NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

Diagnostics Assessment Programme

Artificial intelligence software to help detect fractures on X-rays in urgent care

Final scope

July 2024

1 Introduction

The topic selection oversight panel identified artificial intelligence software to help detect fractures on X-ray in urgent care as potentially suitable for early value assessment (EVA) by the HealthTech Programme based on a topic intelligence briefing.

The software identified for this assessment use algorithms that have been produced using artificial intelligence. The algorithms are fixed but are updated periodically.

A glossary of terms is provided in appendix A.

2 Description of the technologies

This section describes the properties of the diagnostic technologies based on information provided to NICE by manufacturers and experts. NICE has not carried out an independent evaluation of this description.

2.1 Purpose of the medical technologies

Plain film radiography or X-ray is the most common medical imaging approach used to detect fractures in urgent care settings, including the emergency department (ED), urgent treatment centre (UTC), and minor injuries units (MIU). Missed fractures are reported to be the most common diagnostic error in the ED (Hussain et al. 2019). Missed or delayed diagnosis of fractures on radiographs is reported to occur in around 3% to 10% of cases (Kuo et al. 2022). The radiology get it right first time programme national speciality report highlights the increasing demand on radiology services that is not matched by growth in NHS radiology capacity. The result of this increased demand and workforce limitations is that X-rays are often interpreted by healthcare professionals who are not radiology specialists or who are inexperienced at interpreting X-rays, potentially increasing the likelihood of missing a fracture or making unnecessary referrals to fracture clinics. These workforce pressures also lead to delays to diagnosis and longer turnaround times. Experts explained that EDs are busy work environments with high patient numbers and numerous and frequent distractions for staff. Image visualisation facilities may also be suboptimal for accurate interpretation. These factors can all contribute to fractures being missed. Periods of reduced staff numbers, such as outside normal working hours, may also lead to longer turnaround times and increased missed fracture rates (York et al. 2020).

The NHS resolution report on missed fractures suggested that missed fractures can lead to poor patient outcomes and further harms including pain and suffering, loss of function, need for further or prolonged treatments, cosmetic deformity, nerve damage, prolonged recovery and death. Missed and delayed fracture diagnoses can also have an impact on service delivery, for example, increased waiting times, delays in people being discharged, people being recalled, additional medical appointments, surgical procedures and physiotherapy. Examples of clinically significant missed fractures highlighted by clinical experts include fractures of the hip, scaphoid, and foot (Lisfranc fracture). However, clinical experts also suggested that it is typically subtle fractures that tend to be missed, and that these injuries tend not to require any specific fracture management or treatment. The outcomes for missed fractures vary by the type and location of the fracture, mechanism of injury and the characteristics of the person with the fracture.

Artificial intelligence (AI) technologies that can help detect fractures and support healthcare professional interpretation of X-ray images could improve the accuracy of X-ray fracture diagnoses in urgent care settings. The technologies could also speed up the flow of people through the care pathway and reduce the likelihood that fractures are missed before a radiologist or reporting radiographer reviews the X-rays. This could reduce the number of people being recalled and reduce the risk of further injury or harm for people during the period of time between interpretation and initial treatment decision in the ED and the radiology report. Using AI technologies could also reduce the number of people without fractures who are referred to virtual fracture clinics. This could help reduce the workload associated with the remote assessment of fractures by orthopaedic teams, for example. Clinical experts explained that the ionising radiation (medical exposure) regulations (IRMER) state that clinical evaluation of X-rays requires a trained person. Therefore, AI technologies currently can't be used autonomously without human interpretation.

2.2 Product properties

Several companies offer software with artificial intelligence derived algorithms for analysing X-ray images to detect fractures. Some of the companies provide the software directly on their own platforms, whereas others are available via multivendor platforms (for example, the Blackford Platform). The technologies included in this scope assist healthcare professionals in the interpretation of X-ray images to help diagnose fractures. They use X-ray radiographs in DICOM (digital imaging and communications in medicine) format which are stored on the hospital's PACS (picture archiving and communications system). Images are then processed/analysed using proprietary artificial intelligence-derived algorithms.

2.2.1 BoneView (Gleamer)

BoneView is a class IIa CE marked software-as-a-service device designed to assist clinicians in the interpretation of X-ray radiographs. The company states that BoneView detects fractures in X-rays of the appendicular skeleton, ribs and thoracic-lumbar spine. It is compatible with all available X-ray imaging systems. BoneView uses X-ray radiographs in DICOM format. It processes them using artificial intelligence (deep learning) and detects anomalies. The software performs an image analysis step to determine whether the body part is in scope, and it can reject unsupported or unreadable images. Healthcare professionals view the results as DICOM secondary captures with bounding boxes around any abnormalities. A results summary sheet is also provided. The company states that the software identifies fractures, dislocations, effusions and bone lesions and that it is suitable for use in people aged 2 years and over.

2.2.2 qMSK (Qure.ai)

qMSK is a class IIb CE marked technology and is commercially available to the NHS in the UK. The company states that it can detect fractures in the appendicular skeleton and ribs. The company states that qMSK is intended for use in adults only.

2.2.3 Rayvolve (AZ Med)

Rayvolve is a class IIa CE marked AI-based, computer aided diagnosis tool. According to a recent review by Pauling et al. (2024), Rayvolve detects fractures in the appendicular skeleton and ribs. It can also detect dislocations, joint effusions and chest pathologies (pneumothoraces, cardiomegaly, pleural effusions, pulmonary oedema, consolidation, nodules). The company states that Rayvolve identifies fractures and presents the results directly into the clinicians' interpretation console in the existing DICOM series. The tool is integrated into hospitals' existing radiology workflows using Wellbeing's AI Connect gateway. Rayvolve is intended for use in adults only.

2.2.4 RBfracture (Radiobotics)

RBfracture is a class IIa CE marked AI-powered diagnosis software device that assists healthcare professionals in detecting fractures during the review of musculoskeletal X-rays. RBfracture is not intended to replace humans and requires second review by a healthcare professional. The company states that RBfracture detects fractures in X-rays across the entire appendicular skeleton. It also detects effusion of the knee and elbow, lipohaemarthrosis of the knee, rib fractures, and periprosthetic fractures. RBfracture is not trained to detect chronic or healed fractures so is not intended for use in assessing suspected maltreatment and physical abuse of children. The company states that RBfracture is compatible with all available X-ray imaging systems. Healthcare professionals view the outputs of the software in their existing PACS/DICOM viewer. A DICOM study with 1 or more radiographs is sent for analysis and RBfracture returns a summary report overview of analysed images and findings. The summary report uses a red dot to indicate if a fracture or other finding has been detected. It also provides annotated radiographs with bounding boxes around areas of interest and a summary field with the analysis results. Bounding boxes with a dashed line indicate findings with a low confidence score, and a solid line indicates those with a high confidence score. RBfracture provides information in the PACS worklist about whether or not a supported lesion (fracture, effusion, lipohaemarthrosis) is detected. The technology has a built-in image quality module that is able to reject unsupported anatomical regions and images with poor image exposure. RBfracture is approved for use in people above 2 years of age.

2.2.5 TechCare Alert (Milvue)

TechCare Alert is a configuration of Milvue Suite. It is a class IIa medical device intended for the automatic detection and visualisation of multiple bone and chest-related anomalies on conventional radiographs, providing preliminary triage and diagnosis aid to healthcare professionals. The company states that it is a predictive tool that provides preliminary data and that final diagnosis is based on the expertise of the healthcare professional. TechCare Alert is not intended for use in critical care. The company states that TechCare Alert detects fractures in X-rays of the appendicular skeleton and the ribs. It can also detect dislocations, elbow joint effusion, pleural effusion, pulmonary opacity, pulmonary nodules and pneumothorax (Pauling et al. 2024). TechCare Alert uses a classifier to automatically detect potential issues prior to the generation of secondary captures images. If detected, images are either analysed with a warning legend attached to inform readers of the issue or not analysed with an explanation on the summary screen. The company states that TechCare Alert uses the DICOM standard and healthcare professionals view the outputs of the software in their existing PACS/DICOM viewer. Results are provided as DICOM secondary captures, including a summary image and a unified result combining the original image and annotations. Annotations include a bounding box around areas of interest. The company states that there is no age limit for TechCare Alert and it is certified for use in adult and paediatric populations.

3 Target conditions

The target population for this assessment is people attending an urgent care centre, for example, emergency department (ED), urgent treatment centre (UTC) or minor injuries unit with a suspected fracture, for which X-ray imaging is requested. In 2018/19, there were 1,147,822 emergency department attendances where the primary diagnosis was classified as either a dislocation, fracture, joint injury or amputation, accounting for 5.1% of all emergency department attendances (NHS Digital). The frequency of different types of fracture differs between age groups. Fractures are also a common reason for hospital admission, with around 250,000 fracture admissions in England each year according to data from the hospital episodes statistics database (Jennison and Brinsden, 2019). The most common fractures that require hospital admission are hip fractures, followed by distal radius fractures, ankle fractures and hand fractures. Fractures may be broadly categorised as non-complex or complex fractures.

Non-complex fractures account for most of the 1.8 million fractures that occur in England each year. They include a wide range of injuries that can usually be managed in a local hospital rather than a specialist centre. They occur over the complete age range from infancy to old age. Many different bones can be involved and there are numerous mechanisms of injury.

A minority of fractures are classed as complex. They include injuries such as pelvic fractures, open fractures and severe ankle fractures. The treatment of complex fractures is often complicated and usually involves multiple healthcare professionals and specialists. Complex fractures may be associated with major trauma.

Hip fractures occur in the area between the neck of the femur and the femoral head. There are around 65,000 hip fractures each year and they are more common in older people, with average ages of 84 years for men and 83 years for women (National Hip Fracture Database). A clinical expert explained that delayed diagnoses of hip fractures are associated with increased morbidity and mortality.

Fractures of the spinal column may occur in the cervical, thoracic or lumbosacral regions of the spine. They can lead to spinal cord injury resulting in serious neurological damage including paraplegia, tetraplegia or death.

4 Diagnostic and care pathway

The specific imaging approach taken, and care pathway, varies by the site and type of fracture. NICE guidelines on non-complex (NG38) and complex fractures (NG37), hip fracture (CG124), major trauma (NG39), spinal injury (NG41), and head injury (NG232) make recommendations on when different imaging modalities should be considered.

4.1 Initial assessment and X-ray imaging of suspected fractures

Non-complex fractures are initially seen and diagnosed in emergency departments (ED), urgent treatment centres (UTC) or minor injuries units, although a small proportion may present in primary care to GPs and are then referred on. Fracture assessment and diagnosis will typically involve triage where an ED nurse, ACP or doctor will carry out an initial assessment before requesting imaging. The choice of medical imaging requested is determined by factors such as the site of the suspected fracture, the mechanism of injury and characteristics of the person such as age. Depending on these factors, imaging may be either X-ray, CT scan or MRI. X-rays are usually the first line imaging approach for non-complex fractures and are usually performed by a diagnostic radiographer. Clinical experts explained that in some centres, the diagnostic radiographer may also provide some initial interpretation of the X-ray and provide comments to support the healthcare professional's interpretation in the urgent care setting. The images (plain film radiographs), along with any radiography comments are then interpreted by a healthcare professional to decide on the most appropriate treatment option (see section 4.5). Healthcare professionals interpreting the X-ray images may include emergency nurse practitioners (ENP), advanced clinical practitioners (ACP), physiotherapists or ED doctors, who may have varying levels of experience. X-rays are then later reviewed by a radiologist or reporting radiographer.

Although X-ray is usually the first-line imaging approach for most non-complex fractures, other imaging types may be used for some suspected fractures. For example, NG38 recommends that MRI should be considered for first-line imaging in people with a suspected scaphoid fracture. Similarly, the NICE clinical guideline on hip fracture management (CG124) recommends offering MRI or CT if a hip fracture is suspected despite no fracture being detected on X-ray. However, clinical experts explained that these practices may vary between centres based on resources and capacity.

For people with suspected chest trauma, a chest X-ray may be done, but this may be primarily to detect other conditions such as pneumothorax, rather than fractures.

X-ray may also be used for some suspected spinal column fractures. NG41 recommends X-ray as the first-line investigation for people with suspected spinal column injury, without abnormal neurological signs or symptoms in the thoracic or lumbosacral regions. This is followed by CT if the X-ray result is abnormal or there are clinical signs or symptoms of a spinal column injury. It also recommends that X-rays should be considered for children when a clinical suspicion of cervical spinal injury remains after repeated clinical assessment, but the criteria for MRI imaging are not met. Findings of the X-rays should be discussed with a consultant radiologist and further imaging performed if needed.

X-ray imaging may also be used for children (under 16 years), who have a head injury, and neck pain or tenderness but no indications for a CT cervical spine scan. NG232 recommends doing 3-view cervical spine X-rays before assessing range of movement in the neck if there was a dangerous mechanism of injury, safe assessment of range of movement in the neck is not possible, or the person has a condition that predisposes them to a higher risk of injury to the cervical spine (for example, collagen vascular disease, osteogenesis imperfecta, axial spondyloarthritis). Clinical experts explained that although fractures of the skull are typically assessed using CT, fractures of the facial bones may be X-rayed and these can be complex to interpret.

For suspected chest trauma, spinal column fractures and fractures of the head and neck, NICE guidelines (NG39, NG41 and NG232) recommend that X-ray imaging should be done urgently and interpreted immediately by a healthcare professional with training and skills in the area.

4.2 Fractures which require alternative first line imaging approaches

In general, imaging of complex fractures is usually done using CT as it may be indicated for assessing other injuries at the same time. NICE's guideline on the assessment and management of complex fractures (NG37) recommends whole body CT for adults with blunt major trauma and suspected multiple fractures. In children (under 16 years) with suspected multiple injuries clinical judgement should be used to limit CT to the body areas where assessment is needed. NG37 also recommends using CT for first-line imaging in adults with suspected high-energy pelvic fractures and is also preferred for children with clinical judgement to limit CT to the body areas where assessment is needed.

If spinal cord or cervical column injury is suspected, NG41 recommends MRI for children (under 16) and CT for adults.

Suspected fractures of the skull are typically imaged using CT where needed. For people who have a head injury, NG232 recommends that a CT scan or MRI may also be done if there are signs and symptoms suggesting injury to the cervical spine.

4.3 Radiology review and reporting

The NICE guideline on non-complex fractures (NG38) recommends that a radiologist, radiographer or other trained reporter should review the X-rays and provide a definitive written report of suspected fractures before the injured person is discharged. This approach is known as hot reporting and is intended to provide a safety net for people where a fracture is missed at the initial X-ray interpretation in the urgent care setting. Clinical experts have explained that in practice this is not always possible and reporting delays can occur. In some centres X-rays may be reported the same day, however in others there can be significant delays ranging from days to weeks. Both NG37 and NG38 recommend that all patient documentation, including images and reports, travels with people when they are transferred to other departments or centres. Clinical experts explained that this practice is also variable and may vary between centres. Clinical experts said that X-rays are not usually prioritised for radiology reporting, with most centres operating a first-in, first-out system. Delays in radiology reporting present issues with people potentially being discharged with undiagnosed fractures, some of which can present serious risks of further harm or injury (see section 2.1). Depending on the length of delay to radiology review, consequences for people may include pain and discomfort for a period of time, through to the possibility of malunion or incorrect bone healing which can lead to more complicated surgical interventions that otherwise may not have been necessary.

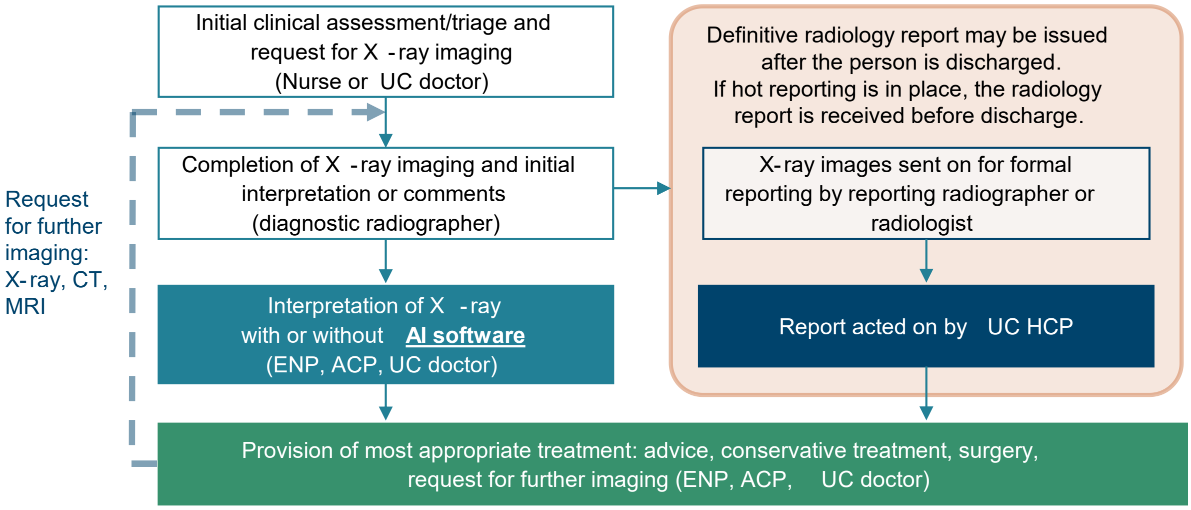
Outside of normal working hours, when there are reduced staff numbers, many centres outsource their radiology reporting to a third-party service provider. Outsourcing may also be used at centres where there is no onsite radiology service. Clinical experts explained that outsourcing is generally more expensive than in-house reporting. They also explained that the quality of these reports can be variable and there are often discrepancies or disagreements with internal opinion.

4.4 Position of AI technologies in the diagnostic pathway

Clinical experts explained that the most likely use case for AI interpretation of X-ray images was as a decision aid used by the healthcare professional in the ED, UTC or minor injuries unit that is interpreting the X-ray and making decisions on the most appropriate treatment or care. AI technologies could also be used by diagnostic radiographers at the time of X-ray acquisition as an aid for commenting.

The proposed position of the AI technologies in the diagnostic pathway is presented in figure 1.

Figure 1. Care pathway for fracture diagnosis as outlined in the NHS resolution report 2022.



ENP: emergency nurse practitioner, ACP: Advanced clinical practitioner, UC: urgent care, HCP: healthcare professional.

4.5 Treatment of fractures

Multiple treatment options are available for fractures including surgical and non-surgical approaches depending on the type of fracture. In most cases, people with a non-complex fracture will be treated in the ED, attend a virtual fracture clinic and follow a patient-led follow up pathway. The decision to admit or discharge a person with a fracture depends on the nature of the fracture and any other related injuries or illnesses that the person may have. NICE NG38 makes recommendations on fracture management and treatment options for different non-complex fractures in adults and children, following initial assessment and diagnostic imaging. These include non-surgical orthopaedic management, manipulation and plaster cast, or surgical fixation. A clinical expert explained that the nature of the injury, local capacity and local policy determines whether surgery is done before the person is discharged or arranged for a later date. Depending on the nature of the fracture, physiotherapy assessment and potentially ongoing physiotherapy review may be offered. The specific treatment approach depends on a range of factors including the age of the person, site of fracture, level of bone displacement, and whether the fracture is complicated (for, example due to an open wound, dislocation or vascular injury).

The management and treatment of complex fractures and hip fractures generally includes surgical intervention as outlined in NG37 and CG124, respectively. Complex fractures and hip fractures may require physiotherapy assessment and regular physiotherapy review. NICE NG41 outlines the early management of traumatic spinal injuries in the emergency department.

4.6 Patient issues and preferences

Timely and accurate detection of fractures in the ED or UTC may reduce waiting times for people with suspected fractures. Technologies that increase the accuracy of fracture detection can reduce the risk of people being discharged with a missed fracture and can help reduce unnecessary pain and discomfort that people may experience while waiting for radiology review and subsequent treatment. Increased accuracy of fracture detection could also reduce the number of people that are recalled following radiology review. A patient expert explained that diagnostic confidence was an important consideration, to provide reassurance that nothing has been missed. Recalls can be particularly inconvenient for fractures in children where parents or carers may need to take time off work. Experts also highlighted that patients may have concerns about use of or sharing of their data when AI technologies are involved, and the need for consent. Patient experts said it was important to be well-informed about the decision-making process, including AI involvement and whether the clinician agrees with the AI output.

5 Comparator

The comparator is standard care for fracture assessment where the urgent care clinician or healthcare professional interprets the X-ray radiograph without AI assistance.

The reference standard or assessment of ground truth is based on the consultant radiologist or reporting radiographer interpretation and report. Although considered the reference standard, fracture detection by a radiologist or reporting radiographer is not 100% accurate as fractures may still be missed.

6 Scope of the assessment

Table 1: Scope of the assessment

Decision question	Does the use of software with artificial intelligence (AI) derived algorithms for analysing X-ray images to detect suspected fractures have the potential to be clinically and cost-effective to the NHS? What evidence is available to support the value proposition outlined in the scope? 1 Improve the accuracy of fracture detection from X-rays in the emergency department, urgent treatment centre or minor injuries unit 2 Service delivery and workflow improvements, for example, reduced waiting times, fewer people being recalled, and a reduction in unnecessary fracture clinic referrals and medical appointments. What are the evidence gaps?
Populations	People presenting to the emergency department, urgent treatment centre or minor injuries unit with a suspected fracture. Depending on the availability of evidence, the following subpopulations may be included: • Children and young people (0 to 16 years of age) • Older people • People with conditions affecting bone health (for example, osteoporosis and osteogenesis imperfecta) Depending on the availability of evidence, the following fracture site subgroups may be included: • Hip • Hand (including wrist), foot (including ankle) • Fractures including the growth plate (Salter-Harris) in children • Fractures of the elbow in children
Interventions	AI used as a decision aid for X-ray image interpretation and fracture assessment prior to radiology review, using any of the following software/platforms: • BoneView (Gleamer) • Rayvolve (AZmed) • Rbfracture (Radiobotics)

	• qMSK (Qure.ai) • TechCare alert (Milvue)
Comparator	ED clinician or healthcare professional interpretation of X-ray radiograph without AI assistance. Reference standard or ground truth based on consultant radiologist or reporting radiographer interpretation and report.
Healthcare setting	Emergency department, urgent treatment centre or minor injuries unit.
Outcomes: intermediate measures	Intermediate measures for consideration may include: • Measures of diagnostic accuracy to detect fractures • Accuracy when used by different healthcare professionals (emergency nurse practitioners, advanced clinical practitioners, urgent care doctors, diagnostic radiographers) • Diagnostic confidence • Healthcare professional X-ray reading time • Time to diagnosis or time to X-ray definitive radiology report • Time spent in the emergency department, urgent treatment centre or minor injuries unit • Time to treatment • Proportion of people that need further imaging • Number of missed fractures • Rate of missed fracture-related further injury • Number of people recalled following radiology review • Number of treatments (plaster casts, surgical procedures, physiotherapy appointments) and extent of treatments (complexity of surgery, length of physiotherapy course) • Number of hospital appointment/visits, including referrals to fracture clinics and orthopaedic assessment • Number of hospital admissions • Length of stay in hospital • Number of further imaging events required • Failure rate or rate of inconclusive AI reports • Healthcare professional user acceptability of AI tools for detecting fractures
Outcomes: clinical	Clinical outcomes for consideration may include: • Morbidity • Mortality
Outcomes: patient-reported	Patient-reported outcomes for consideration may include:

	• Health related quality of life
Outcomes: costs	Costs will be considered from an NHS and Personal Social Services perspective. Costs for consideration may include: • Cost of AI software • Staff costs for X-ray image interpretation • Training costs • Costs of additional medical appointments and further or confirmatory imaging (CT/MRI) • Costs of treatment • Costs of physiotherapy
Measuring cost-effectiveness	The cost-effectiveness of interventions should be expressed in terms of incremental cost per quality-adjusted life year if possible.
Time horizon	The time horizon for estimating clinical and cost effectiveness should be sufficiently long to reflect any differences in costs or outcomes between the technologies being compared.

7 Other issues for consideration

Dislocations and other bone/joint related injuries

Clinical experts highlighted that dislocations can be difficult to detect on X-ray and are likely the most common clinically significant injuries to be missed. Clinical experts suggested that the most common clinically significant non-fracture injuries that may be missed on X-ray include:

- Lipohaemarthrosis
- perilunate dislocation,
- posterior dislocation of the shoulder.

Some of the artificial intelligence technologies are indicated for use in detecting additional bone related injuries and anomalies including dislocations.

Missed fractures in children

X-rays of suspected fractures in children can be more complicated to interpret, particularly at joints. Fractures that include the growth plate (Salter-Harris fractures) are relatively common and can lead to long term biomechanical issues such as limb shortening or abnormal growth. Although NG38 recommends that MRI should be considered for first-line imaging in people with a suspected scaphoid fracture, clinical experts highlighted that in children, X-ray would be the usual approach. A clinical expert also explained that there was a shortage of paediatric radiologists. Some of the AI technologies are approved for use in children aged 2 years and over whereas others are only approved for use in adults.

Child maltreatment/physical abuse

A clinical expert suggested that AI technologies could help identify instances of physical abuse in young children, for example by detecting previous fractures that have healed. However, the technologies may not be approved for this specific use case.

Detection of healed or clinically insignificant fractures by AI technologies

Clinical experts suggested that if AI fracture detection algorithms were overly sensitive and so detected healed fractures, this may lead to an increase in unnecessary referrals and appointments. Healthcare professionals who are not experts in X-ray interpretation may be less confident in overruling these AI decisions.

Incidental findings

Clinical experts explained that AI technologies that only detect fractures could miss incidental findings on X-ray such as a bone metastasis or a pneumothorax and healthcare professionals could be inappropriately reassured by the AI report. Therefore, healthcare professional interpretation of X-rays to assess incidental findings is an important consideration when using AI technologies.

Ongoing studies

AI Assisted Detection of Fractures on X-Rays (FRACT-AI)

The AI Assisted Detection of Fractures on X-Rays (FRACT-AI) (NCT06130397) study is a retrospective multiple-reader multiple-case (MRMC) study which aims to evaluate the impact of an Artificial Intelligence (AI)-enhanced algorithm (Boneview) on the diagnostic accuracy of healthcare professionals in the detection of fractures on plain X-ray. A group of readers will be recruited from 6 distinct clinical groups:

emergency medicine, trauma and orthopaedic surgery, emergency nurse practitioners, physiotherapy, radiology and radiographers, with 3 levels of seniority in each group. Readers will interpret all images in a dataset of 500 plain X-rays involving standard images of all bones other than the skull and cervical spine, with 50% normal cases and 50% containing fractures. Readers will first interpret the images without and then with the assistance of the AI software. The accuracy of the AI software and both the readers' unassisted and assisted interpretations will be compared against a reference standard 'ground truth', which is determined from multiple senior radiology reports for each image. Primary outcome measures include diagnostic accuracy (sensitivity and specificity) of the AI algorithm alone, diagnostic accuracy of readers with and without AI assistance, and reader speed with and without AI. The estimated study completion date is December 2024.

Testing an artificial intelligence tool for childhood fracture detection on X-rays

This is an observational, retrospective, multicentre and multireader cohort study (ISRCTN12921105). It aims to evaluate the impact of BoneView on the diagnostic accuracy, confidence and potential change in management plans of healthcare professionals who routinely review bone radiographs of children (aged 2 to 16 years of age). The study will involve a minimum of 30 readers, including general radiologists, emergency medicine clinicians, reporting radiographers and orthopaedic surgeons who will interpret 500 paediatric limb radiographs (across 4 body parts: ankle, wrist, elbow and knee) without and with the assistance of the AI tool. The scans will include approximately 35% abnormal (fractured) cases and the rest normal to simulate the normal prevalence of injuries in clinical practice. The study will assess the stand-alone performance of the AI tool and its impact on the readers' performance. The ground truth (reference standard) will be set by 2 consultant paediatric radiologists.

AutoRayValid-RBfracture study: evaluating the efficacy of an AI fracture detection system

AutoRayValid-RBfracture study is a multicentre retrospective study evaluating the RBfracture. It aims to assess the AI's impact on diagnostic thinking by analysing consecutive cases with clinical data, providing insights into fracture detection and clinical decision-making. The study will be based in 3 European sites and will include people aged 21 and over with X-ray indications for appendicular fractures. Each site will analyse 500 cases. The reference standard will be based on annotations by 2 experienced readers.

Ongoing real-world data collection (RBfracture)

There are a number of ongoing NHS-based real-world data collection studies that are using RBfracture. These are based at NHS sites including Ayrshire and Arran, Kettering General Hospital, Harrogate and District and Leeds University Hospital. In addition, as part of the NHS England AI lab pilot AI deployment platform (AIDP), East Midlands and Thames Valley radiology services will carry out real-world data collection including post-market surveillance. These studies expect to complete between the end of 2024 and 2025.

Litigation considerations

The NHS resolution missed fractures report states that between 2015/16 and 2017/18 there were 78 successful (closed) claims in which the speciality was 'Accident and Emergency' and there was a missed fracture. This number does not include missed fractures incidents in which there remains an open claim, or no claim has been made. The total cost of missed fracture claims was £1,118,972, including £469,611 paid in damages and £649,361 in legal costs (for both claimants and NHS Resolution). It states that the total annual cost of missed fractures is low as a proportion of the total cost of clinical negligence claims in England and the operational budgets of the organisations that provide ED services. However, it is still an avoidable cost to NHS providers. AI technologies that improve the accuracy of fracture detection and reduce the number of missed fractures could help to reduce the number of claims and associated litigation costs.

8 Potential equality issues

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

Age is a protected characteristic and the scope considers all adults, young people and children who present with a suspected fracture. However, some fractures are more common in certain age groups, for example, hip fractures are more common in older people. Clinical experts explained that bone health can vary widely with age and can be affected by other factors including socioeconomic background. Clinical experts explained that fractures are also more difficult to detect in children (see section 7). People with conditions that affect bone health (for example, osteoporosis and osteogenesis imperfecta) may be more susceptible to fractures. AI technologies may perform differently in people with underlying comorbidities, such as conditions affecting bone health. Clinical experts highlighted that certain drugs can reduce bone density and increase the risk of developing osteoporosis.

Clinical experts also explained that the diversity of populations used to train artificial intelligence algorithms was an important consideration. If the algorithm has been developed, trained and validated in populations in which particular groups (such as people from different ethnic backgrounds, age, or sex) have been underrepresented, they may perform differently in these groups.

9 Potential implementation issues

IT issues

The artificial intelligence technologies will need to integrate into existing hospital PACS systems to ensure there is no disruption or delays to the workflow.

Procurement

Procurement may differ between the technologies with companies offering various pricing options including annual subscriptions and pay per use. Smaller or rural centres that have lower patient numbers and perform fewer X-rays may not have sufficient volume to justify the cost of an annual site licence. Any requirement for the use of specific multi-vendor platforms may limit which trusts can access specific technologies. Any additional, bespoke company software may also be a potential barrier to implementation and may increase the risk of vendor lock-in.

10 Authors

Simon Webster

Topic Lead

Jacob Grant

Technical Adviser

July 2024

Appendix A Glossary of terms

Appendicular skeleton

The portion of the skeleton that includes and supports the limbs. It includes the pectoral girdle and the bony pelvis.

Cervical spine

The upper region of the spine consisting of 7 vertebrae. It extends from the skull base to the first vertebra with a rib attached to it.

Computed tomography (CT)

An imaging test that uses x-rays and a computer to create detailed pictures of the inside of the body. It takes pictures from different angles. The computer puts them together to make a 3-dimensional image.

Lipohaemarthrosis

The escape of fat and blood from the bone marrow into the joint as a result of an intra-articular fracture. It is most frequently seen in the knee.

Lumbosacral spine

Lower region of the spine consisting of the 5 large vertebrae that make up the lumbar spine and 5 fused vertebrae that make up the sacrum.

Magnetic resonance imaging (MRI)

A type of scan that uses strong magnetic fields and radio waves to produce detailed images of the inside of the body.

Major Trauma

Major trauma describes a serious injury that could cause permanent disability or death. Examples of major trauma include serious injuries to the head, the spine or the chest, injuries that cause someone to lose a lot of blood, and complex fractures (such as a broken pelvis or a broken bone that is sticking out through the skin)

Perilunate dislocation

Dislocation of the carpus bones in the wrist relative to the lunate which remains in normal alignment with the distal radius. Often associated with a fracture of the scaphoid.

Periprosthetic fracture

A fracture that occurs around an implanted orthopaedic prosthesis, for example knee or hip replacement

Software-as-a-service (SaaS)

A form of cloud computing in which the provider offers the use of software to a client and manages all the physical and software resources used by the application

Thoracic spine

Middle region of the spine, between the cervical vertebrae and the lumbar vertebrae

Appendix B References

Hussain F, Cooper A, Carson-Stevens A, Donaldson L, Hibbert P, Hughes T, Edwards A (2019) Diagnostic error in the emergency department: learning from national patient safety incident report analysis. BMC Emergency Medicine 19(77)

Jennison T, Brinsden M (2019) Fracture admission trends in England over a ten-year period. The Annals of The Royal College of Surgeons of England 101(3): 208 – 214

Kuo R, Harrison C, Curran TA, Jones B, Freethy A, Cussons D, Stewart M, Collins GS, Furniss D (2022) Artificial Intelligence in Fracture Detection: A Systematic Review and Meta-Analysis. Radiology 304(1): 50 – 62

Pauling C, Kanber B, Arthurs OJ, Shelmerdine SC (2024) Commercially available artificial intelligence tools for fracture detection: the evidence. BJR open 6(1)

York TJ, Jenkins PJ, Ireland AJ (2019) Reporting Discrepancy Resolved by Findings and Time in 2947 Emergency Department Ankle X-rays. Skeletal Radiology 49: 601 – 611