The quest for sensible data analysis in clinical routine: Study case on the new Octavius1500 array and its associated phantoms

Ann Van Esch1,2, Dominique P Huyskens1,2, K Basta2, M Ghislain2 and R Delvaux2
17Sigma, QA-team in radiotherapy physics, Kasteeldreef 2, 3150 Tildonk, Belgium
2Radiotherapy Association, Centre Hospitalier Mouscron - Ste Elisabeth Namur, 7700 Mouscron - 5000 Namur, Belgium
E-mail: ann.vanesch@7sigma.be

Abstract. We have investigated the dosimetric characteristics of the new Octavius1500 array and its associated phantoms (Oct2D and Oct4D) and compared it to the previous systems. In line with the rising amount of publications advocating the retirement of the widely used but overly lenient 3%, 3mm (global dose) gamma analysis criteria, we have also focused on deriving appropriate evaluation criteria to be used with the new measurement systems in clinical routine. Using the Octavius1500 study as an example, we advocate the use of more selective, equipment specific gamma criteria, all starting from the same, very strict acceptance criteria but adding measurement uncertainty to these.

1. Introduction
The purpose of QA is simple: we want to make sure that the dose we have calculated for a radiotherapy treatment plan is the same as the dose we deliver. But how good does the agreement have to be? When looking into research and development, precision is the key and one can be as demanding as one chooses to be. When assessing clinical relevance, however, a more pragmatic approach is desirable and efficiency becomes a primary consideration in the QA process.

To perform QA in clinical routine, most physicists are dependent on the commercially available solutions. As rotational treatments are taking up an ever increasing share of all radiotherapy treatments, the most commonly used verification methods have gradually moved from field-by-field to composite verification through measurements. Composite plan measurement systems produce either planar (2D) or three dimensional (3D) dose information [1-23]. Planar data are obtained from the measurement in a straightforward manner, with data manipulation mostly limited to the application of energy (and sometimes angle) dependent calibration files and correction factors. Only the actually measured datapoints are subjected to gamma evaluation analysis. This is in contrast to most commercial 3D dosimetry systems, for which the 3D dose is not actually measured but reconstructed from a (series of) measurements. These QA systems are considerably more complex than their 2D predecessors and often suffer from a substantial black box-factor. And yet, for most of the radiotherapy centers, the purchased phantom is put to use with limited - if any - validation and the rather universal 3%G,3mm gamma criteria are applied, regardless of the specific dosimetric
characteristics of the measurement system. We would therefore like to join Nelms et al. [24-29] in making a case for the revision of the way in which the gamma analysis is used in clinical practice. Rather than suggesting sufficiently large, universal acceptance criteria to ensure the continued use of the relatively insensitive, binary approach of the gamma evaluation (either pass or fail), we would advocate the use of more selective, equipment specific gamma criteria, all starting from the same, very strict acceptance criteria but adding measurement uncertainty to these. For now, as an initial starting point, we will assume that no clinically relevant changes are expected when the delivered ('real') and calculated doses in the homogeneous phantom agree within 2%L, 2mm for almost the entire irradiated volume (e.g. for 97% of the volume receiving at least 10% of the maximum dose). These starting criteria should be subjected to regular re-evaluation as radiotherapy systems further improve, but with a state of the art, properly commissioned, high quality radiotherapy system (comprised of a dose calculation and dose delivery system), this overall agreement between the calculated and the actual dose should be achievable in a homogeneous phantom for all but the lowest doses.

In the below study, we have used a prototype of the new Octavius1500 array (PTW, Freiburg, Germany) in combination with the Octavius4D phantom as a case-study to obtain the above mentioned, QA-system specific evaluation criteria. A more elaborate description of the below material was published in Medical Physics, September 2014 [30].

2. Methods, Materials and Results

2.1 The Octavius1500 system

The detector is shown in figure 1a. The 1405 ion chambers - of 0.44x0.44x0.30 cm³ each - cover a 27x27 cm² active area and are arranged in a checkerboard pattern. The Octavius1500 array has twice the detector density of its 729 predecessor, resulting in an area coverage of almost 50%. If necessary, 100% coverage can be achieved by performing two measurements, shifting the setup over 5mm in between and merging both dataset within the software. As with every PTW ion chamber array, the array is already calibrated in the 60Co beam at the PTW secondary standard dosimetry laboratory upon delivery. The detector matrix can be used in combination with solid water plates for orthogonal measurement setups as well as with phantoms of the Octavius® series, specifically designed for multidirectional irradiation.

In order to assess the uncertainties in the final 3D dose reconstruction, it is vital to understand the measurement process. The steps in the QA process by means of the Octavius4D rotational phantom are outlined in figure 2. The Octavius4D system (figure 1d) consists of a 2D array, a rotational phantom, a trolley for transport that also holds the electrometer and the control unit, a (wireless) inclinometer and the accompanying Verisoft software. The inclinometer is initialized at gantry zero. The responsibility of the control unit is to make sure that the phantom rotates along with the gantry angle, so the 2D array is always perpendicular to the beam axis and measures the full field output for every angle. To convert the detector signal to absolute dose, a simple energy dependent cross-calibration factor is sufficient. Every 200ms, the software stores the inclinometer readout and the
corresponding 2D array readout in a simple ascii file (*.xcc format). Upon loading this file in the Verisoft software, the latte uses the series of planar measurements to perform the 3D dose reconstruction. In order to be able to do this independently from the TPS, it just needs depth dose information. So before the first use, the user has to measure the depth dose in the water tank for a small selection of open fields (from 3x3 to 27x27°) at SSD=85cm. Then, for every angular measurement point in the xcc file, an individual 3D dose reconstruction is performed based on the gantry angle readout, the 2D measurement orthogonal to the beam axis in mid-phantom and the stored depth dose data. The software deduces the field size from the actual measurement and selects the appropriate depth dose based on this. Then it sums up all individual contributions to obtain the total 3D dose. In order to do so, all data need to be sampled to the same grid, so data are interpolated to a user specified grid (typically the same as the TPS resolution, being 2.5mm) before they are all summed up to result in the final measurement reconstructed 3D dose.

Figure 2. 3D dose measurement based reconstruction with the Octavius4D system.

2.2 Quantifying the measurement uncertainty of the Octavius1500 system

2.2.1 Basic characterisation of the Octavius1500 detector
For the basic characterisation of the detector itself we have simply used a series of open field measurements with the gantry at zero. The detector was first sandwiched between solid water, then slid into to upright rotational phantom. Detector stability, output factor behaviour (down to 3x3cm²), dose rate dependence, linearity and energy dependence were all assessed and found to be most satisfying (all within 1%). The instant measurement stability is heaven, compared to the more cumbersome warm-up and cool-down behaviour of the previous Octavius729 model.

2.2.2 The Octavius1500/Octavius4D tandem
The assessment of the QA process with the rotational phantom is also performed one step at a time. Starting with simple open fields from different gantry angles (static or rotational), we subsequently move on to the validation of a reference sets of pre-validated IMRT & RA plans. This reference set
contains 20 plans on artificial structures as well as clinical IMRT & RA plans (containing 6MV and 18MV treatment beams). We have used Eclipse for generating the plans and calculated the doses with AAA. All plans were delivered on Clinac21IX dual energy machines. All of these plans, we have tediously, elaborately, tirelessly validated before with all other means available to us. So for these reference plans, we are confident that the dose calculation really meets the true delivery within the 97% pass rate for the 2%L 2mm local gamma criteria in a homogeneous phantom. Subsequently, we have submitted 60 routine patient plans to the QA as they came along in clinic.

2.2.3 The gamma evaluation
To make sense of the vast amount of data, the availability of volumetric gamma analysis tools is most welcomed, if not mandatory. First, we did an in-depth evaluation for a variety of gamma criteria from the most tolerant 3 % 3mm global down to the most strict 2%,2mm local evaluation. And for these criteria, we looked at the corresponding pass rates for different threshold isodose levels from 10% up to 95% of the maximum dose. We have found the 10% and 50% dose thresholds useful to assure correct overall dose delivery in the organs at risk as well as in the lower target dose regions. The 95% dose level provides useful insights in the dose coverage of the (high dose) target. Table 1 presents a very compact summary of the whole dataset.

| Reference dataset (20): | \( \gamma_{2\%L}(\%d_{10\%}) \) | \( \gamma_{2\%L}(\%d_{50\%}) \) | \( \gamma_{2\%L}(\%d_{95\%}) \) |
|------------------------|---------------------------------|---------------------------------|---------------------------------|
| 3\%G,3mm              | 99.99 ± 0.02                   | 99.99 ± 0.04                   | 99.92 ± 0.18                   |
| 2\%G,3mm              | 99.92 ± 0.10                   | 99.87 ± 0.15                   | 99.38 ± 1.13                   |
| 2\%G,2mm              | 99.31 ± 0.37                   | 99.15 ± 0.48                   | 97.84 ± 2.50                   |
| 3\%L,3mm              | 99.99 ± 0.02                   | 99.99 ± 0.02                   | 99.92 ± 0.18                   |
| 2\%L,3mm              | 99.72 ± 0.44                   | 99.76 ± 0.31                   | 99.31 ± 1.27                   |
| 2\%L,2mm              | 97.98 ± 1.26                   | 98.48 ± 0.87                   | 97.65 ± 2.71                   |

| RA patients (60)      | \( \gamma_{2\%L}(\%d_{10\%}) \) | \( \gamma_{2\%L}(\%d_{50\%}) \) | \( \gamma_{2\%L}(\%d_{95\%}) \) |
|------------------------|---------------------------------|---------------------------------|---------------------------------|
| 3\%G,3mm              | 99.95 ± 0.11                   | 99.92 ± 0.28                   | 99.47 ± 1.95                   |
| 2\%G,3mm              | 99.77 ± 0.25                   | 99.65 ± 0.58                   | 97.62 ± 4.58                   |
| 2\%G,2mm              | 98.88 ± 0.74                   | 98.59 ± 1.09                   | 94.60 ± 6.67                   |
| 3\%L,3mm              | 99.89 ± 0.21                   | 99.92 ± 0.28                   | 99.47 ± 1.95                   |
| 2\%L,3mm              | 99.69 ± 0.32                   | 99.64 ± 0.59                   | 97.57 ± 4.62                   |
| 2\%L,2mm              | 97.40 ± 1.30                   | 98.09 ± 1.44                   | 94.03 ± 7.07                   |

For the reference dataset, we are confident that the agreement without the measurement uncertainty should be at least 97% for the 2%L,2mm criteria. The fact that this is very close to what we obtain with the actual measurement indicates that, in theory, no considerable measurement uncertainties need to be attributed to the QA system. For the routine patients, we still get very good results, but we do notice that the spread around the mean values is larger and the average pass rates are lower. Reducing the DTA to 2mm makes almost half of the plans fail the 97% pass rate. The reference dataset was acquired in one or two nightly session, with the phantom set up very meticulously and without any real time pressure, apart from the eventual need for sleep. The routine patient dataset was acquired over numerous individual measurement sessions, accompanied by the well known pressure of clinical routine and limited machine availability. The phantom was positioned carefully, but quickly. As it turns out, a 1mm positioning uncertainty very easily slips into the measurement as such.
So rather than insisting on a more meticulous setup, we prefer to go for the pragmatic approach and include this small measurement setup error into the distance to agreement criterion. So in conclusion, we select the 2%L, 3mm gamma with 97% pass rates as a first filter.

Secondly, we did a detailed visual inspection of the data by going through the gamma evaluation for every slice of the 3D dose, and by looking at line profiles. We especially paid attention to the possible impact of the twofold detector density. An example of such an in-detail visual inspection of the data clearly showed the improvement in the 3D data reconstruction because of the twofold detector density. Figure 3 shows two gamma evaluation maps on the same treatment plan. Red and blue points fail the criteria, with red points indicating overdosage, blue points signaling underdosage:

![Figure 3. Impact of the improved detector density of the Octavius1500 (50% coverage) compared to Oct729 (25%).](image)

The upper gamma evaluation map was acquired with the Octavius729 array in the phantom whereas the lower one is with the new Octavius1500 detector. The necessary interpolation in between measurement points created substantially less falls alerts in the Octavius1500 system. The red dotted line is the Octavius729, the green line is the Octavius1500, the black line is the calculation. If desirable, one can even do two measurements with a 5mm shift in between, merge the data and have no more uncovered areas at all.

3. Conclusions
Having characterized the uncertainties we can expect in the measured 3D dose, we have made the following gamma evaluation flowchart (figure 4.) for use in clinical routine for the Octavius1500/Octavius4D 3D dose system.

We apply these strict gamma 3D criteria. If the 97% pass rate is met, all is well and there is no need to even look at the data any further. A green light is given for the treatment. If the 97%PR fails the data are submitted to further visual inspection. First the urgent question regarding the start of the treatment needs to be addressed. As no universal guidelines on the clinical impact of dose deviation are available, we apply a common sense filter, check by how much it failed and look for an explanation for the deviations. It is also wise to check in which dose areas or anatomical regions the deviations are situated (PTV? body dose? OAR?). When in doubt, we quickly run the more tolerant 3%,3mm check to proceed. If it passes, it can proceed to treatment. If the plan still fails this check, the plan is not be treated before the origin of the deviation is found.

Having addressed the imminent clinical importance of the QA, failed plans are investigated in more detail later on for research and development purposes.
As an overall conclusion, we would urge medical physicists to stop hiding behind sloppy criteria by default just because we aim for green lights in the patient files. We are missing out on a lot of possibly useful information if we do so. It pays off to be strict first as long as you are sensible later.

![Sensitive but sensible gamma evaluation flowchart:](image)

**Figure 4.** Proposal of flowchart for gamma evaluation with Octavius1500/Octavius4D QA-system.

4. References

[1] McGarry C K et al. 2013 *Med. Phys.* **40** 091707
[2] Poppe B et al. 2006 *Med. Phys.* **33** 1005-15
[3] Bakhtiari M et al. 2011 *Med. Phys.* **38** 1366-73
[4] Spezi E et al. 2005 *Phys. Med. Biol.* **50** 3361-73
[5] Poppe B et al. 2013 *Med. Phys.* **40** 082106
[6] Poppe B et al. 2007 *Phys. Med. Biol.* **52** 2921-29
[7] Korreman S et al. 2009 *Acta Oncol.* **48** 185-91
[8] Bedford J L et al. 2009 *Phys. Med. Biol.* **54** N167-76
[9] Van Esch A et al. 2007 *Med. Phys.* **34** 3825-37
[10] Chandraraj V et al. 2011 *J. Appl. Clin. Med. Phys.* **12** 338-49
[11] Feygelman V et al. 2011 *J. Appl. Clin. Med. Phys.* **12** 3346
[12] Masi L et al. 2011 *Med. Phys.* **38** 612-21
[13] Garcia-Vincente F et al. 2012 *J. Appl. Clin. Med. Phys.* **13** 111-23
[14] Zhen H et al. 2011 *Med. Phys.* **38** 5477-89
[15] Carrasco P et al. 2012 *Med. Phys.* **39** 5040-9
[16] Stasi M et al. 2012 *Med. Phys.* **39** 7626-34
[17] Olch A J et al. 2012 *Med. Phys.* **39** 81-6
[18] Nelms B E et al. 2012 *Med. Phys.* **39** 4228-38
[19] Feygelman V et al. 2013 *Med. Phys.* **40** 021708
[20] Opp D et al. 2013 *J. Appl. Clin. Med. Phys.* **14** 70-84
[21] Godart J et al. 2011 *Phys. Med. Biol.* **56** 5029-43
[22] Boggula R et al. 2010 *Phys. Med. Biol.* **55** 5619-33
[23] McCurdy B et al. 2008 *Med. Phys.* **35** 2760
[24] Nelms B E et al. 2013 *Med. Phys.* **40** 111722
[25] Cadman P F et al. 2011 *Med. Phys.* **38** 4464
[26] Hussein M et al. 2013 *J. Appl. Clin. Med. Phys.* **14** 274-92
[27] Low D A et al. 1998 *Med. Phys.* **25** 656-61
[28] Depuydt T et al. 2002 *Radiother. Oncol.* **62** 309-19
[29] Ezzell G A et al 2009 Med. Phys. 36 5359-73
[30] Van Esch A et al 2014 Med. Phys. 41 091708-14