Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom

Andrea Molineu,[a] Nadia Hernandez, Trang Nguyen, Geoffrey Ibbott, and David Followill
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030

Purpose: This study was performed to report and analyze the results of the Radiological Physics Center's head and neck intensity-modulated radiation therapy (IMRT) phantom irradiations done by institutions seeking to be credentialed for participation in clinical trials using intensity modulated radiation therapy.

Methods: The Radiological Physics Center's anthropomorphic head and neck phantom was sent to institutions seeking to participate in multi-institutional clinical trials. The phantom contained two planning target volume (PTV) structures and an organ at risk (OAR). Thermoluminescent dosimeters (TLD) and film dosimeters were imbedded in the PTV. Institutions were asked to image, plan, and treat the phantom as they would treat a patient. The treatment plan should cover at least 95% of the primary PTV with 6.6 Gy and at least 95% of the secondary PTV with 5.4 Gy. The plan should limit the dose to the OAR to less than 4.5 Gy. The passing criteria were $\pm 7\%$ for the TLD in the PTVs and a distance to agreement of 4 mm in the high dose gradient area between the PTV and the OAR. Pass rates for different delivery types, treatment planning systems (TPS), linear accelerators, and linear accelerator-planning system combinations were compared.

Results: The phantom was irradiated 1139 times by 763 institutions from 2001 through 2011. 929 (81.6%) of the irradiations passed the criteria. 156 (13.7%) irradiations failed only the TLD criteria, 21 (1.8%) failed only the film criteria, and 33 (2.9%) failed both sets of criteria. Only 69% of the irradiations passed a narrowed TLD criterion of $\pm 5\%$. Varian-Elipse and TomoTherapy-HiArt combinations had the highest pass rates, ranging from 90% to 93%. Varian-Pinnacle[3], Varian-XiO, Siemens-Pinnacle[3], and Elekta-Pinnacle[3] combinations had pass rates that ranged from 66% to 81%.

Conclusions: The head and neck phantom is a useful credentialing tool for multi-institutional IMRT clinical trials. The most commonly represented linear accelerator-planning system combinations can all pass the phantom, though some combinations had higher passing percentages than others. Tightening the criteria would significantly reduce the number of institutions passing the credentialing criteria. Causes for failures include incorrect data entered into the TPS, inexact beam modeling, and software and hardware failures. © *2013 American Association of Physicists in Medicine.* [http://dx.doi.org/10.1118/1.4773309]

I. INTRODUCTION

Since its development in the late 1990s, intensity-modulated radiation therapy (IMRT), which can shape the radiation dose to the tumor while avoiding nearby organs at risk (OARs), has become widespread in its use. Computer-controlled multi-leaf collimators (MLCs) and inverse treatment planning techniques have given radiation oncologists tools to shape the dose distribution and thus improve radiation treatments. As a result, the National Cancer Institute (NCI) has sponsored numerous multi-institution clinical trials since the early 2000s that allow or require IMRT in hopes of improving the clinical outcomes and minimizing late complications.[1]

Since IMRT dose distributions often have complex shapes with steep dose gradients near critical normal tissues, requiring a highly complex delivery process,[2] a major concern with IMRT treatment delivery is quality assurance (QA) that doses will be delivered as planned.[3] The report of the American Association of Physicists in Medicine (AAPM) Task Group 40 provides comprehensive QA guidelines for conventional and 3D conformal radiotherapies[4] but does not include any specific recommendations for the QA of IMRT delivery. However, the recently published AAPM Task Group 142 report, which is intended to supplement the Task Group 40 report, does include specific QA tests and acceptance criteria for IMRT delivery machines.[5] Also, AAPM Task Group 119 suggests commissioning tests for IMRT systems, but does not address patient specific QA.[6] By the nature of its highly modulated, complex treatments, IMRT involves unique problems relating to accurate and reproducible dose delivery and dose measurement for QA. To further complicate matters, institutions have different equipment and protocols they use to plan and deliver IMRT treatments for QA, making direct interinstitutional comparisons difficult. Therefore, the implementation of IMRT in a multi-institution clinical trial requires additional QA to ensure that all the participating institutions deliver comparable and consistent doses. Accordingly, the NCI has mandated that institutions seeking to use IMRT in

clinical trials must first be credentialed by demonstrating the ability to deliver IMRT doses accurately.[7]

The mission of the Radiological Physics Center (RPC), funded by the NCI and located at The University of Texas MD Anderson Cancer Center, is to assure the NCI and the NCI-sponsored clinical trial study groups that institutions participating in their clinical trials can deliver prescribed radiation doses that are clinically comparable and consistent. Credentialing, through the use of benchmark plans or phantom irradiations to assess an institution's ability to interpret a specific protocol's requirements and deliver the specified treatment, plays a necessary role in ensuring that institutions meet the requirements of the protocol and the NCI for using IMRT in clinical trials.[8] The RPC, as a member of the NCI-funded Advanced Technology Consortium, participates in credentialing institutions by requiring institutions to meet certain criteria on knowledge assessments, facility questionnaires, treatment planning benchmark cases, and irradiation of anthropomorphic dosimetry phantoms; these tests are designed to ensure the quality of the complete treatment process, from imaging to planning to treatment delivery.[9] Without credentialing, dose delivery may vary between institutions, and protocol deviations may ultimately occur that may invalidate the trial or, worse yet, lead to flawed conclusions. One of the primary methods the RPC uses for IMRT credentialing is the assessment of an institution's irradiation of the RPC's IMRT head and neck (H&N) phantom. This anthropomorphic phantom was designed in 2000 in collaboration with the Radiation Therapy Oncology Group's medical physics committee specifically for IMRT credentialing.[10] The phantom is mailed to institutions, where physicists are instructed to treat the phantom as they would treat an actual patient. The phantom is imaged, a treatment plan is developed, the treatment is delivered to the phantom, and finally the phantom is returned to the RPC for analysis. Institutions that do not meet the established acceptance criteria cannot use IMRT in NCI-sponsored clinical trials unless the protocol states otherwise.

This study was performed to report and analyze the results of the RPC's H&N IMRT phantom irradiations by institutions seeking to participate in NCI-sponsored clinical trials. Pass/fail rates are reported for various treatment delivery, treatment planning, and institutional variables.

II. MATERIALS AND METHODS

When first developed and used for credentialing, the RPC's H&N phantom housed a block insert containing two target structures, an OAR, and radiation dosimeters. The insert contained a primary planning target volume (PTV) representing a tumor, a secondary PTV representing a lymph node, and an OAR representing the spinal cord (Fig. 1). The primary PTV initially contained two thermoluminescent dosimeters (TLDs) and the secondary PTV and OAR one TLD each; in 2004, the primary PTV was modified to contain four TLDs and the secondary PTV and OAR contained two TLDs each. Institutions are instructed to contour the TLD powder and report the average dose to the TLD as calculated by the treatment planning system (TPS).[10] Sheets of radiochromic film, either MD-55
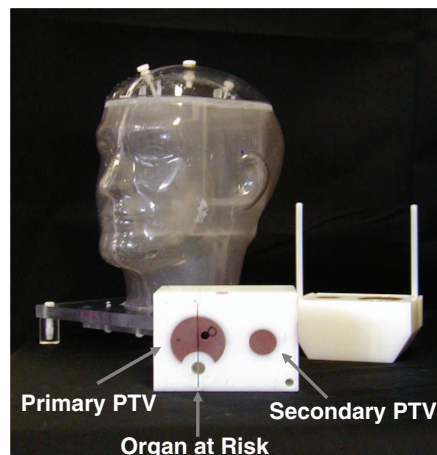

FIG. 1. RPC H&N phantom for IMRT credentialing.

before February 2006 or EBT/EBT2 after February 2006 (International Specialty Products, Wayne, NJ), are placed in the axial planes passing through the centers of the PTVs and in the sagittal plane passing through the center of the primary PTV.

Institutions interested in participating in NCI-sponsored IMRT clinical trials contact the RPC and request that a phantom be shipped to them. Upon receipt of a phantom, the institution fills the phantom with water, images the phantom, designs a treatment plan in accordance with the irradiation instructions accompanying the phantom, and then delivers the treatment just as it would do for an actual patient. A dose of 6.6 Gy is to be administered to at least 95% of the primary PTV, a dose of 5.4 Gy to at least 95% of the secondary PTV, and a maximum dose of 4.5 Gy to the OAR. Institutions are also asked to perform their own routine IMRT QA measurements before returning the phantom to the RPC for analysis.

Upon receipt of an irradiated phantom, the RPC analyzes the TLDs and radiochromic film as described by Molineu *et al.* The TLD-100 capsules (Quantaflux, LLC Oregonia, OH) are read using the same technique as used by the RPC for their TLD beam output verification mail-out dosimetry program.[11,12] The RPC evaluates each batch of TLD-100 powder for dose response, energy dependence, dose linearity, and fading. The TLDs are used to evaluate the absolute dose delivered to the phantom where each TLD is located and to normalize the film dose distributions. The radiochromic film is used to measure dose distributions and treatment field localization. As with the TLDs, the radiochromic film is investigated for dose response and fading as well as energy independence and uniformity. Localization marks are made on each piece of film to specify its unique location and orientation with respect to the targets and the OAR. The details of handling and analyzing the film were previously discussed by Molineu *et al.* and Niroomand-Rad *et al.*[10,13] Dose profiles through the centers of the primary PTVs and the OARs are generated based on the localization marks on each film.

Acceptance criteria for irradiations are that the delivered dose be ±7% of the planned absolute dose to the PTVs and

that the distance to agreement (DTA) be ± 4 mm in the high-gradient region between the primary PTV and the OAR.[10]

Each institution is asked to complete a short questionnaire about which energy and machine was used, the intensity modulation device, the IMRT technique, and the TPS. Since December 2003, the institutions have also been asked to provide the results from their in-house QA measurements. In order for the RPC to compare an institution's planned and measured doses and dose distributions, the institution is required to provide the dose to the TLD volume and isodose data in the form of either hard copy (prior to 2009) or digital imaging and communications in medicine-radiation therapy (DICOM-RT) files, the latter of which may have been submitted to the Image-Guided Therapy Quality Assurance Center and subsequently made available to the RPC. The results from the RPC's examination of the TLDs and radiochromic film were compared with the data provided by the institution as described previously by Molineu et al.[10]

From 2001 through 2011 the RPC mailed IMRT H&N phantoms to 763 distinct institutions and obtained information for 1139 phantom irradiations. Data and results from these irradiations were used for this analysis. We performed an analysis of the phantom pass rate to determine whether it was associated with the IMRT technique, TPS, linear accelerator (linac) manufacturer, or linac-TPS combination. The five IMRT techniques compared were dynamic MLC (sliding window) IMRT, intensity-modulated arc therapy, step-and-shoot MLC IMRT, solid attenuator modulation therapy, and tomotherapy. The analysis of the pass rate versus the TPS used to plan the IMRT treatment compared Varian Eclipse (Varian Medical Systems, Palo Alto, CA), Philips Pinnacle[3] (Philips Healthcare, Andover, MA), TomoTherapy Treatment Planning Station (Accuray, Sunnyvale, CA), XiO (Elekta Inc., Norcross, GA), and other planning systems. The analysis of the phantom pass rate versus the linac manufacturer com-pared Varian Medical Systems, Siemens, Elekta Inc., and TomoTherapy Hi-Art. The results were also analyzed with respect to physicists per machine using the divisions set by the 2004 AAPM Profile of Radiation Oncology Departments.[14]

III. RESULTS

Of the 1139 phantom irradiations in our database, 929 (81.6%) met the RPC's acceptance criteria. Two hundred ten irradiations did not meet the criteria; upon failure, some institutions repeated the irradiation and subsequently passed. All of the repeat irradiations are included in this analysis. Since the introduction of the H&N IMRT phantom, the annual pass rate has continually increased from an initial low of 66% to 93% for the most recent complete calendar year, though the rate for the year of this paper is currently lower (88.5%) than the previous year (Fig. 2).

Of the irradiations that did not meet the acceptance criteria, the majority (156 of 210) failed solely because their irradiations did not meet the dose criterion of $\pm 7\%$ of the planned dose. The other institutions that did not meet the criteria were divided between 21 that did not meet only the DTA criterion of ± 4 mm and 33 that met neither the dose nor the DTA criterion. Three hundred ninety-two repeat phantom irradiations were done. Several institutions reirradiated the phantom either because they wanted to improve their irradiation results after a failure, test different TPS algorithms, or test different treatment delivery systems. The pass rate for subsequent irradiations is 80.9%.

The IMRT H&N phantom results shown in Table I represent the mean ratio of the dose measured by the RPC to the institution calculated dose and the average DTA between the dose distribution measured by the RPC and the dose distribution calculated by the institution. Both the PTV dose ratios and the DTA values show good agreement, with some large
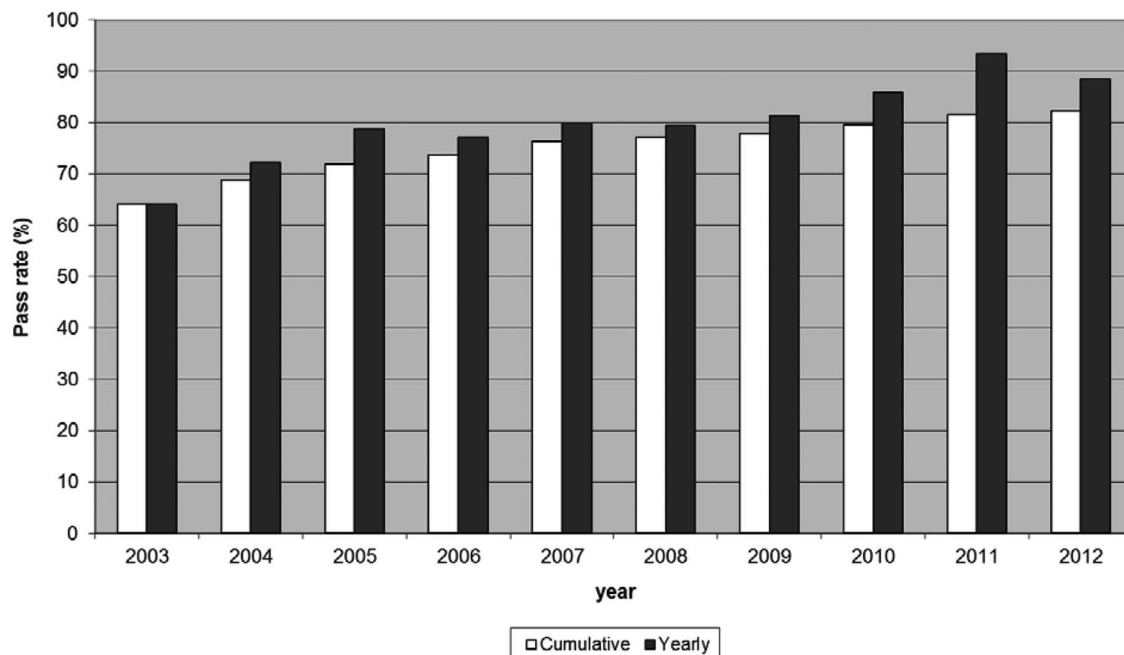

FIG. 2. Cumulative and yearly passing rate over time.

TABLE I. TLD/institution dose ratio and DTA averages.

	Ratio		
	1PTV	2PTV	DTA (mm)
Mean	0.98	0.98	1.6
SD	0.047	0.041	1.9
Range	0.44–1.26	0.40–1.23	0–17

disagreements as indicated by the range of values. Figure 3 shows a histogram of the dose ratios for 6520 TLD point measurements. The distribution appears normal and shows more dose failures in which the measured dose was unacceptably lower than planned (<0.93) than in which the measured dose was unacceptably higher than planned (>1.07). Figure 4 shows a histogram of the DTA displacements in millimeters, with very few falling outside the criterion of ± 4 mm.

As described in Sec. II, we also explored relationships between the phantom pass rate and other parameters. Figure 5 shows the pass rate according to the number of medical physicists per treatment machine. The pass rate was slightly lower than average at institutions with less than 0.76 physicists per machine compared with the pass rate at institutions with at least 0.76 physicists per machine; however, no statistically significant difference was seen between these pass rates ($p = 0.097$). The pass rate was found to be slightly higher for institutions with ≥ 5 treatment machines in their clinic (86%) than for institutions that had only 1 or 2 machines (80%), though no statistically significant difference between these pass rates was seen either ($p = 0.263$) (Fig. 6).

The results of the phantom pass rate analysis are shown in Table II. Most irradiations (930 of 1139) were performed with either the dynamic MLC (sliding window) or the segmental (step-and-shoot) techniques. Pass rates from dynamic

MLC, IMAT, and TomoTherapy deliveries were each found to be statistically significantly different than pass rates from each of the other delivery types, but were not significantly different than each other. Pass rates from segmental and solid attenuator techniques were significantly different from each other as well as from pass rates of each of the other types. However, because only seven irradiations using this technique were performed, it is difficult to compare its pass rate with those of the more widely used techniques and derive any meaningful conclusion.

The analysis showed that pass rates by TPS fell into two groups: The irradiations planned using Eclipse and TomoTherapy TPSs had pass rates of 88% and 93%, respectively, whereas the irradiations planned using Pinnacle[3] and XiO had pass rates of 75% and 76%, respectively. Eclipse and TomoTherapy pass rates were each found to be significantly different than Pinnacle[3] and XiO pass rates ($p \leq 0.001$ for each). Eclipse and TomoTherapy pass rates were not found to be significantly different than each other nor were the pass rates of Pinnacle[3] and XiO. The pass rates for the combined group of all planning systems other than these four were only found to be significantly different than the pass rate for TomoTherapy ($p = 0.003$).

In our analysis of pass rate according to linac manufacturer, three distinct pass rate groupings appeared. As with the TPS, the TomoTherapy (now owned by Accuray) model had the highest pass rate (93%). Varian linacs had the second highest pass rate (85%), and the Siemens and Elekta linacs had pass rates of 70% and 67% pass, respectively. The majority of the irradiations (68%) were performed with Varian linacs, and the remaining irradiations were split roughly evenly between the other three linac manufacturers. Pass rates for Varian and TomoTherapy were significantly different from each other as well as from each of the other two. Elekta and Siemens pass rates were not significantly different from each other.
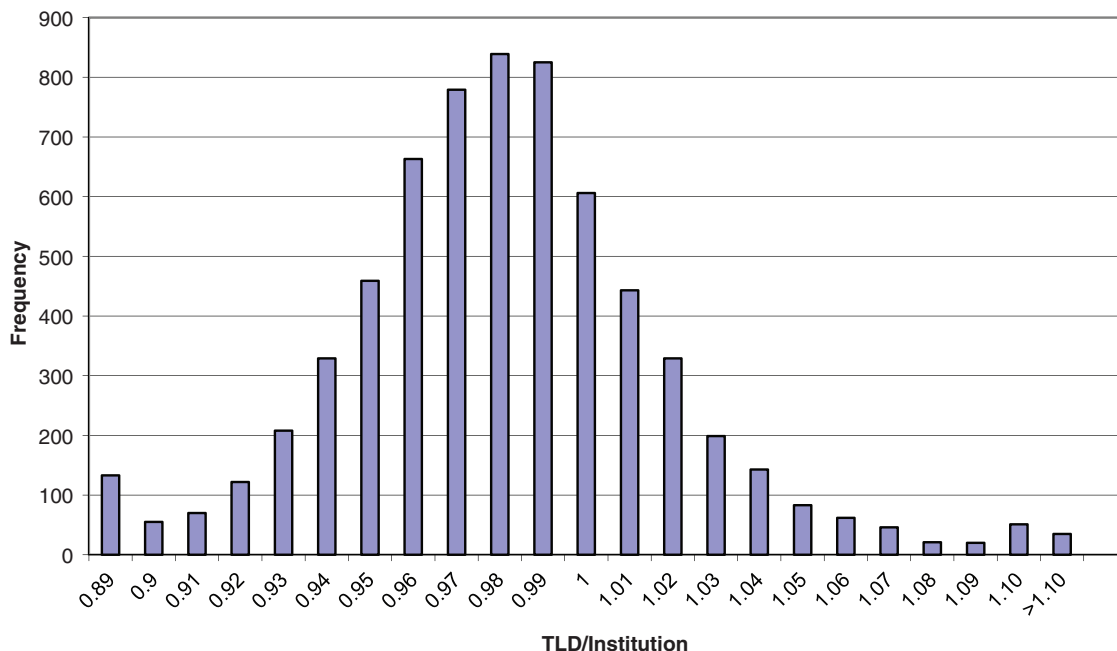

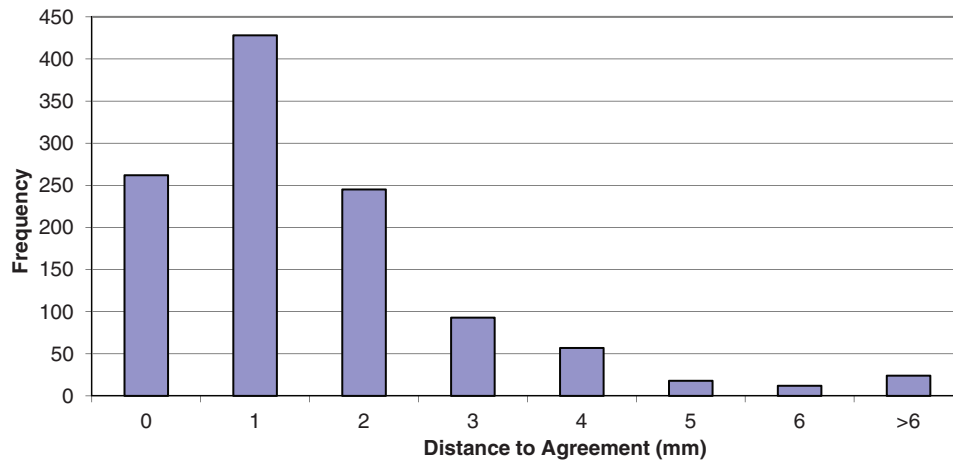FIG. 3. TLD/institution ratio for all PTV TLD results.

FIG. 4. Distance to agreement values for all phantom irradiations.

The bottom portion of Table II shows the numbers and percentages of irradiations done with the linac-TPS combinations that were used for more than 50 phantom irradiations. This dataset comprised 978 of the 1139 irradiations. The two combinations used most often to plan and irradiate the phantom were Varian-Eclipse (372 irradiations) and Varian-Pinnacle[3] (267 irradiations). The Varian-Eclipse and Hi-Art-TomoTherapy combinations had the highest pass rates, with a range of 90%–93%. Their pass rates were not found to be significantly different from each other, but each was significantly different than the pass rates from each of the other combinations. The remaining combinations shown in the table had pass rates in the range of 66%–81%. Pass rates from each combination in this group were significantly different than pass rates from each of the first group. In addition, the pass rate of Varian-Pinnacle[3] was significantly different than the pass rate of Elekta-Pinnacle[3] and Siemens-Pinnacle[3].

IV. DISCUSSION

The RPC's IMRT H&N phantom has been used for more than 10 years as an end-to-end quality audit test of an insti-

tution's ability to image two targets and an OAR and plan and deliver a complex IMRT treatment to those two targets while avoiding the OAR. We found that 81% of institutions' phantom irradiations accurately delivered an IMRT treatment to the H&N phantom that met the RPC's acceptance criteria of ±7%/4 mm. On the basis of these data, we conclude that all major linac-TPS combinations can pass the phantom criteria.

The increase in the phantom pass rate (66% to the current cumulative 82%) since the inclusion of IMRT in clinical trials is likely attributable to several factors. A key factor has been the improvement in the modeling of MLC leaves in TPSs. Specifically, the appropriate modeling of the rounded MLC leaf ends found in Varian linacs in the Pinnacle[3] planning system, as discussed by Cadman *et al.*, is believed to have improved the pass rate for phantoms planned using Pinnacle[3] and irradiated using a Varian linac.[15] In addition, the lower pass rates in the early 2000s were likely due to the inexperience of medical physicists and dosimetrists with IMRT treatment delivery because the technique was new to many clinics. Over time, institutions have become more experienced, dosimetry parameters involving the MLC and small field sizes
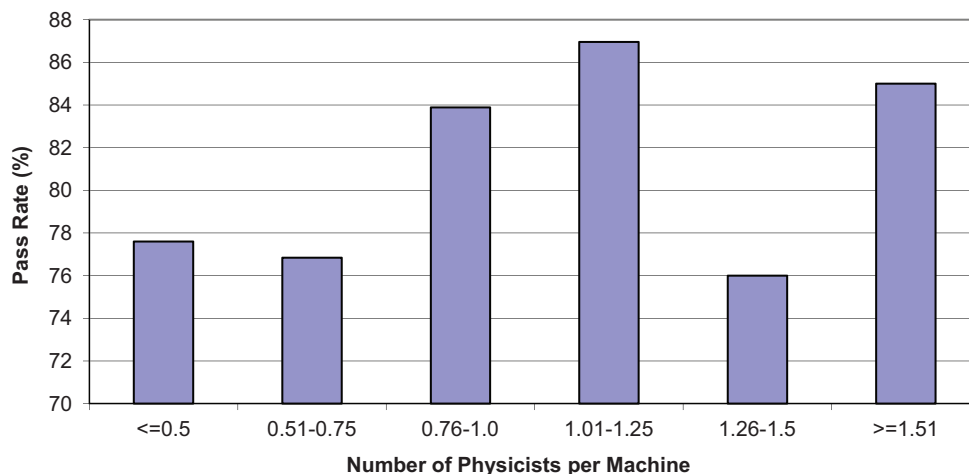

FIG. 5. Pass rate versus the number of physicist per machine.

TABLE II. Pass rate versus IMRT technique, treatment planning system, linear accelerator manufacturer, and linac-TPS combination.

	Pass rate (%)	Attempts	Criterion failed		
			Dose	DTA	Dose and DTA
IMRT technique					
Dynamic MLC	88	296	26	5	5
IMAT	86	103	11	0	3
Segmental	76	634	109	15	25
Solid attenuator	43	7	4	0	0
TomoTherapy	93	99	6	1	0
Treatment planning system					
Eclipse	88	387	30	8	7
Pinnacle[3]	75	425	84	8	13
TomoTherapy	93	99	6	1	0
XiO	76	137	19	4	10
Other	78	91	17	0	3
Linear accelerator manufacturer					
Elekta	67	130	37	4	2
Siemens	70	135	32	3	6
TomoTherapy	93	99	6	1	0
Varian	85	775	81	13	25
Linac-TPS combination					
Elekta-Pinnacle[3]	66	90	28	3	0
Siemens-Pinnacle[3]	67	76	21	0	4
TomoTherapy-HiArt	93	99	6	1	0
Varian-Eclipse	90	372	22	7	7
Varian-Pinnacle[3]	81	267	38	5	9
Varian-XiO	77	74	10	1	6

have been incorporated in dose calculation algorithms, and the training of staff has greatly improved. However, even with these improvements, phantom irradiation failures still occur. Some radiotherapy institutions are still newly implementing IMRT on a larger scale or to different anatomical sites such as the lung and the H&N. Other challenges that may lead to irradiation failures are the adoption of IMRT delivery techniques new to the institution such as tomotherapy and volumetric modulated arc therapy and human error in implementing and delivering IMRT.

The dose ratio and DTA results have wide ranges (Figs. 2 and 3). Despite this, the mean dose ratio was close to 1.00, indicating that a lack of bias in the TLD analysis system. The wide range is something of a concern, however, given that the acceptance criteria themselves are sometimes criticized for being too generous. The criteria of ±7%/4 mm were based on the results from the first 10 institutions that irradiated the H&N phantom and were chosen such that 90% of those 10 institutions should have been able to meet the criteria, assuming a normally distributed data set and with consideration for the uncertainty in TLDs and film dosimeters that were in use at the time. Since the criteria were adopted, however, the pass rate has never met that goal (although it is now being approached). Therefore, regardless of whether some believe the criteria are too generous, increasing their rigidity is not an option at this point since so many institutions were not able to meet the criteria. We conducted an exploratory analysis of the pass rate with the criteria changed to ±5%/4 mm, which are closer to the criteria that most individual institutions use. The results, seen in Table III, indicate that the failure rate would double and pass rates drop from 75%–93% to 54%–79%. Such a decrease in the phantom pass rate would markedly restrict clinical trial participation, which is not the purpose of this credentialing tool. That is, while the H&N IMRT phantom is designed to set a QA standard, it is also meant to ensure that there is consistency between institutions so that their clinical data can be compared; neither of these purposes is superior to the other, and so the credentialing criteria must take both into account.

The RPC has now commissioned software that calculates a gamma index based on a 2D analysis of the film, which should increase the precision with which this criterion can be assessed. This software enables a 2D analysis of the film from the RPC's phantoms that can be compared with the
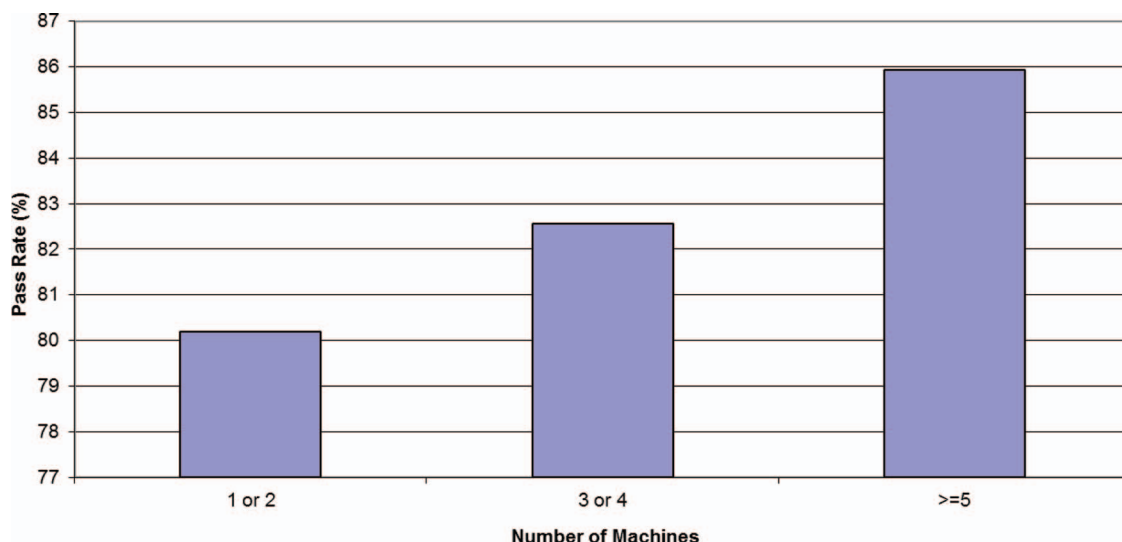
FIG. 6. Pass rate versus the number of machines.

TABLE III. Comparison of pass rates for treatment planning systems with two sets of criteria.

Treatment planning system	Pass rate (%) 5%/4 mm	Pass rate (%) 7%/4 mm
Eclipse	72	88
Pinnacle[3]	56	75
TomoTherapy	79	93
XiO	54	76
Other	56	78

institutions' calculated dose distributions. The RPC and the RTOG agreed to implement this analysis for credentialing purposes beginning in February 2012.

Our analysis of institutional demographics showed that when institutions were divided into categories based on staffing, the distribution was similar to that in the AAPM survey except that the group with the most physicists per machine was over-represented in our database. This variance is likely easy to explain: institutions with more physicists per machine tend to be larger research institutions, almost all of which participate in clinical trials. Because we intuitively thought that more physicists at an institution would indicate a smaller work load per person and thus a lower instance of human error, we were surprised to find that the institutions with a mean of more than one physicist per machine did not have a much higher pass rate (82.0%) than those with only one or fewer physicists per machine (80.9%). These numbers are not significantly different ($p = 0.097$). This is possibly because the number of physicists per machine in an institution may not accurately represent the number of physicists responsible for work on each machine.

Physicists often ask us if higher-quality treatment is provided at large institutions than at smaller institutions. Our results showed that large institutions failed to meet acceptance criteria only slightly less often than smaller institutions with no significant difference ($p = 0.263$).

Our analysis indicated a difference of 10 percentage points between the pass rates for the Varian-Eclipse (90.3%) and Varian-Pinnacle[3] (80.5%) combinations, the two combinations of linac manufacturers and TPSs with the greatest number of irradiations in this study. This is significantly different ($p < 0.001$). The lower pass rate for the Varian-Pinnacle[3] combination may be due in part to an earlier Pinnacle version's inability to adequately model the rounded leaf ends of the MLC.[15] The current version of Pinnacle[3] has improved the leaf end modeling, and the pass rate for the Varian-Pinnacle[3] combination has increased over time. The Varian-Pinnacle[3] pass rate for 2010 and 2011 is 90%. Analysis done near the time that institutions were implementing the improved model showed a pass rate increase of 6 percentage points between old models and newer models. We also note that the combinations of linacs and TPSs from the same manufacturer, such as the Varian linac-Eclipse TPS and the Accuray HiArt TomoTherapy linac-TomoTherapy Treatment Planning Station (listed as TomoTherapy-HiArt in Table II), had the highest pass rates. The use of a planning system and a linac devel-

oped to work together resulted in more passing phantom irradiations. Ranking the different linac and TPS manufacturers by quality is not possible with these data for two reasons. First, the numbers of irradiations performed using linacs from each manufacturer in this study except Varian were so small that that a small change in the number of failures would cause a disproportionately large change in the pass rates. Second, different institutions implement and commission their IMRT programs differently, which can have a large impact on the results.

Reasons for phantom irradiation failures that we have verified with institutions include inaccurate output factors input into the TPS, incorrect percentage depth dose input into the TPS, inadequacies in beam modeling, not adjusting the plan according to QA results from phantom-specific ionization chamber measurements, errors in couch indexing with the Peacock system (Nomos, Pittsburgh, PA), too high a tolerance for MLC leaf positioning, setup errors, and accelerator target malfunctions. If the basic beam model's dose calculations from the TPS do not match dosimetry measurements under simple irradiation conditions, then this disagreement will be even greater when calculating the complex dose distributions found in IMRT treatments. As IMRT delivery becomes more dynamic, as with volumetric modulated arc therapy or tomotherapy, it becomes even more crucial that institutions first verify dosimetry calculations under reference conditions and then compare dose distributions calculated by the TPS and measured in phantom irradiations under more complex conditions. Because the vast majority of the irradiations reported passing their own patient specific QA, it is evident that the use of a simple treatment planning benchmark case would not have detected the discrepancies discovered by these phantom irradiations.

We cannot explain all of the failed irradiations. An institution's physicist sometimes chose to independently determine reasons of failure, and sometimes, despite working with the physicists who carried out the irradiations, we were still unable to determine why the delivery did not go as planned. Of the 210 failures, fewer than 30 were due to gross setup errors that may be less likely to happen with a real patient. Even though the phantom irradiation instructions explicitly state to treat the phantom as if it were a patient, the physicist did not always follow the institutional setup procedures for patients. That leaves >85% of our failures as potential real errors in treatment delivery, not due to one-time physicist error.

It is also notable that many institutions were able to make the necessary changes and pass the phantom irradiation on a subsequent try. Among the strategies employed were inputting new data into the TPS, adjusting models, and updating software and hardware. We note that the improvements were often of the variety that could have been detected at the time of IMRT commissioning, showing the importance of adequate commissioning of the type of treatment to be delivered.

The RPC's IMRT H&N phantom provides a comprehensive evaluation of IMRT for clinical trials that has, for the past 10 years, helped set the standard for IMRT delivery in

NCI-sponsored clinical trials. In addition, numerous errors in IMRT delivery have been detected and resolved, benefiting not only the trial patients but all patients treated at the particular institution. Even though the pass rate continues to climb, vigilance is still needed because many institutions were not able to deliver dose to the phantom within the acceptance criteria of $\pm 7\%/4$ mm and because new IMRT delivery modalities, which require credentialing, are being introduced each year. Therefore, credentialing through the use of an independent end-to-end QA tool that consists of the same dosimetry phantom, same dosimeters, same robust analysis, and same acceptance criteria, such as the RPC's IMRT H&N phantom, is still needed as an important step for ensuring consistency of dose delivery for clinical trials using IMRT.

[a] Author to whom correspondence should be addressed. Electronic mail: amolineu@mdanderson.org

[1] I. J. Das, C. Cheng, K. L. Chopra, R. K. Mitra, S. P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems," J. Natl. Cancer Inst. **100**, 300–307 (2008).

[2] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy, "A technique for the quantitative evaluation of dose distributions," Med. Phys. **25**, 656–661 (1998).

[3] M. Partridge, P. M. Evans, A. Mosleh-Shirazi, and D. Convery, "Independent verification using portal imaging of intensity-modulated beam delivery by the dynamic MLC technique," Med. Phys. **25**, 1872–1879 (1998).

[4] G. J. Kutcher, "Comprehensive QA for radiation oncology: Report of AAPM Radiation Therapy Committee Task Group 40," Med. Phys. **21**, 581–618 (1994).

[5] E. E. Klein, J. Hanley, J. Bayouth, F. F. Yin, W. Simon, S. Dresser, C. Serago, F. Aguirre, C. Ma, B. Arjomandy, C. Liu, C. Sandin, and T. Holmes, "Task Group 142 Report: Quality assurance of medical accelerators," Med. Phys. **36**, 4197–4212 (2009).

[6] G. A. Ezzell, J. W. Burmeister, N. Dogan, T. J. LoSasso, J. G. Mechalakos, D. Mihailidis, A. Molineu, J. R. Palta, C. R. Ramsey, B. J. Salter, J. Shi, P. Xia, N. J. Yue, and Y. Xial, "IMRT commissioning: Multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119," Med. Phys. **36**, 5359–5373 (2009).

[7] J. R. Palta, J. A. Deye, G. S. Ibbott, and J. A. Purdy, "Credentialing of institutions for IMRT in clinical trials," Int. J. Radiat. Oncol., Biol., Phys. **59**, 1257–1259 (2004).

[8] J. Leif, J. Roll, D. Followill, and G. Ibbott, "2893: The value of credentialing," Int. J. Radiat. Oncol., Biol., Phys. **66**, S716 (2006).

[9] G. S. Ibbott, D. S. Followill, H. A. Molineu, J. R. Lowenstein, P. E. Alvarez, and J. E. Roll, "Challenges in credentialing institutions and participants in advanced technology multi-institutional clinical trials," Int. J. Radiat. Oncol., Biol., Phys. **71**, S71–S75 (2008).

[10] A. Molineu *et al.*, "Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the radiation therapy oncology group," Int. J. Radiat. Oncol., Biol., Phys. **63**, 577–583 (2005).

[11] T. H. Kirby *et al.*, "System for photon and electron therapy beams," Int. J. Radiat. Oncol., Biol., Phys. **12**, 261–265 (1986).

[12] T. H. Kirby *et al.*, "Uncertainty analysis of absorbed dose calculations for thermoluminescence dosimeters," Med. Phys. **19**, 1427–1433 (1992).

[13] A. Niroomand-Rad *et al.*, "Radiochromic film dosimetry: Recommendations of AAPM radiation therapy committee task group 55," Med. Phys. **25**, 2093–2115 (1998).

[14] American Association of Physicists in Medicine, "Profile of radiation oncology departments calendar year 2004," Professional Survey 2004.

[15] P. Cadman, R. Bassalow, N. P. S. Sidhu, G. S. Ibbott, and A. Nelson, "Dosimetric considerations for validation of a sequential IMRT process with a commercial treatment planning system," Phys. Med. Biol. **47**, 3001–3010 (2002).