Short Communication

A comprehensive quality assurance program for four-dimensional computed tomography in radiotherapy

Jinane Bakkali Tahiri [a,b,1,*], Martin Kyndt [c], Jennifer Dhont [a], Akos Gulyban [a], Juliane Szkitsak [d], Evelien Bogaert [e], Nick Reynaert [a], Manuela Burghelea [a]

[a] Université libre de Bruxelles (ULB), Hôpital Universitaire de Bruxelles (HUB), Institut Jules Bordet, Department of Medical Physics, Brussels, Belgium
[b] Iridium Netwerk, Radiation Oncology, Antwerp, Belgium
[c] MIM Software Inc, Cleveland, OH, USA
[d] Department of Radiation Oncology, Universitätsklinikum Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[e] Department of Radiotherapy-Oncology, Ghent University Hospital, Gent, Belgium

ARTICLE INFO

ABSTRACT

This study aimed to develop and validate a comprehensive, reproducible and automatic 4DCT Quality Assurance (QA) workflow (QAMotion) that evaluates image accuracy across various regular and irregular breathing patterns. Volume and amplitude deviations, CT number accuracy, and spatial integrity were used as evaluation metrics. For repeatability tests, tolerances were respected with a mean CT number deviation < 10 HU, volume deviation < 2% and diameter and amplitude deviation < 2 mm except for irregular amplitude curves for which an amplitude deviation up to 6 mm was measured. QAMotion was able to flag image artefacts for our clinical 4DCT system.

1. Introduction

Four-dimensional computed tomography (4DCT) has become an essential imaging modality in the radiotherapy treatment planning of thoracic and upper abdominal tumors. Compared to conventional 3DCT, the time-resolved 4D acquisition not only reduces image blurring and motion artifacts but also provides more detailed information on the tumor and organ-at-risk (OAR) motion that can be expected during treatment [1,2,3].

However, artifacts in 4DCT can still occur and originate from different sources such as intra-slice residual motion, finite slice thickness, and binning inaccuracies caused by breathing irregularities[4,5,6]. These artifacts can have a considerable impact on volume reconstruction and subsequently on radiotherapy target volume definition [7]. Moreover, a recent study by Sentker et al.[8] demonstrated that the presence of severe artifacts in 4DCT was a strong negative prognostic factor for local tumor control after stereotactic body radiotherapy (SBRT). These findings strongly support the inclusion of a comprehensive 4DCT performance assessment in the standard radiotherapy Quality Assurance (QA) program.

The Canadian Partnership for Quality Radiotherapy (CPQR) recently published guidelines outlining the performance objectives and corresponding tolerances for 4DCT equipment [9]. These guidelines are intended to be used at the discretion of each individual center to help guide their respective QA program, but do not provide a concrete QA workflow.

Therefore, this study aimed to develop and validate a comprehensive but easily reproducible and automatic 4DCT QA workflow (QAMotion) that evaluates image accuracy across a range of clinically realistic scenarios. The key objectives for QAMotion were identified during an initial brainstorming meeting at the 4th ESTRO Physics Workshop "Clinical Translation of CT innovations in Radiation Oncology" (October 2021). This paper is written to facilitate the introduction of QAMotion into other clinics.

2. Materials and methods

Five requirements were identified to develop QAMotion:

- Identify clinically relevant tests and clinically representative breathing patterns through literature review.
- CT acquisition and reconstruction parameters should be sufficient, easily available, and not deviating from clinical routine.
- QAMotion should be able to identify artifacts and/or deviations from expected results.
- All steps of the evaluation should be relevant, intuitive, and easy to follow.
- The entire workflow should be fully automated, with limited user input, as this would limit the workload and make the QAMotion procedure easy to implement, and the results user-independent.

2.1. QAMotion: phantom, breathing curves and imaging protocols

The Thorax dynamic phantom (CIRS, Sun Nuclear, Norfolk, VA, USA) with imaging insert (2 cm diameter "tumor" sphere) was selected based on in-house availability. The insert was defined as the region of interest (ROI) for all analysis and will be referred to in the remainder of this article as 'Target4D'.

Five regular and three irregular breathing patterns were included in QAMotion. Breathing amplitudes and periods were selected based on a previously published analysis by Szkitsak et al. [10]. A $\cos^6$ function was used to simulate the breathing patterns. All breathing patterns are shown in Supplementary Material Figure S1. In accordance with CPQR recommendation[9], most of the curves were one-dimensional in superior-inferior (SI) direction which is the direction that predominately causes 4DCT image artifacts. However, a 3D motion breathing curve was added to allow for a more comprehensive test.

The amplitudes of the regular patterns ranged from 2.5 mm to 25 mm. The combination of a four seconds breathing cycle with 25 mm amplitude was used to represent a highly moving tumour. The lowest amplitude (2.5 mm) was defined based on the breathing curves of seven pancreatic cancer patients, for which the marker block (i.e. surrogate) was positioned on top of the compression belt.

Irregular breathing patterns were included to analyze possible binning inaccuracies and subsequent artifacts. Breathing irregularities were created by scaling the base breathing cycle in amplitude and/or time and by inserting pauses between two elements (see Supplementary Material Figure S1). For all the tests, the learning curve was always a regular pattern.

To perform QAMotion, a conventional 3DCT without phantom motion (used as ground-truth), and a 10-phase 4DCT, as well as a maximum - (MaxIP) and average intensity projection (AvgIP) are required for each breathing curve. To be able to evaluate the clinical performance of 4DCT equipment, the QAMotion guidelines recommend each center to use its respective clinical 4DCT acquisition protocols for SBRT.

2.2. QAMotion: Image analysis and reporting

The image analysis step was implemented as a fully automated MIM Workflow (MIM Software Inc., Cleveland, OH, USA, v7.3 beta).

Once all images are acquired, the MIM Workflow consists of four steps. A schematic representation of QAMotion is presented in Supplementary Material Figure S2, while a video of the image processing analysis for an irregular breathing curve is provided as Supplementary Material.

2.2.1. Step1: Data input and retrieval

Once the imaging data was acquired and uploaded in MIM Maestro Software, the user was prompted to select the relevant reconstructions (the static 3DCT, the 4DCT phases, MaxIP and AvgIP) and launch QAMotion. The workflow automatically matched the different reconstructions and started the evaluation process. The user also needed to select from a drop down list what type of breathing recording systems was used in their clinic. This information was added to the report,

together with various DICOM tags relevant to the study. The workflow also accessed automatically the recorded RGSC curves (.vxp format) that were saved beforehand into a folder.

2.2.2. Step 2: Target4D auto-contouring

A robust CT number-based thresholding method was developed to automatically contour the target on the ground-truth 3DCT. The image was first contoured with a threshold of -400 HU to identify an air-like region. In the obtained contour, the lungs of the phantom were subsequently identified as they systematically represent a 2D circle on each axial slice. The sphere was then localized by thresholding the left lung. Next, the obtained target volume was compared to the ground-truth sphere insert volume (4.19 ml), and the contour was increased or decreased until an absolute difference between the contoured and ground-truth volume was lower than 0.01 ml. The CT number threshold value was stored to later automatically contour the target on the 4DCT exhale phase and MaxIP reconstruction. A deformable contour propagation method was then used to auto-contour the target on all 4DCT phases using the exhale phase as a moving image. The internal target volume (ITV) was generated by calculating the union of the volume in the ten phases. In the validation/application phase, this thresholding method was verified using three different reconstruction filters (Br40, Sd40 and Qr40) and two slice thicknesses (1 mm and 2 mm).

In the preliminary results, it was noticed that the auto-contouring solution was not applicable to the AvgIP reconstruction, due to the blurring. Therefore another contouring approach was implemented. In the AvgIP reconstruction, the Target4D was contoured using a threshold defined by adding 50 HU to the mean background value of the lungs of the phantom.

2.2.3. Step 3: Image analysis

Once the target was auto-contoured on all reconstructions, the analysis of the 4DCT was performed following 4 criteria: amplitude deviation, volume deviation, CT number variation, and spatial integrity. The evaluation metrics are summarized in Supplementary Material Table S1 along with the corresponding thresholds obtained from the CPQR guidelines (green in the table). The latter were defined specifically for regular breathing patterns and amplitudes lower than 2 cm. Therefore, based on our in-house testing, additional tolerances were introduced for irregular breathing curves and amplitudes larger than 2 cm. In addition to a comprehensive image analysis, QAMotion includes independent verification of the external monitoring system following CPQR recommendations.

2.2.4. Step 4: QA report

The automatic generation of a QA report summarizing the results of the analysis workflow was implemented. To allow the user to efficiently review the results, tolerance-based color-coded feedback was defined as used in Supplementary Material Table S1. An example of the QA report for the Irregular Breathing (IB) pattern is presented in Supplementary Material Figure S3. After each workflow execution, the report was automatically saved in the MIM session attached to the 3DCT and can be exported as a.pdf outside MIM. The results can also be exported in a.csv file format for further statistical analysis if required.

2.3. QAMotion validation and application

To evaluate the applicability, accuracy, and repeatability of QAMotion, the workflow was repeated three times on our clinical 4DCT system. In addition, the fully automated image analysis step was visually inspected by two experts and benchmarked against a manual execution. The required image sets for QAMotion were acquired using the i4DCT algorithm, clinically implemented on our SOMATOM go.Open Pro scanner (Siemens Healthineers, Germany), in conjunction with the Varian Respiratory Gating for Scanners (RGSC) (Varian Medical Systems, Inc, CA, USA) external monitoring system. Scan parameters for

each protocol were: collimation 64.0x0.6 mm^2, 120 kV, slice thickness 2 mm, and kernel Qr40. Amplitude-based and phase-based reconstructions were analyzed.

3. Results

The Target4D auto-contouring method (step 1) was able to detect the insert in the 3DCT with an implemented accuracy of 1% volume difference for all tested reconstruction filters and slice thicknesses.

High consistency was observed for CT number, volume and diameter deviations across repetitions. Fig. 1 presents the results for the phase-based reconstructions, while the amplitude-based are presented in the Supplementary Material Figure S4. The RGSC captured a motion amplitude and period that were within tolerances (2 mm and 1 s) for all tests (see Fig. 2a). Our clinical 4DCT system was able to capture the applied motion with high accuracy, with tolerances respected in all tests (mean < 10 HU, <5% and < 2 mm, respectively) for both amplitude- and phase-based reconstruction. Outliers in volume and diameter deviation revealed an artifact in 2 out of 10 phases of one Irregular Frequency scans, indicated by a volume and diameter underestimation of respectively 20% and 5 mm.

For breathing curves 'Double Amplitude' (DA) and 'Irregular Breathing' (IB), phase-based binning resulted in an amplitude deviation that did not meet CPQR tolerances. Additional tests were conducted by starting the acquisition at the high/low peak, then verifying the

reconstructed amplitude and the images to see if the deviations were due to artifacts. These tests showed that the final amplitude captured in the 4DCT was dependent on when the image acquisition was started (at the high or low amplitude peak), see Supplementary Material Figure S5. However, these curves did not result in any visible artifacts, and volume and CT number metrics were within tolerances. The results of amplitude-based binning were more robust against amplitude variations in the breathing curves, with an amplitude deviation < 2 mm for all curves (see Supplementary Material Figure S6a).

QAMotion also revealed a difference between the MaxIP images reconstructed by Siemens Syngo.Via (Siemens Healthineers) and MIM software. The line profile metrics showed an overestimation of the MaxIP phase-based reconstruction by Siemens Syngo.Via for DA and IB curves (Fig. 2c). The MaxIP amplitude-based reconstruction was within tolerances (see Supplementary Material Figure S6b).

The AvgIP was systematically overestimating the ITV volume by up to 20% for phase-based (mean of 10.58%) and up to 17% for amplitude-based (mean of 8.25%) (see Supplementary Material Figure S7).

4. Discussion

As image quality directly affects not only the ability to identify and delineate target volumes and surrounding OARs but can also potentially cause dosimetric errors, a regular evaluation of 4DCT image quality is an essential part of CT quality assurance.

In this work, a comprehensive 4DCT QA program, named QAMotion, was developed. It provides a fully automated image analysis of CT scans of the dynamic thoracic phantom based on CPQR recommendations which include evaluation of: mean CT number and volume accuracy, amplitude of the moving target, spatial integrity of the moving target at each respiratory phase and the ability of the respiratory monitoring system to accurately monitor the motion. Optimized for efficiency, once the imaging data is imported in MIM, the entire workflow for all eight tests (e.g., eight breathing curves) can be completed in<15 min, making it an ideal tool for periodic QA.

Two separate groups also recently published a detailed overview of their QA workflow for 4DCT commissioning and periodic QA. In the workflow presented by Polizzi et al. [11], 4DCT images of a QUASAR motion phantom were acquired with both regular and irregular breathing curves. Tumor motion and ITV were analyzed to quantify 4DCT performance. Szkitsak et al. evaluated the i4DCT algorithm with phantom measurements in terms of geometric accuracy and image quality [10]. However, both presented solutions are not easily reproducible, and the quality metrics and tolerances were not systematically based on CPQR recommendations. Furthermore, the image analysis was automated but no user interface was provided, making it time-consuming and susceptible to inter-observer variability. Quantitative evaluation of the breathing pattern transfer is an important implementation in QAMotion, which was lacking in previous papers (especially when the breathing signal is further used for gating or tracking the tumor during treatment delivery).

A similar QA workflow was also recently presented and made available for conventional 3DCT (QAMaster) [12]. As both QAMaster and QAMotion are fully automated, their sequential application for a full CT QA procedure could be relatively straightforward.

Validation of the automated workflow showed excellent repeatability across, and robustness against, different imaging protocols and reconstruction methods. Through its implemented range of performance metrics, QAMotion was able to flag image artefacts in its automatically generated report and identify a limitation of the novel i4DCT algorithm related to breathing curves with varying amplitudes. Differences between imaging protocols and reconstruction methods could also be quantified using the proposed workflow. In this respect, QAMotion can also be used during the commissioning of new 4DCT equipment, after a major upgrade, or to perform intra- or inter-institutional comparisons for image quality improvement and standardization.
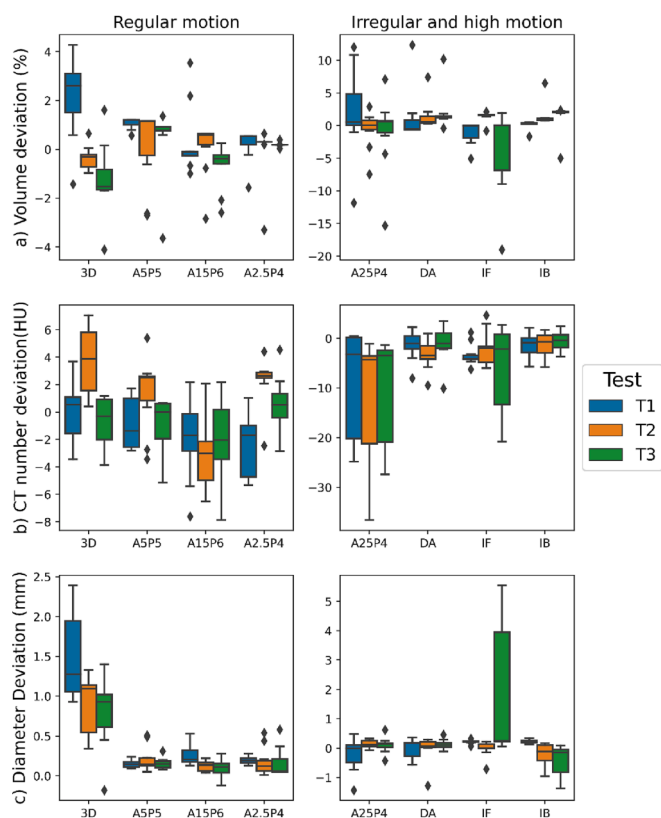

Fig. 1. Results of the QAMotion program applied to the i4DCT system (phase-based reconstruction): (a) Volume deviation expressed in % of the true volume per motion type and per test. (b) CT number deviations between 3DCT and 4DCT expressed in HU per motion type and per test. (c) Diameter deviation in SI direction over all 4DCT phases expressed in mm per motion type and per test. Note that the scale is different for the regular curves (left column) and irregular and highly moving curves (right column) (negative sign means an underestimation). Legend: A = peak-to-peak amplitude (range 2.5–25 mm), P = periods (range 2.8–7 s), 3D = three dimensional movement, DA = Double Amplitude, IF = Irregular Frequency, IB = Irregular Breathing. Tests 1–3 are the measurements repeated 3 times on our clinical 4DCT system.
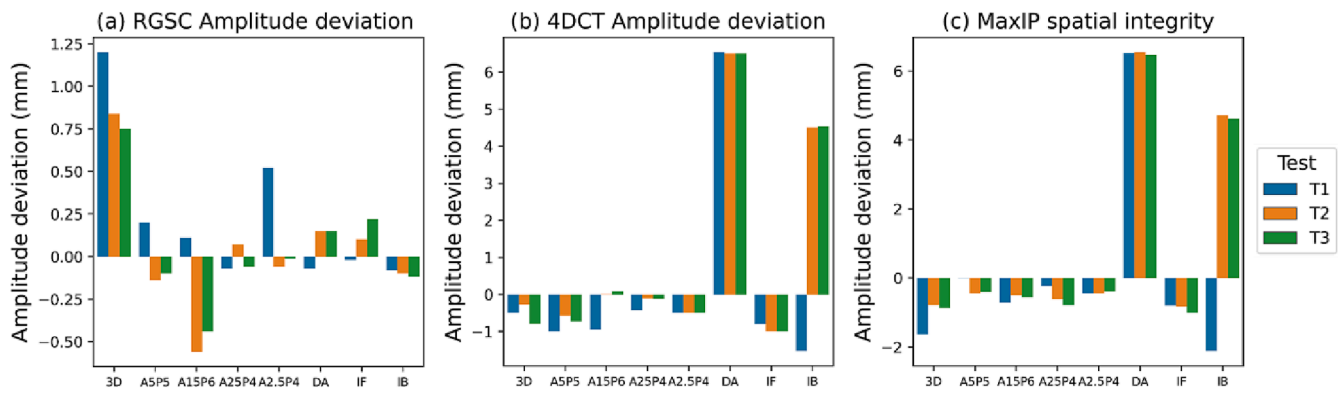
Fig. 2. Results of the QAMotion program applied to the i4DCT system (phase-based reconstruction): (a) Deviation of the external breathing signal amplitude captured by the monitoring system compared to the input amplitude per motion type and per test. (b) Deviation of the motion amplitude on 4DCT compared to the input amplitude per motion type and per test. (c) Evaluation of the motion amplitude on the MaxIP reconstructed by Siemens Syngo.Via per motion type and per test (negative sign means an underestimation). This metric showed an overestimation of the MaxIP for DA and IB curves. Legend: A = peak-to-peak amplitude (range 2.5–25 mm), P = periods (range 2.8–7 s), 3D = three dimensional movement, DA = Double Amplitude, IF = Irregular Irequency, IB = Irregular Breathing. Tests 1–3 are the measurements repeated 3 times on our clinical 4DCT system.

There are a few limitations to QAMotion. The automatic workflow is currently tailored to a specific phantom and imaging insert. Nevertheless, we believe the implemented iterative thresholding method can be easily generalizable to other inserts or even phantoms, making the workflow phantom independent. Secondly, the QAMotion workflow was validated using only one target size and shape. However, Szkitsak et al. [10] already showed that target size has a small influence on the volume accuracy during i4DCT validation.

Lastly, auto-contouring on the AvgIP reconstruction remains a weak point, the CT number threshold is not enough for AvgIP contouring as it represents the blurred motion image over the respiratory cycle. This can lead to untrue results for this reconstruction, thus the systematic overestimation of the ITV by the AvgIP might not be reliable. Therefore AvgIP is usually not used for target delineation (but for dose planning). AvgIP contouring remains a subject to be further investigated.

QAMotion is currently only locally installed and available, however, a collaboration with MIM is ongoing to provide open access to QAMotion via MIMcloud, an internet-based medical image service that provides an easily accessible resource for storing, sharing, and viewing data.

QAMotion (using regular and irregular breathing motion curves) allows us to perform a comprehensive QA workflow of 4DCT imaging and is capable of detecting and quantifying image quality degradation following the CPQR recommendations. The automated workflow allows easy and standardized reporting. Additionally, the repeatability test confirms the robustness of the QA procedure, which initiated an ongoing multi-centric QA audit.

CRediT authorship contribution statement

Jinane Bakkali Tahiri: Conceptualization, Methodology, Data curation, Writing – original draft. **Martin Kyndt:** Software. **Jennifer Dhont:** Writing – review & editing. **Akos Gulyban:** Data curation, Visualization. **Juliane Szkitsak:** Writing – review & editing. **Evelien Bogaert:** Writing – review & editing. **Nick Reynaert:** Writing – review & editing. **Manuela Burghelea:** Supervision, Conceptualization, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2023.100475.

References

[1] Dhont J, Vandemeulebroucke J, Burghelea M, Poels K, Depuydt T, Van Den Begin R, et al. The long- and short-term variability of breathing induced tumor motion in lung and liver over the course of a radiotherapy treatment. Radiother Oncol 2018;126(2):339–46. https://doi.org/10.1016/j.radonc.2017.09.001.

[2] Balter J, Ten Haken R, Lawrence T, Lam K, Robertson J. Uncertainties in CT-based radiation therapy treatment planning associated with patient breathing. Int J Radiat Oncol Biol Phys 1996;36:167–74. https://doi.org/10.1016/S0360-3016(96)00275-1.

[3] Chen G, Kung J, Beaudette K. Artifacts in computed tomography scanning of moving objects. Semin Radiat Oncol 2004;14:19–26. https://doi.org/10.1053/j.semradonc.2003.10.004.

[4] Nakamura M, Narita Y, Sawada A, Matsugi K, Nakata M, Matsuo Y, et al. Impact of motion velocity on four-dimensional target volumes: A phantom study. Med Phys 2009;36(5):1610–7. https://doi.org/10.1118/1.3110073.

[5] Watkins W, Li R, Lewis J, Park J, Sandhu A, Jiang S, et al. Patient-specific motion artifacts in 4DCT. Med Phys 2010;37:2855–61. https://doi.org/10.1118/1.3432615.

[6] Yamamoto T, Langner U, Loo B, Shen J, Keall P. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. Int J Radiat Oncol Biol Phys 2008;72:1250–8. https://doi.org/10.1016/j.ijrobp.2008.06.1937.

[7] Persson G, Nygaard D, Brink C, Westberg J, Rosenschöld P, Specht L, et al. Deviations in delineated GTV caused by artefacts in 4DCT. Radiother Oncol 2010;96:61–6. https://doi.org/10.1016/j.radonc.2010.04.019.

[8] Sentker T, Schmidt V, Ozga AK, Petersen C, Frederic M, Hofmann C, et al. 4D image artifacts affect local control in SBRT of lung and liver metastases. Radiother Oncol 2020;148:229–34. https://doi.org/10.1016/j.radonc.2020.04.006.

[9] Després P, Gaede S. COMP report: CQPR technical quality control guidelines for CT simulators. J Appl Clin Med Phys 2017;19(2):12–7. 10.1002F/acm2.12213.

[10] Szkitsak J, Karius A, Hofmann C, Fietkau R, Bert C, Speer S, et al. Quality assurance of a breathing controlled four-dimensional computed tomography algorithm. Phys Imag Radiat Oncol 2022;23:85–91. https://doi.org/10.1016/j.phro.2022.06.007.

[11] Polizzi M, Kim S, Rosu M. A comprehensive quality assurance procedure for 4D CT commissioning and periodic QA. J Appl Clin Med Phys 2022;23. https://doi.org/10.1002/acm2.13764.

[12] Karius A, Bert C. QAMaster: A new software framework for phantom-based computed tomography quality assurance. J Appl Clin Med Phys 2022;23(4):e13588. https://doi.org/10.1002/acm2.13588.