The NCS code of practice for the quality assurance and control for volumetric modulated arc therapy

To cite this article: Anton Mans et al 2016 Phys. Med. Biol. 61 7221

View the article online for updates and enhancements.
The NCS code of practice for the quality assurance and control for volumetric modulated arc therapy

Anton Mans¹,¹⁰, Danny Schuring², Mark P Arends³, Cornelia A J M Vugts⁴, Jochem W H Wolthaus⁵, Heidi T Lotz⁶, Marjan Admiraal⁷, Rob J W Louwe⁸, Michel C Öllers⁹ and Jeroen B van de Kamer¹

¹ Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands
² Department of Radiation Therapy, Catharina Hospital, Eindhoven, The Netherlands
³ Radiotherapy Institute Friesland, Leeuwarden, The Netherlands
⁴ Radiation Oncology Center, Kantonsspital Aarau-Kantonsspital Baden, Aargau, Switzerland
⁵ Department of Radiotherapy, University Medical Center Utrecht, Utrecht, The Netherlands
⁶ Department of Radiation Oncology, University Medical Center Groningen, Groningen, The Netherlands
⁷ Department of Radiation Oncology, VU University Medical Center, Amsterdam, The Netherlands
⁸ Department of Radiation Oncology, Wellington Blood and Cancer Centre, Wellington Hospital, Wellington 6242, New Zealand
⁹ Maastro Clinic, Maastricht, The Netherlands

E-mail: a.mans@nki.nl

Received 1 July 2016, revised 12 August 2016
Accepted for publication 18 August 2016
Published 19 September 2016

Abstract
In 2010, the NCS (Netherlands Commission on Radiation Dosimetry) installed a subcommittee to develop guidelines for quality assurance and control for volumetric modulated arc therapy (VMAT) treatments. The report (published in 2015) has been written by Dutch medical physicists and has therefore, inevitably, a Dutch focus. This paper is a condensed version of these guidelines, the full report in English is freely available from the NCS website www.radiationdosimetry.org.

After describing the transition from IMRT to VMAT, the paper addresses machine quality assurance (QA) and treatment planning system (TPS) commissioning for VMAT. The final section discusses patient specific QA
issues such as the use of class solutions, measurement devices and dose evaluation methods.

Keywords: volumetric modulated arc therapy, quality assurance, code of practice

(Some figures may appear in colour only in the online journal)

1. Introduction

The introduction of VMAT has led to numerous scientific papers discussing different aspects of quality assurance (QA) for VMAT. Until recently, no comprehensive code of practice (CoP) for VMAT QA has been described that may serve as a guideline to introduce and use VMAT safely in the clinic. Most recent codes of practice concern small-field radiotherapy (NCS 25 (Hermans et al. 2015), AAPM (Das 2015), IPEM (Aspradakis et al. 2010)) or intensity modulated radiation therapy (IMRT) (ESTRO (Alber et al. 2008), NCS (Van der Wal et al. 2013), IPEM (James et al. 2008), AAPM (Ezzell et al. 2003)). Since a consistent guideline helps the user to introduce a new technique and to setup the corresponding QA system, the Netherlands Committee of Radiation Dosimetry (NCS) decided to form a subcommittee to devise a report on QA for VMAT (Mans et al. 2015). This CoP only concerns the application of VMAT using conventional linear accelerators (linacs). It is an extension of previous NCS CoPs concerning general linac QA (Meijer et al. 1996) and IMRT QA (Van der Wal et al. 2013). All reports, including the CoP for VMAT QA, can be downloaded freely from the NCS website: http://radiationdosimetry.org/ncs/publications.

The main difference between IMRT and VMAT is the increased number of degrees of freedom during treatment, including gantry rotation, dose rate changes and in some cases collimator rotation. This poses different demands on the QA compared to IMRT and conventional radiation techniques. The NCS CoP provides guidelines for a reliable introduction of VMAT as well as regular QA to ensure correct VMAT delivery. Both the CoP and this paper deal with VMAT-related QA for the linac, issues regarding the treatment planning system (TPS) and patient-specific QA. Additionally, several tests are described in the machine QA section that may be useful when there are concerns regarding the performance of the linac.

The CoP was originally written to provide guidelines on the introduction of VMAT and on designing a consistent QA program for departments in the Netherlands and Belgium. However, considering lack of published guidelines on this subject, we expect that this CoP will also be valuable for departments outside the aforementioned countries. To our knowledge, this is the first CoP explicitly aimed at VMAT QA.

2. Machine QA

The characterisation and understanding of the dynamic behaviour of the linac is the first step in implementing VMAT (Wang et al. 2013). The suggested machine QA tests will assist the user to assess the dynamic linac properties during VMAT and verify the accuracy of VMAT dose delivery. It is assumed that the linac has already been commissioned for conventional IMRT prior to the introduction of VMAT. Therefore, the guidelines as described in NCS report 22 on the introduction of IMRT (Van der Wal et al. 2013) are used as the starting point for this CoP, and a summary of these guidelines are included.
2.1. Suggested QA tests

2.1.1. Gantry and collimator angle. Both gantry and collimator angle can be accurate within 1°, but it is generally possible to achieve deviations below 0.5°. The gantry and collimator angle can be verified independently for different combinations, for example using an inclinometer or spirit level.

2.1.2. MLC. A picket fence test or a similar test using a static gantry is commonly performed for all cardinal collimator angles for IMRT (Boyer et al 2001, Low et al 2001, Sastre-Padro et al 2004, 2009). The test can be performed with a moving gantry to test VMAT delivery, using a 2D measurement device orthogonal to the beam, for example film or an electronic portal imaging device (EPID). The test includes one VMAT segment with a high dose rate and (almost) static leaves serving as open field, while the next segment with a low dose rate and fast moving leaves serves as a ‘closed’ field. By starting with a central slit and delivering the subsequent slits as shown in figure 1, a maximum MLC speed between the slits is obtained. This test can be performed for different collimator angles and different starting positions of the gantry to test both clockwise and counter-clockwise gantry rotation. The results of the test with the rotating gantry can be compared with the results using a static gantry fixed at 0°. MLC position deviations should be less than 1 mm but deviations less than 0.5 mm are generally achievable. Note that measurement equipment attached to the gantry may displace with respect to the beam due to gravity.

2.1.3. Output. Since the number of monitor units (MUs) can vary considerably between VMAT plans, it is advised to test linearity for a wide range of MUs using a standard reference field configuration. Suggested MU-values are 2, 5, 10, 50, 100, 200, 500, and 999. Preferably, the linearity is verified for different dose rates, including the minimum and maximum allowed dose rate. The output stability with varying dose rates can be tested for each MU level using a test plan which is divided up into three parts with equal MUs. During the first and last parts of the delivery, the lowest allowed dose rate is applied, while the maximum dose rate is applied during the middle part of the delivery. Again, it is recommended to perform all tests for all cardinal gantry angles.

The output accuracy for treatments with more than 1000 MUs (e.g. hypo-fractionated treatments) can be verified using a full arc VMAT delivery over 360° (e.g. a 10 × 10 cm² field) with equal MUs for each control point. It is advised to test linearity up to a high number of MUs within the clinically applied range. The output of the linac is measured at the isocentre using a suitable ionisation chamber, preferably in a cylindrical phantom or by using a gantry-mounted setup. Again, it is possible to extend the test by switching between the minimum and maximum dose rate during the delivery in the same way as described above. The linearity can be assessed by relating all results to the delivery using 100 MU, and should be better than 1%. The reproducibility, defined as the coefficient of variation for repeated measurements, should be less than 0.5%.

2.1.4. Flatness and symmetry. Flatness and symmetry need to be determined at cardinal angles, e.g. using a gantry mounted set-up. Flatness is commonly defined as the dose variation over 80% of the nominal field size at 10 cm depth in a plane perpendicular to the central axis, with a tolerance of 3% (Baily et al 1994, Meijer et al 1996). Symmetry is defined as the maximum difference between dose values symmetrical around the beam axis within the flattened region. This should be within 3%. Under normal circumstances, both flatness and symmetry can be within 1.5%. Additionally, the NCS recommends to test flatness and symmetry with the
gantry at 0° for the allowed minimum and maximum dose rate and three levels between these. For the other cardinal angles, flatness and symmetry need to be determined for the minimum and maximum dose rate are in general within 2% of the reference profile at 0°.

For the dynamic mode, a gantry-mounted 1D or 2D array can be used to obtain an angle-resolved beam profile. For optimal performance, the array should allow a sampling frequency better than one profile per 5°. The NCS advises to perform this test for 5 different dose rates, including the lowest and highest ones allowed. This may be achieved by devising a VMAT plan of the largest field size with a certain number of MUs per 4°, or similar, to obtain the desired approximate dose rate. Finally, flatness and symmetry need to be measured with varying dose rates. For this, several angular sectors with alternating dose rates from maximum to minimum allowed values need to be delivered with the moving system. It is advised to investigate both flatness and symmetry directly after dose rate changes. For this test, a DICOM RT-plan needs to be devised with the proper linac demands. Visual inspection of the results provides adequate insight in the behaviour of the system.

2.2. VMAT performance

2.2.1. Gantry and MLC speed. To characterise the system, the maximum gantry speed and the maximum speed of the slowest leaf can be determined and, if applicable, defined in the TPS. It is recommended to measure the leaf speed both parallel and perpendicular to the

---

Note that for treatment units allowing only for binned dose rates, it is advised to slightly increase the total number of MUs for each arc segment (e.g. 5%) to prevent the system from falling back in a lower dose rate mode.
direction of gravity. Both values may be determined with independent devices or obtained from the linac log files, although this is not an independent assessment.

2.2.2. Dynamic picket fence test. The interaction between gantry speed and dose rate may be determined with a test, similar to the one suggested by Ling et al. (2008) and Bedford and Warrington (2009). In this test, a VMAT plan consisting of seven adjacent dose strips of a 5.6 $\times$ 40 cm$^2$ open fields is used, with a 1 mm gap between them. Each dose strip contains the same number of MUs (e.g. 100 MU) but is delivered using a different combination of dose rate, gantry angle and gantry speed, resulting in a fairly uniform intensity across the film except for the gaps between the strips. The plan exposes a film with build-up material, which is placed on a gantry mounted frame or tray. As a reference, the same plan is delivered without gantry rotation, but with MLC movement to a second film. Profiles obtained in dynamic and static mode usually differ less than 2%. More practical details and alternatives can be found in NCS 24 (Nederlandse Commissie Voor Stralingsdosimetrie 2015).

2.2.3. Inertia. With VMAT, the gantry is frequently accelerated and decelerated and considerable dose rate changes may occur. The linac may issue interrupts due to large field symmetry or flatness errors induced by large gantry speed variation or large dose rate changes. Therefore, the behaviour of the system under extreme circumstances needs to be investigated. A (clinical) VMAT plan with extreme gantry speed and dose rate changes can be used to test the machine performance. NCS 24 (Nederlandse Commissie Voor Stralingsdosimetrie 2015) presents a possible strategy to develop a test plan, but a clinical plan that includes these extremes may be used as well. In addition to delivering these VMAT plans without any problems, the delivery should comply with regular QA demands.

2.2.4. Synchronicity. In the dynamic picket fence test described above, only the dosimetric and MLC synchronicity is investigated since the measurement device is gantry mounted. In that test, however, dose delivery as a function of gantry angle is not taken into account. For this reason the following test, similar to the one described by Van Esch et al. (2011), is recommended in NCS 24. It assesses how the system deals with large variations in gantry speed (inertia). The test consists of a VMAT beam with the MLC leaves remaining stationary during gantry rotation, forming a central gap of a few mm (as small as possible). The full gantry arc is divided into eight equal angular sectors of 2° with high dose delivery and 38° with low dose delivery. During the low dose rate sectors the gantry moves at maximum speed, while during the high dose rate sectors the gantry moves at low speed. If the odd segments are irradiated clockwise and the even segments anti-clockwise, the possible effect of inertia may become even more clear. The delivered dose can be measured with film, sandwiched between a (cylindrical) phantom (figure 2).

The high peaks indicate the entrance beam and the peaks should be separated according to the angular distribution of the high dose segments. To obtain comprehensive insight of the gantry inertia handling, variations to this test are given in NCS 24 (Nederlandse Commissie Voor Stralingsdosimetrie 2015). If the difference between intended and measured angular position is larger than 1°, then the user should investigate its cause. An alternative synchronicity test using an off-axis positioned steel bar and portal imaging is presented by Bedford et al. (2015).

2.2.5. Interrupt handing. It is recommended to determine whether the QA results for a VMAT plan are similar for the uninterrupted delivery and the (multiple) interrupted delivery. The dosimetric deviation measured with an ionisation chamber should be below 1%, preferably <0.5% for a representative point.
2.3. Test frequency

Most tests need only to be performed on indication, for example while investigating erratic behaviour of the linac or poor patient QA results. A complete overview of the frequencies suggested by the NCS for the different tests provided in the previous paragraphs can be found in table 4 of NCS report 24 (Nederlandse Commissie Voor Stralingsdosimetrie 2015).

3. VMAT representation and treatment planning

Most TPSs describe the VMAT delivery as a set of discrete beams, whereas delivery is continuous, see figure 3. This set of beams has to be translated into a DICOM RT plan consisting of a set of control points (CPs), each describing a succeeding linac state (delivered number of MU, leaf positions, jaw positions, gantry and collimator angle, etc).

To calculate the dynamically delivered dose distribution, while making use of discrete beams, the TPS uses a dedicated model (‘VMAT model’ in figure 3). In the simplest model, dose calculation in the TPS is performed using the MUs and MLC shape of the corresponding CP in the DICOM prescription. Evidently, the distribution of MUs and leaf positions along the arc are not taken into account correctly in this approach. A more sophisticated model would distribute either MUs or MLC shapes between CPs. In the optimal model, Monte Carlo dose calculation is used to realistically model the dynamic linac behaviour. Each TPS vendor has its own implementation of such a model. In NCS 24 (Nederlandse Commissie Voor Stralingsdosimetrie 2015) a more elaborate discussion on this topic can be found.

It is worthwhile to note that between the CPs, the machine state is not prescribed and the linac uses linear interpolation to proceed from one CP to the next. The dynamic behaviour of the linac (e.g. dose rate and gantry speed) is ultimately determined by the linac control system. A feedback loop is applied to steer each of the parameters in order to keep deviations small. If this steering would be insufficient and deviations would become too large or a smaller maximum allowed deviation persist for too long, the treatment delivery is interrupted.

3.1. TPS commissioning

Without proper commissioning of the TPS it is difficult to have good results in patient-specific QA. Whenever possible, the practical limitations of the linac, such as minimal field size, maximum leaf speed, are incorporated in the TPS settings in the best way possible. This
minimises the risk of inaccurate VMAT delivery. Generally, not all machine limitations in VMAT delivery can consistently be defined in the TPS. In addition, for accurate VMAT delivery, the parameters in the TPS beam model need not necessarily represent the actual physical properties of the linear accelerators. In actual treatment plans it might be advantageous to limit the extend of e.g. maximum leaf speed, gantry rotation or dose rate to lower values than technically possible to reduce plan complexity. As with other radiotherapy techniques, the user is responsible for validating the parameters of the TPS beam model.

3.1.1. Leaf modelling. In general, the field shapes in VMAT deliveries are solely defined by MLC leaves while the (back-up) jaws are used to minimise the out-of-field dose caused by MLC transmission and inter-leaf leakage. As VMAT deliveries may consist of many highly irregular and small segments, proper modelling of tongue and groove, rounding of the leaf tips and MLC leaf transmission is essential (Feygelman et al 2010). The TPS parameters defining these properties can, in part, be determined through commissioning measurements. However, users should note that the modelling of these properties has a limited accuracy, and that the shape of the calculated beam penumbra depends on the MLC parameter values as well as the modelling of the focal spot. Therefore, fine-tuning may be necessary to improve correspondence between computed and delivered dose. Moreover, independent verification, for instance using patientspecific QC measurements or machine QA tests described previously, is strongly recommended.

3.1.2. Matched beams. It is increasingly common practice to have matched beams for multiple linacs with the same technical specifications. This allows the application of a single beam model in the TPS for these linacs, and transfer of patients between linacs without limitations. However, the accuracy of beam matching needs to be separately validated for VMAT. Specifically, differences in MLC leaf properties within manufacturing tolerances applied by the vendor may lead to relevant changes in the beam penumbra. In addition, jaw tracking or collimator rotation during VMAT may not be available on all machines with matched beams.

3.2. VMAT inverse optimisation

3.2.1. Dose calculation. The additional degrees of freedom in VMAT require efficient computation algorithms for inverse optimisation (Otto 2008, Bzdusek et al 2009). Most
implementations start with a coarse sampling of the arc and increase the number of CPs until the requested gantry spacing is reached. The first iterations of the optimisation are usually first carried out using a fast dose calculation algorithm with limited accuracy, before applying a slower but more accurate dose calculation algorithm during the final iterations. Continuing an optimisation based on previous accurately computed results may help to further improve the VMAT plan. This can be done in combination with adjusting objectives.

3.2.2. Multiple arc optimisation. The use of multiple arcs may enable improved plan conformity when a higher degree of dose modulation is required, but is achieved at the cost of increased calculation and treatment time (Guckenberger et al 2009). However, improved optimisation algorithms may reduce the need for treatment plans with multiple arcs.

Optimisation of multiple arcs is managed in different ways depending on the TPS used, but may result in a situation where each arc delivers dose to a different part of the target volume (Bzdusek et al 2009). This might reduce the robustness of the plan under patient- or organ motion, and users should consider whether this is appropriate. For instance, the impact of the interplay effect during lung treatments may depend on the division of dose delivery between the various arcs. One possibility to avoid a strong geometric separation between the dose delivered by each arc is to sequentially optimise each arc to deliver dose to the whole target volume, and use the dose delivered by the first arc(s) as input base plan for further optimisation of successive arcs. In addition, partial arcs or avoidance regions may help steer the dose delivery of each arc.

3.2.3. Plan complexity. Highly conformal treatment plans can be generated in the TPS using multiple (partial) arcs, high dose rate modulation, and high gantry and/or leaf speed variations. However, the deliverability of high complexity VMAT plans within the departmental tolerances may not be warranted and needs to be investigated carefully (Yang et al 2012, Masi et al 2013). Deliverability can be improved by reducing the VMAT plan complexity. In most TPSs, this is only facilitated in an indirect manner, for example by limiting the anticipated treatment time, minimum segment area, the number of MUs or by limiting changes in dose rate.

3.2.4. Class solutions. Defining settings that serve as a starting point for treatment planning for a certain site will provide more consistent features of the final treatment plans. It also helps in the complex process of finding an optimal solution in the vast solution space of VMAT. Furthermore, having similar treatment plans for a single treatment site will generally result in more consistent plan QA (Yang et al 2012, Masi et al 2013). Such a set of predefined parameters is referred to as a class solution. The quantities defined in a class solution are, for example, optimisation parameters (constraints and weighted objectives), technical requirements like beam settings (start-stop gantry/collimator angle, number of arcs) and computation requirements (gantry angle sampling, allowed complexity).

Although designing a class solution requires considerable effort, the improved efficiency and consistency of treatment planning and plan delivery QA outweigh this investment for larger patient groups. If necessary, individual adaptation of a plan may be considered. This should be done in consultation with the medical physics expert (MPE) responsible for treatment planning and for treatment execution QA.

4. Patient-specific QA for VMAT

The purpose of patient-specific QA is to ensure proper delivery of the intended treatment plan. Although many (human) errors, such as incorrect plan selection, may interfere with this
objective, in this section only technical issues that may cause disagreement between the TPS calculated and actually delivered dose are considered. These issues include TPS beam model limitations, machine limitations and plan transfer errors.

4.1. Efficient QA

It is recommended to devise a method for patient specific QA that is not only safe and accurate but also time efficient. By using class solutions the workload can be reduced significantly. Noting that the best method for an optimal QA program depends largely on the local workflow and available QA devices, several general issues can be discerned.

4.2. Pre-delivery plan checks

4.2.1. Plan transfer validation. Given the ever increasing complexity of the treatment plans it is virtually impossible to manually check correct plan transfer for all VMAT plans. Typically, the user relies on the record-and-verify (R&V) system for this. However, for a new complex treatment technique such as VMAT, it is advised to check data transfer integrity for a number of plans. It is highly recommended to check approximately 5 plans per new class solution to ensure proper data transfer. If the TPS and the R&V system have a shared database, no transfer of the data is required, hence no checking of the data transfer is needed. However, both with and without a shared database, undesired changes to the treatment plan may occur. Such changes can only be found with dosimetric (pre-treatment) plan QA (Mans et al 2010).

4.2.2. Plausibility checks. The use of class solutions will provide VMAT plans with similar characteristics such as the number of MUs. Obviously, there is a certain range of MUs within a specific class solution but treatment plans with values outside this range may require additional attention from the MPE. Computer programs that automatically check for deviation of the plan from the class solution’s MU range may serve as a valuable tool in early detection of outlier plans. It is recommended to perform such a plausibility check on the R&V system data whenever possible. As a bonus, this may serve as a plan transfer check for each patient. Alternatively, many institutes use independent monitor unit calculation programs for a more stringent check. Such programs can be rather elaborate and complete, using beam model data to perform such checks. Interestingly, the better the verification tool is, the less insight it provides regarding compliance with class solutions. In particular, sophisticated verification programs that nearly mimic an independent TPS will provide the same information as the original TPS, especially if the dose kernels are based on the same input data. Having the same answer with different calculation tools does not guarantee conformity to a class solution or successful QA results. Using fairly simple tools will better help in detecting plans that do not conform to the class solution, for example by having more small, off-axis segments resulting in larger errors in the MU calculation. Considering this, the NCS subcommittee prefers simple computational verifications over more elaborate ones.

4.2.3. Treatment interpretation by the linac. The treatment plan is transferred as a set of discrete CPs that are interpreted as a continuous delivery by the treatment machine. In part, this is similar to the delivery of sliding window IMRT but having the additional possibility of rotating gantries and collimators and varying dose rate. Since most TPSs perform dose calculations only at fixed gantry angles, there is a fundamental difference between computed (static) and delivered (dynamic) dose, as described before. As a result, patient-specific QA for VMAT is more complex and more relevant.
Because the machine adapts dose rate, gantry rotation speed and leaf travel speed depending on the delivered number of MUs, it is discouraged to scale the dose to meet possible limitations of the dose measurement devices. Ideally, the plan sent to the linac for QA equals the plan used for clinical treatment.

4.3. Measurement techniques

4.3.1. Low-resolution dosimetry devices. The ionisation and diode chamber are the most well-known and used measurement devices. The main drawback of ionisation chambers is their finite size, causing averaging over the sensitive volume. This can be minimized by using it in rather homogeneous regions, although homogeneous regions in the composite plan do not need to be homogeneous in individual segments. Diode measurements exhibit an angular dependency. This may be counteracted by rotating the device synchronously with the gantry, ensuring orthogonal irradiation for the whole treatment (Dobler et al 2010, Boggula et al 2011, Masi et al 2011). These drawbacks should be considered in the choice of a measurement device and the interpretation of the results.

Using detector arrays, the dose can be measured in several points simultaneously. Still, most arrays have a rather low spatial resolution, with the possible exception of arrays using micro liquid-ionisation chambers (2.5 mm resolution). To detect small leaf positioning errors, a high resolution detection method is needed. In addition, undersampling due to low resolution measurement devices may lead to unreliable gamma evaluation results.

A number of detector array systems offer the possibility of 3D dose reconstruction, based on the 2D measurements, compensating for the low resolution to some extent. The method that is used for this reconstruction should be well understood before interpreting QA results. If data from the TPS is used in the dose reconstruction, some errors in the TPS may not be detected (Steciw et al 2005).

4.3.2. High-resolution dosimetry devices. In contrast to the above-mentioned measurement devices, both film and EPID dosimetry systems have a resolution of about 1 mm or better. Due to development and post-processing, films are used for the verification of integral dose distributions. They can be used for absolute dosimetry if proper calibration, measurement and processing procedures are used. For Gafchromic films, lateral scan artefacts, and inter- and intra-batch variation of the film sensitivity may deteriorate measurement accuracy and require special attention (Van Battum et al 2008, Crijns et al 2013). To avoid dosimetric inaccuracies by parallel film irradiation, the film may be placed slightly off-axis (e.g. 5 mm). Fortunately, there is considerable progress in the use of Gafchromic for dosimetry (Micke et al 2011, Mayer et al 2012).

At present, dosimetry with an EPID is mostly performed in a research setting (Van Elmpt et al 2008), although several vendors already have commercial versions available. The advantages are instantaneous read out and the possibility to obtain portal dose while treating the patient. With this portal dose and a suitable CT image of the patient, the absorbed dose in the patient can be reconstructed. Again, when using the same dose calculation algorithm as the TPS, some errors may go undetected.

4.3.3. Gamma evaluation. The delivered dose distribution can be compared with the calculated one using a gamma analysis (Low et al 1998), for example with reference values of 3% (of the maximum) dose and 3 mm distance to agreement. The NCS advises not to use only $\gamma$ pass rates (percentage of points with $\gamma \leq 1$), since this does not contain any spatial
information and is poorly correlated to actual clinical parameters (Nelms et al. 2011, Stasi et al. 2012). It is advised to include more statistics in this evaluation, such as the average $\gamma$ value and the highest $\gamma$ percentile, along with a visual inspection of the $\gamma$ distribution and dose differences. The achievable pass rates and $\gamma$ criteria depend on the measurement system, the amount of intensity modulation in the treatment plan and the shape of the dose distribution (e.g. highly inhomogeneous dose) (Masi et al. 2011, Yang et al. 2012, Hussein et al. 2013).

Organs at risk close to their tolerance dose deserve extra attention in the verification. Since estimating the clinical consequences of dose deviations in a phantom is difficult, some vendors provide software tools to recalculate the dose on the patient anatomy using measurements in a phantom. Such tools might be helpful for the clinical interpretation of the QA results (Olch 2012), but a thorough validation of these tools is necessary. Still, significant deviations that may not be clinically relevant for one particular patient do require further investigation by the responsible MPE.

4.4. Practical QA recommendations

4.4.1. Gamma evaluation criteria. Whenever possible, use absolute dosimetry for verification. If that is not possible, the required normalization factor should be monitored. Now that many QA devices can reconstruct the measured dose to a 3D dose distribution, 3D verification is recommended. A (3D) gamma evaluation with 3%/3 mm criterion is recommended with a minimum pass rate of 90% for all dose points having a dose of 50% of the prescribed dose or higher. The averaged gamma value can, in general, be below 0.5. For stereotactic treatments exhibiting a large dose inhomogeneity, more stringent settings like 2%/1 mm might be more suitable.

These values may serve as a starting point. Each institute should define their own criteria, possibly different for each class solution, according to their experience and what is technically achievable.

4.4.2. Investigation of deviations. If the above criteria are not met, an MPE should investigate the deviation and decide whether the plan may be used clinically. One possibility is to perform machine QC as described in the previous paragraphs. However, when deviations are due to other causes then machine limitations, these measurements will not reveal the underlying problem. Alternatively, methods that allow time-resolved patient-specific QC can be helpful in tracing the source of deviations through correlation of dose deviations of individual CPs with physical delivery and planning parameters (Nelms et al. 2012, Louwe et al. 2015, Schyns et al. 2016).

4.4.3. Frequency and type of patient-specific QA. Before starting with VMAT treatments, experience with conventional IMRT is highly recommended. The results from IMRT treatments can be interpreted more easily, proving valuable insight in the dosimetric limitations of the TPS for small, irregular fields and highly modulated beams.

If the amount of experience in VMAT is limited, a more extensive dummy run procedure is recommended during the development of a class solution to get more insight in the behaviour of the linac. After acceptance of a class solution, thorough checks need to be performed to see whether the treatment plans conform to the class solution, even when an institution has extensive experience with VMAT. This can be done by examining the planning objectives and other planning parameters, or by comparing the amount of modulation of the treatment plans relative to a reference plan for that class solution. More advanced methods to analyse
the modulation level of a plan have been developed for conventional IMRT (see e.g. McNiven et al 2010, Nauta et al 2011), and a complementary paper concerning VMAT has recently been published (Masi et al 2013). However, this methodology does not take the dynamic parameters into account and only shows limited correlation with dosimetric accuracy.

Once a certain level of experience is gained in VMAT delivery and a sufficient number of pre-treatment verifications have been performed, a plausibility check may suffice as pre-treatment verification. This may be a plausibility check only, consisting of a MU range check or similar. Suggestions when to reduce the intensity of the patient-specific QA are given in table 1.

4.4.4. Evaluating class solutions. Once class solutions have been established, the NCS subcommittee recommends performing dosimetric checks every three months for four plans, randomly selected among clinical patients. In this process, it is wise to ensure that all class solutions are regularly checked on different linacs. A time trend analysis of the resulting gamma distribution (pass rate, average gamma, highest percentile gamma), for example using statistical process control, helps detecting small systematic deviations between planned and delivered dose and may help in improving the overall quality of the entire radiotherapy chain (Pawlicki et al 2005).

**Table 1.** Number of pre-treatment verifications and minimally required class of dosimetric accuracy and resolution of a QA device for different experience levels. Reproduced from Mans et al 2015.

<table>
<thead>
<tr>
<th>Experience level</th>
<th># of patients pre-treatment</th>
<th>Dosimetric accuracy</th>
<th>Spatial resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>No experience with VMAT, development of new class solution</td>
<td>30</td>
<td>Class I</td>
<td>Class I</td>
</tr>
<tr>
<td>Experience with VMAT, development of new class solution similar to existing solutions</td>
<td>5</td>
<td>Class I</td>
<td>Class II</td>
</tr>
<tr>
<td>Experience with VMAT, development of new class solution dissimilar to existing solutions</td>
<td>10</td>
<td>Class I</td>
<td>Class II</td>
</tr>
<tr>
<td>Experience with VMAT (&gt;100 patients total), existing class solution</td>
<td>All patients</td>
<td>Class IV</td>
<td>—</td>
</tr>
<tr>
<td>Re-evaluation of class solution</td>
<td>1</td>
<td>Class I</td>
<td>Class II</td>
</tr>
</tbody>
</table>

Note: The definition of the classes can be found in table 2.

**Table 2.** Dosimetric accuracy and spatial resolution classes as defined in NCS 22 (Van der Wal et al 2013).

<table>
<thead>
<tr>
<th>Dosimetric accuracy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>Highest accuracy, typically ~1% or better, e.g. ionisation chambers</td>
</tr>
<tr>
<td>Class II</td>
<td>Accuracy ~3%, e.g. film or EPID dosimetry</td>
</tr>
<tr>
<td>Class III</td>
<td>Accuracy ~5%, e.g. TLD, MOSFET, diode</td>
</tr>
<tr>
<td>Class IV</td>
<td>No dosimetric accuracy, only independent MU verification</td>
</tr>
<tr>
<td>Class V</td>
<td>No dosimetric accuracy, only data transfer verification</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial resolution</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>Highest resolution, typically 2 mm or better, e.g. film or EPID dosimetry</td>
</tr>
<tr>
<td>Class II</td>
<td>Resolution worse than 2 mm, e.g. multiple ionisation chamber or diode arrays</td>
</tr>
</tbody>
</table>
4.4.5. Additional recommendations. Participating in a dosimetry audit is very useful, also for experienced centres, to determine the quality of the treatments compared to other centres. By comparing the results of different centres with different workflows and/or equipment, more insight can be gained on achievable QA results (Perik et al 2013, Tsang et al 2013).

5. Summary

Early 2015, the NCS published a code-of-practice for VMAT QA. This paper presents its main recommendations, starting with a brief introduction describing the transition from IMRT to VMAT. The second section discusses machine QA for VMAT. The presented tests can be used in the regular QA program or as a means to investigate possible poor QA results. In section 3, the additional requirements for VMAT commissioning in the TPS are discussed. The last section concerns patient-specific QA, covering plan transfer verification, plan plausibility checks, and the value of both high-accuracy and high-resolution measurements. Throughout the paper, special attention is being paid to the added value of class solutions and how to introduce them in the clinic.

References


Nauta M, Villarrreal-Barajas J E and TambaSco M 2011 Fractal analysis for assessing the level of modulation of IMRT fields Med. Phys. 38 5385


Nelms B E, Zhen H and Tomé W A 2011 Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors Med. Phys. 38 1037–44

Olch A J 2012 Evaluation of the accuracy of 3DVH software estimates of dose to virtual ion chamber and film in composite IMRT QA Med. Phys. 39 81

Otto K 2008 Volumetric modulated arc therapy: IMRT in a single gantry arc Med. Phys. 35 310


Sastre-Padro M, Lervåg C, EiIertsom K and Malinen E 2009 The performance of multileaf collimators evaluated by the stripe Test Med. Dosim. 34 202–6


7234
Tsang Y, Ciurlionis L, Clark C and Venables K 2013 Development of a novel treatment planning test for credentialing rotational intensity-modulated radiotherapy techniques in the UK Br. J. Radiol. 86 20120315