Dosimetric assessment of an Atlas based automated segmentation for loco-regional radiation therapy of early breast cancer in the Skagen Trial 1

A multi-institutional study

Eldesoky, Ahmed R; Francolini, Giulio; Thomsen, Mette S; Yates, Esben S; Nyeng, Tine B; Kirkove, Carine; Kamby, Claus; Blix, Egil S; Nielsen, Mette H; Taheri-Kadkhoda, Zahra; Berg, Martin; Offersen, Birgitte V

Technical Note

Dosimetric assessment of an Atlas based automated segmentation for loco-regional radiation therapy of early breast cancer in the Skagen Trial 1: A multi-institutional study

Ahmed R. Eldesoky [a,b,1], Giulio Francolini [a,c,*,1], Mette S. Thomsen [d], Esben S. Yates [d], Tine B. Nyeng [d], Carine Kirkove [e], Claus Kamby [f], Egil S. Blix [g], Mette H. Nielsen [h], Zahra Taheri-Kadkhoda [i], Martin Berg [j], Birgitte V. Offersen [a]

[a] Department of Oncology, Aarhus University Hospital, Norrebrogade 44, DK-8000 Aarhus C, Denmark
[b] Department of Clinical Oncology and Nuclear Medicine, Mansoura University, 60 Elgomhoria st, Mansoura, Egypt
[c] Department of Radiation Oncology, University of Florence, Azienda Ospedaliero-Universitaria Careggi, Largo Brambilla 3, 50134 Firenze, Florence, Italy
[d] Department of Medical Physics, Aarhus University Hospital, Norrebrogade 44, DK-8000 Aarhus C, Denmark
[e] Department of Radiation Oncology, Catholic University of Louvain, 10 Ave Hippocrate, B-1200 Brussels, Belgium
[f] Department of Oncology, Rigshospitalet, Blegdamsvej 9, 2100 Copenhagen, Denmark
[g] Department of Oncology, University Hospital of North Norway, Sykehusvegen 38, 9019 Tromsø, Norway
[h] Department of Oncology, Odense University Hospital, Institute of Clinical Research, University of Southern Denmark, Odense, Winslowparken 19, 3, DK-5000 Odense C, Denmark
[i] Department of Oncology, Zealand University Hospital, Sygeshusvej 10, 4000 Roskilde, Denmark
[j] Department of Medical Physics, Hospital of Vejle, Kabbeltoft 25, 7100 Vejle, Denmark

ARTICLE INFO

ABSTRACT

The effect of Atlas-based automated segmentation (ABAS) on dose volume histogram (DVH) parameters compared to manual segmentation (MS) in loco-regional radiotherapy (RT) of early breast cancer was investigated in patients included in the Skagen Trial 1.
This analysis supports implementation of ABAS in clinical practice and multi-institutional trials.
© 2017 The Authors. Published by Elsevier Ireland Ltd on behalf of European Society for Radiotherapy and Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Recently, an ESTRO delineation guideline-dependent atlas based automated segmentation (ABAS) tool for radiation therapy (RT) of early breast cancer using MIM Maestro software has been developed and adopted at Aarhus University Hospital, Denmark [1–3]. This ABAS tool has shown a significant reduction in segmentation time and a high agreement against a gold standard manual

segmentation (MS), helping to overcome issues related to inter-observer variability and workload burden of conventional manual delineation [4]. Additionally, it maintained its reproducibility and robustness in a multi-institutional clinical validation study [3]. The performance of ABAS against MS has been evaluated geometrically using Dice Similarity Coefficient (DSC), Average Hausdorff Distance and difference in volume. However these geometric parameters have limitations [5], and a more relevant dosimetric analysis is needed to consolidate the contribution of this tool in daily routine.

The purpose of this study was to assess if contouring variations between ABAS and MS significantly affect dose parameters. In a multi-institutional setting, the difference in dose coverage between a manually corrected ABAS and MS of CTVs of the primary (CTVp) and nodal (CTVn) volumes in patients eligible for loco-regional RT of early breast cancer in the Skagen Trial 1 was investigated.

* Corresponding author at: Department of Radiation Oncology, University of Florence, Viale Morgagni 85, 50134 Florence, Italy. Fax: +39 0554379930.
E-mail addresses: ahmed.ramadan@oncology.au.dk (A.R. Eldesoky), francolinigiulio@gmail.com (G. Francolini), mette.skovhus.thomsen@aarhus.rm.dk (M.S. Thomsen), esbeyate@rm.dk (E.S. Yates), tine.bisballe.nyeng@aarhus.rm.dk (T.B. Nyeng), carine.kirkove@uclouvain.be (C. Kirkove), claus.kamby@rh.regionh.dk (C. Kamby), egil.blix@uit.no (E.S. Blix), mette.h.nielsen@rsyd.dk (M.H. Nielsen), ztk@regionsjaelland.dk (Z. Taheri-Kadkhoda), martin.berg@rsyd.dk (M. Berg), birgoffe@rm.dk (B.V. Offersen).
[1] First and second authors contributed equally to the manuscript.

Material and methods

Patient selection

Approved treatment plans of 40 patients were selected from a database of two previous studies investigating quality assurance and ABAS within the Skagen Trial 1 [3]. Data were obtained from 7 institutions in Denmark, Belgium and Norway. To avoid bias related to differences in target volumes or dose prescription, only patients who received treatment of all nodal levels except L1, were without boost administration or breast implants were allowed in the study. Overall, 31 out of 40 treatment plans were also included in the ABAS validation study [3], while the others were part of the Skagen Trial 1 quality assurance protocol (Francolini et al., Quality assessment of clinical target volume delineation and dose planning in the clinically controlled randomized Skagen Trial 1, submitted to radiotherapy and Oncology).

Gold standard manual segmentation (MS)

MS of breast (CTVp_breast), chest wall (CTVp_chest wall), nodal levels except level I (CTVn) and internal mammary (CTVn_IMN) was performed by multiple observers from the participating institutions according to the ESTRO consensus guideline for target volume delineation [1,2]. The immobilization, scanning and use of breath adaptive technique followed the institutional procedures.

Atlas based automated segmentation and manual correction

ABAS was performed using four atlas libraries based on laterality and surgery, previously created on MIM Maestro™ software version 6.5 (MIM Software Inc., Cleveland, OH) [3]. ABAS was exported to the Eclipse™ treatment planning system version 11.0.31 (Varian Medical Systems Inc., Palo Alto, CA, USA) for revision and possible manual correction according to the ESTRO delineation guideline. Manual correction was performed by two research fellows (ARE and GF), blind to the MS, and approved by a breast oncologist (BVO).

Geometric comparison

Both MS and corrected ABAS (ABAS$_{corrected}$) were exported to the MIM software to calculate their spatial overlap (DSC) and volumes for each of the segmented structures. The absolute difference in volume (mL) was also calculated using the following Eq. (1):

$$\Delta V = |V(\text{ABAS}) - V(\text{MS})| \tag{1}$$

Dosimetric comparison

For each patient, the dose plan used for treating the patient was copied to the ABAS$_{corrected}$ structures. DVHs were created for both MS and ABAS of CTVp, CTVn and CTVn_IMN. The DVH parameters determined for both the MS and ABAS$_{corrected}$ dose plans included the V90% (%) for CTVn and CTVn_IMN, the V95% (%) for CTVp either breast or chest wall and the homogeneity index (HI), calculated using the following Eq. (2) [6]:

$$\text{HI} = (D2\% - D98\%)/D50\% [6] \tag{2}$$

The absolute differences in these parameters between MS and ABAS$_{corrected}$ were calculated using the following Eqs. (3–5):

$$\Delta V90\% = |V90\%(\text{ABAS}) - V90\%(\text{MS})| \tag{3}$$

$$\Delta V95\% = |V95\%(\text{ABAS}) - V95\%(\text{MS})| \tag{4}$$

$$\Delta \text{HI} = |\text{HI}(\text{ABAS}) - \text{HI}(\text{MS})| \tag{5}$$

V90% and V95% are expressed as a percentage of volume, thus, ΔV95% and ΔV90% are also expressed as a percentage of volume.

Statistical analysis

Stata® version 12.0 software (StataCorp., Texas, USA) was used for statistical analysis. Descriptive statistics including median, inter-quartile range (IQR) for all parameters were calculated. Shapiro–Wilk normality test showed that compared data did not follow a normal distribution. So, a Wilcoxon signed rank test was used to test the statistical significance of the difference in all parameters and a Spearman's rank correlation test was used for correlation testing. Two sided p-values were provided and p-values <0.05 were considered significant.

Results

Patients' characteristics

Twenty patients included in the study were treated at Aarhus University Hospital and the other 20 were treated at the other 6 institutions (Table 1).

Geometric difference

The median volume of ABAS$_{corrected}$ was larger than MS for CTVp_breast and CTVn. CTVp_chest wall showed a larger median volume of MS compared to the ABAS$_{corrected}$. Both median MS and ABAS$_{corrected}$ volumes were nearly the same for CTVn_IMN. However, the difference was not significant for any of these volumes. A high spatial overlap (median DSC ⩾0.72) was seen between MS and ABAS$_{corrected}$ for all compared structures. CTVp_breast showed the best agreement followed by CTVp_chest wall, CTVn and CTVn_IMN respectively (Table 2).

Dosimetric difference

Overall, HI comparison showed similar dose coverage for MS and ABAS$_{corrected}$; only CTVn and CTVn_IMN showed a minimal, although statistically significant, difference for this parameter. Fig. 1 shows examples of DVH for both MS and ABAS$_{corrected}$ structures.

Both ABAS$_{corrected}$ and MS showed acceptable levels of coverage on all target volumes. Differences were in favor of MS and were statistically significant only for chest wall and CTVn_IMN, with ΔV95% and ΔV90% of 2.5% and <1%, respectively (Table 2).

Table 1
Baseline patients characteristic.

Patients' characteristics	Number
Surgery	
Mastectomy	18
Lumpectomy	22
Laterality	
Right	19
Left	21
Dose	
40 Gy/15 fr	26
50 Gy/25 fr	14
Respiratory gating technique	
Yes	33
No	7
Total	40

Table 2
Results of comparison between manual and automated segmentations as regard volume, DSC, HI, V95, and V90%.

Parameter	CTVp-Breast (n = 22)		CTVp-Chest wall (n = 18)		CTVn-Total (n = 40)		CTVn-IMN (n = 40)	
	Median (IQR)	P-value	Median (IQR)	P-value	Median (IQR)	P-value	Median (IQR)	P-value
Volume (cc)								
Manual	665.64	–	234.95 (220.22)	–	52.39 (23.52)	–	7.85 (3.12)	–
ABAS	(496.26)	–	227.75 (160.34)	–		–	7.40 (1.88)	–
ΔVolume	679.56 (427.42) 33 (56.41)	0.66	28.59 (36.90)	0.99	55.73 (19.13) .90 (7.55)	0.21	0.95 (1.45)	0.96
DSC	0.93 (0.04)	–	0.78 (0.06)	–	0.76 (0.08)	–	0.72 (0.10)	–
HI								
Manual	0.09 (0.06)	–	0.10 (0.05)	–	0.12 (0.04)	–	0.13 (0.04)	–
ABAS	0.09 (0.06)	–	0.12 (0.10)	–	0.12 (0.05)	–	0.14 (0.05)	–
Δ HI	0.01 (0.01)	0.49	0.02 (0.18)	0.23	0.01 (0.02)	**0.01**	0.01 (0.02)	**0.002**
CTVp V95% (%)								
Manual	97.99 (5.57)	–	97.25 (3.50)	–	–	–	–	–
ABAS	98.38 (5.13)	–	95.80 (3.27)	–	–	–	–	–
Δ V95%	0.48 (0.74)	0.61	2.57 (2.74)	**0.02**	–	–	–	–
CTVn V90% (%)								
Manual	–	–	–	–	99.69 (1.28)	–	96.13 (6.79)	–
ABAS	–	–	–	–	99.48 (2.14)	–	96.14 (10.51)	–
Δ V90%	–	–	–	–	0.20 (1.05)	0.12	0.92 (3.42)	**0.001**

CTVp = clinical target volume of the primary tumor site, CTVn-Total = clinical target volume of the nodal levels 2,3,4 and inter-pectoral, CTVn-IMN = clinical target volume of the internal mammary lymph nodes, cc = cubic centimeters, ABAS: Atlas-based automated segmentation, DSC = Dice Similarity Coefficient, HI = Homogeneity Index, V95% = volume of the CTVp covered by 95% of the prescribed dose in percent, V90% = volume of the CTVn covered by 90% of the prescribed dose in percent, Δ = absolute difference, SD = standard deviation, IQR = inter-quartile range.
Statistically significant P values are evidenced in bold in the table.

Correlation between geometric and dosimetric differences

No significant correlation was found between DSC values, ΔVolume values or any of the DVH parameters used for CTVp_breast, CTVp_chest wall and CTVn. A statistical significant correlation between these different parameters was seen for CTVn_IMN. There was a negative correlation between DSC and ΔHI or ΔV90% ($r = -0.60$, $p = 0.00$ and -0.54, $p = 0.0004$, respectively) and a positive correlation between ΔVolume and ΔHI or ΔV90% ($r = 0.40$, $p = 0.01$ and 0.50, $p = 0,001$, respectively). Finally, a negative significant correlation ($r = -0.44$, $p = 0.004$) was found between DSC and ΔVolume for CTVn_IMN.

Discussion

The results of the current study support that ABAS with manual correction can be used safely for dose planning in loco-regional RT of early breast cancer. A DSC >0.7 indicates low inter-observer variability [7], thus, median DSC values above 0.7 for all compared CTVs in the current study reflected high agreement between MS and ABAS_corrected.

Results from the dosimetric comparison showed that dose coverage for both CTVp and CTVn, corresponding to V95% and V90%, respectively, were acceptable not only in clinical practice, but also in the context of the Skagen Trial 1. Indeed, more than 95% of both CTVp and CTVn ABAS_corrected volumes were in median covered by 95% and 90% of prescribed dose, respectively. Furthermore, differences in these parameters were minimal, and only significant for CTVp_chest wall. HI comparison demonstrated overlapping DVH curves for ABAS_corrected and MS (Fig. 1).

No significant differences were found for this parameter in CTVp_breast or CTVp_chest wall reflecting a similar dose distribution for MS and ABAS_corrected volumes. The median difference in HI of CTVn between ABAS_corrected and MS, even if statistically significant, was only 0.01.

Thus, considering the optimal coverage levels for CTVn (V90% > 99% for both ABAS_corrected and MS) this difference was not clinically relevant and under-dosage was not seen.

In a population-based study, irradiation of IMN significantly improves overall survival in node positive breast cancer patients [8]. However, in left sided patients, balance against dose to the heart and left anterior descending coronary artery is critical. In the current study, a median 96% of CTVn_IMN volumes for both ABAS_corrected and MS were successfully covered by 90% of the prescribed dose, and the median differences in V90% and HI between ABAS_corrected and MS for CTVn_IMN were minimal, although statistically significant.

Results of a previous work have shown low contouring variability between ABAS_corrected and MS [3]. However, DSC reliability as an absolute measure of delineation variability testing has been questioned, and geometric analysis used for this purpose may have limits of performances [9]. Therefore, dosimetric comparison is recommended to evaluate the performance of automated segmentation in a more clinically relevant way [10]. Several studies have looked at the difference in DVH parameters between MS and ABAS in different tumor sites [11–16], and dosimetric analysis has been used to quantify the clinical effect of inter-observer variability in breast cancer RT [4,17]. One study has reported that inter-observer variability is responsible for a significant variability in dose coverage of the primary and nodal volumes among dose plans based on nine observers' contouring [4], with a difference in V95% ranging between 10–25%. The current study has shown no significant difference in V95% between manual and automated segmentation of the breast with an inter-quartile range of about 5% for both. Target coverage was not influenced by the use of ABAS_corrected or MS.

Another study has evaluated the performance of ABAS of the breast in patients treated in prone position [17]. Results have shown that a DSC >0.95 against MS has been correlated significantly to better target dose coverage. Conversely, we cannot find such correlation for the breast in the current study.

A significant correlation between variability in contouring and dosimetric differences (DSC and ΔV) between MS and ABAS_corrected has been found for CTVn_IMN only. Therefore, a possible effect of the amount of variation on the dose distribution within this small structure will be expected and a minimal variation will ensure equal coverage. However, the dosimetric difference is clinically acceptable for both MS and ABAS_corrected for CTVn_IMN in this study.
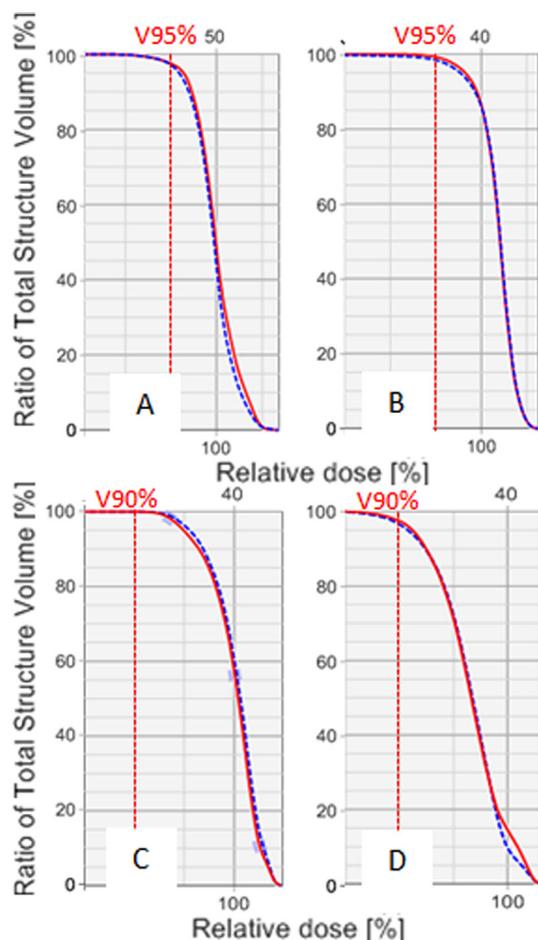
Fig. 1. Example of comparison between DVH curves measured in two patients on ABAS$_{corrected}$ (blue dotted curve) and MS (red continue curve). DVHs are related to chest wall (A) and breast (B) of a post-mastectomy and a post-lumpectomy patient. CTV nodal (C) and CTVn_IMN (D) of a post-lumpectomy patient are represented in the lower part of the figure. V95% and V90% are evidenced by the dotted line. On the higher part of the graph is reported prescribed dose as an absolute value expressed in Gy (50 Gy for A, 40 Gy for B, C and D). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Moreover, the reported dosimetric differences between MS and ABAS$_{corrected}$ are less than that reported in inter-observer variability studies and within the range of clinical acceptance. Therefore, the expected clinical outcome from routine use of ABAS$_{corrected}$ may be considered equivalent to the use of conventional MS for dose planning in loco-regional RT of breast cancer with the advantage of less time and inter-observer variability and more consistency and reproducibility for ABAS$_{corrected}$ compared to MS.

Impact of ABAS$_{corrected}$ on organs at risk was not explored in this analysis, however, it is reasonable to assume even a lower dosimetric impact on these structures.

A potential limitation of the methodology of the current study is the use of the original dose plans based on the MS rather than generating specific plans for ABAS$_{corrected}$. This may theoretically bias the results. If the volume of the ABAS$_{corrected}$ is smaller than the MS volume, it should be covered with the designed plans. However, structures with a comparable coverage level (CTVp_breast and CTVn) between both segmentation methods have shown larger median volumes of ABAS$_{corrected}$, eliminating this bias. Moreover, a better dosimetric coverage is expected if a new plan based on ABAS$_{corrected}$ is created. Therefore, results of the current study may represent the worst-case scenario.

Conclusions

Data from this analysis confirmed the low contouring variability between ABAS and MS.

Overall, comparison in HI and targets coverage showed that dose distribution was similar regardless of the use of ABAS or MS. Furthermore, no relationship was found between DSC and differences in coverage, reflecting that performances of ABAS did not affect dose parameters.

In the context of daily routine practice, ABAS could reduce the time in RT workflow, without meaningful dosimetric impact on treatment plan. This technique can be used in a multi-institutional context. Thus, ABAS is a useful tool and its implementation in clinical activity should be considered.

Conflict of Interest statement

None to declare.

Acknowledgement

References

[1] Offersen BV, Boersma LJ, Kirkove C, Hol S, Aznar MC, Biete Sola A, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer. Radiother Oncol 2015;114:3–10. http://dx.doi.org/10.1016/j.radonc.2014.11.030.

[2] Offersen BV, Boersma LJ, Kirkove C, Hol S, Aznar MC, Sola AB, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer, version 1.1. Radiother Oncol 2016;118:205–8. http://dx.doi.org/10.1016/j.radonc.2015.12.027.

[3] Eldesoky AR, Yates ES, Nyeng TB, Thomsen MS, Nielsen HM, Poortmans P et al. Internal and external validation of an ESTRO delineation guideline - dependent automated segmentation tool for loco-regional radiation therapy of early breast cancer. Radiother Oncol 2016; in press. DOI:10.1016/j.radonc.2016.09.005.

[4] Li XA, Tai A, Arthur DW, Buchholz TA, Macdonald S, Marks LB, et al. Variability of target and normal structure delineation for breast cancer radiotherapy: an RTOG Multi-Institutional and Multiobserver Study. Int J Radiat Oncol Biol Phys 2009;73:944–51. http://dx.doi.org/10.1016/j.ijrobp.2008.10.034.

[5] Jameson MG, Holloway LC, Vial PJ, Vinod SK, Metcalfe PE. A review of methods of analysis in contouring studies for radiation oncology. J Med Imaging Radiat Oncol 2010;54:401–10. http://dx.doi.org/10.1111/j.1754-9485.2010.02192.x.

[6] ICRU. Prescribing, recording, and reporting photon-beam intensity-modulated radiation therapy (IMRT) [ICRU report 83]. J ICRU 2010;10:1–106.

[7] Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Trans Med Imaging 1994;13:716–24. http://dx.doi.org/10.1109/42.363096.

[8] Thorsen LB, Offersen BV, Danø H, Berg M, Jensen I, Pedersen AN, et al. DBCG-IMN: a population-based cohort study on the effect of internal mammary node irradiation in early node-positive breast cancer. J Clin Oncol 2016;34:314–20. http://dx.doi.org/10.1200/JCO.2015.63.6456.

[9] Hanna GG, Hounsell AR, O'Sullivan JM. Geometrical analysis of radiotherapy target volume delineation: a systematic review of reported comparison methods. Clin Oncol (R Coll Radiol) 2010;22:515–25. http://dx.doi.org/10.1016/j.clon.2010.05.006.

[10] Valentini V, Boldrini L, Damiani A, Muren LP. Recommendations on how to establish evidence from auto-segmentation software in radiotherapy. Radiother Oncol 2014;112:317–20. http://dx.doi.org/10.1016/j.radonc.2014.09.014.

[11] Conson M, Cella L, Pacelli R, Comerci M, Liuzzi R, Salvatore M, et al. Automated delineation of brain structures in patients undergoing radiotherapy for primary brain tumors: from atlas to dose-volume histograms. Radiother Oncol 2014;112:326–31. http://dx.doi.org/10.1016/j.radonc.2014.06.006.

[12] Lorenzen EL, Ewertz M, Brink C. Automatic segmentation of the heart in radiotherapy for breast cancer. Acta Oncol 2014;53:1366–72. http://dx.doi.org/10.3109/0284186X.2014.930170.

[13] Pasquier D, Lacornerie T, Betrouni N, Vermandel M, Rousseau J, Lartigau E. Dosimetric evaluation of an automatic segmentation tool of pelvic structures from MRI images for prostate cancer radiotherapy. Cancer Radiother 2008;12:323–30. http://dx.doi.org/10.1016/j.canrad.2008.03.001.

[14] Tsuji SY, Hwang A, Weinberg V, Yom SS, Quivey JM, Xia P. Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2010;77:707–14. http://dx.doi.org/10.1016/j.ijrobp.2009.06.012.

[15] Voet PW, Dirkx ML, Teguh DN, Hoogeman MS, Levendag PC, Heijmen BJ. Does atlas-based autosegmentation of neck levels require subsequent manual contour editing to avoid risk of severe target underdosage? A dosimetric analysis. Radiother Oncol 2011;98:373–7. http://dx.doi.org/10.1016/j.radonc.2010.11.017.

[16] Weiss E, Wijesooriya K, Ramakrishnan V, Keall PJ. Comparison of intensity-modulated radiotherapy planning based on manual and automatically generated contours using deformable image registration in four-dimensional computed tomography of lung cancer patients. Int J Radiat Oncol Biol Phys 2008;70:572–81. http://dx.doi.org/10.1016/j.ijrobp.2007.09.035.

[17] Dipasquale G, Wang X, Chatelain-Fontanella V, Vinh-Hung V, Miralbell R. Automatic segmentation of breast in prone position: correlation of similarity indexes and breast pendulousness with dose/volume parameters. Radiother Oncol 2016;120:124–7. http://dx.doi.org/10.1016/j.radonc.2016.04.041.