



HAL
open science

Artificial intelligence for quality assurance in radiotherapy

L. Simon, C. Robert, Philippe Meyer

► **To cite this version:**

L. Simon, C. Robert, Philippe Meyer. Artificial intelligence for quality assurance in radiotherapy. Cancer/Radiothérapie, 2021, 25 (6-7), pp.623-626. 10.1016/j.canrad.2021.06.012 . hal-03797699

HAL Id: hal-03797699

<https://hal.science/hal-03797699>

Submitted on 16 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Artificial intelligence for quality assurance in radiotherapy

Utilisation de l'intelligence artificielle pour le contrôle de qualité en radiothérapie

L. Simon ^{a*,b}, C. Robert ^{c,d}, P. Meyer ^e

^a Institut Claudius-Regaud, Institut universitaire du cancer de Toulouse-Oncopole (IUCT-O),
Toulouse, France

^b Centre de recherches en cancérologie de Toulouse (CRCT), université de Toulouse, UPS, Inserm
U1037, Toulouse, France

^c Université Paris-Saclay, Gustave-Roussy, Inserm, Radiothérapie moléculaire et innovation
thérapeutique, Villejuif, France

^d Department of Radiotherapy, Gustave-Roussy, Villejuif, France

^e Service d'oncologie radiothérapie, Institut de cancérologie Strasbourg Europe, Strasbourg, France

*Corresponding author

Dr. Luc Simon

Département de physique médicale

Institut universitaire du cancer-Oncopole, 1 avenue Joliot-Curie, 31100 Toulouse, France

Tel.: 05 3115 6109

Email: simon.luc@iuct-oncopole.fr

Congrès SFRO2021, Session SFRO/SFPM, IA et radiothérapie: vision du physicien médical

Abstract

In radiotherapy, patient-specific quality assurance is very time-consuming and causes machine downtime. It consists of testing (using measurement with a phantom and detector) if a modulated plan is correctly delivered by a treatment unit. Artificial intelligence and in particular machine learning algorithms were mentioned in recent reports as promising solutions to reduce or eliminate the patient-specific quality assurance workload. Several teams successfully experienced a virtual patient-specific quality assurance by training a machine learning tool to predict the results. Training data are generally

composed of previous treatment plans and associated patient-specific quality assurance results. However, other training data types were recently introduced such as actual positions and velocities of multileaf collimators, metrics of the plan's complexity, and gravity vectors. Different types of machine learning algorithms were investigated (Poisson regression algorithms, convolutional neural networks, support vector classifiers) with sometimes promising results. These tools are being used for treatment units' quality assurance as well, in particular to analyse the results of imaging devices. Most of these reports were feasibility studies. Using machine learning in clinical routines as a tool that could fully replace quality assurance tests conducted by physics teams has yet to be implemented.

Keywords

machine learning, radiotherapy, quality assurance

Résumé

En radiothérapie, le contrôle de qualité patient est une activité chronophage qui représente un temps important d'inactivité pour les machines de traitement. Il consiste à tester (en réalisant une mesure à l'aide d'un fantôme et d'un détecteur) si un plan de traitement modulé est correctement délivré par la machine de traitement. Il est généralement réalisé pour tous les patients d'un type de traitement donné. L'intelligence artificielle (et plus particulièrement le *machine learning*, est évoquée dans la littérature récente comme une solution qui pourrait permettre de réduire, voire de supprimer la charge de travail des contrôles de qualité patient. Plusieurs équipes ont implémenté avec succès un contrôle de qualité patient virtuel en entraînant un algorithme de *machine learning* à prédire les résultats des contrôles de qualité patient (afin de ne plus les réaliser). Les données d'apprentissage sont généralement composées de plans de traitement passés ainsi que des résultats des contrôles de qualité patient associés. D'autres types de données d'apprentissage ont récemment été introduits comme les positions réelles des lames du collimateur multilames, les indices de complexité du plan ou les vecteurs de gravité. Plusieurs algorithmes de *machine learning* ont été explorés (régression de Poisson, réseaux de neurones, *support vector classifier*) avec quelques résultats prometteurs. Ces outils commencent également à être utilisés pour le contrôle de qualité des machines de traitement, en particulier les dispositifs d'imagerie embarquée, à la fois pour l'étape d'acquisition mais également pour l'analyse des résultats. La majorité de ces travaux sont des études de faisabilité de ces outils et leur utilisation en routine clinique pour remplacer les tests de contrôle de qualité reste à mettre en œuvre.

Mots clés

Machine learning, Radiothérapie, Assurance qualité

1. Introduction

Quality assurance in radiotherapy is mainly managed by physics teams in radiotherapy departments. Quality assurance tasks (management, preparation, delivery, analysis, and reporting) represent a significant workload for these teams but also important machine downtime (from one to two hours per day). However, quality assurance management is a key factor for the security and quality of patient treatment. Quality assurance of radiotherapy can be divided into three main types (we did not consider planning quality assurance in this study).

First, machine quality assurance consists of assessing the performances of different radiotherapy medical devices: linear accelerators, electronic portal imaging devices, onboard imaging, and computed tomography (CT), among others. Image quality, mechanical and dosimetric properties must be regularly checked. Some of these tasks are mandatory and detailed in international and national reference texts. For example, in France, a decree describes the frequencies, tolerances, and modalities of different linear accelerator quality assurance tests (1).

The second type of task is patient-specific quality assurance. Patient-specific quality assurance is not mandatory for modulated plans but strongly recommended by international and national scientific societies. Among patient-specific quality assurance tasks, in vivo dosimetry should be evaluated when it is “technically available” according to Institut du cancer (INCa) (2), monitor units should be independently assessed, and dosimetric measurements of patient treatment plans must be achieved in phantoms using one of the following detectors: ion chamber matrix, films, or electronic portal imaging devices.

The third type of radiotherapy quality assurance detects delivery errors generally based on log files produced during treatment. This quality assurance is less common than the two previously described types, but can be very useful and should soon be generalized.

As with the other elements in the radiotherapy chain, artificial intelligence and in particular machine learning are being introduced to help teams with these quality assurance tasks, for example, by reducing their frequency or even removing them. Although these new algorithms seem very promising, distinguishing what they will really bring in practice remains difficult. For example, the function of deep learning raises new questions about its robustness and trustworthiness.

This study proposes a literature review on what artificial intelligence brings or could bring to patient-specific quality assurance (in the first section) and machine quality assurance (in the second section). We will not focus on the most technical aspects but instead on the results obtained and benefits anticipated.

2. Towards the elimination of the patient-specific quality assurance?

One of the most desired objectives is the potential of decreasing the patient-specific quality assurance workload. In the literature, the term “virtual quality assurance” is used to describe a quality assurance test that is not actually conducted but for which the results are predicted by artificial intelligence. The main goal of the following studies was to teach a machine learning system to detect if a particular patient’s radiotherapy plan will pass a routine patient-specific quality assurance test. Patient-specific quality assurance that is generally achieved for patients treated undergoing intensity-modulated radiotherapy, volumetric-modulated arc therapy, or stereotactic radiotherapy (intra- or extracranial) is generally assumed to detect plans that are too complex to be correctly delivered by a linear accelerator. For example, the plan can use an excessive number of very small multileaf collimator segments or excessively high or low dose rates. Thus, the computed dose can differ from the patient-specific quality assurance measurement. Patient-specific quality assurance can also detect machine errors (generally due to the positions of multileaf collimator leaves) during delivery that can be random or systematic. Before the first patient session, a phantom patient-specific quality assurance is supposed to detect these two possible problems (due to machine delivery errors or the plan’s excessive complexity) and eventually leads to the creation of a new plan if necessary.

Using their patient-specific quality assurance phantom, a Japanese team first controlled 161 intensity-modulated radiation beams (from prostate plans) three times: with no errors, and then after introducing multileaf collimator errors either systematically (all leaves had the same shift) or randomly (shifts were different for each leaf and randomly applied). They used some of these measurements (375) to train a convolutional neural network and the rest (108) to test its capacity to classify the measurements into three groups (no errors, random, and systematic errors). The convolutional neural network’s specificity and sensitivity were always higher than 0.94 and 0.89, respectively. Thus, the convolutional neural network was better than the gamma index pass rate to analyse phantom measurements and classify errors, but did not avoid the measurement itself (3).

To avoid patient-specific quality assurance, several approaches were tested using different machine learning tools to predict patient-specific quality assurance results without actually acquiring them. A first approach consisted of considering only the plan’s characteristics (complexity metrics) to train the machine learning systems. For example, in a study, using a Poisson regression with Lasso regularization, the complexity of 498 intensity-modulated radiotherapy plans was linked to the patient-specific quality assurance results (using a bidimensional detector) (4). A global gamma index pass rate of 3%/3 mm was predicted with an error smaller than 3%. This model was successfully exported to other institutions (5). Using the same dataset, another team trained a convolutional neural network (instead of the Poisson regression) with comparable gamma index pass rate predictions (6). Another team obtained an impressive prediction of the gamma index pass rate using a 15-layer convolutional

neural network and a dataset containing the plans' complexity, the planning target volume and rectum volumes, and the monitor unit values (7).

More recently, the specific performance of a linear accelerator (the results of machine quality assurance) was added to a treatment plan's characteristics in a dataset that was used to train a machine learning tool. A support vector classifier was then used to predict the gamma index pass rate results (8). Using a similar methodology but with a random forest algorithm, Lam et al. obtained 98% of predictions within 3% of the measured 2%/2 mm gamma index pass rate (9).

To go further than simply predicting the patient-specific quality assurance results, these predictions can be used to improve treatment planning system computations and thus no longer have failed patient-specific quality assurance tests. This interesting idea was investigated in the following study in which a dataset composed of 74 volumetric arctherapy plans from three separate institutes was pooled (10). The multileaf collimator's delivery errors were labelled by comparing the planned positions and velocities of the leaves to the real positions and velocities during delivery (obtained in the linear accelerator's dyalog files). These labelled data were used to train a machine learning system (random forest) to predict multileaf collimator errors. This prediction was incorporated into the treatment planning system dose computation (by introducing the predicted errors in the multileaf collimator sequence). The gamma index pass rate was assessed for these plans by comparing the measurements (in a phantom) to the computation including the prediction of multileaf collimator errors. The gamma index pass rate significantly improved (+4.7% for a gamma index pass rate of 1%/2 mm). A similar methodology was recently used by adding the gravity vector in the training data (depending on the gantry rotation) to improve the prediction of multileaf collimator errors (11). A few studies reported the feasibility of predicting the gamma index pass rate and classifying the results (pass/fail) using a random forest (12). In a similar approach, a convolutional neural network was also used to detect known multileaf collimator errors introduced into treatment plans when an electronic portal imaging devices gamma image was achieved (13). Again, these methods decreased the patient-specific quality assurance workload but did not dispense with them. Table 1 summarizes these different machine learning approaches for improving or removing patient-specific quality assurance.

3. Applications for machine quality assurance

Artificial intelligence and machine learning can also help physicists with machine quality assurance. For example, a team developed an automated analysis of the quality assurance of the onboard imaging. In this context, they used a support vector machine to automatically identify image artifacts (14). Another feasibility study used a similar approach based on electronic portal images to develop an automated quality assurance of gantry sag and multileaf collimator offset (15).

To go further, in a preliminary study, a team implemented a convolutional neural network to predict (using time-series modelling) the dosimetric symmetry of beams (16). In the same field, it was possible to simplify linear accelerator commissioning and quality assurance using machine learning: a team implemented a system that was able to predict dose profiles for different field sizes using only a 10 x 10 cm² field as input (17).

Another artificial intelligence challenge is predicting and preventing the facility breakdowns. There is a large amount of data available for a radiotherapy unit. The data include machine quality assurance trends and a huge number of automatically recorded parameters stored by the manufacturer in different log files. A study showed that by visually analysing the data, a physicist might predict a linear accelerator's trends to make proactive actions (18). They concluded that artificial intelligence should be more effective at analysing the data. This approach is already used in non-medical fields (19), and some preliminary studies were investigated in radiotherapy (20).

4. Discussion

As previously detailed, many studies have assessed artificial intelligence for radiotherapy quality assurance, and two important overview articles were recently published (21,22). Among these articles, there were mainly feasibility studies for machine- and patient-specific quality assurance. Using these technologies for clinical routine remains very rare. There are two main reasons. First, to implement and use these tools, very advanced skills in the fields of mathematics, computing science, and data science are needed. This knowledge is not very common in radiotherapy departments, particularly in France, because they are very new. This should change in the future, and the European Federation of Organizations for Medical Physics (EFOMP) recommends introducing these concepts in medical physics training (23). To the best of our knowledge, there are no commercial solutions to artificial intelligence for quality assurance. Moreover, even when artificial intelligence specialists are involved in the clinical implementation of such solutions, it is not obvious to understand the deep functioning of algorithms (in particular the algorithms of deep learning) that are often considered black boxes. In other radiotherapy fields (e.g., organs segmentation presented in a separate article), these machine learning tools' output can be easily assessed by a physicist or physician. However, in the quality assurance field, one of the final goals is to eliminate some quality assurance tests. For example, to replace a patient-specific quality assurance test with a "virtual test" predicted by a machine learning tool, the algorithms must be valid and robust. This is probably why most transparent machine learning tools (Bayesian networks and regression models) are presently preferred to other approaches such as deep learning.

To reduce medical teams' doubts, an interesting issue is the quality assurance of these tools themselves, that is, testing the machine learning output (22). Machine learning users generally split the available data into three separate sets: training, validation, and testing sets. However, in medical

applications in which data are sometimes difficult to obtain, other approaches are preferred, such as “leave one out” that are riskier from a medical perspective.

Thus, using machine learning for quality assurance in radiotherapy is very promising but still in its early stages. It is clearly not as advanced as it is in other medical fields such as radiology or segmentation. Other approaches to reduce the quality assurance workload are envisaged (e.g., using complexity metrics to decide if a patient-specific quality assurance test is needed). It is not obvious if these methods will be generalized in the future.

Conflict of interests

None.

Authors' contributions

LS: supervision, investigation, review and writing; CR, PM: investigation, review and writing.

References

1. Légifrance - Publications officielles - Journal officiel - JORF n° 0209 du 09/09/2007. 2007.
2. Anon. Critères d'agrément pour la pratique de la radiothérapie externe. Boulogne-Billancourt: Institut national du cancer; 2008. Available from: https://www.e-cancer.fr/content/download/58314/531708/file/criteres_radiotherapie_160608.pdf
3. Kimura Y, Kadoya N, Tomori S, Oku Y, Jingu K. Error detection using a convolutional neural network with dose difference maps in patient-specific quality assurance for volumetric modulated arc therapy. *Phys Medica*. 2020;73(April 2020):57–64.
4. Valdes G, Scheuermann R, Hung CY, Olszanski A, Bellerive M, Solberg TD. A mathematical framework for virtual IMRT QA using machine learning. *Med Phys*. 2016;43(7):4323–34.
5. Valdes G, Chan MF, Lim SB, Scheuermann R, Deasy JO, Solberg TD. IMRT QA using machine learning: A multi-institutional validation. *J Appl Clin Med Phys*. 2017;18(5):279–84.
6. Interian Y, Rideout V, Kearney VP, Gennatas E, Morin O, Cheung J, et al. Deep nets vs expert designed features in medical physics: An IMRT QA case study. *Med Phys*. 2018;45(6):2672–80.
7. Tomori S, Kadoya N, Takayama Y, Kajikawa T, Shima K, Narazaki K, et al. A deep learning-based prediction model for gamma evaluation in patient-specific quality assurance. *Med Phys*. 2018;45(9):4055–65.

8. Granville DA, Sutherland JG, Belec JG, La Russa DJ. Predicting VMAT patient-specific QA results using a support vector classifier trained on treatment plan characteristics and linac QC metrics. *Phys Med Biol*. 2019;64(9): 095017.
9. Lam D, Zhang X, Li H, Deshan Y, Schott B, Zhao T, et al. Predicting gamma passing rates for portal dosimetry-based IMRT QA using machine learning. *Med Phys*. 2019;46(10):4666–75.
10. Carlson JNK, Park JM, Park SY, Park JI, Choi Y, Ye SJ. A machine learning approach to the accurate prediction of multileaf collimator positional errors. *Phys Med Biol*. 2016;61(6):2514–31.
11. Chuang K, Giles W, Adamson J. A tool for patient specific prediction of delivery discrepancies in machine parameters using trajectory log files. *Med Phys*. 2021;48(3):978–90.
12. Li J, Wang L, Zhang X, Liu L, Li J, Chan MF, et al. Machine learning for patient-specific quality assurance of VMAT: prediction and classification accuracy HHS public access. *Int J Radiat Oncol Biol Phys*. 2019;105(4):893–902.
13. Nyflot MJ, Thammasorn P, Wootton LS, Ford EC, Chaovaitwongse WA. Deep learning for patient-specific quality assurance: Identifying errors in radiotherapy delivery by radiomic analysis of gamma images with convolutional neural networks. *Med Phys*. 2019;46(2):456–64.
14. Valdes G, Morin O, Valenciaga Y, Kirby N, Pouliot J, Chuang C. Use of truebeam developer mode for imaging QA. *J Appl Clin Med Phys*. 2015;16(4):322–33.
15. El Naqa I, Irrer J, Ritter TA, DeMarco J, Al-Hallaq H, Booth J, et al. Machine learning for automated quality assurance in radiotherapy: A proof of principle using EPID data description. *Med Phys*. 2019;46(4):1914–21.
16. Li Q, Chan MF. Predictive time-series modeling using artificial neural networks for Linac beam symmetry: an empirical study. *Ann NY Acad Sci*. 2017;1387(1):84–94.
17. Zhao W, Patil I, Han B, Yang Y, Xing L, Schüler E. Beam data modeling of linear accelerators (linacs) through machine learning and its potential applications in fast and robust linac commissioning and quality assurance. *Radiother Oncol*. 2020;153(2020):122–9.
18. Chan M, Li MF, Tang Q, Li X, Tang J, Hunt GA, et al. Visual analysis of the daily QA results of photon and electron beams of a trilogy linac over a five-year period. *Clin Eng Radiat Oncol*. 2010;4:290–9.
19. Nadai N, Melani AHA, Souza GFM, Nabeta SI. Equipment failure prediction based on neural network analysis incorporating maintainers inspection findings. In: *Proceedings - Annual Reliability and Maintainability Symposium 2017*; art. no. 7889684. DOI: 10.1109/RAM.2017.7889684.

20. Wojtasik AM, Bolt M, Clark CH, Nisbet A, Chen T. Multivariate log file analysis for multi-leaf collimator failure prediction in radiotherapy delivery. *Phys Imaging Radiat Oncol.* 2020 Jul 1;15:72–6.
21. Kalet AM, Luk SMH, Phillips MH. Radiation therapy quality assurance tasks and tools: the many roles of machine learning. *Med Phys.* 2020;47(5):e168–77.
22. Vandewinckele L, Claessens M, Dinkla A, Brouwer C, Crijns W, Verellen D, et al. Overview of artificial intelligence-based applications in radiotherapy: Recommendations for implementation and quality assurance. *Radiother Oncol.* 2020;153(2020):55–66.
23. Zanca F, Hernandez-Giron I, Avanzo M, Guidi G, Crijns W, Diaz O, et al. Expanding the medical physicist curricular and professional programme to include Artificial Intelligence. *Phys Medica* 2021;83:174–83.

Table 1. Artificial intelligence for quality assurance in radiotherapy: different uses of machine learning for patient-specific quality assurance.

Reference	machine learning type	Nature of training data	Data quantity	Expected outcome
Kimura et al. (3)	Convolutional neural network	Three-dimensional patient-specific quality assurance results	161 prostate plans	Detection of errors in patient-specific quality assurance results
Valdes et al. (4)	Poisson regression	Bidimensional patient-specific quality assurance results Plan complexity metrics	498 intensity-modulated radiotherapy plans	Virtual patient-specific quality assurance (prediction of gamma index pass rate results)
Interian et al. (6)	Convolutional neural network	Same data as (4)	Same data as (4)	Virtual patient-specific quality assurance (prediction of gamma index pass rate results)
Tomori et al. (7)	Convolutional neural network	Bidimensional patient-specific quality assurance results Plan complexity metrics Planning target volume and rectum volume	60 prostate plans	Virtual patient-specific quality assurance (prediction of gamma index pass rate results)
Granville et al. (8)	Support vector classifier	Three-dimensional patient-specific quality assurance results Plan complexity metrics Linear accelerator performance metrics	1620 volumetric arctherapy plans	Virtual patient-specific quality assurance (prediction of gamma index pass rate results)
Lam et al. (9)	Random forest	Bidimensional patient-specific quality assurance results	182 intensity-modulated	Virtual patient-specific quality assurance (prediction of gamma

		Plan complexity metrics Linear accelerator performance metrics	radiotherapy plans	index pass rate results)
Carlson et al. (10)	Random forest	Planned multileaf collimator positions and velocities Delivered multileaf collimator positions and velocities (dynalog)	74 volumetric arctherapy plans	Incorporating predicted multileaf collimator errors into treatment planning system computations
Chuang et al. (11)	Several regression models	Planned multileaf collimator positions and velocities Delivered multileaf collimator positions and velocities (dynalog) Gravity vector and gantry velocity	142 intensity-modulated radiotherapy, 125 volumetric arctherapy	Incorporating predicted multileaf collimator errors into treatment planning system computations