National Institute for Health and Care Excellence

Medical technologies evaluation programme

GID-HTE10015 Artificial intelligence technologies to aid contouring for radiotherapy treatment planning: early value assessment

Consultation comments table

There were 83 comments from 5 groups:

- 34 comments from 6 companies
- 20 comments from 2 NHS trusts
- 13 comments from 3 patient organisations
- 11 comments from 1 healthcare professional
- 5 comments from 2 professional organisations

Some comments have been split because they represented multiple themes. The following themes have been identified:

- Recommendations: comments 1 to 14
- Technologies: comments 15 to 25
- Current management and guidance: comments 26 to 32
- Clinical evidence and effectiveness: comments 33 to 55
- Cost and resource use: comments 56 to 67
- Further evidence generation: comments 68 to 70
- Implementation and managing risks: comments 71 to 77
- Equality considerations: comments 78 to 93

Comment no.	Consultee ID	Group	Section	Comments	NICE response (including changes made to MTCD, if applicable)
Recommendations (n=14 comments)					
1	12	Company		[redacted]**Response to the Early Value Assessment consultation on Artificial Intelligence technologies to aid auto-contouring for radiotherapy treatment planning** **Introduction** The guidance is well-written and comprehensive, covering the current evidence and potential benefits of Artificial Intelligence (AI) technologies for contouring in radiotherapy. We appreciate the recognition of the unmet need and the value proposition of AI technologies to enhance contouring quality, efficiency and consistency, as well as to reduce costs and free up healthcare professional time. We support the recommendations to use nine AI technologies in the NHS while generating more evidence, subject to Digital Technology Assessment Criteria (DTAC) approval and evidence generation agreements. We also concur with the need for further evidence on clinical acceptability, radiation dose, time saving, resource use and adverse events.	Thank you for your comment.
2	14	Patient organisation		At [redacted]we are passionate about the benefits of radiotherapy for metastatic breast cancer. Particularly if systemic treatment is holding most disease steady. I personally have benefited from radiotherapy which controlled an aggressive tumour in my supraclavicular lymph nodes and from stereotactic radiotherapy in the brain. But radiotherapy provision in the NHS is not keeping up with practice in other developed countries. SABR for metastatic disease is confined to around three metastatic sites with no disease elsewhere. New research shows promise in oligo progressive disease, which is where existing metastatic disease is stable or reduced on systemic treatment apart from in isolated areas of progression. With SABR of oligo progression patients can be kept for longer on current line systemic treatments. If artificial intelligence makes radiotherapy quicker and cheaper to plan, so that targeted treatments can reach more patients, we welcome it,	Thank you for your comment. This early value assessment (EVA) guidance focused on the potential risks and benefits of using AI autocontouring technologies to help with radiotherapy treatment planning. Issues around the commissioning and availability of radiotherapy services in the NHS are outside the remit of this assessment.

				assuming toxicity does not increase. I think the bar for caution will partly depend on the tissue. Having brain radiotherapy is frightening, and I have witnessed a close friend get brain necrosis from stereotactic radiotherapy, so am very aware that devastating consequences can follow. To reach more patients, commissioning restrictions need to be loosened so a wider range and larger number of lesions can be treated. [redacted] travelling to Switzerland for SABR on her bone mets because the NHS can only offer her conventional radiotherapy. And her UK NHS oncologist has encouraged her to do so. Doctors are also frustrated that their patients do not meet the criteria for SABR on the NHS when they know they would be good candidates	
3	3	Company	1.1	On behalf of [redacted], thank you for the opportunity to comment on this draft guidance for AI auto-contouring for radiotherapy treatment planning. We are thankful for the Committee's pragmatic conclusions which should result in accelerated access to these technologies that can address staff shortages faced by the NHS.	Thank you for your comment.
4	3	Company	1.1	We would recommend to clarify when and in which context an approval under NHS England DTAC is required, so that NHS providers currently using these technologies would not lose access pending this approval.	Thank you for your comment. All new digital technologies should be assessed using the Digital Technology Assessment Criteria for health and social care (DTAC). NHS England states that DTAC should be completed as part of each new procurement process or contract renewal. Companies should discuss DTAC with future or current NHS providers and NHS England who can assist them in completing this process. No changes have been made to the guidance.

5	13	Professional organisation	1.3	**General comments** We welcome this EVA which will help clinicians to implement AI autocontouring safely and effectively. **Specific comments** 1.3 – 3 years seems a long time in an area where technology is changing rapidly. Could a shorter timescale be considered?	Thank you for your comment. The EVA guidance will be reviewed in 3 years or sooner if sufficient evidence is available. This is in line with the EVA interim statement which states that evidence generation should be for the shortest time necessary to collect the data needed to sufficiently resolve uncertainties in the evidence. The guidance (including section 1.3) has been amended to make this clearer.
6	13	Professional organisation	1.4	1.4 – 'impact on radiation dose' is not clear. Do you mean 'impact of AI autocontouring on radiation dose delivered to the tumour or OARs'?	Thank you for your comment. Section 1.4 of the guidance has been amended to read: 'impact *of AI autocontouring* on radiation dose *to organs at risk (OAR) and the tumour*'.
7	1	NHS trust	Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes, again given the pace of change, but there needs to be a mechanism to allow new entrants into the field technology to be included	Thank you for your comment. The recommendations in the EVA guidance are not intended to limit development or use of other new technologies that may enter the field. The evidence generation plan associated with the recommendations will be published on the NICE

					website alongside the EVA guidance. The EVA guidance will be reviewed after the 3-year evidence generation period or sooner if sufficient evidence is available and may consider and include new technologies developed within this period.
8	2	NHS trust	Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes	Thank you for your comment.
9	4	Company	Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes	Thank you for your comment.
10	5	Patient organisation	Are the recommendations sound and a suitable basis for guidance to the NHS?	The findings of the research into AI auto-contouring technologies for radiotherapy treatment are promising in terms of reducing the time that clinicians need to spend manually contouring. We support the recommendation that all AI contours must be reviewed and edited as needed by a trained healthcare professional before being used in radiotherapy treatment planning. We would encourage clear communication with patients to ensure that they are fully informed about the ongoing clinician involvement in the planning process, so that confidence in the process is maintained.	Thank you for your comment. Section 3.7 of the guidance has been amended to reflect the importance of discussing the use of AI technologies with patients. It reads: '...the committee considered that the risk of AI autocontouring with healthcare professional review and edits is likely to be low. *People having contouring should be made aware that AI technologies are being used, and the role of healthcare professionals in the*

					radiotherapy treatment planning process should be explained."
11	8	Healthcare professional	Are the recommendations sound and a suitable basis for guidance to the NHS?	mostly - comments included in the relevant sections	Thank you for your comment.
12	9	Professional organisation	Are the recommendations sound and a suitable basis for guidance to the NHS?	We are pleased that NICE is producing guidance on AI auto-contouring systems for Radiotherapy, and welcome this. However, we do have some comments on the guidance, outlined below. Firstly, it would be ideal to have guidelines from NICE that indicate how the accuracy of AI based auto-contours can be evaluated, not only as part of commissioning and clinical implementation, but also as part of ongoing quality assurance and surveillance. For example, guidance relating to performing geometric and dosimetric analyses of auto-contours would be extremely beneficial to ensure consistency in practice across the healthcare setting. In a recent survey, 36 UK centres listed the AI auto-contouring software that they were using clinically. With the release of the NICE guidance, several centres will be using software that is currently not recommended in the guidance and are awaiting UKCA approval mark. We feel that there needs to be some guidance from NICE regarding the forward plan for NHS hospitals currently using non-NICE approved software. Furthermore, as this is a rapidly developing technology and the assessment period is 3 years, limiting recommendations to the products specified in the guidance may have the unintended consequence of stifling both innovation and competition in the market. We would recommend that there is a mechanism whereby the products listed in the guidance could be expanded as soon as they become available and meet the assessment criteria in the guidance.	Thank you for your comment. Guidelines on how to evaluate the accuracy of AI autocontours are outside the scope of this EVA guidance. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including appropriate outcome measures and methods of data collection. This plan will be published on the NICE website with the final EVA guidance. The recommendations in the EVA guidance are not intended to limit development or use of other new technologies that may enter the field. The EVA guidance will be reviewed after the 3-year evidence generation period or sooner if sufficient

					evidence is available and may consider and include new technologies developed within this period.
13	11	Company	Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes, please refer back to my response to the first question with regards to and ammendment to the recommendations.	Thank you for your comment. Please see response to comment 25.
14	13	Professional organisation	Are the recommendations sound and a suitable basis for guidance to the NHS?	Yes	Thank you for your comment.
Technologies (n=11 comments)					
15	2	NHS trust	2.1	This section says, '...sometimes the target volumes...'. Some pelvic target volumes, particularly those extending into the para-aortic LNs, take a very long time (extended field arm of PEARLS trial is ~5hrs), so chosen system would ideally contour targets 'all' of the time.	Thank you for your comment. Some AI technologies contour OAR and target volumes while others contour OAR only. Details of the AI technologies can be found in Section 2.2 of the guidance and the final scope on the NICE website. NHS organisations should consider their specific contouring needs when deciding which AI technology to use.
16	2	NHS trust	2.2	I feel less is more here. Over 200 structures purported by Limbus seems excessive. Good quality, requiring minimal editing for the most frequently treated and time-consuming sites should be the priority. These include H&N, pelvic/PALN volumes and bowel OAR. A good bladder volume but poor bowel will not save any time	Thank you for your comment. The number and types of structures contoured varies across the AI technologies included in the guidance. Details of the technologies

					can be found in Section 2.2 of the guidance and the final scope on the NICE website. NHS organisations should consider their specific contouring needs when deciding which AI technology to use.
17	7	Patient organisation	2.1	We would suggest including further detail here on the type of MRI images used. For example, can this technology be adapted to bi-parametric and multi-parametric MRIO technologies. Is it flexible as flexibility may be important. Moreover, we have found a widespread availability of multi-parametric MRI across regions of UK. Therefore, being specific about what type of MRI images was used would help with translatability to NHS setting.	Thank you for your comment. Detail on the type of MRI images used by AI technologies is outside the scope of this EVA guidance. NHS organisations should discuss the type of MRI images used with companies of specific technologies when deciding if they should be used in their organisation.
18	10	Company	2.1 “CT or MRI”	and CBCT?	Thank you for your comment. This has been added to Section 2.1.
19	7	Patient organisation	2.2	Limbus Contour is one of the technologies where multiple sites and imaging modalities can be used. Given that there are issues with data availability and openness from companies holding this technology (pointed out by EAGR), what are the minimum standards required in terms of data access and use from multi-site contouring technologies? Also, what are ramifications of data being close ended or not easily adaptable from a cost perspective and future proofing perspective? Would NICE want to support a technology with open source which can be adapted to various clinical indications, and which can be trained to take in multiple imaging modalities, both current and future ones, which may arise due to innovations? There is a need for a nuanced approach	Thank you for your comment. This EVA guidance assessed the clinical and cost effectiveness of the included AI technologies to determine if they should be used in the NHS while further evidence is generated. The criteria for technologies to be included

				for charity and publicly funded research engaging AI implementation which balances interests of all parties, including companies. Moreover it is imperative that data is high quality and representative for patients with prostate cancer. we are also in agreement with the EAGR in that there should be standards applied for geometric metrics used e.g.DICE or Dice Similarity Coefficient (measure of analysis between two sets of image segmentation data) Furthermore, open protocols for how data (current and future) will be stored locally or centrally would be an important way to consider storage issues and ease of access to past imaging results.	in the assessment is outlined in the final scope on the NICE website. NHS hospitals and trusts should have appropriate information governance policies for using AI technologies. Sections 1.1 and 3.8 of the guidance also states that AI technologies should have national and local DTAC approval before being used in clinical practice. DTAC brings together legislation and good practice in areas of clinical safety, data protection, technical security, interoperability and usability and accessibility. More information on DTAC can be found on NHS England's DTAC website.
20	7	Patient organisation	2.2 "The criteria for including technologies in this assessment are in the topic scope on the NICE website. Nine technologies have regulatory approval for use in the NHS"	What are NICE's guiding principle for AI specific technology? We would want to see appraisal of such new technology with equality, data security and clinical implications to be tied to larger frameworks at National or European level.	Thank you for your comment. NICE early value assessments are conducted in line with NICE's interim process and methods for early value assessment. This involves extensive stakeholder and expert engagement and consideration of national and international guidelines

					relevant to the class of technologies being assessed. More information on early value assessment process and methods can be found on the NICE website. NICE is also playing a leading role in the AI and digital regulations service which supports the development and adoption of AI and data-driven technologies in health and social care. Guidance and advice for developers and adopters can be found on the service website.
21	10	Company	2.2 “standalone software”	My belief was that it's integrated into the Siemens CT scanning process, and doesn't work on other systems, in which case it's not standalone. Happy to be wrong on this.	Thank you for your comment. Technology descriptions in the guidance and final scope were provided by the companies for each included technology and supported by relevant documentation including instructions for use. No changes have been made to the guidance.
22	10	Company	2.2 “classified as class IIa”	I know that some systems are currently class I. Why the difference?	Thank you for your comment. Difference in regulatory classification may be due to how individual companies assess and apply the classification to their device

					and when regulatory approval was completed, for example under EU medical devices directive (MDD) versus EU medical devices regulation (MDR). The Medicines and Healthcare products Regulatory Agency (MHRA) advised that most AI technologies will likely be classified as class IIa or higher under the MHRA guidance on software as a medical device. The government has extended the transition period for CE marked devices in the UK. Companies should contact the MHRA if they are unsure of the regulatory requirements for their technology.
23	4	Company	Has all of the relevant evidence been taken into account?	Yes, however I have two minor amendments regarding the Philips MRCAT: On page 7 of the consultation document (4th bullet point): 1. "It provides automatic contours of the prostate" – please add "associated organs at risk" as there are other co-located organs that may be included. 2. "MRCAT Prostate plus Auto-contouring (Philips) is a clinical application integrated in the Philips Ingenia system for MRI in radiation therapy". Please amend "Philips Ingenia" to "Philips MR-RT Systems" (the technology is not only associated with Ingenia).	Thank you for your comment. These changes have been made to Section 2.2 bullet point 7 of the guidance.
24	6	Company	Has all of the relevant evidence been taken into account?	From a regulatory approval perspective, we have received an update since the initial consultation. We have submitted our technical documentation to our notified body, BSI. BSI have informed us that our review is scheduled for September. Based on this, we anticipate the CE Marking review process to conclude by the end of 2023. We noted that one other vendor is also pending with an expected timeline of 2023, so we hope that this update will allow Radformation's	Thank you for your comment. All technologies listed in Section 1.1 of the guidance currently have regulatory approval. Technologies that are awaiting CE or UKCA

				AutoContour to be included in the list of recommended technologies from the start of the project.	mark approval are encouraged to contact NICE if this is completed within 6 months of guidance publication. Section 1.1 of the guidance may then be updated. No change has been made to the guidance.
25	11	Company	Has all of the relevant evidence been taken into account?	Since starting the process of this MIB, a collaboration has been signed between TheraPanacea and Brainlab for ART-Plan to be offered through the Brainlab Elements Treatment planning system. The driving technology behind the AI contouring is the ART-Plan solution, but it is worth highlighting that this is now also available through Brainlab Elements as well as TheraPanacea and Oncology Systems (which is already noted on the document). We are about to start the process of upgrading all of our customers with extra-cranial contouring to the TheraPanacea solution and once done, the majority of hopsitals who will have that solution available to them will be via Brainlab Elements.	Thank you for your comment. Sections 1.1 and 2.2 of the guidance have been amended to read: 'ART-Plan (TheraPanacea, Oncology Systems; *Brainlab*)'
Current management and guidance (n=7 comments)					
26	10	Company	2.3	This document is stating external beam RT. There is also the clear case for brachytherapy use as this also involves contouring, and AI autocontouring systems are in use for this.	Thank you for your comment. The population for this EVA guidance was determined through extensive scoping and consultation with relevant stakeholders and clinical experts. While this guidance focuses on the use of AI technologies to aid contouring in people having external beam radiotherapy, it does not preclude the use of AI technologies to aid contouring in other populations.
27	8	Healthcare professional	2.4	May also be carried out by medical physicists	Thank you for your comment.

					Medical physicists would be included under the broader category of clinical technologists. No change has been made to the guidance.
28	8	Healthcare professional	2.4	Unaware of RCR OAR contouring guidance - does that exist? We are currently implementing the Global Harmonization Group consensus guidance for OAR naming and definition in the UK (Mir R, et al. Organ at risk delineation for radiation therapy clinical trials: Global Harmonization Group consensus guidelines. Radiother Oncol. 2020 Sep;150:30-39) through clinical trials, ProKnow and national SABR Guidelines. Might be a good idea to mention these?	Thank you for your comment. Section 2.4 of the guidance has been amended to read: 'There are guidelines *and consensus statements on* contouring from organisations such as the European Society for Radiotherapy and Oncology, *the Global Quality Assurance of Radiation Therapy Clinical Trials Harmonization Group* and the Royal College of Radiologists.'
29	8	Healthcare professional	2.4	It is critical that users know what national/international guidance is being used to train these systems as they may well not be in line with what is recommended in the UK. The other aspect is consideration for clinical trials. The UK RT community readily participate in clinical trials, which will include their own protocols for target and OAR delineation. Centres must be aware of differences between these recommendations and their software generated volumes	Thank you for your comment. NHS organisations that are interested in using AI autocontouring are encouraged to ask companies about the specific guidelines used to train their technology as well as information on training datasets.
30	13	Professional organisation	2.4	2.4 – the RCR guidance linked at the end is '…radiotherapy target volume…'. The comma between 'radiotherapy' and 'target' is incorrect and changes the meaning.	Thank you for your comment. This has been amended.

31	8	Healthcare professional	2.5	The RCR is currently working on a guidance document for use of AI for OAR and target definition. Will this be mentioned in the NICE guidance document?	Thank you for your comment. Section 2.4 of the guidance references contouring guidelines from the organisations such as the RCR. This includes any future guidelines or consensus statements on AI autocontouring. No change has been made to the guidance.
32	13	Professional organisation	3.5	3.5 final bullet. ‘Speciality training doctors’ is a more appropriate term than ‘registrars’.	Thank you for your comment. This has been amended.
Clinical evidence and effectiveness (n=23 comments)					
33	7	Patient organisation	2.2	The use of AI technology requires data from multiple studies and populations. Are there any considerations made for UK population representation in these studies?	Thank you for your comment. The EVA guidance recommends that 9 AI technologies can be used in the NHS while further evidence is generated. Further evidence generation will be done alongside the use of these technologies in the NHS and will include a UK population. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including the population, outcome

					measures and methods of data collection. This plan will be published on the NICE website alongside the final EVA guidance.
34	2	NHS Trust	3.1 “They said that healthcare professionals have reported finding it easier to review and edit AI autocontours than to create contours from scratch.”	We have been able to briefly review MVision in our department. I noted that bowel contours were quite poor and actually would've benefitted from being started from scratch.	Thank you for your comment. The use of these technologies while more evidence is generated will help to identify which structures may be most suitable for AI autocontouring. This EVA guidance will be reviewed after the 3-year evidence generation period (or sooner if sufficient evidence is available) to include this evidence and make a recommendation on the routine adoption of these technologies in the NHS.
35	2	NHS trust	3.2 “AI technologies may also produce smoother contours of 3D structures compared with manual contouring.”	Most organic structures are smooth in nature. However, smooth does not always correspond to correct.	Thank you for your comment. This was identified by clinical experts as a potential benefit of AI autocontouring. All AI autocontours must be reviewed and edited as needed before use in radiotherapy treatment planning. This should detect and correct any errors in AI autocontours. No change has been made to the guidance.

36	10	Company	3.2 “smoother contours of 3D structures compared with manual contouring”	Should it therefore be noted that the resulting 3D shape is probably more realistic than with manual contours?	Thank you for your comment. Clinical experts advised that AI technologies may produce smoother contours of 3D structures but did not comment on whether this was more realistic than manual contours. No change has therefore been made to the guidance.
37	3	Company	3.3	We would like to report a factual inaccuracy regarding the design of the Ginn 2023 study for AI-Rad Companion Organs RT: indeed, this study reported on a prospective cohort of 20 patients in which the time savings of AI auto-contouring versus manual contouring were evaluated.	Thank you for your comment. Section 3.3 has been amended to read: ‘• 8 prospective studies (DLCExpert, Limbus Contour, MIM Contour ProtégéAI and MRCAT Prostate plus Auto-contouring) • *4* retrospective studies, (INTContour, MVision Segmentation Service, OSAIRIS, RayStation) • *1 mixed retrospective and prospective study (AI-Rad Companion Organs RT)* • 2 conference abstracts (ART-Plan, AutoContour).’
38	10	Company	3.3	Please describe *why* you chose prospective vs. retrospective vs. abstract studies. It's important to show the reason for different systems being treated differently here, so stating that you chose a specific study for each manufacturer, and that this choice was based on a preference for prospective vs. retrospective, vs abstracts.	Thank you for your comment. The rational for study selection is outlined in section 7.2 of the external assessment group (EAG)’s assessment report on the

					NICE website. A link to this document has been added to Section 3.3 of the guidance.
39	10	Company	3.3	It would be helpful to show how much evidence variou products have.	Thank you for your comment. The evidence landscape for each technology is presented in table 3 of the EAG's assessment report on the NICE website. A link to this document has been added to Section 3.3.
40	2	NHS trust	3.4	The structures described here are very quick to contour, so I would not worry AI struggles with this.	Thank you for your comment.
41	2	NHS trust	3.4 "One expert estimated that for head and neck structures, about 90% to 95% of AI autocontours would be accurate."	In terms of time-saving, this is great news. However, how important is the 'inaccurate' 5-10%? Hopefully, peer-review should resolve these.	Thank you for your comment. All AI autocontours must be reviewed and edited as needed before use in radiotherapy treatment planning. This should detect and correct any errors in AI autocontours. The EVA guidance recommends further evidence generation in the NHS while the technologies are being used. This will help inform the assessment of the effectiveness and usefulness of the technologies in clinical practice when the guidance is reviewed.
42	8	Healthcare professional	3.4	Abdominal OARs are often mis-contoured too and, in particular, where abdominal compression is used	Thank you for your comment.

					The use of these technologies while more evidence is generated will help to identify which structures may be most suitable for AI autocontouring. This EVA guidance will be reviewed after the 3-year evidence generation period (or sooner if sufficient evidence is available) to include this evidence and make a recommendation on the routine adoption of these technologies in the NHS.
43	8	Healthcare professional	3.5	Clinical acceptability also depends on intent. Palliative vs radical RT; conventional vs SABR delivery. In SABR, every mm counts due to the very high dose gradients and high doses per fraction delivered over very few fractions. The usual concept that contouring variations will come out in the wash over a 30 fraction treatment no longer holds with treatments being delivered in 3 or 5 fractions, for example. Dose gradients are very steep and planning is pushed to OAR tolerance so delineation needs to be a lot more precise than when delivering conventional treatments	Thank you for your comment. The use of these technologies while more evidence is generated will help to identify when AI autocontouring may be most accurate and suitable for use. This EVA guidance will be reviewed after the 3-year evidence generation period (or sooner if sufficient evidence is available) to include this evidence and make a recommendation on the routine adoption of these technologies in the NHS.
44	1	NHS trust	Has all of the relevant evidence been taken into account?	Given the fast pace of change for this technology sufficient evidence has been take into account	Thank you for your comment.

45	2	NHS trust	Has all of the relevant evidence been taken into account?	Yes	Thank you for your comment.
46	5	Patient organisation	Has all of the relevant evidence been taken into account?	The evidence covers a broad selection of cancer types, including cervical and other gynaecological cancers. The evidence base appears strong and highlights the potential cost and time saving benefits of this technology, whilst also recognising its limitations and the importance of ongoing, careful review and oversight by clinicians. We would like to see more research and data into the potential benefits of this technology in reducing the impact of radiation toxicity and Pelvic Radiation Disease.	Thank you for your comment. Section 3.15 of the guidance notes that in the longer term, evidence on patient outcomes such as radiation toxicity could become available. But these have not been included as key outcomes because it was considered unlikely that they would be collected within the 3-year evidence generation period of this EVA guidance. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including evidence of adverse effects of treatment and dosimetric analysis. This plan will be published on the NICE website alongside the final EVA guidance.
47	9	Professional organisation	Has all of the relevant evidence been taken into account?	There are some recent references that should be consulted. For example, there has been an extensive comparison of different systems by Doolan et al 2023, https://doi.org/10.3389/fonc.2023.1213068, which tests the efficacy of 5 of the systems investigated here. It is also stated in the final protocol document that the quantitative measure of DICE similarity metric is used, however, some studies (such as	Thank you for your comment. The EAG conducted their search of the clinical and economic evidence in May

				Vaassen et al 2021, https://doi.org/10.1016/j.phro.2019.12.001.) recommend that other quantitative metrics should be utilised. The document also discusses the need for more evidence to be generated regarding the impact of AI autocontouring systems on radiation dose. Dose changes cannot be assessed directly, as a change in contour would require the treatment planner to create a new plan. The guidance highlights that the training datasets used to create the AI model have the risk of inducing bias in the models. This bias has recently been investigated (https://doi.org/10.1016/j.radonc.2023.109747) and should remain under close review.	2023. The EVA guidance therefore does not include any evidence published after this time. Doolan et al. (August 2023) compared 5 AI technologies to manual contouring, with findings seeming to be consistent with the evidence reviewed for this EVA. This EVA guidance will be reviewed after the 3-year evidence generation period (or sooner if sufficient evidence is available) to include new evidence and make a recommendation on the routine adoption of these technologies in the NHS. The EAG's assessment report identified 4 main outcomes from the clinical evidence, specifically geometric analysis, dosimetric analysis, qualitative assessment and timesaving. DICE was the most common geometric metric reported in the evidence but was not the only quantitative measure used. Details of geometric analysis in the evidence is reported in Appendix D of the EAG's assessment report which can be found on the NICE website.
48	10	Company	Has all of the relevant evidence	We are happy that your have taken into account all relevant evidence we have submitted.	Thank you for your comment.

			been taken into account?		
49	13	Professional organisation	Has all of the relevant evidence been taken into account?	Yes	Thank you for your comment.
50	1	NHS trust	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	Yes	Thank you for your comment.
51	3	Company	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	Yes, with the exception of a factual inaccuracy listed below in a comment on section 3.3.	Thank you for your comment. Please see response to comment 37.
52	4	Company	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	Yes	Thank you for your comment.
53	8	Healthcare professional	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	Yes	Thank you for your comment.
54	9	Professional organisation	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	There are some potential benefits that have not been listed in the guidance. In particular, these systems will enable a greater consistency in contouring, and will allow the generation of contours for anatomical structures that are not routinely contoured, which could potentially lead to new insights.	Thank you for your comment. The potential for AI technologies to improve consistency of contours was reported in Section 3.2 of the guidance. This section has been amended

					to include other potential benefits: “One expert said that AI technologies helped improve how they were defining structures and may produce smoother contours of 3D structures than manual contouring. *AI technologies may also produce contours for structures not routinely contoured in standard care. This could improve treatment planning and quality of care*.’
55	11	Company	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	yes	Thank you for your comment.
Cost and resource use (n=12 comments)					
56	2	NHS trust	Section 1 textbox ‘Managing the risk of early access’	Why such a discrepancy between technology price?	Thank you for your comment. Technology prices are determined by the individual companies.
57	2	NHS trust	Section 1 textbox ‘Managing the risk of early access’	A simple prostate volume, including target, OARs and review can be done in 45 mins. Therefore, using the highest priced AI would need to reduce the time taken to zero to be cost neutral.	Thank you for your comment. NHS organisations should consider their contouring needs when deciding to use a specific AI technology. The NICE resource impact assessment team has

					produced a resource impact tool which is available on the NICE website. This can be used to help NHS organisations determine the potential resource impact and/or cost savings from using AI technologies in their organisation. NHS organisations are encouraged to discuss technology costs with companies when deciding if a specific technology is appropriate for use.
58	10	Company	3.4	Surely there is also an important "resource only" impact? As the NHS is understaffed, anything that involves less time for a process will provide greater staff bandwidth into other aspects of their roles. Regardless of cost, this becomes significant when there aren't enough staff and a hospital is unable to recruit.	Thank you for your comment. Section 3.1 of the guidance outlines the potential value of AI autocontouring in releasing staff pressures including allowing more time for patient-facing tasks. No change to the guidance has been made.
59	2	NHS trust	3.5 “The clinical experts estimated time saving of 10 minutes to 30 minutes depending on the amount of editing needed.”	If cost savings are priority (presumably they are as this is NHS), these savings suggest dismissing the highest priced technologies from the discussion.	Thank you for your comment. The clinical evidence showed timesaving ranging from 3 minutes to 80 minutes. Clinical experts advised that the potential time saving will depend on the complexity of contours and the degree of edits needed. NHS organisations should consider their

					contouring needs when deciding to use a specific AI technology. The NICE resource impact assessment team has produced a resource impact tool which is available on the NICE website. This can be used to help NHS organisations determine the potential resource impact and/or cost savings from using AI technologies in their organisation.
60	8	Healthcare professional	3.5	staff groups involved in OAR contouring vary across hospitals, with dosimetrists, clinical technologists and medical physicists also involved a part from radiographers or registrars	Thank you for your comment. Radiographers and registrars were an example to demonstrate the impact of timesaving and staff grade on cost savings. Costs and resource use will vary depending on who is doing the contouring. The NICE resource impact assessment team has produced a resource impact tool which is available on the NICE website. This can be used to help NHS organisations determine the potential resource impact and/or cost savings from using AI technologies and can be adjusted to include different staff groups including clinical technologists,

					dosimetrists and consultants.
61	2	NHS trust	3.6 “band 7 radiographer (£65 per hour based on PSSRU 2021)”	I'm a radiographer at the top of band 7. My hourly rate is £25 per hour! If £65 per hour is the figure being used to quantify cost savings, these need to be revised.	Thank you for your comment. The EAG used healthcare professional costs from the Personal Social Services Research Unit (PSSRU 2021). These costs include total expenditure incurred to produce one unit of output in health and social care, such as the cost of one hour of GP time. PPSRU costs are usually higher than just the salary cost because they include overheads such as training. Costs reported in the EVA guidance may therefore differ from NHS agenda for change or hourly rates. PSSRU costs are routinely used in NICE guidance development. The NICE resource impact assessment team has produced a resource impact tool which is available on the NICE website. This can be used to help NHS organisations determine the potential resource impact and/or cost savings from using AI technologies and can be adjusted to include different staff groups and technology costs.

62	10	Company	3.6	What about the differential between the lowest cost staff member using the £50 per plan software and the highest cost staff member using the £4 per plan software? That shows the variability even more.	Thank you for your comment. The EAG's assessment showed that potential cost saving of AI technologies was impacted by technology costs, timesaving and healthcare professional grade. The NICE resource impact assessment team has produced a resource impact tool which is available on the NICE website. This can be used to help NHS organisations determine the potential resource impact and/or cost savings from using AI technologies and can be adjusted to include different staff groups and technology costs.
63	3	Company	3.9	We recommend for the guidance to recognise that skill loss is a secondary risk compared to that of staff shortages, which is indeed addressed by the technologies being assessed, through time savings, in that shortages could result in increased waiting times and delayed access to radiotherapy for patients with cancer.	Thank you for your comment. Please see response to comment 58.
64	2	NHS trust	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	No. Document describes band 7 radiographer costing £65 per hour. Correcting this will change cost savings quite dramatically.	Thank you for your comment. Please see response to comment 61.
65	5	Patient organisation	Are the summaries of clinical and cost effectiveness reasonable	The summaries of clinical and cost-effectiveness recognise that the technology would need to save a significant amount of time in order to result in financial savings. When gynaecological cancers were auto-contoured using AI, half of the contours needed to receive minor or major adjustments before treatment (Coughlan et al). It is important to	Thank you for your comment. The use of these technologies while more

			interpretations of the evidence?	factor in the time a clinician needs to spend adjusting AI auto-contours when identifying clinical and cost effectiveness, though – as identified by NICE - healthcare professionals have reported finding it easier to review and edit AI auto-contours than to create contours from scratch. It will be important to ensure that adequate review time is ensured and protected for clinicians. Ongoing evaluation will be needed to establish whether certain cancer types are better suited to AI auto-contouring.	evidence is generated will help to identify which structures may be most suitable for AI autocontouring. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including timesaving across different anatomical structures. This plan will be published on the NICE website alongside the final EVA guidance. This EVA guidance will then be reviewed after the 3-year evidence generation period (or sooner if sufficient evidence is available) to include this evidence and make a recommendation on the routine adoption of these technologies in the NHS.
66	10	Company	Are the summaries of clinical and cost effectiveness reasonable interpretations of the evidence?	I have noted in specific responses that the business effectiveness is heavily weighted to financial only. There is little comment on how effective AI contouring technologies can be to relieving some of the staff-shortage related workload. We wholly agree that a solution should be cost effective, however the NHS also needs to note the benefit of what a technology enables Trusts to achieve that they would not manage otherwise.	Thank you for your comment. Please see response to comment 58.
67	13	Professional organisation	Are the summaries of clinical and cost effectiveness reasonable	Yes, but cost-effectiveness estimates focus on radiographer time not oncologist time. The EVA acknowledges AI can produce contours for OARs which are usually contoured by radiographers or target volumes (often lymph node regions) which are usually contoured by clinical oncologists. Target volumes usually take longer to contour than OARs.	Thank you for your comment. Please see response to comment 60.

			interpretations of the evidence?	The cost-saving of AI autocontouring may therefore be greater where AI is used to contour target volumes (more time saved, clinical oncologists' time more expensive than radiographers).	
Further evidence generation (n=3 comments)					
68	3	Company	1.4	On the proposed outcome 'adverse events associated with AI autocontouring and contouring errors': We recommend for the committee to note that adverse events (AEs) experienced by patients are indeed distinct from contouring errors, and argue that it would be impractical to document AEs associated with AI auto-contouring in the proposed evidence generation plan. It is very unlikely that AI auto-contouring errors would result into AEs, because contouring is only an initial step in radiotherapy treatment planning and because contours are always reviewed and used by a multidisciplinary team of qualified health professionals for the purpose of treatment planning. Contours would not be used by these professionals unless they do meet appropriate quality standards, and therefore contouring errors can only result in a potential loss of the time savings associated with auto-contouring, in the cases when these professionals have to spend time correcting these errors or re-contouring before using the contours for treatment planning. In addition, we recommend to the Committee to note that contours are a precursor to actual radiotherapy treatment plans, which involve trade-offs and potentially compromises between planned radiation doses to the target volume(s) (i.e. tumours) versus doses to organs at risk. Clinical expertise and decision-making is essential in defining a treatment plan, with some degree of variation depending on the intent of treatment, patient characteristics, etc. As a result, the relationship between the quality of contours and patient-experienced AEs is only indirect. For these reasons, we recommend for the Committee to (1) consider removing 'adverse events' from the evidence generation plan for these technologies, and (2) consider that the incidence and extent of contouring errors would be best captured by the time saving outcome elsewhere recommended in the evidence generation plan.	Thank you for your comment. The committee considered that the risk of using AI technologies was low because all AI autocontours must be reviewed and edited as needed by a trained healthcare professional. However, it is still important to report and record adverse events from using AI technologies compared with standard care during the evidence generation period. This is standard practice for evidence generation in the NHS and was an important committee consideration. In Section 1.4 bullet point 5, contouring errors was intended to be presented as a separate outcome to adverse events associated with AI autocontouring. For clarity, this has been amended to read: '*contouring errors* and adverse events associated with AI autocontouring'.
69	3	Company	1.4	On the proposed outcome 'impact on radiation dose': As commented regarding AEs below, it should be noted by the Committee that contours are a prerequisite for radiotherapy treatment	Thank you for your comment.

				plans, and that the latter may vary between clinicians and according to the intent of radiotherapy. Evidence on the impact on radiation dose of using AI auto-coutouring could therefore only be reliably developed in a 1:1 comparison of AI-derived auto-contours versus manual contours for a given set of patients. However, this would not seem practical considering the draft evidence generation plan considered for these technologies, comparing outcomes between NHS providers using versus not using AI auto-contouring. Furthermore, this outcome may not be relevant for decision-making, because the relationship between the quality of contours and radiation doses to target volume(s) or organs at risk is unclear. For instance, an algorithm that would systematically overestimate contours by some margin around organs at risk would, generally and all things being equal, be expected to result in artificially flattened dose-volume histograms, ie a theoretically lower exposure of OARs, without any impact on doses actually delivered. Conversely, if these overestimated OAR contours were indeed used for treatment planning, they could result in lower target volume doses being necessary to spare artificially larger OARs. Although appropriate contours are indeed necessary to generate an optimal treatment plan, it does not seem practical to infer the quality of contours from changes in observed radiation doses, as these may result from variations in patient characteristics, clinical guidelines, therapy intent, other technological progress for radiotherapy planning or delivery systems, or other variations associated with healthcare professionals. We therefore recommend for this outcome to be removed from the proposed evidence generation plan.	The clinical evidence included dosimetric analysis as one of the 4 main outcomes, alongside geometric analysis, qualitative assessment and timesaving. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including appropriate outcome measures and methods of data collection. This plan will be published on the NICE website alongside the final EVA guidance. No change has been made to the guidance.
70	7	Patient organisation	2.5	Would it be possible to request data which compares auto segmentation/contouring to manual contouring in terms of oncological control and treatment outcomes? Is AI more likely to maximize efficacy from OAR targeted treatment? Can we tie to 5 year survival outcomes? linking to long term outcome data is a key way for us to assess impact on patients as a result of new technological implementation	Thank you for your comment. Section 3.15 of the guidance notes that in the longer term, evidence on patient outcomes such as radiation toxicity could become available. But these have not been included as key outcomes because it was considered unlikely that they would be

					collected within the 3-year evidence generation period. This section has been amended to add survival outcomes to the patient outcomes that may become available in the longer term.
Implementation and managing risks (n=7 comments)					
71	12	Company		**Recommendations** We propose some possible areas for improvement or clarification in the guidance: The guidance could address how AI technologies could be integrated into existing workflows and systems, and what challenges or barriers may exist for implementation. Some examples or best practices from early adopters of AI technologies in radiotherapy could be useful. The guidance could discuss how AI technologies could be updated or retrained over time to reflect changes in guidelines, imaging protocols or patient populations. This could affect the performance and validity of AI models, and may require some mechanisms or standards to ensure quality assurance. Other trends that could be considered include: • Vendors offering clear insights into their algorithm training and validation processes, such as the demographics, geographic regions, manufacturers of the imaging equipment, and diversity of cases (including rare ones) covered by the datasets. • Groups advocating for governance frameworks with clear parameters for auditing, quality checks, and oversight. • Public reporting via clear dissemination of performance audits/attributes and validation outcomes to improve trust.	Thank you for your comment. Additional considerations for implementing AI technologies to aid contouring in the NHS can be found in the assessment report, overview and final scope on the NICE website. Implementation considerations for AI technologies more generally are also included in the supporting documentation on the NICE website. Section 3.12 of the guidance states that companies should provide information on training datasets as part of their product information pack. While the other trends mentioned are important to consider, they fall outside the scope of this guidance. No change to the guidance has therefore been made.
72	7	Patient organisation	Section 1 text box	We would suggest re-wording this sentence as it implies that individual trusts should create a framework for AI technologies. Rather, we would	Thank you for your comment.

			‘Managing the risk of early access’ “Local NHS hospitals and trusts should have appropriate information governance policies for using AI technologies.”	recommend there being a nationwide or Integrated Care Board level framework so that the responsibility does not fall on individual Trusts. Although technologies and in particular access to patient data may be locked in by Trusts, I think general information governance should be shared (best practice) more widely/nationally.	Information governance policies for AI technologies should be both at the national and local level. For clarity, Section 1 text box has been amended to remove reference to ‘local’ NHS hospitals and trusts.
73	7	Patient organisation	2.6 “Comparators may also include no contouring for cases when AI technologies may produce contours for structures not routinely contoured in standard care.”	Past surveys on AI from [redacted] have indicated that some of our supporters would like to see human oversight being part of the process. would it be possible to develop a protocol that clearly defines how much time will be spent on use of AI vs manual to enable safe use of technology? Also, would AI tech be able to have a feature that recommends instances where it is appropriate for manual contouring/AI contouring?	Thank you for your comment. This EVA guidance focuses on the use of AI technologies to aid contouring with healthcare professional review. This means that all AI autocontours must be reviewed by a trained healthcare professional and edited as needed before being used in radiotherapy treatment planning. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies including appropriate outcome measures and methods of data collection. This will likely include evidence on time for healthcare professional review and edits as outlined in Section 1.4. The plan will

					be published on the NICE website alongside the final EVA guidance.
74	10	Company	3.7 “The committee concluded that there should be ongoing reporting of any errors in AI autocontouring and adverse events associated with these technologies”	Healthcare providers are responsible for checking all contours before approving for clinical use. This covers this.	Thank you for your comment. Please see response to comment 68.
75	2	NHS trust	3.9 “In the future, more widespread use of these technologies could result in a skill loss in the workforce.”	I believe the opposite - I think this is an exciting opportunity to be involved with implementing and the ongoing management of AI.	Thank you for your comment.
76	8	Healthcare professional	3.9	Very much disagree. Editing a structure is not the same as learning to contour it yourself on many different anatomies. It is very likely there will be skill loss unless radiotherapy centres maintain training programmes for registrars (and other clinical staff). It is very likely that automation bias will be introduced and with loss of skill there will be less and less editing being done resulting in further skill loss... I think autocontouring is indeed necessary but we must recognise skill loss	Thank you for your comment. Section 3.9 of the guidance has been amended to read: “In the future, more widespread use of these technologies could result in a skill loss in the workforce. Clinical experts advised that healthcare professionals would nearly always do some editing as part of their review of the autocontours. *The committee considered that it is important for healthcare professionals to develop and to maintain contouring skills so they could*

					adequately review and edit AI autocontours. Some technologies provided training packages for healthcare professionals to *develop and* practice their skills."
77	9	Professional organisation	Are the recommendations sound and a suitable basis for guidance to the NHS?	In terms of the 'managing risk' section of the guidance, we feel that there are some changes that should be made. The guidance highlights that deskilling of the workforce is a risk, but that this is unlikely to be significant. However, we feel that the risk is greater than implied in the guidance. As time passes, the workforce may be unable to produce manual contours to the same skill level, and may not have the appropriate knowledge to understand how to correct the contours. Potentially, the workforce may come to trust the AI contours more than any manual contours. Furthermore, it is possible that automatically generated contours are 'waved through' by busy, over-stretched staff, without the proper review/editing process. This could lead to automation bias becoming a high, ongoing risk and the guidance should highlight how this risk can be managed. Additionally, there is a data security risk through the sharing of patient data with a third party, which is not identified in the guidance and should be highlighted. Furthermore, there is a large number of AI auto-contouring businesses, and this number is growing. We are concerned that some companies may not continue to trade, due to being small, start-up enterprises. Therefore, there is a risk of a department implementing an AI solution and becoming reliant on this solution for patient workflows, and then the company ceases to trade. The department would need to decommission one AI auto-contouring solution, and then select a new one and implement it. We feel this risk should be highlighted as something to consider when selecting and implementing an AI auto-contouring solution.	Thank you for your comment. Please see response to comment 76. Digital health technologies are rapidly developing which may result in changes in the technology marketplace. This falls outside the scope of this guidance. No changes to the guidance have therefore been made.
Equality considerations (n=16 comments)					
78	12	Company		We applaud the inclusion of equality considerations and the awareness of possible algorithmic bias in AI models depending on the training population. We agree with the suggestion to provide information on training datasets and to evaluate the performance of AI technologies in local populations.	Thank you for your comment.

				We value the National Institute for Health and Care Excellence's (NICE) recognition of open-source solutions, which aligns with NHSE's published commitment to increase the use and sharing of open-source code in the NHS.	
79	2	NHS trust	3.11 "Experts advised that there may be a lack of representation of female pelvis"	Those technologies have potentially missed a trick here. Female pelvis are one of the most time-consuming volumes we have to do. Appropriate representation should be considered in ongoing development/teaching of the models.	Thank you for your comment.
80	7	Patient organisation	3.11	Ethnicity should also be a consideration as well as for controlling for other biases as described by Equality of Opportunity Difference metric. This requires adversarial training method to reduce bias. Concept explained by OECD: https://oecd.ai/en/catalogue/metrics/equality-of-opportunity-difference-eod	Thank you for your comment. The committee carefully considered all the equality issues related to the use of these technologies in the NHS. It considered that AI technologies may be subject to algorithmic bias because the population demographic used in the training may differ from the population in clinical practice. The clinical experts advised that accuracy of AI autocontouring was most likely to be affected by age or anatomical differences from the training dataset. Section 3.14 bullet point 1 has been amended to highlight the need for more information on population demographics. It reads: "The committee considered that the demographics of datasets used for training

					an algorithm may differ from populations in clinical settings. It highlighted the need for evidence generation in how AI technologies work in clinical practice in local populations, *including information on population demographics such as age, sex, disability and ethnicity*"
81	10	Company	3.12	A more positive outlook in these cases is that, with AI Autocontouring covering the majority of patients well, the timesavings allow the healthcare professionals to spend more thorough time on those unusual patient cases. In this mindset, AI contouring provides increased resource to work on the difficult patients.	Thank you for your comment. This potential benefit is stated in Section 1 (textbox) of the guidance.
82	10	Company	3.13	We would advise against training bespoke local trust data as it can jeopordise one of the goals of AI contouring which is to increase consistency. Any input contours must be curated before use.	Thank you for your comment. Clinical experts extensively discussed the potential benefits and risks of training bespoke local trust data. The committee considered that in the future AI technologies could be trained on a representative national population.
83	10	Company	3.13 "Companies should provide information on training datasets as part of their product information pack"	Difficut to require as this is considered by most to be proprietary information. Regulatory requirements for this are already met. Is NICE proposing additional regulation?	Thank you for your comment. The committee considered that information on AI technology datasets, such as population demographics, would help NHS organisations to decide if a technology is

					suitable for use in their local population. Several companies agreed that this information could be provided as part of the technology information pack. This is not related to regulatory requirements and does not require companies to share proprietary information.
84	10	Company	3.13 “Ideally in the future AI technologies could be trained on a representative national population.”	Fair enough as an "ideal", however the level of agreement required between competing parties in choosing the input data would be very hard to achieve in practice. Additionally who would be responsible for curating the data quality and consistency that would allow high-enough quality input data?	Thank you for your comment. These are important considerations. Details of how this would be achieved falls outside the scope of this guidance.
85	10	Company	3.14 “to ensure no specific group is disadvantaged”	This is an unfair comment. I refer to my comment on the advantage AI contouring provides in that it enables healthcare professionals greater ability to focus on the less common patient populations when they present for treatment. Are you suggesting that it is a disadvantage for a child, young person or person with atypical anatomy to need more focussed attention from the healthcare professionals?	Thank you for your comment. This statement has been removed from Section 3.14.
86	1	NHS trust	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	It willl always need to be modelled separately for male and female, not necessarily just in the pelvis Thre may be a reason on some sites to separate by ethnicity, and the ethnicity mix of the training population should be disclosed	Thank you for your comment. Please see response to comment 80.
87	2	NHS trust	Are there any equality issues that need special consideration and are not covered in	Yes. It sounds like more female pelvic and atypical anatomical data needed.	Thank you for your comment.

			the medical technology consultation document?		
88	4	Company	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	No	Thank you for your comment.
89	5	Patient organisation	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	As mentioned in the Equality impact assessment, AI models can contain algorithmic bias depending on the population used in training. Populations used in training datasets may not be representative of patient populations in clinical practice which can cause potential age, sex, and disability bias. NICE highlights that more evidence needs to be generated on: • clinical acceptability of contours and amount of edits needed • impact on radiation dose • time saving including time for healthcare professional review and edits • resource use defined by healthcare professional grade and time • adverse events associated with AI auto contouring and contouring errors. We would encourage the monitoring and recording of these data to include information on the age, sex, race, and disability status of the patients involved. This may help identify where – if any – there are patterns of outcomes which are related to patient demographics. In the Multiple Technologies Evaluation Programme, very few of the technologies have data on treating cervical cancer. We would encourage concerted efforts to identify or commission further studies into using AI intelligence to aid auto-contouring for radiotherapy treatment for cervical cancer.	Thank you for your comment. Please see response to comment 80. The NICE data and analytics team is developing an evidence generation plan based on the guidance which will guide further evidence generation for these technologies. This will likely include evidence across anatomical sites and cancers. The plan will be published on the NICE website alongside the final EVA guidance.

90	8	Healthcare professional	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	no	Thank you for your comment.
91	9	Professional organisation	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	We feel that there needs to be more transparency from the manufacturers regarding the characteristics of the populations used to train auto-contouring algorithms. We understand that there is an element of patient confidentiality to consider, however there needs to be a broad indication of the range of patient characteristics within population datasets. A lack of transparency on whether minority and other under-represented groups have been included in training and validation has the potential to impact public trust in AI auto-contours. We suggest that the guidance in 3.13 is amended, such that information on the demographics of population datasets used in the development of the AI algorithm must be provided within their product information pack. Ideally, this information should indicate characteristics such as age range, gender ratios, race and inclusion of disabilities. This would allow NHS Trusts and other healthcare providers to ensure that cancer models used in these systems are relevant to their patient populations. Furthermore, there needs to be a mechanism to permit equity of access to this software during this period of evaluation. Some NHS Trusts are able to fund the implementation of this technology, however this may not currently be possible for other Trusts. Consideration should be given towards a funding stream for Trusts willing to take up this technology, as this will enable all patients and organisations to benefit. This will also ensure that the required evidence is generated rapidly and without introducing bias, which may occur if only a small and unrepresentative sample of Trusts use this technology during this period of evaluation.	Thank you for your comment. Section 3.13 has been amended to read: “Technology developers or companies should provide information on training datasets as part of their product information pack, *including demographics of population datasets*.”
92	11	Company	Are there any equality issues that need special consideration and are not covered in	No	Thank you for your comment.

			the medical technology consultation document?		
93	13	Professional organisation	Are there any equality issues that need special consideration and are not covered in the medical technology consultation document?	No	Thank you for your comment.