi:10.5858/arpa.2020-
49	32 33		Learning for Pathology. Arch Pathol Lab Med 2021; <b>145</b> :1228–54. doi:10.5858/arpa.2020- 0541-cp
50	22		0341-cp
51 52	34	12	Ström P, Kartasalo K, Olsson H, et al. Artificial intelligence for diagnosis and grading of
53	35		prostate cancer in biopsies: a population-based, diagnostic study. Lancet Oncol 2020; <b>21</b> :222–
54	36		32. doi:10.1016/S1470-2045(19)30738-7
55 56	37	13	Steiner DF, Nagpal K, Sayres R, et al. Evaluation of the Use of Combined Artificial Intelligence
57	38		and Pathologist Assessment to Review and Grade Prostate Biopsies. JAMA Netw Open
58	39		2020; <b>3</b> :1–14. doi:10.1001/jamanetworkopen.2020.23267
59 60			
50			

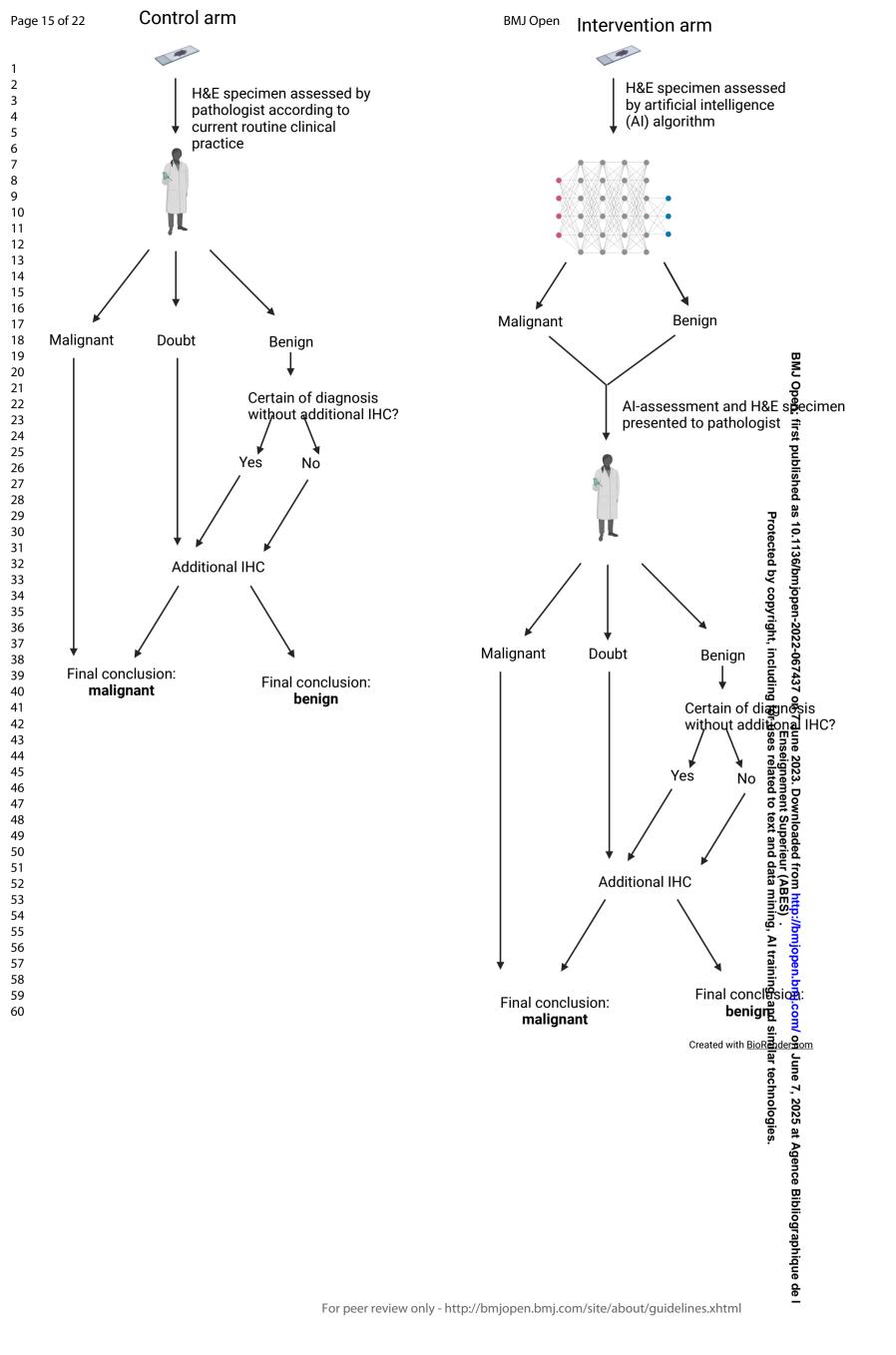
2			
3	1	14	Bulten W, Balkenhol M, Belinga JJA, et al. Artificial intelligence assistance significantly
4	2		improves Gleason grading of prostate biopsies by pathologists. Modern Pathology
5 6	3		2021; <b>34</b> :660–71. doi:10.1038/s41379-020-0640-y
7	4	15	Muehlematter UJ, Daniore P, Vokinger KN. Approval of artificial intelligence and machine
8	5	15	learning-based medical devices in the USA and Europe (2015–20): a comparative analysis.
9 10	6		Lancet Digit Health 2021; <b>3</b> :e195–203. doi:10.1016/S2589-7500(20)30292-2
11			
12	7	16	Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of
13	8		Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin
14 15	9		2021; <b>71</b> :209–49. doi:10.3322/caac.21660
16	10	17	Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions
17	11		involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020; <b>2</b> :e549–60.
18	12		doi:10.1016/S2589-7500(20)30219-3
19 20	10	10	Companylla C. Harrow MC. Concellant L. et al. Clinical and a computational methology using
21	13 14	18	Campanella G, Hanna MG, Geneslaw L, <i>et al</i> . Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. <i>Nat Med</i> 2019; <b>25</b> :1301–9.
22	14 15		doi:10.1038/s41591-019-0508-1
23 24	15		001.10.1038/341391-019-0308-1
25	16	19	da Silva LM, Pereira EM, Salles PGO, et al. Independent real-world application of a clinical-
26	17		grade automated prostate cancer detection system. <i>Journal of Pathology</i> 2021; <b>254</b> :147–58.
27	18		doi:10.1002/path.5662
28 29	19	20	Visiopharm. https://visiopharm.com/app-center/app/metastasis-detection-ai/.
30			
31	20	21	Raciti P, Sue J, Ceballos R, <i>et al.</i> Novel artificial intelligence system increases the detection of
32 33	21 22		prostate cancer in whole slide images of core needle biopsies. <i>Modern Pathology</i>
34	22		2020; <b>33</b> :2058–66. doi:10.1038/s41379-020-0551-y
35	23	22	Bejnordi BE, Veta M, Van Diest PJ, et al. Diagnostic assessment of deep learning algorithms for
36 37	24		detection of lymph node metastases in women with breast cancer. JAMA - Journal of the
38	25		American Medical Association 2017; <b>318</b> :2199–210. doi:10.1001/jama.2017.14585
39	26	23	R Core Team. R: A Language and Environment for Statistical Computing. 2022.https://www.r-
40 41	27		project.org/
41	20	24	Manuta LA Mulo C Mulo Mulating the relative risk is each studies and aliginal trials
43	28	24	McNutt LA, Wu C, Xue X, <i>et al.</i> Estimating the relative risk in cohort studies and clinical trials
44 45	29		of common outcomes. <i>Am J Epidemiol</i> 2003; <b>157</b> :940–3. doi:10.1093/aje/kwg074
45 46	30	25	Lumley T, Kronmal R, Ma S. Relative risk regression in medical research: models, contrasts,
47	31		estimators, and algorithms UW Biostatistics Working Paper Series 293 [http://www bepress
48	32		com/uwbiostat/paper293] 2006.
49 50	33	26	Zou G. A Modified Poisson Regression Approach to Prospective Studies with Binary Data. Am J
50	34	20	<i>Epidemiol</i> 2004; <b>159</b> :702–6. doi:10.1093/aje/kwh090
52			
53	35	27	Castor Electronic Data Capture. 2019.
54 55	36	28	Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for
56	37		interventions involving artificial intelligence: the CONSORT-AI extension. Lancet Digit Health
57	38		2020; <b>2</b> :e537–48. doi:10.1016/S2589-7500(20)30218-1
58 59			
60			

1 2			
3 4 5 6	1 2 3	29	Cruz Rivera S, Liu X, Chan AW, <i>et al.</i> Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Lancet Digit Health. 2020; <b>2</b> :e549–60. doi:10.1016/S2589-7500(20)30219-3
7 8 9 10	4 5 6	30	Colling R, Pitman H, Oien K, <i>et al.</i> Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. <i>Journal of Pathology J Pathol</i> 2019; <b>249</b> :143–50. doi:10.1002/path.5310
11 12 13 14	7 8 9	31	Kuijpers CCHJ, Sluijter CE, von der Thüsen JH, <i>et al.</i> Interlaboratory variability in the histologic grading of colorectal adenocarcinomas in a nationwide cohort. <i>American Journal of Surgical Pathology</i> 2016; <b>40</b> :1100–8. doi:10.1111/his.12923
15 16 17 18	10 11 12	32	Kuijpers CCHJ, Sluijter CE, von der Thüsen JH, <i>et al.</i> Interlaboratory variability in the grading of dysplasia in a nationwide cohort of colorectal adenomas. <i>Histopathology</i> 2016; <b>69</b> :187–97. doi:10.1111/his.12923
19 20 21 22 23	13 14 15	33	Flach RN, Willemse P-PM, Suelmann BBM, <i>et al.</i> Significant Inter- and Intralaboratory Variation in Gleason Grading of Prostate Cancer: A Nationwide Study of 35,258 Patients in The Netherlands. <i>Cancers (Basel)</i> 2021; <b>13</b> . doi:10.3390/cancers13215378
24 25 26 27 28	16 17 18 19	34	van Dooijeweert C, van Diest PJ, Willems SM, <i>et al.</i> Significant inter- and intra-laboratory variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901 patients in the Netherlands. <i>Breast Cancer Res Treat</i> 2019; <b>174</b> :479–88. doi:10.1007/s10549-018-05082-y
29 30 31 32 33	20 21 22	35	van Dooijeweert C, van Diest PJ, Willems SM, <i>et al.</i> Significant inter- and intra-laboratory variation in grading of invasive breast cancer: A nationwide study of 33,043 patients in the Netherlands. <i>Int J Cancer</i> 2020; <b>146</b> :769–80. doi:10.1002/ijc.32330
33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 7 58 59 60	23		
			13



om/ on June 7, 2025 at Agence Bibliographique de l d similar technologies.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml



# Reporting checklist for protocol of a clinical trial.

Based on the SPIRIT guidelines.

## Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the SPIRITreporting guidelines, and cite them as:

Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin J, Dickersin K, Hróbjartsson A, Schulz KF, Parulekar WR, Krleža-Jerić K, Laupacis A, Moher D. SPIRIT 2013 Explanation and Elaboration: Guidance for protocols of clinical trials. BMJ. 2013;346:e7586

30								
31 32	Reporting Item							
<ul> <li>33</li> <li>34</li> <li>35</li> <li>36</li> <li>37</li> <li>38</li> <li>39</li> <li>40</li> <li>41</li> <li>42</li> <li>43</li> <li>44</li> <li>45</li> <li>46</li> <li>47</li> </ul>	Administrative information			Number no ora ana mining, Ai training, 1				
	Title	<u>#1</u>	Descriptive title identifying the study design, population, interventions, and, if applicable, trial acronym	Al training				
	Trial registration	<u>#2a</u>	Trial identifier and registry name. If not yet registered, name of intended registry	na ang				
	Trial registration: data set	<u>#2b</u>	All items from the World Health Organization Trial Registration Data Set	na na 1				
48 49 50	Protocol version	<u>#3</u>	Date and version identifier	gles. 1 s.				
50 51 52 53 54 55 56 57 58 59	Funding	<u>#4</u>	Sources and types of financial, material, and other support	1				
	Roles and responsibilities: contributorship	<u>#5a</u>	Names, affiliations, and roles of protocol contributors	1				
60	For	peer revi	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml					

### BMJ Open

1 2 3 4 5 6	Roles and responsibilities: sponsor contact information	<u>#5b</u>	Name and contact information for the trial sponsor	na
7 8 9 10 11 12 13 14 15 16	Roles and responsibilities: sponsor and funder	<u>#5c</u>	Role of study sponsor and funders, if any, in study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication, including whether they will have ultimate authority over any of these activities	na
17 18 19 20 21 22 23 24 25	Roles and responsibilities: committees	<u>#5d</u>	Composition, roles, and responsibilities of the coordinating centre, steering committee, endpoint adjudication committee, data management team, and other individuals or groups overseeing the trial, if applicable (see Item 21a for data monitoring committee)	na
26 27	Introduction			
28 29 30 31 32 33	Background and rationale	<u>#6a</u>	Description of research question and justification for undertaking the trial, including summary of relevant studies (published and unpublished) examining benefits and harms for each intervention	3
34 35 36 37 38 39 40 41	Background and rationale: choice of comparators	<u>#6b</u>	Explanation for choice of comparators	3
	Objectives	<u>#7</u>	Specific objectives or hypotheses	4
42 43 44 45 46 47 48	Trial design	<u>#8</u>	Description of trial design including type of trial (eg, parallel group, crossover, factorial, single group), allocation ratio, and framework (eg, superiority, equivalence, non-inferiority, exploratory)	4
49 50	Methods:			
51 52	Participants,			
53 54	interventions, and outcomes			
55 56 57 58 59	Study setting	<u>#9</u>	Description of study settings (eg, community clinic, academic hospital) and list of countries where data will	4-5
60	For	peer revie	w only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de I Enseignement Superieur (ABES) . Protected by copyright, including for uses related to text and data mining, Al training, and similar technologies.

1 2			be collected. Reference to where list of study sites can be obtained	
3 4 5 6 7 8 9	Eligibility criteria	<u>#10</u>	Inclusion and exclusion criteria for participants. If applicable, eligibility criteria for study centres and individuals who will perform the interventions (eg, surgeons, psychotherapists)	5
10 11 12 13 14 15	Interventions: description	<u>#11a</u>	Interventions for each group with sufficient detail to allow replication, including how and when they will be administered	6
16 17 18 19 20 21	Interventions: modifications	<u>#11b</u>	Criteria for discontinuing or modifying allocated interventions for a given trial participant (eg, drug dose change in response to harms, participant request, or improving / worsening disease)	na
22 23 24 25 26 27	Interventions: adherance	<u>#11c</u>	Strategies to improve adherence to intervention protocols, and any procedures for monitoring adherence (eg, drug tablet return; laboratory tests)	na
28 29 30	Interventions: concomitant care	<u>#11d</u>	Relevant concomitant care and interventions that are permitted or prohibited during the trial	na
31 32 33 34 35 36 37 38 39 40 41 42	Outcomes	<u>#12</u>	Primary, secondary, and other outcomes, including the specific measurement variable (eg, systolic blood pressure), analysis metric (eg, change from baseline, final value, time to event), method of aggregation (eg, median, proportion), and time point for each outcome. Explanation of the clinical relevance of chosen efficacy and harm outcomes is strongly recommended	6, 8
43 44 45 46 47 48	Participant timeline	<u>#13</u>	Time schedule of enrolment, interventions (including any run-ins and washouts), assessments, and visits for participants. A schematic diagram is highly recommended (see Figure)	6
49 50 51 52 53 54 55	Sample size	<u>#14</u>	Estimated number of participants needed to achieve study objectives and how it was determined, including clinical and statistical assumptions supporting any sample size calculations	7
56 57 58 59 60	Recruitment	<u>#15</u> peer revie	Strategies for achieving adequate participant enrolment to reach target sample size ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	5

BMJ Open

1 2 3 4 5 6	Methods: Assignment of interventions (for controlled trials)			
7 8 9 10 11 12 13 14 15 16 17	Allocation: sequence generation	<u>#16a</u>	Method of generating the allocation sequence (eg, computer-generated random numbers), and list of any factors for stratification. To reduce predictability of a random sequence, details of any planned restriction (eg, blocking) should be provided in a separate document that is unavailable to those who enrol participants or assign interventions	5
<ol> <li>18</li> <li>19</li> <li>20</li> <li>21</li> <li>22</li> <li>23</li> <li>24</li> <li>25</li> <li>26</li> <li>27</li> <li>28</li> <li>29</li> <li>30</li> <li>31</li> <li>32</li> <li>33</li> <li>34</li> <li>35</li> <li>36</li> <li>37</li> <li>38</li> <li>39</li> <li>40</li> <li>41</li> <li>42</li> <li>43</li> <li>44</li> <li>45</li> <li>46</li> <li>47</li> </ol>	Allocation concealment mechanism	<u>#16b</u>	Mechanism of implementing the allocation sequence (eg, central telephone; sequentially numbered, opaque, sealed envelopes), describing any steps to conceal the sequence until interventions are assigned	na
	Allocation: implementation	<u>#16c</u>	Who will generate the allocation sequence, who will enrol participants, and who will assign participants to interventions	na
	Blinding (masking)	<u>#17a</u>	Who will be blinded after assignment to interventions (eg, trial participants, care providers, outcome assessors, data analysts), and how	na
	Blinding (masking): emergency unblinding	<u>#17b</u>	If blinded, circumstances under which unblinding is permissible, and procedure for revealing a participant's allocated intervention during the trial	na
	Methods: Data collection, management, and analysis			
48 49 50 51 52 53 54 55 56 57 58 59	Data collection plan	<u>#18a</u>	Plans for assessment and collection of outcome, baseline, and other trial data, including any related processes to promote data quality (eg, duplicate measurements, training of assessors) and a description of study instruments (eg, questionnaires, laboratory tests) along with their reliability and validity, if known.	8
60	For	peer revie	ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	

			BMJ Open	Page 20 of 22
1 2 3			Reference to where data collection forms can be found, if not in the protocol	BMJ Ope
4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25	Data collection plan: retention	<u>#18b</u>	Plans to promote participant retention and complete follow-up, including list of any outcome data to be collected for participants who discontinue or deviate from intervention protocols	BMJ Open: first published a na
	Data management	<u>#19</u>	Plans for data entry, coding, security, and storage, including any related processes to promote data quality (eg, double data entry; range checks for data values). Reference to where details of data management procedures can be found, if not in the protocol	as 10.1136/bmjopen-2022-067437 on 7 June Ens Protected by copyright, including for uses 99 8 <sup>-9</sup> 7 <sup>-8</sup> 7 <sup>-</sup>
	Statistics: outcomes	<u>#20a</u>	Statistical methods for analysing primary and secondary outcomes. Reference to where other details of the statistical analysis plan can be found, if not in the protocol	2-067437 on 7 June Ensi , including for uses 7-8 7
26 27 28	Statistics: additional analyses	<u>#20b</u>	Methods for any additional analyses (eg, subgroup and adjusted analyses)	2023. Dow related to 7-8 7
29 30 31 32 33 34 35	Statistics: analysis population and missing data	<u>#20c</u>	Definition of analysis population relating to protocol non- adherence (eg, as randomised analysis), and any statistical methods to handle missing data (eg, multiple imputation)	Downloaded from http:/ nent Superieur (ABES) t o text and data mining 7-8 7-8
36 37	Methods: Monitoring			'bmjop 3, Al tr
38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 55 55	Data monitoring: formal committee	<u>#21a</u>	Composition of data monitoring committee (DMC); summary of its role and reporting structure; statement of whether it is independent from the sponsor and competing interests; and reference to where further details about its charter can be found, if not in the protocol. Alternatively, an explanation of why a DMC is not needed	ttp://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de l :S) . ining, Al training, and similar technologies. ග ක
	Data monitoring: interim analysis	<u>#21b</u>	Description of any interim analyses and stopping guidelines, including who will have access to these interim results and make the final decision to terminate the trial	, na Bibliograp
56 57 58 59 60	Harms For	<u>#22</u> peer revie	Plans for collecting, assessing, reporting, and managing solicited and spontaneously reported adverse events ew only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	8 8

Page 21 of 22 BMJ Open				
1 2 3			and other unintended effects of trial interventions or trial conduct	
4 5 6 7 8 9 10 11 12	Auditing	<u>#23</u>	Frequency and procedures for auditing trial conduct, if any, and whether the process will be independent from investigators and the sponsor	na
	Ethics and dissemination			Protec
13 14 15	Research ethics approval	<u>#24</u>	Plans for seeking research ethics committee / institutional review board (REC / IRB) approval	9 9
16 17 18 19 20 21 22 23 24	Protocol amendments	#25	Plans for communicating important protocol modifications (eg, changes to eligibility criteria, outcomes, analyses) to relevant parties (eg, investigators, REC / IRBs, trial participants, trial registries, journals, regulators)	Enseignement Superieur (AB) Protected by copyright, including for uses related to text and data m 9 a a 2 8
$\begin{array}{c} 24\\ 25\\ 26\\ 27\\ 28\\ 29\\ 30\\ 31\\ 32\\ 33\\ 34\\ 35\\ 36\\ 37\\ 38\\ 39\\ 40\\ 41\\ 42\\ 43\\ 44\\ 45\\ 46\\ 47\\ 48\\ 49\\ 50\\ 51\\ 52\\ 53\\ 54\\ 55\\ 56\\ 57\\ 58\\ 59\\ 60\\ \end{array}$	Consent or assent	<u>#26a</u>	Who will obtain informed consent or assent from potential trial participants or authorised surrogates, and how (see Item 32)	nseignement Su na na
	Consent or assent: ancillary studies	<u>#26b</u>	Additional consent provisions for collection and use of participant data and biological specimens in ancillary studies, if applicable	uperieur (ABES (t and data min 8
	Confidentiality	<u>#27</u>	How personal information about potential and enrolled participants will be collected, shared, and maintained in order to protect confidentiality before, during, and after the trial	ES) . nining, Al training, and similar technologies ⊗
	Declaration of interests	<u>#28</u>	Financial and other competing interests for principal investigators for the overall trial and each study site	9 9
	Data access	<u>#29</u>	Statement of who will have access to the final trial dataset, and disclosure of contractual agreements that limit such access for investigators	chnologies. 8-9
	Ancillary and post trial care	<u>#30</u>	Provisions, if any, for ancillary and post-trial care, and for compensation to those who suffer harm from trial participation	na
	Dissemination policy: trial results		Plans for investigators and sponsor to communicate trial results to participants, healthcare professionals, the w only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	9

BMJ Open: first published as 10.1136/bmjopen-2022-067437 on 7 June 2023. Downloaded from http://bmjopen.bmj.com/ on June 7, 2025 at Agence Bibliographique de l Enseignement Superieur (ABES).

her relevant groups (eg. via publication, sults databases, or other data sharing ), including any publication restrictions gibility guidelines and any intended use of rriters or granting public access to the full cipant-level dataset, and statistical code t form and other related documentation ipants and authorised surrogates ction, laboratory evaluation, and storage of cimens for genetic or molecular analysis in al and for future use in ancillary studies, if ' is distributed under the terms of the Creative checklist was completed on 12. August 2022 using he EQUATOR Network in collaboration with	er relevant groups (eg, via publication, sults databases, or other data sharing , including any publication restrictions	
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).		9
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).	0 01	
is distributed under the terms of the Creative checklist was completed on 12. August 2022 using the EQUATOR Network in collaboration with (ABES).		/ copyright, 8
data mining,	imens for genetic or molecular analysis in	na na na
	checklist was completed on 12. August 2022 using	ur (ABES) . data mining,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

**BMJ** Open

Page 22 of 22

liographique de l