



Contouring in transition: perceptions of AI-based autocontouring by radiation oncologists and medical physicists in German-speaking countries

Samuel M. Vorbach^{1,9} · Florian Putz^{2,9} · Ute Ganswindt¹ · Stefan Janssen^{3,9} · Maximilian Grohmann^{4,9} · Stefan Knippen^{5,9} · Felix Heinemann^{6,9} · Rami A. El Shafie^{7,9} · Jan C. Peeken^{8,9,10,11}

Received: 18 November 2024 / Accepted: 31 March 2025
© The Author(s) 2025

Abstract

Background Artificial intelligence (AI)-based autocontouring software has the potential to revolutionize radiotherapy planning. In recent years, several AI-based autocontouring solutions with many advantages have emerged; however, their clinical use raises several challenges related to implementation, quality assurance, validation, and training. The aim of this study was to investigate the current use of AI-based autocontouring software and the associated expectations and hopes of radiation oncologists and medical physicists in German-speaking countries.

Methods A digital survey consisting of 24 questions including single-choice, multiple-choice, free-response, and five-point Likert scale rankings was conducted using the online tool umfrageonline.com (enuvo GmbH, Pfäffikon SZ, Switzerland).

Results A total of 163 participants completed the survey, with approximately two thirds reporting use of AI-based autocontouring software in routine clinical practice. Of the users, 92% found the software helpful in clinical practice. More than 90% reported using AI solutions to contour organs at risk (OARs) in the brain, head and neck, thorax, abdomen, and pelvis. The majority (88.8%) reported time savings in OAR delineation, with approximately 41% estimating savings of 11–20 min per case. However, nearly half of the respondents expressed concern about the potential degradation of resident training in sectional anatomy understanding. Of respondents, 60% would welcome guidelines for implementation and use of AI-based contouring aids from their respective radiation oncology societies. Respondents' free-text comments emphasized the need for careful monitoring and postprocessing of AI-delivered autocontours as well as concerns about overreliance on AI and its impact on the development of young physicians' contouring and planning skills.

Conclusion Artificial intelligence-based autocontouring software shows promise for integration into radiation oncology workflows, with respondents recognizing its potential for time saving and standardization. However, successful implementation will require ongoing education and curriculum adaptation to ensure AI enhances, rather than replaces, clinical expertise.

Keywords Artificial intelligence · Questionnaires · Autocontouring · Perceptions · Radiation oncology

✉ Samuel M. Vorbach
Samuel.Vorbach@i-med.ac.at

¹ Department of Radiation Oncology, Medical University of Innsbruck, Innsbruck, Austria

² Department of Radiation Oncology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

³ Department of Radiation Oncology, University of Lübeck, Lübeck, Germany

⁴ Department of Radiotherapy and Radiation Oncology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

⁵ Department of Radiation Oncology, Helios Hospitals Schwerin, Schwerin, Germany

⁶ Department of Radiation Oncology, University Medical Center Freiburg, Freiburg, Germany

⁷ Radiotherapy and Radiation Oncology, University Medical Center Göttingen, Göttingen, Germany

⁸ Department of Radiation Oncology, Klinikum rechts der Isar, Technical University of Munich (TUM), Munich, Germany

⁹ Digitalization and Artificial Intelligence Working Group, German Society for Radiation Oncology, Berlin, Germany

¹⁰ Partner Site Munich, German Consortium for Translational Cancer Research (DKTK), Munich, Germany

¹¹ Institute of Radiation Medicine (IRM), Helmholtz Zentrum München (HMGU), Munich, Germany

Introduction

The rapid development of artificial intelligence (AI) offers great promise in various aspects of healthcare, from drug discovery to cancer diagnosis [1–3]. Within the field of radiation oncology, AI has the potential to transform several aspects of the radiotherapy workflow [4, 5].

A critical step in radiation oncology is the precise graphical delineation of the target volumes to be irradiated and the organs at risk (OARs) to be spared. Errors in this process can have serious consequences: insufficient tumor contouring can lead to underdosing, which can affect tumor control and survival; conversely, overestimation of the target volume or inadequate contouring of OARs can lead to excessive radiation exposure to healthy tissue, which increases the risk of toxicity and impairs the patient's quality of life [6, 7].

Manual contouring, traditionally performed by radiation oncologists, is time consuming [8] and subject to inter- and intraobserver variability both within and across radiotherapy centers [9–12]. Early approaches to autosegmentation mainly used conventional techniques such as intensity analysis, shape modeling, and atlas-based methods [13]. These then-innovative approaches are limited in terms of availability, accuracy, and adaptability, especially for anatomical variations and different cancer types, and still require significant manual effort [14–17].

The emergence of deep learning techniques, particularly convolutional neural networks (CNNs), has led to significant improvements in the performance of autocontouring algorithms. These multilayer feed-forward neural networks and, in particular, specialized variants based on the U-Net architecture, are optimized for the segmentation of image datasets. Their architecture consists of a contracting and an expanding path. The encoder or contracting path in a U-Net is responsible for progressively downsampling the input image to capture high-level features. At the deepest layer, often referred to as the bottleneck, the image is represented in its most abstract form. The decoder or expansion path then gradually upsamples this representation to generate a detailed segmentation map. Skip connections between corresponding layers of the encoder and decoder allow the model to transfer spatial and contextual information from the input image, thereby preserving important macroscale details in the final segmentation map [18].

As AI-based autocontouring tools move from research projects into routine clinical use, several vendors now offer commercially available AI-based autocontouring software tools [19, 20]. Their integration into clinical routine has represented a remarkable advancement. Recent studies have shown the benefits of AI-based autocontouring software for various tumor sites: the authors highlighted the potential to increase efficiency (time savings when contouring OARs)

[21, 22] as well as to reduce contouring and dose inconsistencies, thereby contributing to standardization and quality assurance [23].

However, the introduction of AI-powered autocontouring software into clinical practice also raises significant concerns about quality assurance, education and training, potential deskilling of clinicians, and overreliance on automated systems as well as perceived risks and barriers to implementation.

Few studies have attempted to capture the views of radiation oncologists on these factors. For example, Zhai et al. [24] developed and tested a model to investigate factors influencing the acceptance of AI contouring technology in China. At the same time, Hindocha et al. [25] conducted a survey among clinical oncologists in the UK in which 78% reported that AI would have a positive impact on radiation oncology. The presented study aimed to assess the perceptions of radiation oncology professionals in German-speaking countries regarding AI-based autocontouring software and its current use, making it the first survey of its kind in this region.

Methods

To assess the current use and potential benefits and risks of AI-based autocontouring software, the survey questions were grouped into different categories. The categories included sociodemographic data, such as age, gender, or country of residence, as well as more specific data such as experience in using AI-based autocontouring software, the need for an implementation guideline, or the perception of AI-based autocontouring software as a potentially dangerous tool. All respondents first answered the questions on sociodemographic data. All participants who reported already using an AI-based autocontouring software in clinical routine answered questions about their experience with the software. Those who do not use AI-based autocontouring software solutions skipped these questions and were redirected to questions about their general perceptions and opinions on AI-based autocontouring software.

The anonymous survey included single- and multiple-choice questions, five-point Likert scale questions (scale points ranging from 1=disagree to 5=fully agree), and the possibility to add further comments in a free-text box (see the Appendix for the full questionnaire). The questions were selected in a multistep process by members of the Digitalization and Artificial Intelligence Working Group of the German Society for Radiation Oncology (DEGRO). The commercially available online survey tool *umfrageonline.com* (enuvo GmbH, Pfäffikon SZ, Switzerland) was used for this study. The corresponding survey link was sent out via the professional mailing lists of the DEGRO and

the Austrian Society for Radiation Oncology (ÖGRO). The survey was open from June 11 to August 3, 2024. An initial invitation and a reminder email were sent during this period. Participation in the survey was both voluntary and anonymous, and all respondents agreed to publication of the study results. Ethical approval was not required for an anonymized questionnaire without patient data.

Data analysis

Raw data were obtained directly from the online tool umfrageonline.com (in Excel V16; Microsoft, Redmond, WA, USA) and then exported to SPSS Statistics (V26; IBM Corporation, Armonk, NY, USA). Responses were analyzed using descriptive statistics. The Wilcoxon rank-sum test was used to compare responses on ordinal scales between the two subgroups. Nominal data were analyzed using the chi-squared test. When the expected number of observations in more than 20% of the cells was less than 5, Fisher's exact test was used. A value of $p < 0.05$ was considered as statistically significant. To account for multiple testing in pairwise comparisons, Bonferroni correction was applied. Free-text comments were reviewed by two authors (SMV, JCP) and grouped into topic categories for further analysis.

Results

Study participants

The DEGRO professional mailing list contained 1440 email addresses and the ÖGRO mailing list contained 282 email addresses. After removing duplicate addresses, nonfunctional addresses, and those belonging to retired colleagues, 1568 valid email addresses were identified. Of these, 188 participated in the survey, resulting in a response rate of 12.0%. In total, 163 out of 188 (86.7%) questionnaires were completed in full. Personal characteristics of these participants are summarized in Table 1. Overall, 89.6% of the participants were physicians, of whom more than 80% were specialized in radiation oncology. Of the participants, 85 (52.2%) were female and 46.0% were between 30 and 49 years of age. The current place of employment was a university hospital, a non-university hospital, an ambulatory health center, and a medical practice for 84 (51.5%), 23 (15.3%), 34 (20.9%), and 20 (12.3%) respondents, respectively, with Germany being the primary country (67.5%) of employment.

Current use of AI-based autocontouring software

In 114 responses (69.9%), physicians are responsible for contouring the OARs, with an almost even split between

Table 1 Characteristics of the participants

Characteristic	Participants
<i>Profession</i>	
Resident	28 (17.2%)
Radiation oncology specialist	118 (72.4%)
Physicist	17 (10.4%)
<i>Age (years)</i>	
20–29	7 (4.3%)
30–39	38 (23.3%)
40–49	37 (22.7%)
50–59	53 (32.5%)
≥ 60	28 (17.2%)
<i>Gender</i>	
Female	85 (52.2%)
Male	78 (47.8%)
<i>Country in which profession is practiced</i>	
Germany	110 (67.5%)
Austria	49 (30.0%)
Switzerland	4 (2.5%)
<i>Institution type</i>	
University hospital	84 (51.5%)
Non-university hospital	25 (15.3%)
Ambulatory health center	34 (20.9%)
Medical practice	20 (12.3%)

residents ($n = 54$) and radiation oncology specialists ($n = 60$). While 107 respondents (65.7%) indicated that AI-based autocontouring software is already in use in clinical routine, 22.1% of respondents do not use it and do not currently plan to implement it. The proportion of respondents reporting AI-based autocontouring software use varied across institutional types: 72.6% (61 users of 84 respondents) in university hospitals, 36.0% (9 of 25) in non-university hospitals, 64.7% in ambulatory health centers (22 of 34), and 75.0% (15 of 20) in medical practices. Chi-squared test revealed a significant association between institution type and the use of AI-based autocontouring software ($p = 0.006$), with significantly more respondents of university hospitals, ambulatory health centers, and medical practices reporting its use in clinical routine ($p < 0.001$) compared to those reporting from non-university hospitals.

The top three AI-contouring-based software products used comprised Limbus AI Inc. (30.8%), three slightly differing software solutions of Siemens Healthineers (27.1%), and ART-Plan. One respondent indicated that they use Elements (Brainlab AG), which is an atlas-based and not an AI-based autocontouring software. The total list of eight AI-based autocontouring software solutions used by the 107 user respondents are listed in Table 2. More than half of them (54.2%) have been using the software for 1–3 years, while 36.5% have been using it for less than a year. In 71% of all cases, the software runs on a local server. Be-

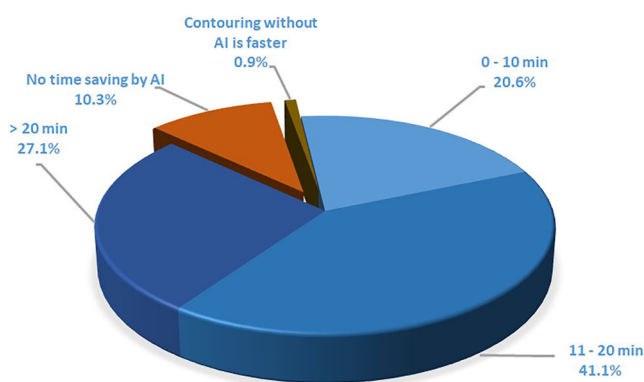
Table 2 State of artificial intelligence-based autocontouring software

Characteristic	
<i>Responsible for OAR contouring are</i>	
Residents	54 (33.1%)
Radiation oncology specialists	60 (36.8%)
Physicists	11 (6.8%)
Radiation therapists	38 (23.3%)
<i>Current use of AI-based autocontouring software</i>	
Clinical routine	107 (65.7%)
Research	3 (1.8%)
None, implementation planned	17 (10.4%)
None, not planned to be implemented	36 (22.1%)
<i>AI-based autocontouring software used*</i>	
MVision (MVision AI Oy, Helsinki, Finland)	11 (10.3%)
Limbus AI (Limbus AI Inc., Regina, SK, Canada)	33 (30.8%)
ART-Plan (TheraPanacea, Paris, France)	16 (15.0%)
AI Rad Companion Organs RT/Syngo.via/Direct Organs (Siemens Healthineers AG, Erlangen, Germany)	29 (27.1%)
Ray Station (RaySearch Laboratories AB, Stock- holm, Sweden)	9 (8.4%)
Contour Protégé AI (MIM Software Inc., Cleve- land, OH, USA)	2 (1.9%)
Elements (Brainlab AG, Munich, Germany)	1 (0.9%)
Name not recalled	6 (5.6%)
<i>AI-based autocontouring software is operated*</i>	
On a local server	76 (71.0%)
Cloud-based	31 (29.0%)
<i>Duration of use of AI-based autocontouring* software</i>	
< 1 year	39 (36.5%)
1–3 years	58 (54.2%)
> 3 years	10 (9.3%)
<i>Evaluation of different products before purchase*</i>	
Yes	54 (50.5%)
No	53 (49.5%)
<i>For which OARs are AI-based autocontouring software used?*, **</i>	
Brain	98 (91.6%)
Head and neck	100 (93.5%)
Thorax	104 (97.2%)
Abdomen	101 (94.4%)
Pelvic	102 (95.3%)
Spine	87 (81.3%)
<i>Do you use automatically generated structures for CTV definition?*</i>	
Not possible	47 (43.9%)
No benefit	36 (33.7%)
Saves time in the contouring process	24 (22.4%)

OAR organs at risk, CTV clinical target volume

*100% corresponds to 107 answers

**multiple answers possible

**Fig. 1** Estimation of time savings per organ at risk contouring with artificial intelligence (AI)-based autocontouring software

fore purchasing an AI-based autocontouring software, half (50.5%) of the respondents had tried several products. Artificial intelligence-based autocontouring was used for OAR contouring of the brain, head and neck, thorax, abdomen, and pelvis by more than 90% of the respondents. 56.1% of the participants dispose of an AI solution that offers automatic segmentation of target volumes. However, only 40% of these users stated that it actually saves time in the contouring process. The responses regarding current utilization of AI-based autocontouring software are summarized in Table 2.

Time savings in OAR delineation were reported by 88.8% of participants, with only one person (0.9%) stating that the use of AI-based autocontouring software increased the time required to delineate OARs. The time saved using AI-based autocontouring software was estimated to be between 11 and 20 min per case for 41.1% of the respondents, while 27.1% described time savings of even more than 20 min (Fig. 1). There were no significant differences in terms of time savings depending on the profession (specialists vs. residents, $p=0.162$; and physicians vs. physicists, $p=0.917$) or on the type of institution (university hospitals vs. other, $p=0.299$). Finally, there was no significant difference in the time saved during contouring depending on the type of software used (Limbus AI vs. other, $p=0.825$; Siemens Healthineers software solutions vs. other, $p=0.213$; and ART-Plan vs. other, $p=0.541$).

The frequency with which AI-based OAR contouring can be accepted without correction varies. For example, 16.8% of respondents stated that a correction is required in four out of five cases, while 11.2% stated that 80–100% of self-segmented OARs can be accepted without correction (Fig. 2). Around a quarter of respondents (26.2%) estimated that no correction is required in 41–60% of cases, while a further quarter (27.1%) stated that 61–80% of self-segmented OARs can be accepted without manual correction. There were no significant differences in the acceptance rate of

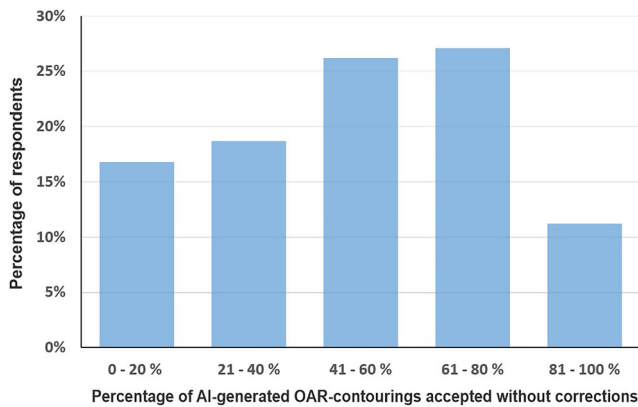


Fig. 2 Proportion of respondents reporting different acceptance rates of artificial intelligence (AI)-autosegmented organs at risk (OARs) without manual correction

AI-based OARs for different professional groups (specialists vs. residents, $p=0.111$; and physicians vs. physicists, $p=0.989$) or types of institution (university hospitals vs. other, $p=0.322$). There were no significant differences in the acceptance rate of OARs of different AI software solutions between the respective user groups (Limbus AI vs. other, $p=0.142$; Siemens Healthineers software solutions vs. other, $p=0.161$; and ART-Plan vs. other, $p=0.426$).

A minority of respondents (24.3%) stated that using their current software solution reduces the need for staff in their department. Conversely, a significant proportion (64.5%) rather disagreed or disagreed with this statement, indicating widespread skepticism about its impact on human resources. However, 92% of respondents agreed that the AI-based autocontouring software is helpful in clinical practice.

Of all respondents who use an AI solution, 87.7% do not consider it dangerous in principle. When asked whether

they believe that use of an AI-based autocontouring software will lead to greater standardization and quality assurance, 79.1% of participants agreed or somewhat agreed.

However, almost half (46.6%) rather agreed or agreed that the increasing use of AI-based autocontouring software will degrade resident training in understanding sectional anatomy. Even with the increasing use of AI-based autocontouring software solutions, the vast majority (93.2%) do not believe that the *raison d'être* of radiation oncologists will be threatened. Guidelines for the implementation and use of AI-based autocontouring software solutions provided by the respective radiation oncology societies would be welcomed by 60.1% of participants (Fig. 3).

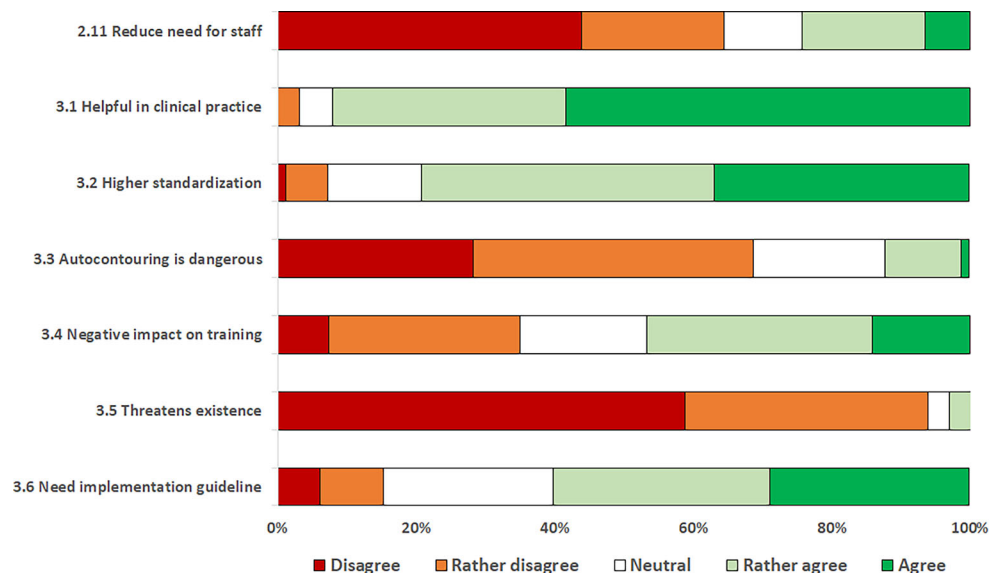
Perceptions of the AI-based autocontouring software were homogeneous within the physician professional group: no significant difference was found between specialists and residents across any of the Likert scale responses (p -values ranging from 0.116 to 0.810).

Physicists, on the other hand, were more likely than physicians to agree that the use of AI-based autocontouring software will decrease the demand for radiation oncologists in the future, even if they still tend to disagree with this statement overall (median Likert scaling [LS]: 2 vs. 1, $p=0.037$). In addition, physicists are more likely to agree that the increasing use of AI-based autocontouring tools will degrade resident training in understanding sectional anatomy compared to physicians (median LS: 4 vs. 3, $p=0.037$).

Participants working in Germany are more likely than participants from other countries to state that the increasing use of AI-based autocontouring software will degrade resident training in understanding sectional anatomy (median LS: 4 vs. 3, $p<0.001$).

Although agreement was high in both groups, participants working at a university hospital were more likely than those working at a non-university institution to agree that

Fig. 3 Perceptions of artificial intelligence (AI)-based autocontouring software based on a five-item Likert scale. For graphical purposes, the questions were paraphrased. The original questionnaire is reported in the Appendix



the respective radiation oncology societies should establish guidelines for the implementation and use of AI-based autocontouring software (median LS: 4 vs. 4, $p=0.015$).

We asked respondents to comment on their views or experiences with AI-based autocontouring software in a free-text box. Free-text comments revealed several key aspects, including the need for careful oversight and manual correction of AI-based contours as well as concerns about the high cost of implementation and its potential impact on education and training. While the respondents acknowledged the potential of AI to improve efficiency and standardization, they also highlighted risks such as overreliance on automation and its potential impact on the skill development of younger clinicians. A detailed analysis of the free-text responses is provided in the Appendix.

Discussion

This study offers insights into the perspectives, current applications, and obstacles of AI-based autocontouring software in radiation oncology across DEGRO and ÖGRO members. The results support the implementation and acceptance of AI-based autocontouring software, which is of utmost importance given the huge potential of AI in radiation oncology.

Previous studies have highlighted various aspects of AI-based autocontouring. Zhai et al. [24] reported on a self-developed model to assess the acceptance of AI-based autocontouring software in China. Among the 307 respondents, technical resistance was low, and the overall perception of AI was high. However, current usage, fears, and expectations were not captured in this study, which may be due to the fact that almost 60% of the respondents had not yet used AI-based autocontouring software and almost a quarter of the respondents were still medical students. Mugabe [26] reported the views of a multidisciplinary group including 15 radiation oncologists from New Zealand. However, they focused on the impact of AI in general, and only 35% reported using AI tools for autosegmentation. Brouwer et al. reported on the perception of 213 medical physicists from 202 radiation oncology centers across Europe regarding AI applications in general [27]. Wong et al. [28] reported on perceptions of Canadian radiation oncologists, physicists, radiation therapists, and radiation trainees regarding the general impact of AI. To date, only two surveys have been conducted that truly focus on AI-based autocontouring software: Hindocha et al. [25] reported on the responses of 51 clinical oncologists in the UK and Bourbonne et al. [20] reported on the French perspective of young radiation oncologists (85% residents).

In our survey, 65.7% of respondents reported using AI-based autocontouring software in routine clinical practice,

which is higher than the 45% of respondents in the UK survey [25] and closer to the 60.7% of French respondents [20]. Like our study, these two surveys are not representative because they report on different numbers of respondents per center and did not include all centers in each country, so there are no data on the true prevalence of AI use.

Nevertheless, these studies suggest that the clinical use of AI technologies in radiation oncology is still in its early stages. A study conducted in New Zealand [26] reported that “AI usage was low” but overall, respondents had “a high likelihood to adopt AI.” Similarly, nearly 90% of Turkish radiation oncologists surveyed believed that adopting AI would improve their work [29]. While there is optimism regarding the potential of AI, several barriers and concerns might slow the widespread adoption of AI in clinical routine. A key challenge might be represented by the lack of AI expertise. One survey reported that a quarter of radiation oncologists rated their knowledge of AI as “very poor” and 94% expressed a need for further training [29]. The study conducted in New Zealand identified lower familiarity with AI as a barrier, which correlated with a lower intention to use AI.

In addition, a Canadian survey found that while most healthcare professionals recognize the potential of AI to improve patient care, concerns about the threat of job displacement and about changing professional roles contribute to some reluctance [28]. Addressing these psychological barriers is critical; raising awareness of AI as a collaborative tool, rather than a replacement threat, can help foster trust and acceptance. Rosenbacke et al. highlight that so-called explainable AI (XAI), which provides clear, clinically relevant explanations, increased clinicians’ trust. Their findings emphasize the nuanced role of comprehensive explanations [30]. Consequently, one of the most important strategies for driving AI adoption is adequate education and training, as explicitly requested by 94% of surveyed radiation oncologists [29]. Professional workshops and hands-on training may thus help to demystify AI tools. An acceptance study conducted in China reported that clinicians are more likely to adopt AI if they believe it will significantly improve patient care or their workflow efficiency [24]. Successfully integrating AI into radiation oncology requires addressing both technical limitations and human factors. Overcoming skepticism requires a multifaceted approach; education, training, functional transparency, and guided institutional support are all crucial to promote AI adoption. While early adopters pave the way, late adopters can gain confidence as the benefits of AI become increasingly evident in clinical practice.

In our study, AI-based autocontouring was reportedly used in over 90% of cases for OAR contouring of the brain, head and neck, thorax, abdomen, and pelvis, compared to only 43–67% in the UK survey [25], suggesting

that AI-based autocontouring is now increasingly used. An overwhelming 88.8% of our participants reported time savings in OAR delineation, with 41.1% estimating savings of 11–20 min per case and 27.1% reporting even greater time savings of over 20 min. These results are comparable to the 88.7% of young French radiation oncologists [20] who reported savings of 25–100% in segmentation time, highlighting the great potential of AI for revolutionizing the time-consuming task of manual segmentation.

While AI-based autocontouring has demonstrated significant time savings in OAR delineation, its application to target volume segmentation remains limited. In our study, only 56.1% of participants had access to an AI solution capable of automatic target volume segmentation. More notably, among those who used such a software, only 40% reported actual time savings in the contouring process. These findings underscore a critical limitation: despite advancements in AI-based autocontouring, its effectiveness and efficiency in target volume segmentation are still lacking. The fact that only about half of the respondents have access to an AI solution for this task—and the fact that the majority of users do not experience meaningful time savings—highlights an unmet need for more reliable and clinically useful AI-driven target volume segmentation tools. Further development and validation of AI models tailored to target volume contouring are necessary to fully harness the potential of automation in radiotherapy planning. Irrespective of these limitations, an overwhelming 92% of all respondents already consider AI-based autocontouring software solutions helpful, underscoring the technology's perceived value and its promising role in clinical practice.

Given the widespread appreciation of AI-based software benefits, it is not surprising that respondents who added free-text comments highlighted its potential for improving the clinical workflow, addressing staffing shortages, and facilitating the implementation of advanced technologies such as adaptive planning. Others advocated the expansion of AI applications into additional areas of clinical practice. However, while the integration of AI-based autocontouring software has been largely well received, it is important to recognize its limitations and potential risks. Accordingly, respondents raised concerns about quality assurance, education, and training, and warned of the potential deskilling of clinicians and overreliance on automated systems (for more detailed analysis of free-text commentaries, see Appendix B).

To address these challenges, 60% of respondents would welcome guidelines for the implementation and use of AI-based autocontouring software solutions. Indeed, already in 2020, Vandewinckele et al. published recommendations for implementation and quality assurance regarding AI-based applications in radiotherapy [19]. They recommend, as one of our interviewees also noted, the formation of a dedi-

cated multidisciplinary team to ensure safe and appropriate AI use and to educate the entire team on the use and limitations of AI-based autosegmentation. They proposed a two-stage workflow: in the “commissioning phase,” the AI model should be evaluated using an internal dataset. During the “implementation and quality assurance phase”, the implementation team should train and educate all future users in the correct application and interpretation of AI output. Ongoing documentation of the necessary changes, regular meetings between the implementation team and users, and regular quality assurance (QA) of AI output performance following successful implementation have been recommended. Importantly, specific QA runs should specifically address changes in the overall imaging workflow, e.g., after changes in CT scanners or acquisition protocols, as suggested by Vandewinckele et al. In parallel to our study, Hurkmans et al. elaborated “A joint ESTRO and AAPM guideline for development, clinical validation and reporting of artificial intelligence models in radiation therapy” in 2024 [31]. They emphasize the difficulty of validating AI-based segmentation, especially since defining a gold standard or ground truth segmentation is challenging. They also recommend that once an appropriate ground truth has been established, a qualitative (e.g., Likert scale) and a quantitative (e.g., Hausdorff distance [32]) metric should be used as well as a time trial to evaluate the usefulness of the model.

In our view, both reports address in detail aspects relevant to reliably developing and clinically validating AI models, e.g., by implementation of skilled teaching and quality control teams. In light of the very recently published cohesive guideline by the joint European and American expert group [31], the development of a valid and reliable work guide for the implementation of clinically used AI tools, as desired by the majority of our study participants, has already made encouraging progress. We thus further encourage all clinicians already using AI-based autocontouring software solutions to share their experiences and concerns in existing and newly formed national and international expert panels. The thereby-supported continuous improvement of consensus guidelines will then help radiation oncologists considering the implementation of such automated tools in their clinical routine and ensure widespread acceptance and safe implementation of AI-based autocontouring software.

Limitations

An important limitation of online surveys is response bias: those who favor AI may be more likely to complete the questionnaire. Responses are inherently self-reported and may not reflect the true usage of AI-based autocontouring. The survey was not designed to be representative in terms of providing a complete documentation of the use of AI-

based autocontouring solutions in German-speaking radiotherapy clinics and practices, so it is possible that several respondents reflect clustered experiences and the opinions of larger centers. Thus, our study does not provide representative data on the prevalence of actual AI use and acceptance. In addition, topics related to the data security of AI-based autocontouring software solutions were not explicitly addressed in the questionnaire. Cross-professional comparisons are limited by the very unequal numbers of answers from physicists and physicians. Furthermore, although radiation therapists were invited to participate, no DEGRO or ÖGRO representative of this professional group responded, thus further limiting the generalizability of the findings to all relevant professions using AI-based autocontouring software solutions.

Conclusion

Our survey assessment supports the potential for AI-based autocontouring software to become an integral part of the clinical workflow in radiation oncology. While the majority of respondents are positive about AI, especially concerning the achievable time savings, and see its potential for improving standardization, there is a clear need for ongoing education and thoughtful integration of AI tools into clinical practice. As AI continues to evolve, adaptation of core curricula will be crucial to ensure that AI enhances rather than replaces clinical expertise and skills.

Appendix

Appendix A: survey full text

Literal translation from German to English. The original questionnaire was distributed in German.

- 1.1 Which professional group do you belong to? *Drop-down menu.*
 - A: Resident, radiation oncology specialist, physicist, radiation therapist.
- 1.2 How old are you? *Drop-down menu.*
 - A: 20–29 years, 30–39 years, 40–49 years, 50–59 years, ≥ 60 years.
- 1.3 Gender? *Drop-down menu.*
 - A: Female, male, other.
- 1.4 Which country do you work in? *Drop-down menu.*
 - A: Germany, Austria, Switzerland, other (free-text response).
- 1.5 Where do you work? *Drop-down menu.*
 - A: University hospital, non-university hospital, ambulatory health center, medical practice.
- 1.6 Primary contouring of organs at risk (independent of specialist control) is the responsibility of the? *Drop-down menu.*
 - A: Residents, radiation oncology specialists, physicists, radiation therapists.
- 2.1 Do you use AI-based autocontouring software at your workplace? *Drop-down menu.*
 - A: Yes, in clinical routine; Yes, but only for research; No, but implementation is planned within the next year; No, no implementation is currently planned.

Questions 2.2 through 2.11 were asked only of respondents who already use AI-based autocontouring software in routine clinical practice.
- 2.2 Which AI-based autocontouring software do you use? *Drop-down menu.*
 - A: MVision (MVision AI), Limbus AI (Limbus AI Inc.), ART-Plan (TheraPanacea), AI Rad Companion Organs RT (Siemens Healthineers), Ray Station (RaySearch Laboratories), Contour Protégé AI (MIM Software Inc.), Elements (Brainlab AG), other (free text).
- 2.3 How is the AI-based autocontouring software operated? *Drop-down menu.*
 - A: On a local server, cloud-based.
- 2.4 How long have you been using AI-based autocontouring software? *Drop-down menu.*
 - A: < 1 year, 1–3 years, > 3 years.
- 2.5 Did you test different products before purchasing the software solution? *Drop-down menu.*
 - A: Yes, No.
- 2.6 For which OAR do you use AI-based auto-contouring software? *Drop-down menu. Multiple answers possible.*
 - A: Brain, head and neck, thorax, abdomen, pelvic, spine.
- 2.7 How much time do you save on average per contouring by using your autocontouring Software? *Drop-down menu.*
 - A: Extends contouring, None (neither speeds up nor slows down), 0–10 min, 11–20 min, > 20 min.
- 2.8 In how many cases can the autosegmented OARs be adopted without correction? *Drop-down menu.*
 - A: 0–20%, 21–40%, 41–60%, 61–80%, 81–100%.
- 2.9 Do you use automatically generated structures for CTV definition? *Drop-down menu.*
 - A: Not offered by my software solution, No benefit for the contouring time, Saves time during the contouring process.
- 2.10 Is it possible to individualize the CTV definition to the patient's clinical situation with the available software? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.

- 2.11 Do the current software solutions reduce the need for staff in your department? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 3.1 Do you think the use of AI-based autocontouring software is helpful in your clinical practice? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 3.2 Do you think that the use of AI-based autocontouring software leads to higher standardization and quality assurance? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 3.3 Do you consider the use of AI-based autocontouring software to be dangerous? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 3.4 Do you think that using AI-based autocontouring software will degrade resident training in understanding sectional anatomy? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 3.5 Do you think that the increasing use of AI-based software threatens the existence of radiation oncologists? *Drop-down menu.*
 - A: Disagree Rather disagree, Neutral, Rather agree, Agree.
- 3.6 Would you like to see guidelines for the implementation and use of AI-based autocontouring software by the respective radiation oncology societies? *Drop-down menu.*
 - A: Disagree, Rather disagree, Neutral, Rather agree, Agree.
- 4.0 Do you have any comments on this subject?
 - A: Optional free-text answers.

Appendix B: analysis of free-text comments

The free-text comments were evaluated and assigned to six different subject areas, whereby a single comment could contain content that could be assigned to several subject areas.

(I) Control

Respondents expressed the need for careful oversight of AI-generated structures, emphasizing the importance of manual correction by specialists to ensure accuracy ($n=3$). Concerns were raised about the potential for errors if AI output is not properly reviewed, especially given the differences in contouring by different specialists. In addition, some high-

lighted the importance of maintaining standardization and routine corrections to mitigate risk ($n=2$).

(II) Economics

Four respondents highlighted the high cost of AI-based autocontouring software, with some stating that the technology is prohibitively expensive and not yet widely accessible ($n=3$). Others pointed out that while AI has the potential to improve workflow and reduce overtime work, it has not yet led to significant cost savings or a reduced need for labor.

(III) Education and training

Others highlighted the impact of AI on the education and training of medical staff. Concerns were raised about the potential for younger physicians to become overly reliant on AI, leading to a decline in manual skills and critical examination ($n=3$). Some suggested that AI might require a complete restructuring of assistant training programs to ensure that professionals can identify and correct AI errors. Others noted that training guidelines should be developed to ensure effective use of AI in clinical practice. In addition, the role of radiation therapists in supporting contouring was identified as a potential area for workforce development ($n=3$).

(IV) Implementation

Four respondents emphasized the need for proper implementation strategies when introducing AI into clinical workflows. This includes the creation of dedicated teams of physicians, physicists, and IT specialists with the resources to create best practice protocols for the use of AI-based autocontouring. AI was also seen as essential for addressing staffing shortages and incorporating advanced technologies such as adaptive planning. There were calls for the development of guidelines on how AI should be integrated into routine practice, with some acknowledging that AI is already helping to reduce workload and improve work-life balance.

(V) Future outlook

Five respondents were optimistic about the potential of AI, predicting that it might eventually outperform human specialists and provide significant relief as patient numbers continue to grow. The need to fully realize the potential of AI beyond contouring was emphasized, with suggestions to expand its use into different areas of clinical practice. Despite concerns about training, the respondents highlighted AI's ability to improve quality assurance and standardization. AI was also seen as having a positive impact on the

quality of follow-up care and other digital applications. Experienced professionals noted that AI will likely streamline processes in the future, making routine tasks easier and faster.

(VI) Limitations

Respondents acknowledged the current limitations of AI ($n=5$). Some noted that AI still makes serious mistakes and needs to improve significantly before it can be fully relied upon. Concerns were raised that younger professionals might become overly reliant on AI, potentially reducing skill development.

Acknowledgements The authors thank all participants for their willingness to participate and share their experiences and perceptions.

Funding Open access funding provided by University of Innsbruck and Medical University of Innsbruck.

Conflict of interest S.M. Vorbach, F. Putz, U. Ganswindt, S. Janssen, M. Grohmann, S. Knippen, F. Heinemann, R.A.E. Shafie, and J.C. Peeken declare that they have no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Gurmesa DK, Jimma W (2024) Explainable machine learning for breast cancer diagnosis from mammography and ultrasound images: a systematic review. *Bmj Health Care Inform*. <https://doi.org/10.1136/bmjhci-2023-100954>
- Sarkar C, Das B, Rawat VS, Wahlang JB, Nongpiur A, Tiewsoh I et al (2023) Artificial Intelligence and Machine Learning Technology Driven Modern Drug Discovery and Development. *Int J Mol Sci*. <https://doi.org/10.3390/ijms24032026>
- Sharma M, Savage C, Nair M, Larsson I, Svedberg P, Nygren JM (2022) Artificial Intelligence Applications in Health Care Practice: Scoping Review. *J Med Internet Res* 24(10):e40238. <https://doi.org/10.2196/40238>
- Luk SMH, Ford EC, Phillips MH, Kalet AM (2022) Improving the Quality of Care in Radiation Oncology using Artificial Intelligence. *Clin Oncol* 34(2):89–98. <https://doi.org/10.1016/j.clon.2021.11.011>
- Parkinson C, Matthams C, Foley K, Spezi E (2021) Artificial intelligence in radiation oncology: A review of its current status and potential application for the radiotherapy workforce. *Radiography* 1(s8):63. <https://doi.org/10.1016/j.radi.2021.07.012>
- Chen AM, Chin R, Beron P, Yoshizaki T, Mikaelian AG, Cao M (2017) Inadequate target volume delineation and local-regional recurrence after intensity-modulated radiotherapy for human papillomavirus-positive oropharynx cancer. *Radiother Oncol* 123(3):412–418. <https://doi.org/10.1016/j.radonc.2017.04.015>
- Riegel AC, Vaccarelli M, Cox BW, Chou H, Cao Y, Potters L (2019) Impact of Multi-Institutional Prospective Peer Review on Target and Organ-at-Risk Delineation in Radiation Therapy. *Pract Radiat Oncol* 9(2):e228–e35. <https://doi.org/10.1016/j.prro.2018.10.016>
- Teguh DN, Levendag PC, Voet PW, Al-Mamgani A, Han X, Wolf TK et al (2011) Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head and neck. *Int J Radiat Oncol Biol Phys* 81(4):950–957. <https://doi.org/10.1016/j.ijrobp.2010.07.009>
- Growcott S, Dembrey T, Patel R, Eaton D, Cameron A (2020) Inter-Observer Variability in Target Volume Delineations of Benign and Metastatic Brain Tumours for Stereotactic Radiosurgery: Results of a National Quality Assurance Programme. *Clin Oncol* 32(1):13–25. <https://doi.org/10.1016/j.clon.2019.06.015>
- Guzene L, Beddok A, Nioche C, Modzelewski R, Loiseau C, Salleron J et al (2023) Assessing Interobserver Variability in the Delineation of Structures in Radiation Oncology: A Systematic Review. *Int J Radiat Oncol Biol Phys* 115(5):1047–1060. <https://doi.org/10.1016/j.ijrobp.2022.11.021>
- Bridge P, Fielding A, Rowntree P, Intraobserver Variability PA (2016) Should We Worry? *J Med Imaging Radiat Sci* 47(3):217–220. <https://doi.org/10.1016/j.jmir.2016.06.004>
- van der Veen J, Gulyban A, Willems S, Maes F, Nuyts S (2021) Interobserver variability in organ at risk delineation in head and neck cancer. *Radiat Oncol* 16(1):120. <https://doi.org/10.1186/s13014-020-01677-2>
- Erdur AC, Rusche D, Scholz D, Kiechle J, Fischer S, Llorián-Salvador Ó et al (2024) Deep learning for autosegmentation for radiotherapy treatment planning: State-of-the-art and novel perspectives. *Strahlenther Onkol*. <https://doi.org/10.1007/s00066-024-02262-2>
- Huynh E, Hosny A, Guthrie C, Bitterman DS, Petit SF, Haas-Kogan DA et al (2020) Artificial intelligence in radiation oncology. *Nat Rev Clin Oncol* 17(12):771–781. <https://doi.org/10.1038/s41571-020-0417-8>
- Samarasinghe G, Jameson M, Vinod S, Field M, Dowling J, Sowmya A et al (2021) Deep learning for segmentation in radiation therapy planning: a review. *J Med Imaging Radiat Oncol* 65(5):578–595. <https://doi.org/10.1111/1754-9485.13286>
- Cui Y, Chen W, Kong FM, Olsen LA, Beatty RE, Maxim PG et al (2015) Contouring variations and the role of atlas in non-small cell lung cancer radiation therapy: Analysis of a multi-institutional preclinical trial planning study. *Pract Radiat Oncol* 5(2):e67–75. <https://doi.org/10.1016/j.prro.2014.05.005>
- Kim Y, Patwardhan K, Beichel R, Smith B, Ulrich E, Mart C et al (2016) Impact of contouring accuracy on expected tumor control probability for head and neck cancer: semiautomated segmentation versus manual contouring. *Int J Radiat Oncol Biol Phys* 96(2):E545
- Shaukat Z, Farooq QUA, Tu S, Xiao C, Ali S (2022) A state-of-the-art technique to perform cloud-based semantic segmentation using deep learning 3D U-Net architecture. *Bmc Bioinformatics* 23(1):251. <https://doi.org/10.1186/s12859-022-04794-9>
- Vandewinckele L, Claessens M, Dinkla A, Brouwer C, Crijns W, Verellen D et al (2020) Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. *Radiother Oncol* 153:55–66. <https://doi.org/10.1016/j.radonc.2020.09.008>
- Bourbonne V, Laville A, Wagneur N, Ghannam Y, Larnaudie A (2023) Excitement and Concerns of Young Radiation Oncologists over Automatic Segmentation: A French Perspective. *Cancers (basel)*. <https://doi.org/10.3390/cancers15072040>
- Strolin S, Santoro M, Paolani G, Ammendolia I, Arcelli A, Benini A et al (2023) How smart is artificial intelligence in organs delin-

- eration? Testing a CE and FDA-approved Deep-Learning tool using multiple expert contours delineated on planning CT images. *Front Oncol* 13:1089807. <https://doi.org/10.3389/fonc.2023.1089807>
22. Zabel WJ, Conway JL, Gladwish A, Skliarenko J, Didiodato G, Goorts-Matthews L et al (2021) Clinical Evaluation of Deep Learning and Atlas-Based Auto-Contouring of Bladder and Rectum for Prostate Radiation Therapy. *Pract Radiat Oncol* 11(1):e80–e9. <https://doi.org/10.1016/j.prro.2020.05.013>
 23. Thor M, Apte A, Haq R, Iyer A, LoCastro E, Deasy JO. Using Auto-Segmentation to Reduce Contouring and Dose Inconsistency in Clinical Trials (2021) The Simulated Impact on RTOG 0617. *Int J Radiat Oncol Biol Phys* 109(5):1619–1626. <https://doi.org/10.1016/j.ijrobp.2020.11.011>
 24. Zhai H, Yang X, Xue J, Lavender C, Ye T, Li JB et al (2021) Radiation Oncologists' Perceptions of Adopting an Artificial Intelligence-Assisted Contouring Technology: Model Development and Questionnaire Study. *J Med Internet Res* 23(9):e27122. <https://doi.org/10.2196/27122>
 25. Hindocha S, Zucker K, Jena R, Banfill K, Mackay K, Price G et al (2023) Artificial Intelligence for Radiotherapy Auto-Contouring: Current Use, Perceptions of and Barriers to Implementation. *Clin Oncol* 35(4):219–226. <https://doi.org/10.1016/j.clon.2023.01.014>
 26. Mugabe VK (2021) Barriers and facilitators to the adoption of artificial intelligence in radiation oncology: A New Zealand study. *Tech Innov Patient Support Radiat Oncol* 18:16–21. <https://doi.org/10.1016/j.tipsro.2021.03.004>
 27. Brouwer CL, Dinkla AM, Vandewinckele L, Crijns W, Claessens M, Verellen D et al (2020) Machine learning applications in radiation oncology: Current use and needs to support clinical implementation. *Phys Imaging Radiat Oncol* 16:144–148. <https://doi.org/10.1016/j.phro.2020.11.002>
 28. Wong K, Gallant F, Szumacher E (2021) Perceptions of Canadian radiation oncologists, radiation physicists, radiation therapists and radiation trainees about the impact of artificial intelligence in radiation oncology—national survey. *J Med Imaging Radiat Sci* 52(1):44–48. <https://doi.org/10.1016/j.jmir.2020.11.013>
 29. Kirakli E, Gürsel Ş, Yirmibeşoğlu EE, Etiz D, Özyar E (2023) A national survey on the awareness of artificial intelligence technologies among radiation oncologists-Turkish Society of Radiation Oncology Artificial Intelligence and Information Technology Group. Study *Türk Onkoloji Dergisi-turkish J Oncol* 38:
 30. Rosenbacke R, Melhus Å, McKee M, Stuckler D (2024) How Explainable Artificial Intelligence Can Increase or Decrease Clinicians' Trust in AI Applications in Health Care: Systematic Review. *Jmir Ai* 3:e53207. <https://doi.org/10.2196/53207>
 31. Hurkmans C, Bibault JE, Brock KK, van Elmpt W, Feng M, Fuller DC et al (2024) A joint ESTRO and AAPM guideline for development, clinical validation and reporting of artificial intelligence models in radiation therapy. *Radiother Oncol* 197:110345. <https://doi.org/10.1016/j.radonc.2024.110345>
 32. Huttenlocher DP, Klanderman GA, Rucklidge WJ (1993) Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Machine Intell* 15(9):850–863

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.