Development and validation of an automatic commissioning tool for the Monte Carlo dose engine in myQA iON

M. Cohilis[a], L. Hong[b], G. Janssens[c], S. Rossomme[c], E. Sterpin[a,d], J.A. Lee[a] and K. Souris[a]

[a] *Université catholique de Louvain, Institut de Recherche Expérimentale et Clinique, MIRO Lab, Brussels, 1200, Belgium*

[b] *University of Florida Proton Therapy Institute, Jacksonville, Florida, USA*

[c] *Ion Beam Applications, Louvain-la-Neuve, Belgium*

[d] *KU Leuven, Department of Oncology, Laboratory of Experimental Radiotherapy, , Leuven, 3000, Belgium*

Corresponding author: M. Cohilis

E-mail: marie.cohilis@uclouvain.be

(Dated: February 24, 2022)

Keywords: PBS proton therapy, commissioning, independent dose calculation, Monte Carlo, automation

ABSTRACT

Independent dose verification with Monte Carlo (MC) simulations is an important feature of proton therapy quality assurance (QA). However, clinical integration of such tools often generates an additional and complex work load for medical physicists. The preparation of the necessary clinical inputs, such as the machine beam model, should therefore be automated. In this work, a methodology for automatic MC commissioning has been devised, validated, and developed into a MATLAB tool for the users of *myQA iON*, the recent QA platform of IBA Dosimetry. With this workflow, all necessary parameters can easily be tuned using dedicated optimization methods. For the optical parameters, the assumption of a single or double Gaussian is made. To model the energy spectrum, a Gaussian function is assumed and parameters are optimized using either MC simulations or a library of pre-computed Bragg peaks. For the absolute dose calibration, commissioning fields can be reproduced in an accelerated way with the dose engine to retrieve the necessary parameters. We discuss in a first time the tool efficiency and show that one can optimize all parameters in less than 4 minutes per energy with excellent accuracy. We then validate a beam model obtained with the tool by simulating homogeneous spread-out Bragg peaks (SOBPs) and patient QA plans previously measured in water. An average range agreement of 0.29 ± 0.34 mm is achieved for the SOBPs while 3%/3mm local gamma passing rates reach 99.3% on average over all 62 measured patient QA planes, which is well within clinical tolerances.

I. INTRODUCTION

Proton therapy is made more and more widely available as it allows for better dose conformation than conventional radiotherapy [1, 2]. Pencil beam scanning (PBS), specifically, is a delivery technique that becomes popular due to its excellent dose sculpting and its ease of delivery. As PBS enables complex intensity-modulated proton therapy (IMPT), it requires extensive quality assurance (QA) to ensure patient safety. This process generally consists of comparing, for each beam,

the calculated dose distribution with measurement in a phantom [3, 4], which can be cumbersome. Moreover, it requires extra beam time beyond patient treatment. These drawbacks justify the increasing use of numerical simulations to replace and complement part of the QA measurements by providing a verification of the TPS planned dose. A 3D dose can also be reconstructed in the patient geometry from the spot positions and intensities recorded in the delivery log-files [5–8]. For this purpose, independent Monte Carlo (MC) dose calculation is considered as the gold standard as it is more accurate than analytical algorithms [9, 10] and allow clinicians to verify dose distributions within a framework that differs from their treatment planning system (TPS).

Although independent MC dose calculation has the potential to greatly enhance patient-specific QA (PSQA) workflows [9, 10], it also raises additional difficulties. The first concern is the ease of integration and use in conjunction with other tools of the clinical workflow, which the developers of the tool should optimize. The second issue is the generation of necessary clinical inputs such as the CT calibration curve and the machine beam model. These inputs provide the dose calculation algorithms with a description of some center-specific hardware devices like the CT scanner or the proton beam line. Modeling such a device in independent or open-source MC simulation software is often out of the area of expertise of both the tool developers and the hospital physicists. The clinicians' enthusiasm for independent MC tools can therefore get quickly curbed. To solve this problem, we believe that the commissioning procedure, including beam modeling and CT calibration, should be:

- Fully automated and user-friendly;
- Requiring no other measurements than those already performed for the initial commissioning of the TPS;
- Fast enough to comply with hardware limitations that might be imposed in clinical or corporate environments;
- Able to reproduce within tight tolerances commissioning data and the performance of the primary dose calculation in commissioning conditions.

This paper describes a complete methodology for automatic commissioning, as well as its implementation into a tool that allows inexperienced users to commission a MC algorithm for PBS

treatment machines of most leading proton therapy system manufacturers. We developed this tool in MATLAB to optimize beam models for the fast open-source MC algorithm MCsquare [11, 12], which is also the MC dose engine integrated into the new QA platform of IBA Dosimetry, *myQA iON*. We thus report an extensive assessment of the commissioning methodology used by *myQA iON* users, which is tested and validated for the compact gantry (Proteus®ONE, IBA) of the University of Florida Proton Therapy Institute (UFPTI).

II. MATERIALS AND METHODS

A. General presentation

MCsquare is a fast open-source MC tool dedicated to the simulation of PBS proton therapy treatments. In MCsquare, beam characteristics are interpolated from a parameters look-up table. These energy dependent parameters can be classified as follows [12–14]:

- **Phase space:**

  Optical parameters of the beam geometry and defined at the nozzle exit, in both lateral directions x and y: spot size (σ_x, σ_y), which describes the spatial distribution, beam divergence (σ_θ, σ_ϕ), which models the angular distribution and correlation between spatial and angular spreads ($\rho_{x\theta}$, $\rho_{y\phi}$);

- **Energy spectrum:**

  Parameters determining the shape of the integrated depth dose (IDD): mean energy (E_m) and energy spread (σ_E) that reproduce the IDD curve;

- **Absolute dose:**

  Parameters ensuring correct absolute doses: number of protons per monitor unit (P_{MU}).

Regarding CT calibration, MCsquare requires two input files: HU to mass density and HU to material conversion files. Here a classical stoichiometric method [15] was implemented and is therefore not discussed further.

The commissioning tool itself is released as an open-source code [16] and comes with a graphical user interface to help with the optimization and validation of the beam model, the addition of

range shifters, and the generation of a CT calibration curve. As inputs, the tool directly reads measurement files extracted from the RayStation or Eclipse TPS to avoid time losses and errors.

B. Commissioning procedure

This section briefly describes the measurements needed for beam modeling along with the methods implemented in the commissioning tool of *myQA iON*. More detailed descriptions of the measurement process can be found elsewhere [12–14, 17].

1. Phase space

In order to model the beam geometry in three dimensions, pencil beams are measured in transverse planes at several distances from the isocenter. One-dimensional profiles, in X and Y directions, are then generally extracted from the 2D acquisitions as required by the TPS.

Laterally, a proton beam fluence is typically modeled with a 2D Gaussian function. In some cases, the halo can also become significant [18] and the beam model then may require a second, wider Gaussian function. Recently, it was also suggested that stable distributions, i.e., a generalization of the normal distributions, could slightly improve accuracy [19].

In this commissioning tool, single or double Gaussian modeling are possible.

For every chosen commissioning energy, post-processing of the measurements goes through two steps:

1. *Lateral modeling:* One or two Gaussian functions are fitted to the X and Y lateral profiles. Spot sizes at various distances from isocenter are thus determined as the standard deviation of the Gaussian function(s).

2. *3D modeling:* Relying on Courant-Snyder theory [20], modeling along the beam direction can be a function of previously extracted spot sizes:

$$\begin{cases} \sigma_x^2(z) = \sigma_x^2(0) - 2\rho_{x\theta}(0)\sigma_x(0)\sigma_\theta(0)z + \sigma_\theta^2(0)z^2 \\ \rho_{x\theta}(z) = \dfrac{\rho_{x\theta}(0)\sigma_x(0) - \sigma_\theta(0)z}{\sigma_x(z)} \\ \sigma_\theta(z) = \sigma_\theta(0) \end{cases}$$

and similarly for the Y axis. A MATLAB built-in least-squares fit is used. Note that Courant-Snyder theory concerns purely beam geometry and does not account for scattering in air.

2. *Energy spectrum*

The delivered beam is rarely mono-energetic. This energy spread must be taken into account when computing the dose, especially when dealing with such particles as protons, which have a well defined range. The beam energy spectrum therefore needs to be modeled. MCsquare assumes a Gaussian spectrum for proton energies, which is a widely accepted approximation shown to be efficient [13, 14]. The parameters to be tuned are thus the mean energy and its standard deviation.

To determine the energy spectrum parameters, integral depth doses (IDD) of pencil-beams with one given nominal energy are measured in a water phantom, for each selected commissioning energy.

As the energy spectrum parameters will affect the shape of the Bragg peak, the difference between simulated and measured IDDs must be minimized by adjusting the energy spectrum parameters. This process is often manual, which can quickly become long and tedious. In the proposed commissioning tool, two different methods can automate the energy spectrum tuning:

- *Fast mode.* In a similar way as Kimstrand et al. [21], the optimization involves a library of precomputed mono-energetic IDDs. MCsquare generated these IDDs for generic phase space parameters, to span the whole range of clinical energies between 50 and 280 MeV, with steps of 0.05 MeV. The energy spectrum can thus be optimized by fitting a weighted sum of the pre-computed IDDs to the measured IDD. This method has the drawback that the generic phase space does not take into account the actual optical parameters found at the previous step (spot sizes, angular spread, correlation). Instead, these are either constant or linearly interpolated over the whole energy range between fixed extreme values with spot sizes between 9 and 1 mm, divergence of 0.004 rad, and correlation of 0.5.

- *Accurate mode.* The optimization runs the Nelder-Mead algorithm [22], which offers the

advantage to be robust to the stochastic noise of MC simulations. The downside of this method is its slowness, as each iteration entails a MC simulation. To speed it up, an accurate guess is provided as initial condition by preliminary optimization of the energy spectrum with the pre-computed IDD of the fast mode.

In both cases, the objective function to be optimized involves several metrics: the range 80 error $E_{r_{80}}$ in mm, the dose-to-peak error E_{dtp} in %, the Bragg peak width error E_{BPw} in mm, and the range 20 error $E_{r_{20}}$ in mm. A weighting system favors essential metrics, yielding the objective function

$$f(x) = 10E_{r_{80}}^2 + E_{r_{20}}^2 + (1.2E_{\mathrm{BPw}} - E_{\mathrm{dtp}})^2 \ .$$

3. Protons per MU

The monitor unit (MU) generally varies from one treatment machine to the other and is defined by a given dose measured in reference conditions. In order to obtain a correct absolute dose with the commissioned MC algorithm, the number of protons per MU must be calculated. Two different measurement setups can achieve this goal: either the dose is measured in the center of a large homogeneous field [23], or a single pencil beam is delivered and measured with an appropriate, i.e., sufficiently large, IC.

In order to tune the number of protons per MU for each commissioning energy, Gomà's method is used [24]. By re-simulating the commissioning experience with the MC tool, it is possible to retrieve the dose per proton $D_{\mathrm{MC,p}}$ obtained by the detector. The number of protons per MU is then

$$P_{\mathrm{MU}} = \frac{D_{\mathrm{meas,MU}}}{D_{\mathrm{MC,p}}} \ ,$$

where $D_{\mathrm{meas,MU}}$ is the dose per MU obtained through the commissioning measurement.

The commissioning workflow provides two possible methods to simulate the commissioning experiment with the MC algorithm:

- *Accurate mode:* The full commissioning measurement is simulated. For example, if a large

field of N spots was measured during commissioning, then the N spots are simulated with the MC code. Computing until the MC noise reaches the desired accuracy level can take a long time.

- *Fast mode:* Only one pencil beam is simulated. If the measurement consists of a large field, the simulated beamlet is duplicated and rotated multiple times within MATLAB in order to obtain the full field, using the information on the source-to-axis distances. This allows the user to spare a lot of simulation time while keeping excellent accuracy.

C. Model assessment

The commissioning tool was used to tune the beam model of the new compact gantry of UFPTI in order to prepare the use of *myQA iON* in the clinic. Both accurate and fast modes have been benchmarked. The efficiency of the methodology is evaluated in two steps. First, we comment on calculation speed and optimization accuracy based on the commissioning measurements only. Next, we assess the quality of the UFPTI beam model with a classical clinical validation scheme.

1. Workflow efficiency

The calculation time was measured for all three different beam commissioning phases, i.e., phase space, energy spectrum, and protons/MU tuning. For the last two, both fast and accurate modes were considered. Calculations were performed on a Linux system with two Intel Xeon Gold 6248 (total 40 CPU cores at 2.5GHz).

The accuracy of the phase space parameters calculated at the nozzle exit was assessed by propagating them to the measured positions with the Courant-Snyder theory. Calculated values were then compared to initial spot size measurements.

For the energy spectrum, the optimization was assessed with various error metrics enabling a comparison between measured and final simulated IDDs. Range, Bragg peak width, and dose-to-peak errors were considered.

Finally, fast and accurate modes for the commissioning of the absolute dosimetry were compared in terms of relative differences of the beam model parameters, i.e., protons/MU.

2. *Clinical validation*

Clinical validation of the UPFTI Proteus®ONE machine beam model went through two different steps. First, various spread-out Bragg peaks (SOBP) were measured in water using the Zebra multi-layer ionization chamber (IBA Dosimetry). The obtained depth-dose profiles were compared to MCsquare simulations. Cubic targets of 3x3x3, 5x5x5, 7x7x7 and 10x10x10 cm^3 were delivered at various isocenter depths, resulting in 18 different SOBPs. The isocenter was placed at 5, 10, 15, 20, and 25 cm depth for the two smallest cubes and at 5, 10, 15, and 20 cm depth for the two largest cubes. These SOBPs were also delivered on a MatriXX 2D array detector (IBA Dosimetry) placed in water, at isocenter depth. Doses were recomputed in *myQA iON* with a resolution of 0.59x0.59x5 mm^3 and then resampled to 2x2x2 mm^3. A noise level below 1% was achieved by simulating 10^8 particles. A 2D local gamma index (3%/3mm and 2%/2mm) compared *myQA iON* simulations to MatriXX measurements, with the latter as reference.

As a second step, the beam model was evaluated by recalculation of PSQA fields. The MatriXX detector measured dose in transverse planes at two specific depths for each field, in water. Breast, prostate, brain, head and neck (H&N), and cranio-spinal (CSI) cases were studied, for a total of 10 patients and 62 plane measurements. Breast, H&N, CSI, and some of the brain cases featured a range shifter. Doses were computed by MCsquare at a resolution of 2x2x2 mm^3 and with an uncertainty of 1%. A 2D gamma local index (3%/3mm and 2%/2mm) was used to compare *myQA iON* results to measurements, with measurements set as reference.

III. RESULTS

A. Workflow efficiency

Table I reports mean computation times over all energies for each of the three beam commissioning steps. Both fast and accurate modes are considered for the tuning of the energy spectrum and protons/MU. Fast modes show huge time gains, especially for the energy spectrum.

The comparison between measured and recalculated spot sizes yielded good results. In X, the error remained below 0.06 mm while, for Y, it reached up to 0.23 mm.

Average computation times				
Phase space	Energy spectrum		Absolute dosimetry	
	Fast	*Accurate*	*Fast*	*Accurate*
0.16 s	0.80 s	2225.22 s	229.44 s	5230.49 s

Table I. Beam commissioning: computation times averaged over all commissioned energies for all three parts of the tuning process.

Figure 1 shows the evaluation of various depth-dose metrics for the accurate and the fast modes. As it can be seen, the accurate mode for the energy spectrum allows for slightly better accuracy, especially for the physical range. Range 80 is systematically too short when performing a MC simulation with the fast-generated beam model. However, this error remains below 0.05 mm for each commissioned energy. For the accurate mode, the range error is 0.004 ± 0.006 mm.

Figure 2 shows the distribution of the relative error on the protons/MU computed between the accurate and the fast modes. The maximum absolute error is below 0.4%, with a time gain fluctuating between 15 and 32 with respect to nominal energy.

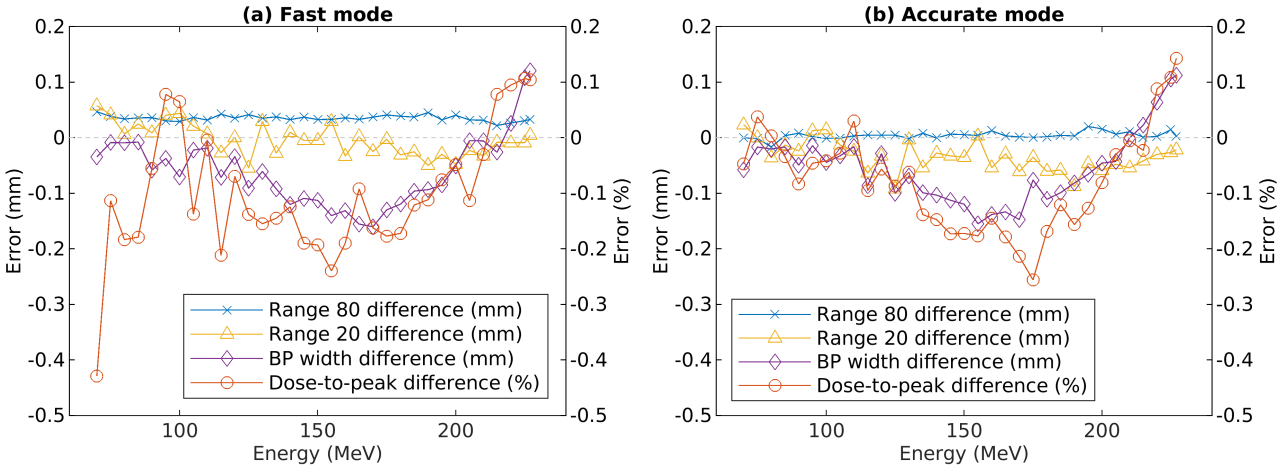

Figure 1. Evaluation of the accuracy of the energy spectrum tuning using (a) the fast mode; (b) the accurate mode. Error metrics are obtained after running a MCsquare simulation of a single pencil beam in water and comparing with commissioning measurements.

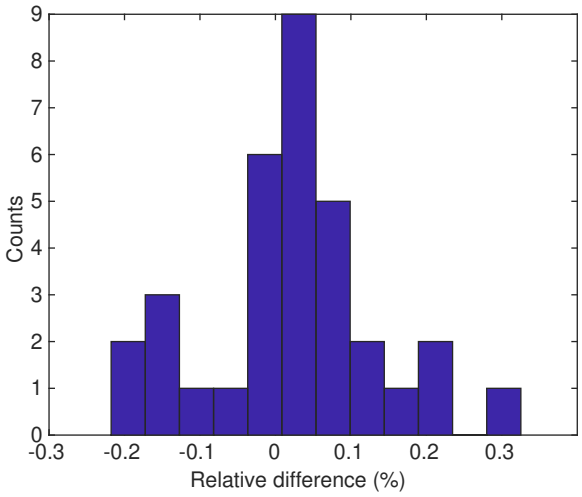

Figure 2. Distribution of the relative difference between protons/MU obtained with the fast and the accurate modes.

B. Clinical validation

1. SOBPs in water

Figure 3 compares depth-dose curves measured by the Zebra and simulated using *myQA iON* for 5x5x5 cm^3 SOBP beams. Overall a very good agreement is observed, except in the plateau of longest range SOBP where the nuclear reaction model used in the MC dose engine may cause some discrepancies. Most points in the plateau and SOBP regions remain nevertheless within 2% of the Zebra reference, as shown in Figure 4 for two high energy cubes. The figure also compares Zebra with the RayStation MC simulation, that performs similarly to *myQA iON*. Over all SOBPs, an average clinical range agreement of 0.29 ± 0.34 mm was obtained.

Table II reports the results of the 2D gamma index for SOBP transverse planes. In each case, the results for both 3%/3mm and 2%/2mm criteria are shown.

2. Patient-specific QA

Figure 5 shows the distribution of gamma passing rates obtained for all 62 PSQA transverse planes studied. Criteria of 3%/3mm and 2%/2mm were used.

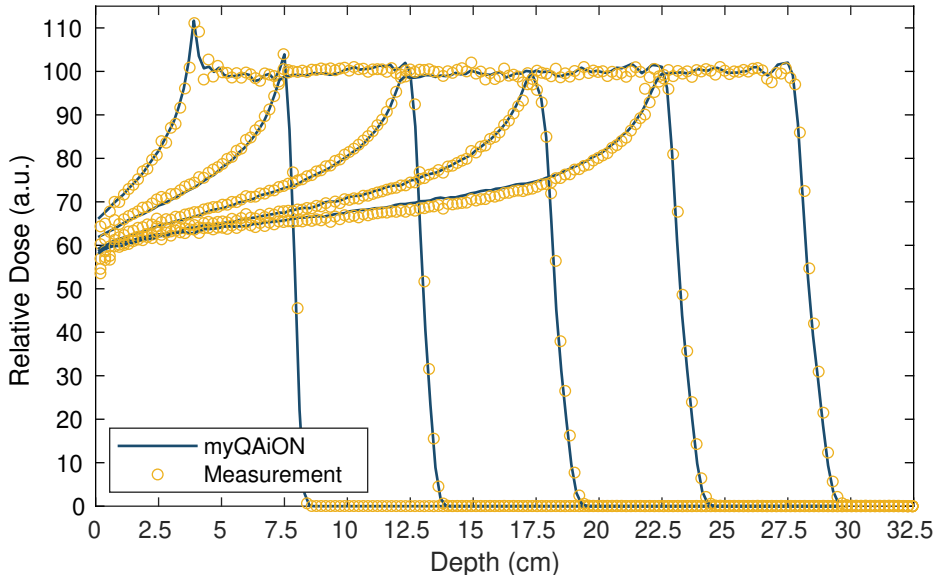

Figure 3. SOBPs of 5x5x5 cm^3 in water: comparison between Zebra measurements and *myQA iON*.

	3x3x3 cm^3		**5x5x5 cm^3**		**7x7x7 cm^3**		**10x10x10 cm^3**	
	3%/3mm	*2%/2mm*	*3%/3mm*	*2%/2mm*	*3%/3mm*	*2%/2mm*	*3%/3mm*	*2%/2mm*
5 cm depth	97.0	90.9	100	94.6	100	97.3	100	94.6
10 cm depth	100	98.3	100	100	100	96.8	100	95.0
15 cm depth	98.3	93.3	100	91.9	100	94.8	100	96.0
20 cm depth	100	96.7	99.2	93.7	100	95.7	100	96.0
25 cm depth	100	96.0	100	95.5	N.A.	N.A	N.A.	N.A.

Table II. Resulting passing rates (%) of 2D gamma index comparisons between *myQA iON* and Matrixx measurements for cubic SOBPs.

IV. DISCUSSION

Simulation-based QA has raised more and more interest lately as it saves time in the clinical workflow of radiotherapy. In particular, the use of a dose engine, independent from the main TPS, becomes frequent in proton therapy facilities. Integration of these tools, however, is long and complex for the hospital physicists.

The IBA Dosimetry QA software for proton therapy, *myQA iON*, makes no exception. It features a MC dose engine named MCsquare that needs to be commissioned in every facility acquiring this software. For this reason, we collaborated with IBA Dosimetry to develop a methodology for fast and reliable commissioning of MCsquare with a minimal work load [16]. The source code

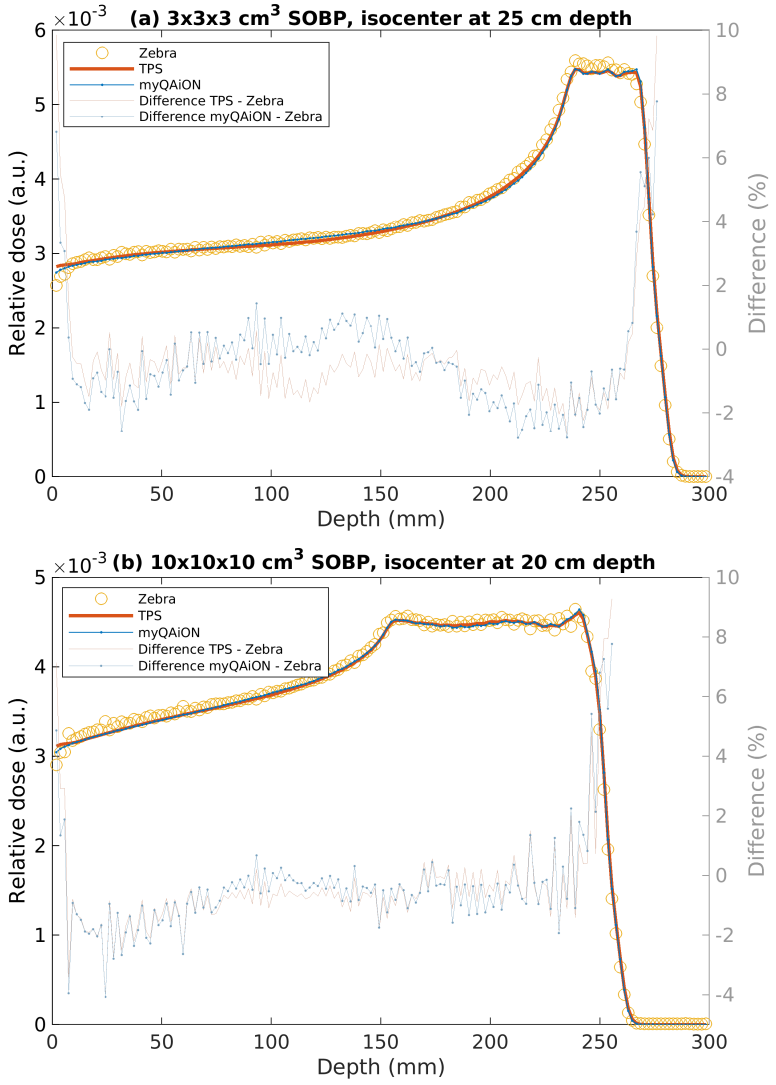

Figure 4. SOBPs in water: comparison of Zebra, TPS and *myQA iON* for high energy SOBPs: (a) 3x3x3 cm^3, (b) 10x10x10 cm^3.

of MCsquare is open, publicly available and several institutions use it independently from *myQA iON* for research and clinical developments [11, 25].

In this paper, we describe the methods implemented in the commissioning tool, benchmark its performance, and validate a beam model obtained from the tool for clinical use. To tune the phase space, no MC simulations are required and computational performance is therefore not an issue. With the chosen model, we were able to predict spot sizes within 0.06 mm in X and 0.23 mm in Y. In comparison, Grevillot et al. report errors below 0.15 mm [13]. The slightly lower accuracy in the Y direction may originate from an outlier in our measured data. The spot size measured at the

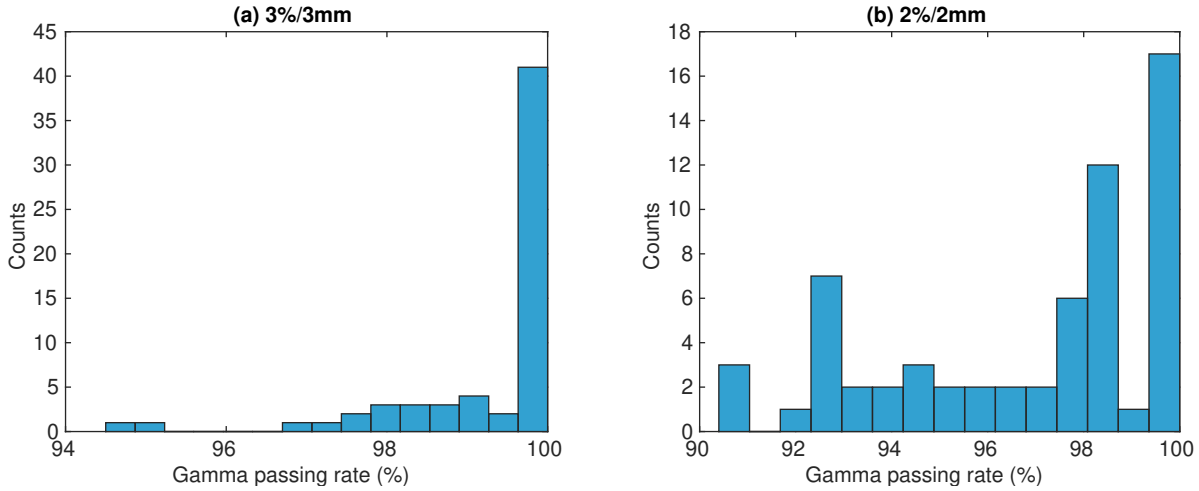

Figure 5. Gamma passing rates distribution for the 62 PSQA transverse planes: (a) 3%/3mm, (b) 2%/2mm.

closest position to the nozzle exit (400 mm) seems actually too small in comparison with others and the equations cannot reproduce it correctly. Figure 6 illustrates it, showing the variation of the spot size in Y at different distances from isocenter for an energy of 120 MeV. Although our main goal is normally to predict the spot size close to isocenter, the use of a range shifter is frequent in PBS and correct modeling of the phase space is thus also necessary ahead of isocenter. Therefore, it remains important to account for this spot size in the fitting process.

To tune the energy spectrum, two different methods were implemented. The fast method allows

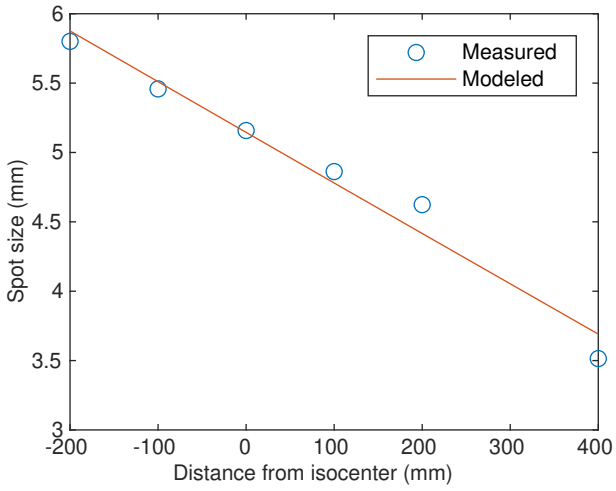

Figure 6. Spot sizes in Y direction at various distances from isocenter, for a proton beam of 120 MeV.

the user to optimize the beam model with no specific hardware requirements due to the use of a mono-energetic Bragg peaks library, while the second one can be long due to the amount of MC simulations to be run repeatedly. However, we found that the fast mode is accurate enough, except if the distance between the nozzle exit and the isocenter differs from the one assumed when creating the generic library of Bragg peaks (see Figure 1). This problem might originate from an incorrect length of air crossed by protons, resulting in slightly shifted ranges in the library. To address this issue and to make the fast mode available to any user, we are currently enlarging the library. In terms of accuracy, ranges were reproduced within 0.05 mm and 0.02 mm for the fast and accurate modes, respectively, over all commissioned energies. The width of the Bragg peaks, evaluated at 60% of peak dose, were always within 0.16 mm for the fast mode and 0.15 mm for the accurate modes. For the dose-to-peak metric, the error remained within 0.43% with the fast mode and 0.26% with the accurate mode. This is globally an improvement compared to what was reported elsewhere [12], [14], [13], although not always the exact same metrics were evaluated. For example, Huang et al. report range errors within 0.1 mm, as well as average FWHM and peak-to-plateau ratio differences of 1 mm and 0.4%, respectively [12]. Fracchiolla et al. obtained clinical ranges within 0.1 mm and dose-to-peak differences below 2% [14]. It is noteworthy, however, that the accuracy of the reference measurement itself is limited, due to a possible imprecise setup or a small error in the WET of the chamber wall. Paganetti reports an uncertainty of 0.3 mm on the measured ranges [26], while Bäumer et al. even suggest 0.5 mm [27]. These numbers being already quite high with respect to the range uncertainties that are considered in PBS treatment planning, it highlights the necessity to limit additional uncertainties stemming from the least-squares fit.

For the tuning of the number of protons per MU, fast and accurate modes were also implemented. A large field must often be simulated for each commissioning energy in order to adjust this parameter and very low MC statistical noise is necessary to obtain accurate results. For this purpose, we proposed a fast method that simulates a single pencil beam and builds the rest of the field by successive rotations of the simulated beamlet. Alternatively, we could have implemented a convolution [14] or a dose area product approach [28], assuming parallel beams. Nevertheless, generality was a key feature we sought for this tool and we chose to avoid all kinds of approximation. This still resulted in a mean gain factor of over 22 compared to the traditional method for a reference field of 10x10 cm^2.

As a proof-of-concept, we then validated the beam model obtained for the compact gantry of UFPTI. A range agreement of 0.29 ± 0.34 mm was observed with Zebra measurements of SOBPs in water (see Figures 3 and 4), well within the accuracy of the Zebra reported by IBA Dosimetry (an accuracy inferior or equal to 0.5 mm compared to measurements in water). The largest discrepancies were observed in the plateau region of the IDD. Several reasons can explain this. First, MCsquare nuclear cross sections are based on ICRU-63 report [29] and might not reproduce entrance dose as accurately as in slower general-purpose codes such as Geant4 [30]. Uncertainties on the cross sections are reported to be between 5 and 10% by ICRU [29]. Second, we clearly see from Figures 3 and 4 that non-negligible noise affects the Zebra measurements, especially in the first half of the plateau. This could be explained partly by inaccuracies in the MLIC calibration and especially its uniformity calibration. Finally, as the curves are only relative measurements, normalization will also impact the exact reported errors. The comparison between various measured and simulated transverse planes of cubic SOBPs yielded reasonable results from the gamma index analysis. The AAPM Task Group 185, which provides guidelines as to clinical commissioning of IMPT systems, mentions in its recent report that dose measurements and simulations "should be brought to within an agreement of 2%/2mm 98% for uniform fields in homogeneous media" [31]. This is not the case here as only two of the fields comply with this recommendation (cf. Table II). However, this is not surprising for the following reasons. First, a resolution of 2x2x2 mm^3 was used for the MC dose in the gamma index. As the current implementation of the gamma index in *myQA iON* does not interpolate the evaluated dose, this means that the 2%/2mm gamma index actually behaved more like a stringent dose difference metric than like a 2%/2mm gamma index. Besides, the AAPM TG 218 report recommends using a spacing below one third of the gamma index distance-to-agreement (DTA) criterion for the evaluated dose distribution [32], and Cohilis et al. show the negative impact of a low resolution on a practical case [33]. Second, the resampling that occurs before computing the gamma index could also affect the results of the dose comparison. It is suggested in the literature that resampling the evaluated MC dose would lead to underestimated passing rates compared to a direct calculation with the desired resolution [33]. All the achieved 2%/2mm passing rates remain nevertheless above 90% and we believe that they are mainly the result of the resolution choices and that an improved

resolution in MC simulations, complying with the recommendation of AAPM TG 218 [32], would have considerably enhanced the results.

This statement is also supported by the results achieved for the PSQA plans. The recommendation of AAPM TG 185 for PSQA inhomogeneous fields is to use a criterion of 3%/3mm, and they mention that we should obtain 95% of the points passing [31]. All 3%/3mm dose comparisons scored here above 95%, except one that scored 94.5%, with an excellent average of 99.3% (cf. Figure 5). Moreover, all 2%/2mm gamma index yielded passing rates above 90%, with an average of 96.9%. These results might actually still be underestimated, considering the resolution of 2x2x2 mm^3 used for MC simulations and the recommendation of AAPM TG 218 to use for the evaluated dose a spacing below one third of the DTA criterion [32] to avoid negative effects such as previously described.

Although the automatic commissioning workflow was tested here on the data of a single facility, it has been used already to commission many treatment rooms, equipped with PT systems of different manufacturers, raising no particular issue and proving to be robust.

V. CONCLUSIONS

An easy and fast commissioning procedure could open the path to an increased use of independent Monte Carlo algorithms in clinics, as well as in research institutions. Multicenter studies, for example, could benefit from this. We thus propose an automatic commissioning tool that is user-friendly, fast and generic enough to be used in the commissioning of most proton therapy systems. The tool is now used to commission *myQA iON* but its source code remains open and accessible to any MCsquare user.

FUNDING

Marie Cohilis is supported by the Télévie Grant from the Belgian *Fonds National pour la Recherche Scientifique* F.R.S.-FNRS (Grant No. 7450517F).

Kevin Souris is funded by the Walloon region (MECATECH/BIOWIN, grant number 8090).

John A. Lee is a Senior Research Associate with the Belgian *Fonds National pour la Recherche*

Scientifique F.R.S.-FNRS.

DECLARATION OF INTEREST

This project was conducted in close collaboration with IBA Dosimetry.

---

[1] Clark CH, Bidmead AM, Mubata CD, Harrington KJ, Nutting CM. Intensity-modulated radiotherapy improves target coverage, spinal cord sparing and allows dose escalation in patients with locally advanced cancer of the larynx. Radiother Oncol 2004;70(2):189–198.

[2] Zhang X, Li Y, Pan X, Xiaoqiang L, Mohan R, Komaki R, Cox JD, Chang JY. Intensity-Modulated Proton Therapy Reduces the Dose to Normal Tissue Compared With Intensity-Modulated Radiation Therapy or Passive Scattering Proton Therapy and Enables Individualized Radical Radiotherapy for Extensive Stage IIIB Non-Small-Cell Lung Cancer: A Virtual Clinical Study. Int J Radiat Oncol Biol Phys 2010;77(2):357–366.

[3] Arjomandy B, Sahoo N, Ciangaru R, Zhu R, Song X, Gillin M. Verification of patient-specific dose distributions in proton therapy using a commercial two-dimensional ion chamber array. Med Phys 2012;37(11):5831–5837.

[4] Lin L, Kang M, Solberg T, Mertens T, Baeumer C, Ainsley C, McDonough J. Use of a novel two-dimensional ionization chamber array for pencil beam scanning proton therapy beam quality assurance. J Appl Clin Med Phys 2015;16(3):270–276.

[5] Meier G, Besson R, Nanz A, Safai S, Lomax A. Independent dose calculations for commissioning, quality assurance and dose reconstruction of PBS proton therapy. Phys Med Biol 2015;60(7):2819–2836.

[6] Winterhalter C, Fura E, Tian Y, Aitkenhead A, Bolsi A, Dieterle M, Fred A, Meier G, Oxley D, Siewert D, Weber D, Lomax A, Safai S. Validating a Monte Carlo approach to absolute dose quality assurance for proton pencil beam scanning. Phys Med Biol 2018;63(17).

[7] Guterres Marmitt G, Pin A, Ng Wei Siang K, Janssens G, Souris K, Cohilis M, Langendijk J, Both S, Knopf A, Meijers A. Platform for automatic patient quality assurance via Monte Carlo simulations

in proton therapy. Phys Med 2000;70:49–57.

[8] Toscano S, Souris K, Gomá C, Barragan Montero A, Puydupin S, Vander Stappen F, Janssens G, Matic A, Geets X, Sterpin E. The calibration of CT Hounsfield units for radiotherapy treatment planning. Phys Med Biol 2019;41:111–124.

[9] Yepes P, Adair A, Grosshans D, Mirkovic D, Poenish F, Titt U, Wang Q, Mohan R. Comparison of Monte Carlo and analytical dose computations for intensity modulated proton therapy. Phys Med Biol 2018;63(4).

[10] Huang S, Souris K, Li S, Kang M, Barragan Montero AM, Guillaume J, Lin A, Garver E, Ainsley P Christopher Taylor, Xiao Y, Lin L. Validation and application of a fast Monte Carlo algorithm for assessing the clinical impact of approximations in analytical dose calculations for pencil beam scanning proton therapy. Med Phys 2018;45(12):5631–5642.

[11] Souris K, Lee JA, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. Med Phys 2016;43(4):1700–1712.

[12] Huang S, Kang M, Souris K, Ainsley C, Solberg TD, McDonough JE, Simone II CB, Lin L. Validation and clinical implementation of an accurate Monte Carlo code for pencil beam scanning proton therapy. J Appl Clin Med Phys 2018;19(5):558–572.

[13] Grevillot L, Bertrand D, Freud N, Sarrut D. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4. Phys Med Biol 2011;56:5203–5219.

[14] Fracchiolla F, Lorentini S, Widesott L, Schwarz M. Characterization and validation of a Monte Carlo code for independent dose calculation in proton therapy treatments with pencil beam scanning. Phys Med Biol 2015;60:8601–8619.

[15] Schneider W, Bortfeld T, Schlegel W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. Phys Med Biol 2000;45(2):459–478.

[16] `https://gitlab.com/openmcsquare/commissioning`, accessed in 2021.

[17] Schneider U, Pedroni E, Lomax A. The calibration of CT Hounsfield units for radiotherapy treatment planning. Phys Med Biol 1996;41:111–124.

[18] Harms J, Chang C, Zhang R, Lin L. Nuclear halo measurements for accurate prediction of field size factor in a Varian ProBeam proton PBS system. J Appl Clin Med Phys 2019;21(1):197–204.

[19] Van den Heuvel F, George B, Schreude N, Fiorini F. Using stable distributions to characterize proton pencil beams. Med Phys 2018;45(5):2278–2288.

[20] Courant E, Snyder H. Theory of the alternating-gradient synchrotron. Ann Phys 1958;3:1–48.

[21] Kimstrand P, Traneus E, Ahnesjö A, Grusell E, Glimelius B, Tilly N. A beam source model for scanned proton beams. Phys Med Biol 2007;52:3151–3168.

[22] Nelder J, Mead R. A Simplex Method for Function Minimization. The Computer Journal 1965; 7(4):308–313.

[23] Andreo P, Burns D, Hohlfeld K, Huq M, Kanai T, Laitano F, Smyth V, Vynckier S. TRS-398 - Absorbed dose determination in external beam radiotherapy: an international code of practice for dosimetry based on standards of absorbed dose to water. IAEA, 2006.

[24] Gomà C, Lorentini S, Meer D, Safai S. Proton beam monitor chamber calibration. Phys Med Biol 2014;59:4961–4971.

[25] `https://openmcsquare.org`, accessed in 2021.

[26] Paganetti H. Range uncertainties in proton therapy and the role of Monte Carlo simulations. Phys Med Biol 2012;57(11):99–117.

[27] Bäumer C, Koska B, Lambert J, Timmermann B, Mertens T, Takoukam Talla P. Evaluation of detectors for acquisition of pristine depth-dose curves in pencil beam scanning. J Appl Clin Med Phys 2015;16(6):151–163.

[28] Gomà C, Vörös S, Safai S. Reference dosimetry of proton pencil beams based on dose-area product: a proof of concept. Phys Med Biol 2017;62:4991–5005.

[29] International Commission in Radiation Units and Measurements. Nuclear data for neutron and proton radiotherapy and for radioprotection (ICRU report 63). ICRU Publications, Oxford, United Kingdom, 2000.

[30] Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce M P Asai, Axen D, Banerjee S, Barrand G, Behner F, Bellagamba J, Boudreau J, Broglia L, Brunengo A, Burkhardt H, Chauvie S, Chuma J, Chytracek R, Cooperman G. Geant4 — a simulation toolkit. Nucl Instrum Methods Phys Res A 2003;506(3):250–303.

[31] Farr J, Moyers M, Allgower C, Bues M, Hsi W, Jin H, Mihailidis DN, Lu H, Newhauser W, Sahoo N, Slopsema R, Yeung D, Zhu X. Clinical commissioning of intensity-modulated proton therapy

systems: Report of AAPM Task Group 185. Med Phys 2020;48(1):e1–e30.

[32] Miften M, Olch A, Mihailidis D, Moran J, Pwlicki T, Molineu A, Li H, Wijesooriya K, Shi J, Xia P, Papanikolaou N, Low D. Tolerance limits and methodologies for IMRT measurement-based verification QA: Recommendations of AAPM Task Group No. 218. Med Phys 2018;45(4):e53–e83.

[33] Cohilis M, Sterpin E, Lee J, Souris K. A noise correction of the gamma index method for Monte Carlo dose distribution comparison. Med Phys 2020;47(2):681–692. https://doi.org/10.1002/mp.13888.