Artificial intelligence in orthopaedic service delivery: From referral to follow-up

Andrew Coppola, Andreas Fontalis, Fares S Haddad and Deiary F Kader

Andrew Coppola is an Orthopaedic Specialist Registrar in the Wessex Deanery and an AI Research Fellow at the South West London Elective Orthopaedic Centre, Epsom. He is a PhD candidate at University College London, researching artificial intelligence in orthopaedics. He holds an MSc in Surgical Innovation from Imperial College London, focused on predictive AI in hip and knee replacement, and was awarded an NHS Topol Digital Fellowship. He serves on the Operational Board of the OrthoAI Alliance (ORAIA).

Andreas Fontalis is a Specialist Registrar in Trauma and Orthopaedics on the Percivall Pott Rotation in London and Honorary Associate Professor at University College London. His doctoral and subsequent research has focused on enhanced technologies in hip and knee arthroplasty, including randomised trials of robotic surgery and machine-learning applications to outcome prediction. He serves on the AAHKS Digital Health and Social Media Committee and the Operational Board of the OrthoAI Alliance (ORAIA).

Artificial intelligence (AI) is likely to influence orthopaedics by reshaping clinical pathways. From referral and triage to patient selection and follow-up, it may support more responsive, personalised and efficient care if implemented with clinical leadership, clear governance and patient-centred decision-making.

Orthopaedic practice is entering another period of technological change. Previous advances have often centred on implants, fixation, arthroscopy, navigation, robotics and perioperative care. The next phase may be less visible in theatre and more closely related to how patients move through the system, including how they are referred, assessed, selected for treatment, monitored while waiting and followed up after intervention.

At the end of March 2026, 7.1 million referral-to-treatment pathways, representing six million people, were waiting to start treatment in England[1]. Musculoskeletal disease is common, chronic and associated with pain, disability and loss of independence[2]. Delays matter because among patients awaiting arthroplasty, poorer function and psychological distress are well described, with some reporting states considered 'worse than death'[3,4]. AI should therefore be considered as an enabler of more responsive service delivery.

The term AI covers several overlapping technologies. Natural language processing (NLP) and large language models (LLMs) can support referral analysis, clinical summaries, documentation and patient communication. Retrieval-augmented generation (RAG) can link language models to local guidance and evidence. Computer vision can assist with radiograph interpretation, wound assessment and implant surveillance. Predictive analytics can estimate outcomes, complications, length of stay and deterioration risk. They can be combined in workflows, but have different data requirements, risks and validation needs.

The aim is not to replace clinical judgement, but to redesign pathways around better information, earlier risk identification and more personalised allocation of clinical attention. The sections below follow the patient journey, highlighting current applications and evidence gaps. This framework is summarised in Figure 1, and the main applications, benefits and safeguards are outlined in Table 1.

1. Referral and triage

The orthopaedic pathway commonly begins with referral from primary care, emergency departments, community musculoskeletal services or other specialties. Referral quality is variable, with information spread across free-

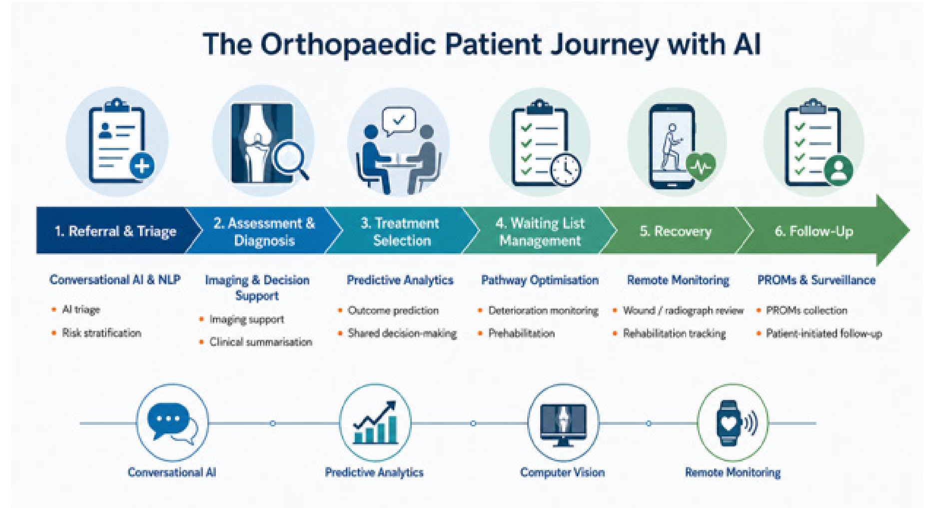

Figure 1: Potential applications of AI across the orthopaedic patient journey, from referral and triage through assessment, treatment selection, waiting list management, recovery and long-term follow-up.

Fares Haddad is Professor of Orthopaedic and Sports Surgery and Divisional Clinical Director of Surgical Specialties at UCLH, and Director of the Institute of Sport, Exercise and Health (ISEH) at University College London. He is Medical Director for Orthopaedics at HCA Healthcare UK, Editor in Chief of the Bone and Joint Journal and President of the OrthoAI Alliance (ORAIA).

Deiary Kader is a Consultant Orthopaedic Surgeon specialising in knee surgery at the South West London Elective Orthopaedic Centre, SWLEOC, where he serves as Director of Research. He is committed to integrating AI in orthopaedics and is currently Vice President of the OrthoAI Alliance, ORAIA. He is Founder of NGMV and the Life Academy charity, and previously served as a war trauma reconstruction surgeon with the ICRC in Lebanon and Swisscross in Kurdistan, Iraq.

text letters, imaging reports, physiotherapy notes, medication records and previous correspondence. This creates triage burden and may contribute to variation in access and pathway allocation.

NLP and conversational interfaces could summarise symptoms, prior treatment, imaging, comorbidities and red flags, while identifying missing information such as radiographs or documentation of conservative management[5]. The ARCHERY project is a useful cautionary example from arthroplasty. It applied NLP to referral letters, radiology reports and clinic letters to predict selection for total hip replacement, but performance was limited and deteriorated on external validation6. This negative result is valuable because it shows that orthopaedic referral data are heterogeneous and that surgical selection is difficult to infer from free text alone.

Outside orthopaedics, NICE has conditionally recommended DERM for urgent suspected skin cancer triage, with early evidence suggesting fewer dermatologist referrals[7]. These examples suggest AI-assisted routing is most plausible when the clinical task, escalation route and audit standards are explicitly defined.

2. Assessment and diagnosis

The outpatient consultation integrates symptoms, imaging, function, expectations, comorbidity and patient preference into a management plan. It also carries administrative work, including record review, documentation, shared decision-making and communication with primary care.

AI may reduce this burden through pre-consultation summaries, structured extraction of comorbidities and ambient scribing that produces draft clinic notes or letters[8]. Systems such as Nuance DAX, Heidi and Suki illustrate the direction of travel, although evidence in routine UK orthopaedic clinics remains limited. Clinical value will depend on whether documentation is accurate, editable and genuinely reduces work rather than adding another task.

Orthopaedics is also image rich. NICE early value assessment guidance supports selected AI fracture-detection tools, including TechCare Alert, BoneView, RBfracture and Rayvolve, in urgent care while further evidence is generated[9]. Similar computer vision approaches may support osteoarthritis grading, implant surveillance and deformity analysis[5]. Imaging findings must still be interpreted in context, as severe radiographic osteoarthritis may be minimally symptomatic while disabling symptoms may occur with modest radiographic change.

3. Treatment selection

Patient selection is one of the most complex aspects of orthopaedic practice. In arthroplasty, the decision to operate requires integration of pain, function, radiographic severity, comorbidity, frailty, expectations and willingness to undergo rehabilitation. In sports surgery, predicted return to activity, graft failure risk and patient goals may be more relevant. In spinal surgery, symptom pattern, imaging correlation, neurological deficit and psychosocial factors are central. In trauma reconstruction, decisions may depend on soft-tissue status, fracture pattern, physiological reserve and the risk of complications. These examples illustrate why patient selection is rarely reducible to a single variable or guideline threshold. Machine learning models may assist by estimating outcomes such as symptom improvement, complications, length of stay, revision, dissatisfaction or return to activity.

Treatment selection is where predictive analytics and personalised data-driven practice become most relevant. A recent study of 5,720 patients with knee osteoarthritis found that models using clinical data and imaging showed good >>

Pathway stage	AI technology	Example application	Main benefit	Key risk / safeguard
Referral & Triage	NLP, LLMs, RAG	Referral summaries, red flags, pathway guidance	Faster triage	Missed urgency, audit required
Assessment & Diagnosis	Computer vision, NLP	Fracture detection, OA grading, note drafting	Less diagnostic/documentation burden	Automation bias, external validation needed
Treatment Selection	Predictive analytics / ML	PROMs, complications, LOS	Personalised counselling and PDDP	Poor calibration, centre-level validation needed
Waiting List Management	Predictive analytics, LLMs	Deterioration flags, optimisation prompts	Active pathway management	Rationing risk, transparent prioritisation needed
Recovery	Remote monitoring, computer vision	Wounds, pain scores, rehabilitation progress	Earlier complication detection	Digital exclusion, escalation routes needed
Follow-Up	PROM analytics, remote monitoring, NLP	PIFU support, implant surveillance	Fewer low-value appointments	Safety-net failure, maintain access

AI (artificial intelligence); NLP (natural language processing); LLMs (large language models); RAG (retrieval-augmented generation); ML (machine learning); OA (osteoarthritis); PROMs (patient-reported outcome measures); LOS (length of stay); PDDP (personalised data-driven practice) and PIFU (patient-initiated follow-up).

Table 1: AI applications across the orthopaedic pathway, including opportunities and safeguards.

discriminative performance for predicting dissatisfaction after total knee replacement, supporting a potential role in counselling while highlighting the need for external validation[10]. PDDP describes the use of multimodal data – including records, imaging, PROMs and wearable data – to generate patient-specific insights that complement evidence-based medicine[11]. Evidence remains mixed: a systematic review found that AI did not consistently outperform orthopaedic surgeons, reinforcing the need for careful evaluation rather than assumption of benefit[12].

Prediction should not be conflated with judgement. Lower predicted benefit may reflect modifiable risk factors, psychological distress, social deprivation, frailty or comorbidity. Predictive analytics should improve preparation and personalise care, rather than exclude patients from treatment. The clinical question is not simply whether a patient is suitable for surgery, but what support is needed to maximise the chance of a good outcome.

4. Waiting list management

For many orthopaedic patients, waiting is not neutral. Pain may worsen, mobility may decline and patients may decondition. Comorbidities may also change, affecting perioperative risk and recovery. This is particularly important for patients already in severe pain or functional limitation[3,4].

AI could help convert passive waiting lists into more active clinical pathways. Systems could combine PROMs, comorbidity markers, prescribing data, healthcare contacts and attendance patterns to identify deterioration, optimisation needs, changed symptoms or patients who no longer need surgery. This could support validation and prioritisation, but needs governance to avoid rationing.

AI may also improve waiting-list efficiency by addressing operational causes of lost capacity. For example, NHS England reported that AI-supported appointment non-attendance prediction reduced did-not-attend appointments by almost one third in a Mid and South Essex pilot, preventing 377 missed appointments and enabling 1,910 additional patients to be seen[13]. Although this is not orthopaedic clinical prioritisation, it demonstrates how AI can support pathway performance when applied to a clearly defined service problem.

Clinically, a patient awaiting knee arthroplasty with escalating opioid use, repeated healthcare contacts and declining mobility may benefit from earlier review. Conversely, improvement after physiotherapy, injection or weight loss may justify reassessment before surgery.

There is, however, an important distinction between prioritisation and rationing. Prioritisation identifies clinical need, deterioration and modifiable risk, whereas rationing uses prediction to restrict access. Prioritisation should account for pain, disability, work, caring responsibilities, frailty, deprivation and patient preference, not simply predicted utilisation or administrative convenience.

5. Recovery

Recovery makes pathway redesign visible to patients. Current follow-up often relies on scheduled appointments, radiographs and patient-initiated contact. This familiar model may miss deterioration between appointments while generating low-yield reviews for patients recovering well.

AI-enabled monitoring could make surveillance more continuous by combining PROMs, pain scores, wound images, mobility data and red flags. The value would be earlier escalation for wound leakage, increasing pain or stalled rehabilitation, and reassurance for those following an expected trajectory.

A multimodal model using postoperative wound images and PROMs performed comparably to clinician triage in gastrointestinal surgery and simulated an 80% reduction in staff time[14]. Although not orthopaedic, it illustrates how image-based monitoring could translate to orthopaedic wound surveillance if validated in relevant pathways.

6. Follow-up

Follow-up concerns longer-term access, surveillance and reassurance after the initial surgical episode. Patient-initiated follow-up already moves routine review towards need-based access[15]. In a multicentre randomised trial after hip and knee replacement, electronic PROM monitoring with automated alerts produced small improvements across several health dimensions[16]. This is not AI in itself, but provides the digital substrate for AI-enabled pattern detection and escalation.

Remote follow-up must address equity and safety. Digital exclusion, language barriers, health literacy and access to technology affect participation[17]. Some patients may under-report symptoms or misinterpret questions. AI-supported follow-up should therefore sit alongside accessible routes back to clinical review, so that patients most in need of care are not least visible to the technology.

Governance and implementation

AI benefits should not obscure implementation challenges. Orthopaedic data are often messy. Routine datasets may contain missing PROMs, inconsistent coding, variable imaging quality and incomplete follow-up. Referral information is often unstructured, while treatment decisions may reflect local practice, surgeon thresholds, implant availability and rehabilitation pathways. These limitations directly affect model performance, fairness and transferability.

Validation must go beyond headline accuracy. A model may separate higher- from lower-risk groups yet give unreliable individual probabilities. For counselling, prioritisation or follow-up, calibration, external validation and prospective evaluation are essential before AI tools are embedded in care pathways[18,19]. A model developed in one centre may not transfer safely to another with different case mix, referral patterns, coding practices, surgical techniques or outcome collection.

Tools that influence access to care should be auditable, explainable and monitored for unintended consequences, including delayed treatment, inequitable prioritisation or inappropriate exclusion from specialist review. Evaluation should include safety, usability, workflow burden, equity, cost-effectiveness and patient outcomes.

Implementation should be clinically led and tied to defined service problems, not enthusiasm for innovation alone. The NHS AI Lab Ethics Initiative supports ethical adoption of AI-driven technologies, and the NICE Evidence Standards Framework sets evidence expectations for digital health technologies[20,21]. Specialty-led collaboration is also essential. Networks such as the Orthopaedic Artificial Intelligence Alliance (ORAIA) can help clinicians, researchers, data scientists and patients define relevant questions, support multicentre validation and promote responsible implementation[22].

Conclusion

Some applications are closest to practice now, including ambient documentation, AI-assisted fracture detection, PROM collection and PIFU-style monitoring. More contentious uses – including referral triage, treatment selection and waiting-list prioritisation – require stronger evidence, calibration and external validation before routine adoption.

Research should move beyond proof-of-concept models towards multicentre studies using real-world orthopaedic data. These studies should report negative as well as positive findings, examine equity, and quantify whether AI improves workflow rather than adding another layer of work. The priority is not AI for its own sake, but pathway redesign that makes orthopaedic care safer, more timely and more personal. ■

References

References can be found online at www.boa.ac.uk/publications/JTO.