ISRCTN16251539 https://doi.org/10.1186/ISRCTN16251539

Evaluating how artificial intelligence affects clinicians' decisions when interpreting wrist X-rays

Submission date
31/12/2025

Recruitment status
No longer recruiting

Registration date
29/01/2026

Overall study status
Completed

Last Edited
20/03/2026

Condition category
Injury, Occupational Diseases, Poisoning

[X] Prospectively registered
[_] Protocol
[_] Statistical analysis plan
[_] Results
[_] Individual participant data
[X] Record updated in last year

Plain English summary of protocol

Background and study aims
Artificial intelligence (AI) systems are increasingly being developed to help clinicians interpret medical images such as X-rays. Some AI tools can detect fractures with a level of accuracy similar to clinicians in technical testing. However, much less is known about how AI affects clinicians' real decision making, confidence, and trust, and particularly how the way AI information is presented, influences these factors.
The aim of this study is to understand how AI assistance affects clinicians' ability to detect wrist fractures on X-rays, and how different styles of AI output (more or less information) influence diagnostic decisions. The findings will help inform the safe and effective design of AI decision-support tools for use in clinical practice.

Who can participate?
We are inviting clinicians who routinely interpret wrist or hand X-rays as part of their clinical work to take part. This includes doctors working in emergency departments, trauma, orthopaedics, or radiology training roles, and advanced nurse practitioners working in emergency departments or minor injury units. Participants must be able to give informed consent and attend the study session.

What does the study involve?
Participants will attend a single study session. During the session, they will review a set of anonymised wrist X-ray cases using a computer-based system. For each case, participants will first make a diagnosis without any AI support. They will then be shown an AI output for the same case and given the opportunity to confirm or revise their diagnosis. The AI support will be presented in one of two formats: a low-information display, showing a suggested diagnosis, or a high-information display, showing a short structured report with additional information such as the AI's confidence and feedback on image quality. Participants will also record their confidence in each decision. They will additionally be invited to take part in a short "think-aloud" exercise and interview, where they explain their reasoning while reviewing selected cases.

What are the possible benefits and risks of participating?
There is no direct clinical benefit to participants. However, participation may be professionally interesting and contribute to research aimed at improving how AI tools are designed and used in healthcare.
The study involves minimal risk. Participants may experience mild fatigue from reviewing multiple cases. There is no assessment of individual clinical competence, and all data are anonymised for analysis.
Participation is voluntary, and participants may withdraw for any reason.

Where is the study run from?
Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford (UK)

When is the study starting and how long is it expected to run for?
January 2026 to December 2026

Who is funding the study?
National Institute of Health and Care Research (NIHR) (UK). Rachel Kuo is supported by an NIHR Doctoral Research Fellowship (NIHR302562)

Who is the main contact?
Rachel Kuo, rachel.kuo@ndorms.ox.ac.uk

Contact information

Type(s)
Public, Principal investigator, Scientific

Contact name
Ms Rachel Kuo

Contact details
Botnar Research Centre
Nuffield Department of Orthopaedics, Rheumatology, and Musculoskeletal Sciences,
Windmill Road
Oxford
United Kingdom
OX3 7LD
+44 (0)1865 227374
rachel.kuo@ndorms.ox.ac.uk

Additional identifiers

Integrated Research Application System (IRAS)
296655

National Institute for Health and Care Research (NIHR)
302562

Study information

Scientific Title
A multi-reader multi-case study of clinician performance with, and without artificial intelligence assistance for wrist fracture detection, with embedded qualitative study

Study objectives
Primary objective:
To evaluate the impact of artificial intelligence (AI) assistance on clinicians' diagnostic performance when interpreting wrist radiographs, measured using area under the receiver operating characteristic curve (AUC).

Secondary objectives:
1. To assess the effect of AI assistance on clinician diagnostic confidence and time-to-decision.
2. To compare clinician performance and decision behaviour between low-information and high-information AI output formats.
3. To evaluate patterns of agreement and disagreement between clinician decisions and AI outputs, including agreement when AI is correct and when AI is incorrect.
4. To explore whether the effects of AI assistance differ by clinician characteristics, such as professional role or level of experience.

Exploratory objectives:
1. To examine how clinicians interact with different AI outputs and how these interactions influence diagnostic reasoning.
2. To identify qualitative themes relating to clinician trust in, reliance on, or scepticism toward AI decision support using think-aloud and interview data.
3. To generate design insights to inform the safe and effective implementation of AI decision-support tools in clinical practice.

Ethics approval required
Ethics approval required

Ethics approval(s)
approved 22/12/2025, Camden and Kings Cross Research Ethics Committee (2 Redman Place, Stratford, London, E20 1JQ, United Kingdom; +44 (0)2071048112; CamdenandKingsCross.REC@hra.nhs.uk), ref: 22-LO-49

Primary study design
Observational

Secondary study design
Multi-reader multi-case study with embedded qualitative study

Study type(s)

Health condition(s) or problem(s) studied
Wrist fracture

Interventions
This is a multi-reader multi-case (MRMC) study evaluating clinician interpretation of wrist radiographs with and without artificial intelligence (AI) decision support.

Participating clinicians will interpret a fixed set of anonymised wrist radiograph cases using a secure, computer-based research application. For each case, clinicians will first provide an unaided diagnostic assessment and confidence rating. They will then be shown an AI output for the same case and asked to confirm or revise their diagnostic assessment and confidence.

AI assistance will be presented in one of two predefined formats at the case level: a low-information display (diagnostic label and visual heatmap) or a high-information display (structured report including diagnostic label, model confidence, image quality feedback, and heatmap). Each case will be assigned to one AI display format, and all readers will view the same cases under the same format.

Updated 20/03/2026: AI assistance will be presented in one of two predefined formats at the case level: a low-information display (diagnostic label) or a high-information display (structured report including diagnostic label, model confidence, image quality feedback, and heatmap). Each case will be assigned to one AI display format, and all readers will view the same cases under the same format.

No clinical treatment or patient intervention is administered. All cases are retrospective and de-identified, and participant responses do not influence patient care.

Quantitative data collected will include diagnostic scores, categorical decisions, confidence ratings, time-to-decision measures, and agreement between clinician decisions and AI outputs. Diagnostic performance will be assessed using area under the receiver operating characteristic curve (AUC), with comparisons made between unaided and AI-assisted interpretation, and between low- and high-information AI formats, accounting for the MRMC study design.

An embedded qualitative component will involve participants completing a think-aloud exercise and brief semi-structured interview while interpreting selected cases. These sessions will explore how clinicians interact with the AI system and how AI outputs influence reasoning, confidence, and decision-making. Qualitative data will be analysed using thematic analysis.

Intervention Type
Other

Primary outcome(s)
1. Clinician diagnostic discrimination measured using area under the receiver operating characteristic curve (AUC) derved from reader diagnostic scores against the reference standard during unaided and AI-assisted interpretation of wrist radiographs at within the study session

Key secondary outcome(s))
1. Clinician diagnostic confidence measured using self-reported confidence rating at the time of unaided and AI-assisted interpretation of wrist radiographs at within the study session

2. Time to diagnostic decision measured using system-recorded timestamps (seconds per case) at the time of unaided and AI-assisted interpretation of wrist radiographs at within the study session

3. Clinician–AI agreement measured using binary agreement between clinician AI-assisted decision and AI recommendation at the time of AI-assisted interpretation of wrist radiographs at within the study session

4. Clinical management decision measured using categorical management choice (e.g. discharge vs follow-up) at the time of unaided and AI-assisted interpretation of wrist radiographs at within the study session

5. Interaction with high-information AI features measured using system-recorded usage logs (access and time spent) during AI-assisted interpretation of wrist radiographs at within the study session

6. Clinician reasoning and decision-making processes measured using audio-recorded think-aloud protocols analysed with reflexive thematic analysis during AI-assisted interpretation of selected wrist radiograph cases at within the study session

Completion date
26/02/2026

Eligibility

Key inclusion criteria
Wrist imaging:
1. Adult patients (18 years or older) who have had a x-rays of the hand/wrist for suspected fracture at Oxford University Hospitals NHS Foundation Trust, 2010-2022.

Readers:
1. Clinicians who interpret wrist and/or hand radiographs as part of routine clinical practice.
2. Healthcare professionals (advanced nurse practitioners and doctors) who currently work, have worked, or will work (within 1 year of the study) in emergency departments, minor injuries units, radiology, musculoskeletal trauma services in the United Kingdom.
3. Able to provide informed consent.
4. Able to attend the study session and use the study software.

Healthy volunteers allowed
Yes

Age group
Mixed

Lower age limit
18 years

Upper age limit
70 years

Sex
All

Total final enrolment
28

Key exclusion criteria

1. Clinicians who do not routinely interpret wrist or hand radiographs.
2. Inability to provide informed consent.
3. Prior involvement in the development, training, validation, or evaluation of the AI system used in the study.
4. Prior access to, or knowledge of, the reference standard labels for the study cases.
5. Any conflict of interest or role that could reasonably compromise independent interpretation of cases (e.g. involvement in case selection or ground-truth annotation).

Date of first enrolment
18/02/2026

Date of final enrolment
26/02/2026

Locations

Countries of recruitment
United Kingdom

England

Study participating centre
University of Oxford
University Offices
Oxford
England
OX1 2JD

Sponsor information

Organisation
University of Oxford

ROR
https://ror.org/052gg0110

Funder(s)

Funder type

Funder Name

National Institute for Health and Care Research

Alternative Name(s)
National Institute for Health Research, NIHR Research, NIHRresearch, NIHR - National Institute for Health Research, NIHR (The National Institute for Health and Care Research), NIHR

Funding Body Type
Government organisation

Funding Body Subtype
National government

Location
United Kingdom

Results and Publications

Individual participant data (IPD) sharing plan
The dataset generated and analysed during this study will be made available in a publicly accessible repository (e.g. GitHub) to support transparency and reproducibility of the multi-reader multi-case (MRMC) analyses. All data will be anonymised prior to sharing, with readers labelled using non-identifiable codes (e.g. R1, R2, …). No personally identifiable information or linked demographic data will be included in the shared dataset. All code will be shared.

Transcripts from the think-aloud study will not be publicly available.

IPD sharing plan summary
Stored in publicly available repository