NICE National Institute for Health and Care Excellence

Evidence generation plan for digital technologies for managing mild to moderate symptoms of hip or knee osteoarthritis: early value assessment

Implementation support

Published: 22 January 2026

www.nice.org.uk

Contents

1 Purpose of this document 3

2 Evidence gaps 4

2.1 Essential evidence for future committee decision making 4

2.2 Evidence that further supports committee decision making 5

3 Approach to evidence generation 7

3.1 Evidence gaps and ongoing studies 7

3.2 Data sources 8

3.3 Evidence collection plan 8

3.4 Data to be collected 9

3.5 Evidence generation period 11

3.6 Following best practice in study methodology 11

4 Monitoring 13

5 Minimum evidence standards 14

6 Implementation considerations 15

Evidence generation 15

Equalities 15

Adverse events 16

1 Purpose of this document

NICE's early value assessment of digital technologies for managing mild to moderate symptoms of hip or knee osteoarthritis recommends that more evidence is generated while the following technologies are being used in the NHS:

- getUBetter
- Good Boost
- Hinge Health
- Joint Academy
- Phio Engage
- re.flex
- Sword Thrive
- TrackActive Me.

Pathway Through Arthritis can only be used in research and is not covered in this plan.

This plan outlines the evidence gaps and what data needs to be collected for a NICE review of the technologies again in the future. It is not a study protocol but suggests an approach to generating the information needed to address the evidence gaps. For assessing comparative treatment effects, well-conducted randomised controlled trials are the preferred source of evidence. Evidence generated through other study approaches will also be considered.

The companies are responsible for ensuring that data collection and analysis takes place.

NICE will withdraw the guidance if the companies do not meet the conditions in section 4 on monitoring.

After the end of the evidence generation period (3 years), the companies should submit the evidence to NICE in a format that can be used for decision making. NICE will review all the evidence and assess if the technologies can be routinely adopted in the NHS.

2 Evidence gaps

This section describes the evidence gaps, why they need to be addressed and their relative importance for future committee decision making.

The committee will not be able to make a positive recommendation without the essential evidence gaps (see section 2.1) being addressed. The companies can strengthen the evidence base by also addressing as many other evidence gaps (see section 2.2) as possible. This will help the committee to make a recommendation by ensuring it has a better understanding of the patient or healthcare system benefits of the technologies.

2.1 Essential evidence for future committee decision making

Clinical effectiveness

The impact of the technologies on intermediate and longer-term clinical outcomes in comparison with conventional management of osteoarthritis is uncertain. The effect on clinical effectiveness of changing from paper-based to digital programmes is also uncertain.

Evidence on intermediate clinical outcomes should include:

- health-related quality of life (HRQoL) and patient-reported outcomes (including pain and stiffness, activity impairment, physical function, self-efficacy and psychological outcomes)
- referrals for corticosteroid injections
- medication use
- adverse events.

More evidence is needed on the longer-term clinical effects (for example, referrals to specialist services and clinic visits). Information on the longevity of any clinical benefits will provide a clearer indication of the accumulated benefits over time and support cost-

effectiveness modelling. Follow ups should include the duration of the self-management programme (which is technology specific) and be at least 12 months (ideally 18 months) later.

Resource and service impact

Early cost-effectiveness modelling was driven by the impact on HRQoL of digital technologies compared with standard care, the digital technology costs, the reduction in physiotherapy and GP resource use, and the duration of the impact of digital technology on resource use. In addition to evidence on the health improvements offered by the technology, further evidence is needed on resource and service use when using the technologies compared with conventional management pathways. This should include overall costs, and the broader resource impact that osteoarthritis management has on the healthcare system over at least 12 months, and ideally 18 months after using the technology.

Key areas that will help to address this evidence gap are:

- the number and cost of clinical contacts, visits and referrals to specialist services
- the costs for implementing, integrating, maintaining and using the digital technologies.

User engagement and acceptability

More evidence on intervention uptake, adherence, satisfaction, acceptability, completion and attrition rates (including reasons for stopping treatment) help the committee assess the real-world uptake of the technologies and support future cost-effectiveness modelling.

2.2 Evidence that further supports committee decision making

Clinical effectiveness and resource use in different subgroups

The impact of the technologies on clinical effectiveness and resource and service use in different subgroups is unknown. These subgroups may include:

- people who have osteoarthritis that affects different parts (hip or knee) or different

musculoskeletal comorbidities

- people who may not be able to attend daytime in-person physiotherapy sessions (for example, people with work or caring responsibilities or people living in rural communities with long travel times to clinics)
- people who stop using the technology
- people who access the technology through different referral pathways.

3 Approach to evidence generation

3.1 Evidence gaps and ongoing studies

There are 2 ongoing studies evaluating Phio Engage that may contribute some clinical-effectiveness data. These are due to end in December 2025.

Table 1 summarises the evidence gaps and ongoing studies that might address them. Information about evidence status is derived from the external assessment group's report; evidence not meeting the scope and inclusion criteria is not included. The table shows the evidence available to the committee when the guidance was published.

Table 1 Evidence gaps and ongoing studies

Evidence gap	Clinical effectiveness	Resource use and service impact	User engagement and acceptability	Clinical effectiveness and resource use in different subgroups
getUBetter	Limited evidence	No evidence	No evidence	No evidence
Good Boost	Limited evidence	No evidence	No evidence	No evidence
Hinge Health	Limited evidence	No evidence	Limited evidence	No evidence
Joint Academy	Limited evidence	Limited evidence	Limited evidence	No evidence
Phio Engage	Limited evidence Ongoing studies	No evidence	No evidence	No evidence
re.flex	Limited evidence	No evidence	Limited evidence	No evidence
Sword Thrive	Limited evidence	No evidence	No evidence	No evidence

Evidence gap	Clinical effectiveness	Resource use and service impact	User engagement and acceptability	Clinical effectiveness and resource use in different subgroups
TrackActive Me	Limited evidence	No evidence	No evidence	No evidence

3.2 Data sources

NICE's real-world evidence framework provides detailed guidance on assessing the suitability of a real-world data source to answer a specific research question. The NHS England Secure Data Environment (SDE) service could potentially support this research. This platform provides access to high-standard NHS health and social care data that can be used for research and analysis. SDEs are data storage and access platforms that bring together many sources of data, such as from primary and secondary care, to enable research and analysis. They could be used to collect data to address the evidence gaps. Subnational SDEs are designed to be agile and can be modified to suit the needs of new projects. Within an SDE, the data can be linked to other useful data, such as that from primary care, and could provide information on important confounders (for example, comorbidities).

The Osteoarthritis Initiative (OAI) could potentially provide additional information for comparison. The quality and coverage of real-world data collections are of key importance when used in generating evidence. Active monitoring and follow up through a central coordinating point is an effective and viable approach of ensuring good-quality data with broad coverage.

3.3 Evidence collection plan

NICE suggests a mixed-methods approach to address the identified evidence gaps: a prospective real-world comparative cohort study combined with a qualitative survey. The qualitative component should explore user experience, engagement and barriers to access in more depth.

Data could be collected through a combination of:

- primary data (for example, outcome measures and surveys)
- data generated through the technology itself (for example, engagement metrics and session completion)

- routinely collected real-world data sources (for example, Clinical Practice Research Datalink [CPRD] and Hospital Episode Statistics [HES]).

Prospective real-world comparative cohort study

In this type of study, data should be collected from healthcare services where the digital technology is offered and compared with services where it is not. People in both groups should be followed from the point at which they would typically be offered the technology. The comparison group should include people from similar services with comparable patient populations and standard care pathways but without access to the digital technology. Ideally, the study should be conducted across multiple centres to reflect the diversity of the NHS service provision. Non-random assignment to interventions introduces a risk of confounding bias. So, appropriate methods, such as matching or adjustment (for example, propensity score methods), should be used to minimise selection bias and balance confounding factors between groups. High-quality data on patient characteristics will be essential to support these methods. The identification of key confounders should be informed by expert input during protocol development.

Qualitative survey

Qualitative data could be generated through appropriate methods, such as surveys, focus groups or interviews. This should include reported outcomes (acceptability, usability and preferences) from people using the technologies. The robustness of the findings will depend on:

- broad and inclusive distribution across eligible users
- the sample of respondents being representative of the population of potential users.

3.4 Data to be collected

Data collection should follow a predefined protocol, and quality assurance processes should be put in place to ensure the integrity and consistency of data collection. NICE's real-world evidence framework provides guidance on the planning, conduct and reporting of real-world evidence studies.

Study criteria

At recruitment, eligibility criteria for the suitability of using the digital technologies and inclusion in the real-world study should be reported. Detailed descriptions of the technologies should include their training requirements, digital-safety assurance and the specific products and versions, as well any optional features of the products that are being used (including any artificial intelligence add-ons for program selection).

Service-user characteristics and clinical outcomes

- Information about individual characteristics at baseline, for example, sex, age, ethnicity, first language, medicines, diagnosis and affected body area, comorbidities, socioeconomic status and location, with other important covariates chosen with input from clinical specialists. Characteristics should include those needed for adjustment to address confounding and for subgroup analysis.
- Measures recorded at baseline and follow up (at least 12 months, ideally 18 months), of:
  - health-related quality of life (EQ-5D)
  - patient-reported outcomes, including pain and stiffness, psychological outcomes, physical function, activity impairment and self-efficacy (for example, Musculoskeletal Health Questionnaire, Numeric Pain Rating Scale (0 to 10) for the target joint, Patient Specific Functional Scale, Work Ability Index or Arthritis Self-Efficacy Scale).
- Number of referrals for corticosteroid injections.
- Adverse events (including reporting detail around whether or not they are intervention related).

Engagement and acceptability

- Usability, satisfaction and acceptability of the technologies.
- Intervention adherence, uptake, completion and attrition rates (including reasons for not using the technology).

Resource and system

- Number and cost of face-to-face physiotherapy sessions (and details about profession and banding of staff leading or supporting the sessions).
- Number and cost of appointments in primary, secondary and community care.
- Medication use.
- Any additional interactions with healthcare professionals outside of appointments (for example, time to review data provided by the technologies, where relevant).
- Referrals to secondary care and other specialist services.
- Costs of digital technologies for supporting management of osteoarthritis, including:
  - licence and maintenance fees with subscription duration
  - healthcare professional staff time and training costs to support the service
  - integration with digital NHS systems
  - implementation costs
  - other technology costs (including additional hardware or software).

It is also important to report and specify if any optional features of the technologies are being used during evidence generation.

3.5 Evidence generation period

This will be 3 years to allow for setting up, implementing the test, data collection, analysis and reporting.

3.6 Following best practice in study methodology

Following best practice in conducting studies is paramount to ensuring the reliability and validity of the research findings. Adhering to rigorous guidelines and established standards is crucial for generating credible evidence that can ultimately improve patient care. NICE's real-world evidence framework details some key considerations.

For early value assessments, a key factor to consider as part of the informed consent process is to ensure that patients (and their carers, as appropriate) understand that data will be collected to address the evidence gaps identified in section 2. Where applicable, this should take account of NICE's guidance about shared decision making.

4 Monitoring

NICE will contact the companies:

- within 6 months of publication of this plan to confirm agreements are in place to generate the evidence
- annually to confirm that the data is being collected and analysed as planned.

The companies should tell NICE as soon as possible of anything that may affect ongoing evidence generation, including:

- any substantial risk that the evidence will not be collected as planned
- new safety concerns
- the technology significantly changing in a way that affects the evidence generation process.

If data collection is expected to end later than planned, the companies should contact NICE to arrange an extension to the evidence generation period. NICE reserves the right to withdraw the guidance if data collection is delayed, or if it is unlikely to resolve the evidence gaps.

5 Minimum evidence standards

During the evidence generation period new technologies may become available. This section summarises the minimum evidence requirements that a new technology would need to meet to be considered in NICE's evaluation after the evidence generation period.

The committee noted that the digital technologies have comparable clinical effectiveness for managing mild to moderate symptoms of hip or knee osteoarthritis in comparison with conventional management of these conditions, but evidence for this is uncertain. It heard that the technologies may increase access to treatment in some populations and reduce the demand for in-person GP or physiotherapy appointments, but evidence for this was also limited.

For new technologies, the committee has indicated that it may, in future, be able to recommend technologies in this topic area that have UK-based evidence for:

- non-inferiority of the digital platforms compared with conventional management of the condition in terms of clinical effectiveness
- user engagement with the technology, including intervention acceptance, usability and completion rates
- cost savings resulting from resource use associated with the technologies.

Companies can strengthen the evidence base by also having evidence for uptake rates in different subpopulations.

6 Implementation considerations

The following considerations around implementing the evidence generation process have been identified through working with system partners.

Evidence generation

- Users can access the digital technologies through a variety of referral pathways, including self-referral without medical oversight. It is important to capture the referral pathway of the study participants. This will ensure like-for-like comparisons of various populations to enable future subgroup analyses.
- Conventional management for mild to moderate symptoms of hip or knee osteoarthritis may be offered by various healthcare professionals and may involve multiple allied healthcare professionals and various medical interventions. Details around conventional management should be reported during evidence generation to minimise the risk of bias and accurately assess costs for cost comparisons.

Equalities

- Face-to-face management sessions for mild to moderate symptoms of hip or knee osteoarthritis should be available for people with conditions that are not indicated for use with the digital technologies or who decline using the technologies.
- People who are eligible to use a technology in line with its intended use but who are excluded for any other reason should be described in the reporting of future evidence.
- There is a risk that using digital technologies could widen the gap in access to managing these conditions. Support and resources may be needed for:
  - people who are less familiar with using digital technologies or have limited access to equipment or the internet
  - neurodivergent people
  - people with learning disabilities
  - people with visual, hearing or cognitive impairments

- people who have problems with manual dexterity
- people who have difficulties reading, writing or understanding health-related information (including people who cannot read English).
- people who are experiencing homelessness
- people who are living in a multiple occupancy household
- people who are living in residential care.

Adverse events

- Reporting intervention-related adverse events (for example, worsening pain) is essential to assess any risk associated with the use of the technologies in the NHS.

ISBN: 978-1-4731-8159-5