National Institute for Health and Care Excellence

HealthTech programme

GID-HTE10043: Robot-assisted surgery for orthopaedic procedures

Draft guidance – comments

There were 60 comments from 9 consultees:

- 26 comments from 3 company representatives
- 23 comments from 3 specialist organisations
- 11 comments from 3 unaffiliated consultees

The following themes were identified, and the comments arranged as follows:

- The recommendations, comments 1-5
- Approach to the evidence review, comments 6-17
- Economic evidence, comments 18-21
- Equality considerations, comments 22-26
- Further evidence generation, comments 27-29
- Clarity of the wording, comments 30-34
- General, comments 35-37
- Evidence generation plan, comments 38-60

Comment number	Consultee type	Section number	Comment	Response
Theme 1 – the recommendations				
1	Consultee 3 Stryker	Consultation question - Are the recommendations sound and a suitable basis for guidance to the NHS?	1. Page 3, section 1.1: The recommendations state that these technologies can only be used: “if the evidence outlined in the evidence generation plan is being generated” and “once they have appropriate regulatory approval including NHS England's Digital Technology Assessment Criteria (DTAC) approval”. We are currently engaging with the DTAC process, but more clarity will be required on how the DTAC process will be implemented for Robotic Surgery. In its current format it is not entirely applicable as it designed for fully digital technologies. B) Further, it should be noted that additional procedures are likely to become available on robotic platforms in the future. This should be considered in any final recommendations made.	Thank you for your comment. Further information regarding the DTAC process is available through NHS England. The recommendations in the guidance document are intended to prompt this process. Additional indications may be considered in future assessments when evidence becomes available. The recommendations in this EVA are based on those where evidence is currently available.
2	Consultee 5 Johnson and Johnson	Consultation question - Are the recommendations sound and a suitable basis for guidance to the NHS?	Despite the methodological limitations identified, we understand the need for further evidence generation and believe that a recommendation for use on that basis is reasonable.	Thank you for your comment.
3	Consultee 7 The British Orthopaedic Association		The British Orthopaedic Association (BOA) welcomes the recommendations outlined in the Early Value Assessment document for robot-assisted surgery in orthopaedic procedures. The BOA supports the proposed approach to continue the use of this innovative technology, whilst further evaluating its impact through a full assessment.	Thank you for your comment.

Comment number	Consultee type	Section number	Comment	Response
4	Consultee 7 The British Orthopaedic Association		The BOA fully supports the recommendations for continued use and ongoing evaluation of robot-assisted surgery in orthopaedics. The BOA is committed to working alongside policymakers, researchers, and healthcare providers to ensure these technologies deliver meaningful benefits for patients while maintaining the highest standards of value-based care.	Thank you for your comment.
5	Consultee 9 MicroPort Orthopedics		The early value assessment document recommends that Skywalker be used in research setting only which seems to be mainly based on the lack of evidence and information on the system. The supporting document also stated that the NICE committee was not able to obtain information on the Skywalker. We would like to request the opportunity to provide all evidence available on the Skywalker system to ensure the early value assessment reflects the latest clinical use and evidence of the Skywalker system.	Thank you for your comment. The evidence submitted has been described in an addendum to the EAR that was presented to the committee at the second Medical Technologies Appraisal Committee meeting. The committee decided to change the recommendation for the SkyWalker technology to a recommendation for further evidence generation in the NHS.
Theme 2 – approach to evidence review				
6	Consultee 3 Stryker	Consultation question - Has all of the relevant evidence been taken into account?	1. Page 14, section 3.8: As previously stated, we feel that the scope of evidence considered was too narrow. There were pieces of evidence that were excluded such as RWE from other countries such as the US. Evidence from the AOANJRR was also discredited. Additionally, the selected evidence was restricted to those published in the last 5 years. This may limit the inclusion of early / foundational research that would not have been repeated since that time. 2. The report states, “The committee agreed that robot-assisted surgery broadly showed non-inferiority with conventional surgery in primary outcomes. These included length of hospital stay, complications, patient-reported outcome measures (PROMs), utilities and surgical revisions. The committee noted that alignment, which was a secondary outcome, was consistently better with robot-assisted surgery. But the evidence did not suggest that this resulted in better PROMs or clinical outcomes.” We are not in agreement that the broad statement above should be applied to all robotic systems. Mako has evidence demonstrating	Thank you for your comment. 1. Within the published protocol the EAG acknowledged a large body of evidence for robot-assisted surgery in orthopaedics. It defined a publication date range and evidence prioritisation hierarchy to identify the evidence most relevant to the decision problem. Older evidence (including older generation robotics, older software) may

Comment number	Consultee type	Section number	Comment	Response
			improvements in all the parameters noted however the data was discredited from the assessment. 3. Page 15, section 3.9: The report states, “Most PROMs showed no difference between robot-assisted surgery and conventional surgery. When statistically significant differences were seen, the benefit tended to be from robot-assisted surgery. But, the committee noted that many of these differences were below the minimally clinically important difference…” The committee recommends additional PROMs collection on the basis that the evidence in their scope showed no significant difference (or that which did not reach MCID). The recommendation for further PROMS collection does not acknowledge that most of the available PROMs have ceiling effects and may not be sensitive enough to detect appropriate / meaningful difference in patient outcomes. Please see a recent editorial (https://boneandjoint.org.uk/article/10.1302/0301-620X.106B10.BJJ-2024-0876). 4. Page 16, section 3.10: This section states “A clinical expert highlighted that the Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) contains separate revision data for robot-assisted surgery and conventional surgery. But this showed no statistically significant difference between surgical methods.” However, while a 2023 analysis of the AOANJRR found no difference between revision rates of robotic surgery and non-robotic assisted TKA when adjusting for age, gender, ASA, BMI, bearing surface, patella component usage and stability, a 2024 AOANJRR analysis (https://aoanjrr.sahmri.com/annual-reports-2024) found significant improvements in Stryker’s Triathlon implant when used for TKA implanted using Mako. A) Triathlon CR (with and without patella) shows a significantly lower six-year CRR when implanted with Mako compared to manual: 2.1 vs 2.6. This is a 19% relative improvement. B) Triathlon CR with patella shows a significantly lower six-year CRR when implanted with Mako compared to manual: 1.6 vs 2.3. This is a 30% relative improvement).	not be fully generalisable to the current UK NHS. A pragmatic approach to EVA is permitted as per the interim process and methods (PMG39). 2. Within the EAG report (section 5.10) it is highlighted that Mako had the highest quality evidence in TKA and that from the randomised evidence conducted in the UK (as outlined in the results extracted for Mako summarised in section 5.7 of the EAG report) no significant difference in length of hospital, complications, PROMs, utilities or revisions were found between RAS and conventional surgery. No statistical different in range of motion at 5 years (secondary outcome) was reported from the randomised UK evidence. It is unclear from the consultation comment which additional randomised evidence was excluded by the EAG which would be deemed higher quality than those summarised in the EAG report that would counter these findings.

Comment number	Consultee type	Section number	Comment	Response
				3. Thank you for sharing the editorial from Clement et al. 2024 (Bone Joint J, 1033-1035) which discussed whether the current minimal clinically important differences (MCIDs) are fit for purpose. The mean difference between surgical methods and MCIDs were not consistently reported across studies and outcomes, introducing uncertainty that means the EAG cannot rule out important differences existing. The Committee acknowledged the potential ceiling effect and discussed alternative PROMs. It broadly recommended further evidence generation to reduce uncertainties associated with PROMs data currently collected in the NJR, which were deemed to be the most appropriate measures to inform future decision making. Collection of more data through linkage of the NJR with NHSE's PROMs data could address ceiling effects and reduce uncertainty through obtaining more precise point estimates and narrower confidence intervals.

Comment number	Consultee type	Section number	Comment	Response
				4. Evidence from the AOANJRR is acknowledged in the EAG report (section 5.9) and the draft guidance. The EAG and clinical experts provided reasons why available data from other settings may not be generalisable to the NHS. These included differences in age and American Society of Anesthesiologists risk score between populations, and between the UK and Australian healthcare systems. As stated in the EAR, no statistical difference in revision rate was found for total or unicompartmental knee arthroplasty between RAS and conventional methods when adjusting for other patient baseline characteristics. A & B. Thank you for highlighting the AOANJRR analysis in their 2024 report which is specific to Stryker's Triathlon implant. This found that cumulative revision rate with RAS versus conventional surgery (using the Triathlon CR/Triathlon) was statistically different across the entire period (from 2016 onwards) with a hazard ratio of 0.86

Comment number	Consultee type	Section number	Comment	Response
				(95%CI 0.75 to 0.99); p=0.033 when adjusting for age, gender, ASA, BMI, patella component usage and tibial fixation. However, as stated above differences in patient characteristics and orthopaedic practice between UK and Australia, mean that the generalisability of these results is uncertain
7	Consultee 4	Consultation question - Has all of the relevant evidence been taken into account?	No. The 5 year cut off has excluded a large volume of literature. The omission of papers due to the specific robot not being obvious limits the interpretation of the findings. The exclusion of meta-analysis and systematic reviews (of current robotic systems) limits the analysis.	Thank you for your comment. As outlined in section 4.1 of the EAG report: systematic reviews were included and restricted to those exclusively including technologies in scope only and reporting meta-analysis of primary outcomes (as listed in the final scope). It is unclear from the consultation comment which specific systematic review or meta-analysis has been excluded, and whether it was in scope of the decision problem of this EVA. Within the published protocol the EAG acknowledged a large body of evidence for robot-assisted surgery in orthopaedics and outlined time limits (date of publication) and evidence hierarchy to include

Comment number	Consultee type	Section number	Comment	Response
				evidence that was of best quality and most pertinent to the objective to the decision problem. Older evidence (including older generation robotics, older software) may not be fully generalisable to the current UK NHS. A pragmatic approach to EVA is permitted as per the interim process and methods (PMG39).
8	Consultee 4	Consultation question - Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	No They have been limited by the literature review as all the evidence has not been considered in the modelling.	See response to **comment 7** above.
9	Consultee 5 Johnson and Johnson	Consultation question - Has all of the relevant evidence been taken into account?	No. As per our response to the EAG, there are 101 studies that were defined as in-scope by the EAG but have been disregarded; we feel that in order to complete a comprehensive assessment of these novel medical technologies, these studies should have been evaluated more thoroughly. We believe that time and resource constraints should not be prioritized over quality when it comes to NICE assessments, and are therefore not in alignment with the EAG's position that "due to the size of the evidence base and time/resource constraints the EAG focused on highest quality evidence (prioritising UK, prospective designs with the largest sample size) and primary outcomes." As stated on the NICE website an "EVA is for promising medical technologies that meet a national unmet need. Technologies suitable for EVA are: in need of further data collection or evidence generation before they can be recommended for use in the NHS". With this in mind, the evidence included in the review of these medical technologies should not be restricted to only the highest quality of evidence, namely RCTs. The purpose of the EVA is to provide an early evaluation on promising technologies that may address an unmet need, but are acknowledged to need more evidence,	Thank you for your comment. Within the published protocol the EAG acknowledged a large body of evidence for robot-assisted surgery in orthopaedics and outlined time limits (date of publication) and evidence hierarchy to include evidence that was of best quality and most pertinent to the objective to the decision problem. A pragmatic approach to EVA is permitted as per the interim process and methods (PMG39). As stated

Comment number	Consultee type	Section number	Comment	Response
			as such, criteria for the inclusion of evidence should be broad, and diverse study designs and sources of evidence should be considered.	in the published protocol, prioritisation of clinical and economic evidence was based on the following criteria (in descending order): 1. Studies conducted in the UK. 2. Studies reporting data for the prioritised outcomes 3. Prospective comparative studies followed by retrospective comparative and non-comparative studies analysing the highest number of patients. 4. The most recent evidence. Older evidence (including older generation robotics, older software) may not be fully generalisable to the current UK NHS. A total of 26 studies were prioritised in the EAG report, including: 8 RCTs, 12 prospective comparative cohorts and 6 retrospective cohorts. Therefore, randomised evidence only represented 30% (8/26) of the

Comment number	Consultee type	Section number	Comment	Response
				key studies for which outcomes were extracted. An additional 101 studies were considered as within scope but not prioritised (using the above evidence hierarchy) and acknowledged within Appendix B2 of the EAG report.
10	Consultee 5 Johnson and Johnson	1.6 What evidence generation and research is needed	We agree that a key benefit of robot-assisted surgery is precise implant positioning and the ability to perform individualised placement that would not be possible without robotic assistance, and that such an approach may improve patient satisfaction and reduce the demand for revision surgery. We therefore question the decision by the EAG to exclude evidence comparing mechanically-aligned manual TKA with alternatively-aligned robotic TKA, when such alternative alignment is consistently achievable only with the use of technology, and is likely, therefore, to be a significant differentiator between the outcomes.	Thank you for your comment. It is unclear which specific excluded evidence this consultation is referring to. However, the EAG notes that the retrospective cohort study by Morrisey (Cureus, 2023; e38872) compared TKA using VELYS with both kinematically aligned (n=66) and with conventional mechanically aligned (n=99) surgery. This study was included in the EAG report (see Table 5, Table 18, Appendix B1). However, within section 5.5.3 of the EAG report, it is stated that no statistical difference in outcomes was found between arms of that study. The EAG note that one study provided academic in confidence by the company was prioritised for this device based on the evidence

Comment number	Consultee type	Section number	Comment	Response
				hierarchy outlined in the published protocol.
11	Consultee 5 Johnson and Johnson	3 Evidence gaps and ongoing studies	A significant body of evidence that did meet the scope and inclusion criteria was also not included.	Thank you for your comment. Within the published protocol the EAG acknowledged a large body of evidence for robot-assisted surgery in orthopaedics and outlined time limits (date of publication) and evidence hierarchy to include evidence that was of best quality and most pertinent to the objective to the decision problem. Older evidence (including older generation robotics, older software) may not be fully generalisable to the current UK NHS. A pragmatic approach to EVA is permitted as per the interim process and methods (PMG39).
12	Consultee 6 The Royal College of Surgeons of England	Consultation question - Has all of the relevant evidence been	No. The scope of the evidence appears to be too narrow. There are some fundamental studies that have been missed out including some RCTs as below. A prospective randomized controlled trial comparing CT-based planning with conventional total hip arthroplasty versus robotic arm-assisted total hip	Thank you for your comment. Within the published protocol the EAG acknowledged a large body of evidence for

Comment number	Consultee type	Section number	Comment	Response
		taken into account?	arthroplasty. Fontalis A, Kayani B, Plastow R, Giebaly DE, Tahmassebi J, Haddad IC, Chambers A, Mancino F, Konan S, Haddad FS. Bone Joint J. 2024 Apr 1;106-B(4):324-335. doi: 10.1302/0301-620X.106B4.BJJ-2023-1045.R1.PMID: 38555946 The LANCET robotic system can improve surgical efficiency in total hip arthroplasty: A prospective randomized, multicenter, parallel-controlled clinical trial. Xu Z, Chai S, Chen D, Wang W, Dai J, Zhang X, Qin J, Song K, Li X, Han J, Chang Q, Zhang M, Xue C, Lu J, Wu L, Yao Y, Li L, Jiang Q. J Orthop Translat. 2024 Apr 5;45:247-255. doi: 10.1016/j.jot.2023.12.004. PMID: 38601198 Comparison of Surgical Time, Short-term Adverse Events, and Implant Placement Accuracy Between Manual, Robot-assisted, and Computer-navigated Total Hip Arthroplasty: A Network Meta-analysis of Randomized Controlled Trials. Kunze KN, Bovonratwet P, Polce EM, Paul K, Sculco PK. J Am Acad Orthop Surg Glob Res Rev. 2022 Apr 1;6(4):e21.00200. doi: 10.5435/JAAOSGlobal-D-21-00200. PMID: 35472191 Robotic trials in arthroplasty surgery. Khatri C, Metcalfe A, Wall P, Underwood M, Haddad FS, Davis ET. Bone Joint J. 2024 Feb 1;106-B(2):114-120. doi: 10.1302/0301-620X.106B2.BJJ-2023-0711.R1. PMID: 38295854 Patients undergoing robotic arm-assisted total knee arthroplasty have a greater improvement in knee-specific pain but not in function. Clement ND, Galloway S, Baron J, Smith K, Weir DJ, Deehan DJ. Bone Joint J. 2024 May 1;106-B(5):450-459. doi: 10.1302/0301-620X.106B5.BJJ-2023-1196.R1. PMID: 38688485 We are also concerned external registry evidence has not been accepted. We do not understand the rationale for restricting accepted evidence to the last five years. Some of the common RAS systems such as Mako have been in existence for 15 years or more so there is a risk that a study limit of five years rejects important foundational research relevant to current systems.	robot-assisted surgery in orthopaedics and outlined time limits (date of publication) and evidence hierarchy to include evidence that was of best quality and most pertinent to the objective to the decision problem. Older evidence (including older generation robotics, older software) may not be fully generalisable to the current UK NHS. A pragmatic approach to EVA is permitted as per the interim process and methods (PMG39). With regards to the specific papers highlighted in the consultation comment: - Fontalis et al. 2024; 324-355: was identified during EAG literature search and excluded at title and abstract sift as it did not explicitly report the robotic system used. This study was also not included in the company RFI, nor was it raised at stakeholder consultation or MTAC1. This represents the only randomised evidence in THA across any of the technologies and does offer

Comment number	Consultee type	Section number	Comment	Response
				additional evidence which would have been prioritised based on its relevance to the UK, study design, and reporting of primary outcome (PROMs) up to 1 year. This study is an UK RCT (n=60 patients, single-centre, multiple surgeons) which was statistically powered to detect restoration of the native horizontal centre of hip rotation (effect size of 0.8 SD 1.0, as determined from pilot study) between robotics (Mako) and conventional total hip arthroplasty. The EAG note that this is considered a secondary outcome in the final scope. The study reported that the robotic arm group was associated with statistically significant superior accuracy in restoring the native horizonal centre of hip rotation (mean absolute error in robotics arm of 1.4 mm, IQR 0.87 to 3.42) and in conventional arm 4.3 mm, IQR 3 to 6.8; p<0.001). Vertical centre of hip rotation was 0.91 (SD 0.73) in robotic and 2.3 (SD

Comment number	Consultee type	Section number	Comment	Response
				1.3) in conventional arm; p<0.001). The study also reported acetabular component orientation, combined offset, and leg length; all of these outcomes (including the primary outcome of which the study was powered) were considered secondary outcomes in the final scope. Total operating theatre time (patient entering to leaving) was longer in robotic arm (85.2 min (SD 28) compared with 69.7 min (SD 28.4); p=0.038); differences with 95% CI between arms were not reported. But, no evidence of a statistical difference in surgical time (skin open to skin closure) between arms (69.4 min (SD 29) with RAS, compared with 60.6 min (SD 29.6) with conventional surgery; p=0.254); differences with 95% CI between arms were not reported. No statistically significant difference in WOMAC, OHS, UCLA activity scale, EQ5D, Total Health, HHS or HOOS were observed at 1 year.

Comment number	Consultee type	Section number	Comment	Response
				However, the authors acknowledged that *"the trial lacked adequate power to identify the minimally clinically important difference in any of the PROMs that were studied"*. The study also reported that no serious adverse events, stem revisions or dislocations occurred in either group. Therefore, this study adds to the evidence base such that the availability of evidence for primary outcomes for Mako THA should be updated to state: ○ PROMs: **GREEN** ○ Complications: **GREEN** ○ Learning curve: **GREEN** ○ Revision: **GREEN** ○ Operating time: **GREEN** The EAG acknowledges that the study is small and not powered to detect differences across these areas. Therefore, the overall conclusions of the EAG remain. This evidence was presented to the

Comment number	Consultee type	Section number	Comment	Response
				committee during the second Medical Technologies Appraisal Committee meeting, but no change in recommendations was made. - Xu et al. 2024; 247-255: was identified during EAG literature search and excluded at the title and abstract sift excluded as the LANCET robotic system was not listed in the NICE Final Scope. - Kunze et al. 2022; e21.00200: was identified during EAG literature search and excluded at full paper review (see row 76 of Appendix B4 in EAG report) as all studies used the ROBODOC/ORTHODOC robotic system which is not available in the UK and not listed in the NICE Final Scope. - Khatri et al. 2024; 114-120: was identified during EAG literature search and excluded at title and abstract sift due to study design: *"This annotation outlines the need to assess these technologies and discusses the design and challenges when*

Comment number	Consultee type	Section number	Comment	Response
				conducting such trials, including surgical workflows, isolating the effect of the operation, blinding, and assessing the learning curve." - Clement et al. 2024; 450-459: was included in the EAG report and prioritised as key evidence (see Appendix B1 of the EAG report). Evidence from the AOANJRR is acknowledged in the EAG report (section 5.9) and the draft guidance. Reasons provided for the limits of generalising data to the NHS included differences in age and American Society of Anesthesiologists risk score between populations, and between the UK and Australian healthcare systems. As stated in the EAG report, no statistical difference in revision rate was found for total or unicompartmental knee arthroplasty between robotic and conventional methods when adjusting for other patient baseline characteristics.

Comment number	Consultee type	Section number	Comment	Response
13	Consultee 6 The Royal College of Surgeons of England	Consultation question - Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	1. There is a danger that all robotic systems have been grouped together for some of the analysis, despite the different levels of evidence associated with each. There is evidence from Mako that there is decreased hospital stay, fewer complications and fewer interventions, for example for instability and total hip arthroplasty. The evidence seems to have been underplayed in an effort to group all the systems together. 2. The NICE medical technologies advisory committee recommends the collection of additional patient-reported outcome measures (PROMs) on the basis that the evidence in scope showed no significant difference (or that which did not reach the level of minimal clinically important difference (MCID)). The recommendation for further PROMS collection does not acknowledge that most of the available PROMs have ceiling effects and may not be sensitive enough to detect appropriate / meaningful difference in patient outcomes. The effect size of arthroplasty interventions is large and increments should not need to be as high as MCID and other advantages such as reduction in complications or revisions will be key. Are the current minimal clinically important differences fit for purpose? Clement ND, Haddad FS. Bone Joint J. 2024 Oct 1;106-B(10):1033-1035. doi: 10.1302/0301-620X.106B10.BJJ-2024-0876. PMID: 39348900 3. The draft guidance states that Australian Orthopaedic Association National Joint Replacement Registry (AOANJRR) showed no statistically significant difference in revisions between surgical methods. The 2024 AOANJRR found significant improvements in Stryker's Triathlon implant when used for total knee arthroplasty implanted using Mako. The Triathlon CR (with and without patella) shows a significantly lower six-year cumulative revision rate when implanted with Mako compared to manual. 4. We feel it is premature to estimate QALYs and perform a cost-effectiveness analysis at this stage of the EVA process. As we have discussed in our response, NICE is not currently reviewing the full range of available evidence, and it would therefore be appropriate to require a more robust evidence base before performing a cost-effectiveness analysis comparing individual robotic systems to conventional surgery. The economic evaluation in the current report should not be	Thank you for your comment. 1. Evidence for all technologies was individually prioritised and analysed in the external assessment report and considered individually by the committee. The committee used its clinical expertise when interpreting the evidence. The outcome for this early value assessment is that the committee recommended that all devices could be used whilst further evidence is generated. The committee were aware that some devices had fewer evidence gaps than others but it concluded that all devices have potential benefits to the NHS. 2. Statistical significance of difference in PROMs were reported in the EAG report, separated by robotic system used (see sections 5.2.1, 5.3.1, 5.4.1, 5.5.1, 5.6.1, 5.7.1 of the EAG report). The EAG note that the only randomised evidence which demonstrated a difference in PROMs was the change in Knee injury and Osteoarthritis Outcome Score (KOOS) at 1 year between NAVIO and conventional

Comment number	Consultee type	Section number	Comment	Response
			used to inform decisions regarding the uptake of robotics in NHS as further (robust) evidence is required prior to estimating the ICERs for robotic technology.	surgery, and also between CORI and conventional surgery, as reported by Adamska et al. 2023 (n=215 across 3 arms). Due to the lack of statistical significance, the EAG did not refer to the MCID for each PROM. Events including complications and revisions are included in the evidence generation alongside PROMs. 3. See response to **comment 6, section A & B,** above. 4. Early economic modelling is permitted within the context of EVAs (see PMG39).
14	Consultee 7 The British Orthopaedic Association		1. While the BOA supports the recommendations for continued use and further evaluation of robot-assisted surgery, the BOA should highlight opportunities to enhance the robustness of the assessments. 2. Firstly, the literature review underpinning this Early Value Assessment adopts a five-year cut-off, which while practical, excludes a significant portion of foundational evidence on which these technologies are built. The inclusion of older but still relevant studies could provide valuable insights into the evolution and efficacy of these systems over time. 3. Secondly, it is noted that some literature has been excluded due to a lack of specificity regarding the robotic platform or company. In the context of total hip replacement, for example, there is an implication that all robotic arms are associated with the MAKO system. Leveraging this association to include additional relevant studies may have significantly informed the evaluation, particularly the Markov model underpinning the health economics analysis.	1. Thank you for your comment. 2. See response to **comment 11** above. 3. Exclusion criteria were stated in the published protocol, which included: *"Studies not explicitly reporting the robotic system used, unless one of the companies submitted the study and confirmed that their technology was used."* The EAG conducted an independent literature review of clinical and

Comment number	Consultee type	Section number	Comment	Response
				economic evidence which was supplemented by evidence provided by each of the companies (see section 4.1, Appendix A2 of the EAG report). The EAG note that multiple systems are indicated for total hip replacement, and therefore could not robustly attribute all published results to Mako.
15	Consultee 7 The British Orthopaedic Association		The BOA appreciates the emphasis placed on registry data within the Early Value Assessment, particularly the use of the UK and Australian joint registries. However, it is considered that future evaluations could benefit from incorporating data from a wider range of registries. Leveraging international datasets could provide a broader perspective on the real-world effectiveness and outcomes of robot-assisted surgery. Additionally, while the integration of UK registry data with NHS Digital PROMs is supported, concerns remain about the potential limitations of these PROMs in evaluating the nuances of robotic versus conventional surgery. Specifically, ceiling effects in these measures may reduce their ability to differentiate between these approaches, which could impact the accuracy of outcome assessments. Exploring more sensitive tools or augmenting existing PROMs with complementary measures may address this limitation.	Thank you for your comment and for your suggestions of further evidence which could address committee recommendation for further evidence (see the published Evidence Generation Plan).
16	Consultee 1	Consultation question - Has all of the relevant evidence been taken into account?	Yes	Thank you for your comment.
17	Consultee 3 Stryker	Consultation question - Are the summaries of clinical and cost effectiveness	1. Page 17, section 3.13: “The committee concluded that the benefits seen in the evidence for Mako could be similar in the other 4 technologies that have less mature evidence. So it decided to make a conditional recommendation for use during the evidence generation period for the 5 technologies.” Whilst we welcome the committee’s conditional recommendation given the current absence of	Thank you for your comment. 1. Within the EAG report, clinical evidence was presented for each robotic

Comment number	Consultee type	Section number	Comment	Response
		reasonable interpretations of the evidence?	evidence across the various platforms in scope, we would encourage that any future assessments (after the evidence generation period) consider the evidence for each platform separately. Each of these platforms differ in their use, the implants that they are used with and even in the overall process – for example Mako utilises CT imaging prior to surgery where the others do not. These factors are highly likely to lead to differences in efficacy. Further, costs are likely to differ greatly between the platforms as well. We welcome the committee's recommendation in section 3.17 "The committee acknowledged that further evidence generation should focus on collecting utility data for each individual technology to better understand if there are differences between them" and would reiterate that future assessments should not assume that evidence on one technology is transferable to the others (see Vermue, H. et al. (2022) 'The evolution of robotic systems for total Knee Arthroplasty, each system must be assessed for its own value: A systematic review of clinical evidence and meta-analysis', Archives of Orthopaedic and Trauma Surgery, 143(6), pp. 3369–3381. doi:10.1007/s00402-022-04632-w.). 2. Pages 18-21, costs and resource use: We believe that it is vital that any future economic modelling makes efforts to calculate the cost per procedure. A) We also feel it is premature to estimate an ICER and perform a cost-effectiveness analysis at this stage of the EVA process. It would be more appropriate for the EAR to report that a more robust evidence base is required prior to performing a cost-effectiveness analysis comparing Mako to conventional surgery. The economic evaluation in the current EAG report should not be used to inform decisions regarding the uptake of robotics in NHS as further (robust) evidence is required prior to estimating the ICERs for robotic technology. 3. Page 19, section 3.16: "The committee acknowledged that the results were from a conceptual economic model that was built around several assumptions and highly uncertain utility inputs". We agree that the conceptual model was built around many assumptions and that this model may be difficult to draw conclusions from. We would ask at this stage if the expectation is to use the same model structure for any future assessments? 4. Page 21, section 3.21: "The NHS England steering group advised that	system separately (see sections 5.2, 5.3, 5.4, 5.5, 5.6, 5.7 of the EAG report). Each individual technology provider that has been recommended for further evidence generation will attend follow-up meetings to monitor progress of the evidence generation process. It is the responsibility of the individual company to collect evidence that will address the gaps identified. This will increase the likelihood that appropriate data is available for each technology at that next assessment of this topic, allowing the committee to make technology-specific recommendations if it deems this to be appropriate. The study by Vermue (Arch Orthop Trauma, 2023; 3369-3381) referred to the in consultation comment, was identified by the EAG literature search and excluded due to the intervention including ROBODOC (which is not available in the UK and not listed in the NICE Final Scope). 2 & 3. NICE acknowledge the uncertainties in the economic model built for

Comment number	Consultee type	Section number	Comment	Response
			procurement is within its scope and may have a role in negotiating implant prices at a national level." National pricing already exists through NHS supply chain, where price is linked to volume and provides clear and consistent pricing information. 5. Page 22, section 3.22: "The committee agreed that there were evidence gaps in all technologies assessed in this early value assessment. It noted in particular that, for THA, Mako only had limited non-randomised evidence and CORI had no evidence within scope". We would suggest that this statement be reworded. Other platforms (ROSA and VELYS) are not currently indicated for THA. The statement as currently written is worded in a way that suggests that Mako has less evidence for THA but it is in fact the only platform with evidence in this procedure. A possible minor wording change could be "The committee agreed that there were evidence gaps in all technologies assessed in this early value assessment. It noted in particular that, for THA, only Mako had limited non-randomised evidence. CORI had no evidence within scope and the other platforms are not currently indicated for this procedure".	this evaluation. The early value assessment interim statement (PMG39) section 3.22 mentions the model is built to explore uncertainty with the aim of identifying key drivers of the model results to inform decision making about further evidence generation. This model was developed based on published evidence and may be used as a framework for future assessments of the cost-effectiveness of RAS when further evidence becomes available. 4. Clinicians on the committee and contacts within the NHS England steering group have advised that pricing is inconsistent. Furthermore, company representatives noted that costs were negotiable. The NHS England steering group are working to establish a framework for costs which will be beneficial when conducting future economic analyses. 5. Thank you for your comment. This statement has been amended following the identification of randomised

Comment number	Consultee type	Section number	Comment	Response
				evidence for THA during public consultation.
Theme 3 – economic evidence				
18	Consultee 5 Johnson and Johnson	Consultation question - Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	1. No. As communicated to the EAG, J&J has been unable to critically assess the work because we have received only a fully redacted document from them. The information and results pertaining to our own technology have not been made visible to us so we have been unable to check for factual inaccuracies in the model. 2. Further, we question the validity of the utility values used. The EAG have noted that the lack of cost-effectiveness for TKA and UKA is likely due to the utility values used. While we accept the uncertainty of the evidence around utility values, the EAG have noted that increased precision may result in improvements in activity levels and lower revision rates. Additionally, there are several studies for VRAS that demonstrate a reduction in pain (Alton et al, 2023) a reduction in morphine use (Severson et al, 2023), improvement in functional scores (Alton et al, 2023), improvement in walking scores (Spitzer et al, 2024), reduction in adverse events requiring intervention (Alton et al, 2023), reduction in revisit and readmissions (Huang et al, 2024) and reduction in length of stay (Severson et al, 2024). Potential short-term benefits of VRAS in TKA identified in these studies have not been appropriately captured; we believe it is reasonable to expect an improvement in quality of life with VRAS compared to conventional surgery in the initial months following primary TKA, during the acute healing period. Given the high sensitivity of the model results to utilities, incorporating the short-term improvement in quality of life following primary TKA would likely significantly impact the model results. We do not believe that the use of inferior utility values for robotic systems is supported by the evidence considering that none of the randomised evidence reported a statistically significant difference at 1 or 5 years. Given the lack of probabilistic sensitivity analysis, using these point estimates favouring conventional surgery has a significant impact on the model's deterministic results. We request that utility values at different time points, including those during the acute recovery period, are considered and that the	Thank you for your comment. 1. Individually redacted documents are not shared, as per section 5.4.17 of NICE health technology evaluations: the manual, if a technical engagement happens, all information marked as confidential will not be released to stakeholders even though they have signed a confidentiality agreement. All technology specific information used in the economic model has been taken from the request for information forms submitted by the companies and has undergone quality assurance checks as part of the development process. 2. Regarding the evidence referred to in the consultation comment: • Alton et al. 2023 was considered and excluded by the EAG (see Appendix A4, EAG report)

Comment number	Consultee type	Section number	Comment	Response
			utility values are reflective of what has been found to be statistically significant in randomised and observational studies. With equivalence or no statistically significant difference in PROMs being a plausible worst-case scenario, this would likely result in robotic surgery not being dominated by conventional surgery. 3. There are other important factors that should also be taken into consideration and included in the economic model, such as length of stay, reduction in adverse events requiring intervention, reduction in revisit and readmissions, number of trays used and sterilisation costs. Not including these data from manufacturers means that an incomplete assessment has been undertaken on the robots in scope and that the full cost-savings associated with the use of the robots has not been realised.	• Unclear what the reference Severson et al. 2023 is. However, Severson et al. 2024 was included in the EAG report but not prioritised (provided academic in confidence) • Spitzer et al. 2024 was considered and excluded by the EAG (see Appendix A4, EAG report) • Two studies by Huang et al. 2024 were included in the EAG report as in scope but not prioritised (provided academic in confidence) The EAG considers it unlikely that incorporating a greater QALY gain for robotic surgery in the short term (that is, in the months following surgery) would change the direction of the results. Over the lifetime time horizon of the model, any short-term gains would likely be too small to counteract the similarity in utilities over the longer term. That said, the EAG has stated the limitations of the modelling in their report and note that an important part of the EVA process is to use

Comment number	Consultee type	Section number	Comment	Response
				the economic evaluation modelling to highlight key uncertainties and areas for further evidence generation. 3. The EAG did not find any evidence for differences in length of stay (although we did explore the effect of reducing length of stay in sensitivity analysis, this is reported in section 9.3.2 of the EAG report), adverse events, or readmissions, either between conventional surgery and robotic surgery, or indeed, between robotic technologies. The EAG has also suggested a detailed micro-costing exercise to better understand the need for accessories (noting that this may differ between robotic systems), as this was not feasible within an EVA. The EAG notes that costs for sterilisation, for example, could be taken from the shared references. The economic model could be updated to reflect such differences in the future, and this has been highlighted for future evidence generation in Table 44 of the EAG report.

Comment number	Consultee type	Section number	Comment	Response
19	Consultee 5 Johnson and Johnson	Not specified	We believe there are significant limitations in the methodology. J&J have been unable to verify the factual accuracy of the data used by the EAG as we have only been provided with a fully redacted document from them, including redaction of details pertaining to our own system. Substantial relevant evidence has not been considered and we believe an incomplete assessment has been conducted without realising the full cost-savings of robotic surgery. Specifically, we contend that the utility values used in the economic model are unlikely to be clinically plausible and are not supported by the evidence generally, and this has been a major determinant of the model outcome.	Thank you for your comment. Individually redacted documents are not shared, as per section 5.4.17 of NICE health technology evaluations: the manual, if a technical engagement happens, all information marked as confidential will not be released to stakeholders even though they have signed a confidentiality agreement. The uncertainties in the economic model built for this evaluation are acknowledged. The early value assessment interim statement (PMG39) section 3.22 mentions the model is built to explore uncertainty with the aim of identifying key drivers of the model results to inform decision making about further evidence generation. This model was developed based on published evidence and may be used as a framework for future assessments of the cost-effectiveness of RAS when further evidence becomes available.

Comment number	Consultee type	Section number	Comment	Response
20	Consultee 5 Johnson and Johnson	1.6 What evidence generation and research is needed	Given the methodological limitations, we believe that the statement that "for total or partial knee arthroplasty, the results suggest the technologies in this guidance are not likely to be cost-effective" is overstated and misleading. "May not be cost-effective" would seem to be a fairer assessment of the evidence, and the limitations and uncertainties should be more clearly stated.	Thank you for your comment. The wording of this statement has been changed to reflect the uncertainty in the model.
21	Consultee 8	Additional evidence	I have attached two accepted papers for the BJJ and BJO that may be of some use. They certainly support robotic total and partial knee arthroplasty from a cost effectiveness point of view. 1. https://pubmed.ncbi.nlm.nih.gov/39740684/ 2. Cost-effectiveness analysis of robotic-arm assisted versus manual total knee arthroplasty in the United Kingdom (pre-print)	Thank you for your comment. Summaries of the evidence have been provided in Appendix A and were presented to the committee for consideration in the second Medical Technology Appraisal Committee meeting. No changes to the recommendations were made, but the committee agreed that these papers provided additional evidence supporting the potential cost-effectiveness of RAS for TKA and PKA.
Theme 4 – equality considerations				
22	Consultee 1	Consultation question - Are there any equality issues that need special consideration and are not covered in the medical technology	No	Thank you for your comment.

Comment number	Consultee type	Section number	Comment	Response
		consultation document?		
23	Consultee 4	Consultation question - Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	Patients with "mental or neuromuscular disorders " should not be excluded.	Thank you for your comment. This issue was raised at the second Medical Technologies Appraisal Committee meeting. Clinicians on the committee advised that these conditions are not specific to RAS. The mention of mental or neuromuscular disorders has been removed from the guidance. The guidance now prompts clinicians to refer to individual technology instructions for use and apply their clinical judgement to determine if RAS or conventional surgery is most appropriate for the individual.

Comment number	Consultee type	Section number	Comment	Response
24	Consultee 5 Johnson and Johnson	Consultation question - Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	We do not believe there are any equality issues that are not already covered. We do feel that some of the identified issues are overstated insofar as they are equally pertinent to conventional manual total knee replacement and not specifically robotic total knee replacement.	Thank you for your comment.
25	Consultee 5 Johnson and Johnson	1.6 What evidence generation and research is needed	Some of the equality scenarios described as limiting access to robotic joint arthroplasty are equally relevant to any other technique of joint arthroplasty using standard primary implants, including manual surgery. They are therefore limitations to surgery with standard primary implants, rather than robotics per se, since manual arthroplasty with those same implants would be equally unsuitable.	Thank you for your comment.
26	Consultee 6 The Royal College of Surgeons of England	Consultation question - Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	1. Equity of access to RAS systems for surgeons and trainees remains a key issue and should be a particular focus of this early value assessment. 2. To ensure equitable access and sustainability, the following factors should be considered: A. Healthcare Equity: Implementation strategies must address potential disparities in access to robotic technologies across regions and institutions, and NICE should provide guidance on this rather than merely highlight the issue. Software and system design should account for any differences in usability between different groups of surgeons (by gender or other protected characteristic), to avoid disadvantage for surgeons and patients. B. Clinical Indications: NICE should provide clear guidance on the specific indications where robot-assisted approaches offer a demonstrable advantage. C. Real-World Evidence: Continuous monitoring of outcomes via national registries will be crucial to refining practice and ensuring patient benefit. It is important for NICE to set out how it will keep pace with developments in the	Thank you for your comment. 1. NICE are working with the NHS England Robotically assisted Surgery steering committee to promote equitable access to the technologies. NICE supports technology developments that reduce inequalities and increase usability for both clinicians and patients. 2A & 2B. It was beyond the scope of this EVA to investigate specific subgroups within those undergoing orthopaedic procedures. Further evidence generation

Comment number	Consultee type	Section number	Comment	Response
			evidence base for RAS. RCS England is keen to guide and advise on this based on 30 years of experience in supporting and facilitating good quality surgical research and through our robotics network.	may identify those in whom RAS is most beneficial. 2C. NICE thanks the RCS for their support of using national registry data and welcomes the expertise they may offer.
Theme 5 – further evidence generation				
27	Consultee 6 The Royal College of Surgeons of England	Not specified	The Royal College of Surgeons of England welcomes the opportunity to contribute to the ongoing NICE early value assessment of robot-assisted surgery (RAS) for orthopaedic procedures. RCS England is committed to supporting surgical innovations that enhance patient care while upholding the highest standards of safety, efficacy, and equity. As a royal college, our primary focus is to ensure that emerging surgical innovations and technologies such as RAS align with the principles of equity of access, patient benefit and safety, clinical efficacy, and value for the surgical workforce and overall healthcare system. RCS England supports the aims of this early value assessment, and recommends addressing the following points to facilitate a more robust methodology for appraising emerging technologies: Clinical Effectiveness: Orthopaedic RAS platforms may be image-based or image-less, and use methods of passive, semi-active, and active control, leading to significant differences between the platforms. Data on each type should be included. Safety and Training: The introduction of robotic systems must include strategies including the use of simulation to mitigate the risks associated with the surgical learning curve, supported by standardised training pathways for surgeons. The REINVENT study, supported by RCS England cross-examines trainees' views on training and should be considered. Economic Impact: While upfront costs of robotic systems are significant, these	Thank you for your comment. These factors may be considered in future assessments.

Comment number	Consultee type	Section number	Comment	Response
			must be weighed against potential long-term savings from improved patient outcomes, reduced complication rates, reductions in length of stay, and resource utilisation efficiencies.	
28	Consultee 7 The British Orthopaedic Association		As highlighted in the Early Value Assessment, there is evidence that robot-assisted surgery enhances precision in some orthopaedic procedures. The BOA considers that the next critical step is to demonstrate how these advancements translate into improved patient outcomes, such as better functional recovery and reduced complications.	Thank you for your comment.
29	Consultee 3 Stryker		3. Page 14, section 3.7: This section states that “Experts told the committee that robotic technologies are most commonly obtained through volume-based contracts, whereby NHS trusts commit to a number of procedures each year. This approach to purchasing means that robotic technologies are more likely to be cost-effective in high-volume orthopaedic centres. The committee was also aware that the high cost of the technologies means that robot-assisted surgery is more widely available in the private sector. The committee noted that limiting access to robot-assisted surgery to these hospitals may exacerbate existing inequalities. The committee also noted that robot-assisted surgery may be more beneficial in complex surgical cases. These cases are typically done in lower volume centres, with more prehabilitation and rehabilitation, as well as more advanced planning because of the associated surgical risks.” More evidence and analysis should be undertaken on the benefits of RAS in both high-volume low complexity centres and centres that focus on complex patients.	Thank you for your comment. This has been considered in the evidence generation plan.
Theme 6 – clarity of wording				
30	Consultee 2 National Joint Registry	1.6 What evidence generation and research is needed	Care should be taken with the language used. 'This allows the surgeon to position and align the implants in the correct position for each person' implies that a 'correct' position cannot be achieved through conventional surgery	Thank you for your comment. This word ‘allows’ has been changed to ‘assists’ to reflect the potential for RAS to aid the surgeon in aligning the implant.

Comment number	Consultee type	Section number	Comment	Response
31	Consultee 2 National Joint Registry	3.1 Unmet need and potential benefits	We are not convinced that these satisfaction rates should be characterised as 'low' given the percentages.	Thank you for your comment. The word 'low' has been replaced by 'moderate'.
32	Consultee 5 Johnson and Johnson	1.6 What evidence generation and research is needed	We would caution against the use of the term "better alignment" since we would contend that the alignment is only "better" if it can be shown to have a positive impact on clinical outcomes; implant alignment is consistently more precise with robot-assisted surgery than with conventional surgery, but the impact on clinical outcomes is likely related to the target alignment (and how precisely that is achieved) which is why we believe it is an oversight to reject evidence assessing alternative target alignments consistently achievable only with technology.	Thank you for your comment. The term 'better alignment' has been replaced with 'more precise alignment' to reflect the outcome of interest. The following sentence acknowledges the uncertainty in the link between alignment and clinical outcomes. Alternative target alignments achieved with technology were out of scope due to not representing standard care in the NHS.
33	Consultee 5 Johnson and Johnson	2.8 VELYS Robot-assisted Solution (Johnson & Johnson)	FACTUAL INNACURACY – the VELYS Robot-assisted Solution robotic arm is not attached to a base station. The arm is stored on a satellite station and mounted on the operating table for use. Further, while the robotic arm maintains the saw within the planned cutting plane, and will stop working if it leaves that plane, there is no boundary control provided by the system within the plane itself. This information was also conveyed to the EAG previously.	Thank you for your comment. This factual inaccuracy has been amended.
34	Consultee 1	Consultation question - Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes- but the section on training should also include an understanding of not only the technology but also safety principles specific to each system. This could be misinterpreted by the reader by understanding the set up and use of the device is needed but not a proficient understanding of tackling technical challenges in emergency scenarios.	Thank you for your comment. The wording in the guidance has been amended to 'All members of the surgical team must be trained on each robotic technology that they use including its safety principles'.
Theme 7 – general				

Comment number	Consultee type	Section number	Comment	Response
35	Consultee 1	Consultation question - Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	Yes - clearly explained to the reader	Thank you for your comment.
36	Consultee 7 The British Orthopaedic Association		The BOA has taken a proactive stance on the integration of robotics into orthopaedic care. This is reflected in the production of four dedicated documents addressing the role of robotics in orthopaedics. These efforts align closely with the Department of Health and Social Care's *"The Medical Technology Strategy: One Year On"* (April 2024), which emphasises safe and effective adoption of transformative healthcare technologies.	Thank you for your comment.
37	Consultee 7 The British Orthopaedic Association		The BOA recognises the importance of generating robust evidence to support the widespread adoption of robotic technologies. The BOA is encouraged by initiatives like the REINFORCE and RACER studies, which represent key steps in bridging the gap between innovation and evidence. These studies exemplify the UK's leadership in advancing this field, whilst addressing critical questions about clinical and economic impacts.	Thank you for your comment.
Theme 8 – evidence generation plan				
38	Consultee 2 National Joint Registry	3.7 Equality considerations	How will the mapping of robotic surgery provision be published? It would be useful for the NJR too be looped into this exercise.	Thank you for your comment. The mapping of RAS activity is being carried out by NHS England robot-assisted surgery steering group. NICE can provide links to the NHS England guidance once it is published.
39	Consultee 2 National Joint Registry	3 Data sources	We would recommend exploring opportunities to link intraoperative metadata from the robotic systems with the NJR to maximise the utility of the dataset	Thank you for your comment. Section 6 has been amended to highlight the opportunity of linking data to the National Joint Registry.

Comment number	Consultee type	Section number	Comment	Response
40	Consultee 1	Consultation question - Could the period while surgeons are learning to use the technologies have a significant impact on the clinical and cost-effectiveness of them?	Potentially- as whilst the surgeon and theatre team are on their learning curve, they are more likely to perform more errors which could affect clinical outcomes by leading to adverse events.	Thank you for your comments. The Committee discussed the evidence and concluded that the learning curve for RAS is similar to that for manual surgery. It decided that the learning curve is therefore unlikely to significantly impact on the cost-effectiveness of the technology.
41	Consultee 2 National Joint Registry	3.13 Clinical effectiveness	To note, the NJR does not cover all parts of the UK, so any procedures done in Scotland will be be included.	Thank you for your comment. Section 6 on system considerations has been amended to include “the National Joint Registry does not cover all parts of the UK, for example procedures done in Scotland are not included”.
42	Consultee 2 National Joint Registry	3.17 Economic modelling	We support the desire to obtain more PROMs data, but it is unclear how this will be achieved in practice. It would be preferable for this to be a central model via national PROMs collection rather than a manufacturer led process	Thank you for your comments. The evidence generation plan highlights the minimum data items that should be collected by anyone using the plan. Although it is primarily intended to be a guide for manufacturers collecting data, it is not exclusive to them. The plan specifies PROMs already collected by the National Joint Registry linked to the NHS England's national patient-reported outcome measures (PROMs) programme. The plan also states that “active monitoring and follow up

Comment number	Consultee type	Section number	Comment	Response
				through a central coordinating point is an effective and viable approach for ensuring good-quality data with broad coverage".
43	Consultee 2 National Joint Registry	5	One factor that should be carefully accounted for in any analysis is the software version of the robotic system. It is not advisable to determine that a robotic system is 'the same' if software versions have changed over time.	Thank you for your comment. A section on technologies has been added to section 3 of the evidence generation plan under "Data to be collected", it now includes "a detailed description of the RAS technologies and the specific versions."
44	Consultee 2 National Joint Registry	6 Evidence generation	NJR data may not give a complete picture of trainee numbers as lead surgeon data is often recorded as being by a consultant regardless of whether a trainee is operating	Thank you for your comment. NICE acknowledges that the National Joint Registry does not collect all of the data specified in the plan. The evidence generation plan suggests linking to other data sources where data is available, and carrying out an audit where it is not. Section 6 has been amended to "The impact of RAS on the development of the manual skills of trainees is unknown. Competency is reviewed through peer-review and reported in National Joint Registry audits, but may not detail if the operation was performed by a trainee."

Comment number	Consultee type	Section number	Comment	Response
45	Consultee 3 Stryker	Consultation question - Could the period while surgeons are learning to use the technologies have a significant impact on the clinical and cost-effectiveness of them?	Evidence has shown that it takes around 7 to 14 procedures to gain full competency in the technique. As such, we do not believe the training period would have a significant clinical or cost effectiveness impact.	Thank you for your comment.
46	Consultee 4	Consultation question - Are the recommendations sound and a suitable basis for guidance to the NHS?	No. The time frames are too short.	Thank you for your comments. The evidence generation period has been extended to 3 years, which is the maximum time period for evidence generation within the remit of the Early Value Assessment program.
47	Consultee 4	Consultation question - Could the period while surgeons are learning to use the technologies have a significant impact on the clinical and cost-effectiveness of them?	Yes We are still learning on the optimum orientation (target that we should achieve). The learning curve will alter the time taken for surgery which will impact on the HE modelling on the number of cases per operating session.	Thank you for your comment. Please see response to comment 40.

Comment number	Consultee type	Section number	Comment	Response
48	Consultee 5 Johnson and Johnson	Consultation question - Could the period while surgeons are learning to use the technologies have a significant impact on the clinical and cost-effectiveness of them?	Our experience with VRAS would suggest that any impact would be relatively minor. We would expect to see outcomes and resource use improve over time during the learning curve, but evidence supports that even during the early stages of adoption there is non-inferiority to manual TKA (Alton et al 2024) and we would anticipate an improvement being realised after a relatively short learning period (Pagen et al 2022, Lall et al 2023, Morrisey et al 2023).	Thank you for your comments.
49	Consultee 5 Johnson and Johnson	Not specified	It is unclear how the evidence generation plan is expected to be implemented. Further guidance is required for the companies responsible for ensuring that data collection and analysis takes place to ensure that the methodology is sufficiently robust and comprehensive to satisfy the requirements of NICE, while also being consistent between companies and achievable within the specified timeframe. The practical details of data collection remain unclear. Such details are required by the companies in sufficient time to prepare for and appropriately resource the data collection, especially where outcomes are not currently collected by existing sources. To ensure consistency we would recommend the methodology is agreed with all manufacturers; clear central direction would facilitate this.	Manufacturers of the technologies included in the guidance will be contacted by NICE following publication of the guidance to provide further details on how NICE monitors evidence generation. The evidence generation plan acts a guide for those collecting outcomes to fulfil the evidence gaps. The plan is not intended to be used as a study protocol for collecting data, as a result NICE acknowledges that there may be variation in the data collected for different technologies. Assistance on individual protocols may be sought from NICE Advice.

Comment number	Consultee type	Section number	Comment	Response
50	Consultee 5 Johnson and Johnson	1 1 Purpose of this document	NICE notes that for assessing comparative treatment effects, well-conducted randomised controlled trials are the preferred source of evidence and in the current EVA observational studies are not included. It is unclear if retrospective evidence generation will allow for a positive technology specific recommendation to be made.	Thank you for your comments. The evidence generation plan reflects the outcomes which have been prioritised by the NICE committee, and proses a pragmatic approach to collecting those. The methodology is suited for assessing real-world data where randomised control trials are not practical, this is outlined in NICE's Real-World Evidence Framework.
51	Consultee 5 Johnson and Johnson	2 Resource use	It is important to include not only the cost of consumables, but the sterilization requirements for different robotic systems and for manual surgery, as there may be significant differentiation here, which can have a large impact on total procedure cost.	Thank you for your comment. The committee discussed the potential impact of sterilisation costs and decided that it was unlikely to significantly impact the cost-effectiveness modelling.
52	Consultee 5 Johnson and Johnson	3 REINFORCE trial	The referenced REINFORCE trial is for general surgery robotics and we would suggest not relevant to this document unless it is framed as an example of study design.	Thank you for your comment. Several specialist committee members confirmed that the REINFORCE trial is collecting data in orthopaedic centres and using RAS technologies within the scope of this topic.

Comment number	Consultee type	Section number	Comment	Response
53	Consultee 5 Johnson and Johnson	3 Real-world historical control study with propensity score methods	It is unclear how the historical controls would be appropriately matched with the robotic cases. Over time it is likely that differences in treatment care pathways/protocols for arthroplasty have likely been implemented within the NHS. These differences would also likely have a significant impact on patient outcomes and would complicate the ability to identify the impact of robotics when compared to historical controls. In addition, the endpoints that are not included in the current real world evidence data sets would not be available for these historical controls.	Thank you for your comment. NICE acknowledges the limitations of historical control data and will consider these in the new assessment at the end of the evidence generation period. The methodology is described as a pragmatic method for assessing real-world data and is outlined in NICE's Real-World Evidence Framework. NICE acknowledges that the proposed historical control method for data collection will not obtain all the specified outcomes, therefore the evidence generation plan suggests a prospective audit for outstanding data items.
54	Consultee 5 Johnson and Johnson	3 Prospective audit	For the prospective audit that has been recommended as a mechanism to collect the data not captured by real world evidence datasets, it is not clear if the expectation is that these datapoints are captured for all NHS patients or if a sample would be acceptable. Either way, we would recommend a consistent methodology is agreed with all manufacturers for this.	Thank you for your comment. Researchers will need to decide on an appropriate study size based on statistical calculations for their individual studies. Assistance on individual protocols may be sought from NICE Advice.
55	Consultee 5 Johnson and Johnson	3 Evidence generation period	The 2 years evidence generation period is too short to generate a reasonable sample size and follow up for a system like VELYS RAS that is only just entering the market.	Thank you for your comment. The evidence generation period has been extended to 3 years.

Comment number	Consultee type	Section number	Comment	Response
56	Consultee 6 The Royal College of Surgeons of England	Consultation question - Are the recommendations sound and a suitable basis for guidance to the NHS?	With regards to suggested evidence generation, it appears that the committee is recommending that some studies that have already been published are repeated. Given that the NJR was initially designed to detect issues with implants, it remains to be seen whether this is the optimal tool to answer the open questions suggested. Two years may not be enough for all the systems to get permissions and collect the data that they wish or need to collect. We recommend the time parameters be extended.	Thank you for your comments. Please see response to comment 6 regarding the literature that was reviewed. NICE acknowledges that the National Joint Registry does not collect all of the data specified in the plan. The evidence generation plan suggests linking to other databases where data is available, and carrying out an audit where it is not. The evidence generation period has been extended to 3 years.
57	Consultee 6 The Royal College of Surgeons of England	Consultation question - Could the period while surgeons are learning to use the technologies have a significant impact on the clinical and cost-effectiveness of them?	The learning curve has been studied, particularly for the Mako system but also increasingly for others while they are introduced, and there does not appear to be a harm element to it. Surgical time is often extended slightly, which could affect cost effectiveness, but a number of studies suggest this is only for a very short time period. With 3D-based planning and robotic arm assistance, there is no harm to the patient in terms of operative accuracy; the learning curve relates to surgeon and operating team comfort and seems to be resolved fairly quickly.	Thank you for your comments.
58	Consultee 7 The British Orthopaedic Association		The BOA supports the outlined evidence generation recommendations, as these are crucial for validating the clinical and economic impacts of robot-assisted surgery. However, it is noted that several key research questions—particularly questions one, two, three, and seven—are likely to face significant time lags in generating robust data. While the proposed two-year timeframe is an ambitious and commendable goal, the complexity and scope of these questions suggest that a three-to-five-year	Thank you for your comments. The evidence generation period has been extended to 3 years.

Comment number	Consultee type	Section number	Comment	Response
			horizon may be more realistic for capturing meaningful outcomes. To address this challenge without disrupting progress, the BOA should suggest incorporating flexibility into the timelines. For example, the evaluation period could be extended if interim results indicate that additional time is required to gather high-quality evidence. This approach ensures that evidence generation remains thorough and reflective of real-world outcomes, while maintaining momentum in the adoption of innovative technologies that can improve patient care.	
59	Consultee 7 The British Orthopaedic Association		The BOA recognises that the use of robotic surgery in revision hip and knee replacements and shoulder procedures remains at an early stage. These are inherently more complex surgeries, and it is likely that the two-year timeline proposed in the Early Value Assessment may not provide sufficient time to fully evaluate their potential impact. Given the complexities of these procedures, robotic technology offers considerable promise in addressing unique surgical challenges and improving outcomes. The BOA considers that it is important to allow sufficient flexibility in timelines to ensure that the development and adoption of these innovative approaches are fully supported.	Thank you for your comments. The evidence generation period has been extended to 3 years.
60	Consultee 3 Stryker	6 Evidence generation	2. Page 4, section 1.6: We welcome the evidence generation recommendations and largely agree with the needs identified. We would suggest that capturing data on casemix would also be very useful. A) Section 3.20 states that the committee agreed that assuming a single procedure volume for all centres was a limitation while section 3.13 discusses that different centres and different robots offer different procedures. Therefore, it is important to understand differences in casemix between centres and platforms as these will affect the economic modelling completed after the evidence is generated. 4. Evidence generation plan, page 2: The plan states that “After the end of the evidence generation period (2 years), the companies should submit the evidence to NICE in a form that can be used for decision making”. NICE should be clear	Thank you for your comments. A. The Committee agreed that stratifying data by casemix would be useful for identifying scenarios where RAS offers best value. Section 3.4 has been amended to include “surgery indication” to allow for casemix analyses. The other patient characteristics as well as the measures of around

Comment number	Consultee type	Section number	Comment	Response
			what this form should be – does this suggest that the evidence must be published in a peer-reviewed journal? 5. Evidence generation plan, table 1, page 5: This table currently shows that for total hip arthroplasty that there is no ongoing study on Mako, however, the RACER-hip trial on (https://www.isrctn.com/ISRCTN13374625) will gather Mako data on this procedure – this should be reflected in the table. 6. Evidence generation plan, page 5, Data sources: The plan states that "The National Joint Registry (NJR) is the data source that is most likely to be able to collect the real-world data necessary to address the essential evidence gaps". We would note at this stage that the data is not currently available to industry or the public, and this should be factored into future timelines. The evidence generation timelines proposed by the committee may not be realistic. While it is reasonable to expect an evidence plan within 6 months, contracts can take much longer and are not in the full control of the manufacturer / study sponsor. Two years are also an expeditious timeframe to design, contract, execute, analyse and publish a study. We recommend the time parameters be extended. Further, given that the registry was initially designed to detect issues with implants, it remains to be seen whether this is the optimal tool to answer the open questions.	resource use should enable these analyses. The volume of procedures has also been added to the list of subgroups analyses in section 3.4. 4. Evidence in the form of a peer-reviewed journal publication (accepted or published) is preferred. Pre-prints and any evidence that follows the format of a manuscript are also suitable. 5. The table has been amended to read "Ongoing" rather than "No evidence". 6. The evidence generation period has been extended to 3 years, which is the maximum time period for evidence generation within the remit of the Early Value Assessment program. Please note the evidence does not need to be published by this date, but must be in a form that is suitable for decision making such as a manuscript. NICE acknowledges that the NJR does not collect all of the data specified in the plan. The evidence generation plan suggests linking to other databases where data is

Comment number	Consultee type	Section number	Comment	Response
				available, and carrying out an audit where it is not.

Appendix A – Additional health economic evidence summary (response to comment 21)

Two economic studies are referenced in the comment:

1. Blyth, M. J. G., Clement, N. D., Choo, X. Y., Doonan, J., MacLean, A., & Jones, B. G. (2025). Robotic arm-assisted medial compartment knee arthroplasty is a cost-effective intervention at ten-year follow-up. *The bone & joint journal*, *107-B*(1), 72–80. https://doi.org/10.1302/0301-620X.107B1.BJJ-2024-0245.R2

2.Sagoo, S. S. (preprint). Cost-Effectiveness of Robotic-Arm Versus Manual Total Knee Arthroplasty in the UK http://dx.doi.org/10.2139/ssrn.4718214. *Bone and joint open.*

Blyth et al., (2025) is a Scottish-based cost-utility analysis from a RCT of medial compartment knee arthroplasty with the Mako device (n=45) compared to conventional surgery (n=40). This analysis contains extended EQ-5D follow-up of the 5-year economic analysis included in the external assessment report (Clement et al., 2023). This study was recently published in January 2025.

Over the 10-year period, 10 participants in the conventional surgery group underwent revision procedures; 5 were revised to total knee arthroplasty, 3 had a knee arthroscopy and 2 had a debridement, antibiotics and implant retention procedure compared with 0 revision procedures in the robot-assisted surgery (RAS) group. The study reported a mean difference in EQ-5D of 0.050 (95%CI -0.024 to 0.122) at 10 years indicating a non-statistically significant improvement in EQ-5D with RAS at 10 years.
The overall QALY gain per patient over the 10-year follow-up period was greater in the RAS group compared with the conventional surgery group ([7.410 SD 2.315 vs 7.223 SD 2.347]; mean difference 0.186 (95% CI -0.626 to 0.999). The costs for the conventional surgery group included the costs of the 10 revisions and resulted in a per patient cost of £1,850. The RAS per patient cost of £1,051 did not include any revisions but included the cost of CT scans, robot hire and surgical consumables. This analysis showed with 10-year follow-up RAS was a dominant intervention.

Sensitivity analyses explored the impact of removing the different types of revision costs (septic costs and arthroscopy costs) but keeping the aseptic revision costs. This analysis, resulted in conventional surgery becoming less costly than RAS, but in all scenarios the ICER was below the £20,000 willingness to pay (WTP) threshold. The base case assumed an annual case load of 400 patients and sensitivity analysis indicated that both with and without septic revision costs a minimum of 50 cases per year were required for the ICER to remain below the WTP threshold. Cost threshold analysis including septic revision costs showed that RAS was cost saving when more than 100 cases per year were done however if septic revision costs were removed RAS only became cost saving when more than 800 cases per year were done.

The study may not be generalisable to the English and Welsh NHS because it was conducted in a Scottish hospital, which is associated with different costs. The key limitation of this study was the small sample size for detecting differences in EQ-5D, with the author stating that 2,500 participants per study group would be required. This number is unlikely to be feasible for a single study because of recruitment and timing constraints and reinforces the importance of collecting PROMs data in the evidence generation

period to inform future recommendations. This small sample size may have also resulted in a random difference in revisions, with no events in the RAS arm and 10 in the conventional surgery arm, resulting in greater overall costs for conventional surgery. Limited power was also a limitation of other studies assessed by the EAG and so the EAG relied on data from the National Joint Registry (NJR) to inform the economic model. This again reinforces the importance of the evidence generation period to collect additional data on revision rates through collaboration with the NJR.

This study provides useful data to inform future economic modelling but is limited by the small number of participants providing EQ-5D data. The EAG's model used 5-year QALY data from Clement et al., 2023 that favoured conventional surgery, resulting in RAS being dominated by conventional surgery since it also had higher costs. The data from the present study suggests QALYs are improved with RAS and the per patient costs for RAS was less over a 10-year time horizon, and this supports RAS becoming undominated. However, this study is unlikely to affect the conclusions of the EAG's economic modelling, which stated that the model 'is sensitive to changes in utilities, where applying the upper and lower confidence interval from RCTs led to the ICER for RAS changing from being dominated to almost being dominant, as the ICER generated is low'. A further evidence generation period remains critical to collect more PROMs data to reduce uncertainty in the QALY data used in the model.

The study by Sagoo et al., (preprint) describes a cost utility analysis using data from the 12-month ROAM RCT which was included in the EAG's clinical evidence assessment. This study was conducted in England and compared total knee arthroplasty with the Mako system (n=50) to conventional surgery (n=50). The cost utility analysis used a Markov model and extrapolated the data over a 10-year time horizon.

The trial measured EQ-5D, reporting unadjusted QALYs of 0.5149 (0.3719 to 0.6579, unclear variance reporting) with RAS and 0.5008 (0.422 to 0.5796, unclear variance reporting) with conventional surgery at 12 months, based on complete case analysis of

76 participants. Baseline-adjusted QALYs gained were higher in the RAS group, with a difference of 0.014 additional QALYs (95%CI -0.0504 to 0.0785), but RAS was more expensive than conventional surgery (£2,236, 95%CI £425 to £4,046), resulting in an ICER at 1-year of £158,785. This ICER was reduced to £123,770 with imputation of missing QALY data but remained above the £20,000 WTP threshold.

The Markov model constructed was similar to the EAG's, but did not include a post-surgery state to account for utility decrements in the first year following revision surgery, instead including a disutility of -0.1 for the first cycle (1-year) in the 'well post-revision state. Similarly to the EAG's model, the present study assumed no difference in revision rates but had a starting age of 60 years compared to 72 years in the EAG's model. Base case results with the 10-year time horizon study showed that RAS was £1,833 more costly than conventional surgery and gained 0.17 QALYs more, resulting in an ICER of £11,109. Several sensitivity analyses explored other scenarios, including a time horizon of 5 years (ICER= £19,597) and 20 years (ICER= £6,116), indicating that the benefits of RAS may increase with time, and intervention-specific revision rates based on data showing higher rates with conventional surgery in the AOANJRR (ICER= £10,177).

This study is directly applicable to the NHS and provides utilities that can inform future economic modelling. The Markov model used a similar structure to the EAG's model, but used different inputs and starting points which may explain why the results differ. The main input that was different was utilities, with the EAG's model applying a disutility of -0.022 compared to a gain of 0.17 in the present study, which was within the 95%CI used by the EAG. This disutility in the EAG's model resulted in RAS being dominated by conventional surgery since it also was more costly. The authors concluded that RAS may plausibly be cost-effective over a 10-year time horizon, but that further data was necessary to inform the model. This is similar to the results of the EAG's model which concluded that the model was highly sensitive to changes in utilities. Further evidence generation is important to collect more PROMs data to reduce uncertainty in the QALY data used in the model.