Published in final edited form as:

Int J Radiat Oncol Biol Phys. 2014 December 1; 90(5): 1195–1201. doi:10.1016/j.ijrobp.2014.08.334.

Institutional patient-specific intensity-modulated radiation therapy quality assurance does not predict unacceptable plan delivery as measured by IROC Houston's head and neck phantom

Stephen F. Kry, PhD^{1,2,*}, Andrea Molineu, MS¹, James Kerns, MS^{1,2}, Austin Faught^{1,2}, Jessie Y. Huang^{1,2}, Kiley Pulliam, MS^{1,2}, Jackie Tonigan, MS^{1,2}, Paola Alvarez, MS¹, Francesco Stingo, PhD^{2,3}, and David S. Followill, PhD^{1,2}

¹IROC Houston, Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, Texas

²The University of Texas Health Science Center Houston, Graduate School of Biomedical Sciences, Houston, Texas

³Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas

Abstract

Purpose—To determine whether in-house patient-specific IMRT QA results predict the Imaging and Radiation Oncology Core (IROC)-Houston phantom results.

Methods and Materials—IROC Houston's IMRT head and neck phantoms have been irradiated by numerous institutions as part of clinical trial credentialing. We retrospectively compared these phantom results with those of in-house IMRT QA (following the institution's clinical process) for 855 irradiations performed between 2003 and 2013. The sensitivity and specificity of IMRT QA to detect unacceptable or acceptable plans was determined relative to the IROC Houston phantom results. Additional analyses evaluated specific IMRT QA dosimeters and analysis methods.

Results—IMRT QA universally showed poor sensitivity relative to the head and neck phantom i.e., poor ability to predict a failing IROC Houston phantom result. Depending on how the IMRT QA results were interpreted, overall sensitivity ranged from 2% to 18%. For different IMRT QA methods, sensitivity ranged from 3% to 54%. Although the observed sensitivity was particularly poor at clinical thresholds (e.g., 3% dose difference or 90% of pixels passing gamma), receiver

© 2014 Elsevier Inc. All rights reserved.

*Corresponding author: Stephen F. Kry, PhD, IROC Houston, Department of Radiation Physics, Unit 547, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Houston, TX 77030, USA; phone: 713-745-8989; fax: 713-794-1364; sfkry@mdanderson.org.

Conflict of interest: None

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

operator characteristic analysis indicated that no threshold showed good sensitivity and specificity for the devices evaluated.

Conclusions—IMRT QA is not a reasonable replacement for a credentialing phantom. Moreover, the particularly poor agreement between IMRT QA and the IROC Houston phantoms highlights surprising inconsistency in the QA process.

Keywords

IMRT QA; RPC; phantom; IROC Houston

Introduction

The mission of the Imaging and Radiation Oncology Core at Houston (IROC Houston), formerly the Radiological Physics Center, is to ensure that institutions deliver accurate and consistent radiotherapy in support of clinical trials. This includes consistency of the intended dose distribution (which is challenging in itself [1]), as well as accurate delivery of the intended distribution. The importance of IROC Houston's mission was highlighted in TROG 02.02, where the quality of radiotherapy had a greater effect on overall survival than even the hypothesized effect of the pharmaceutical agent under evaluation (2). Quality radiotherapy is important for patient outcome, and to maximally power clinical trials.

Phantoms are one of IROC Houston's central tools for ensuring agreement between calculated and delivered dose distributions for intensity-modulated radiotherapy (IMRT). However, phantom irradiations are time-consuming and add (albeit very marginally) to the overall cost of clinical trials. Therefore, internal evaluation, such as an institution's patient-specific IMRT quality assurance (QA), has been suggested as an alternative to phantoms (3). This suggestion makes conceptual sense because both IROC Houston phantoms and IMRT QA are designed to verify that radiotherapy is delivered as intended. However, the methodology differs substantially between these 2 approaches: the phantom measures the dose in the geometry used to plan the treatment, whereas patient-specific IMRT QA measures the dose recalculated on a hybrid phantom (with or without using the planned gantry angles). In addition, the RPC system uses consistent dosimeters and analyses, whereas IMRT QA methods are widely varied.

Therefore, we sought to determine whether IMRT QA can replace IROC Houston phantoms. We compared the results of 855 head and neck phantom irradiations with the results of each institution's in-house IMRT QA of the same plan.

Methods

Phantoms

IROC Houston's head and neck phantom (Figure 1) includes a primary and secondary target and an avoidance structure, as described previously (4–6). Institutions are instructed to treat the phantom as a patient, including scanning, contouring, planning, setting up, and delivering the treatment plan.

The delivered dose is measured by thermoluminescent dosimeters (TLDs) and film. Four TLDs are in the primary target and 2 are in the secondary target. TLDs are processed according to standard IROC Houston practice (7), accounting for fading, nonlinearity in the dose response, energy dependence, and variability in the TLD reader. Considering the uncertainty in this process (7) and variations in the energy spectra from different treatments (8), the uncertainty in the dose is $\pm 2.5\%$ (1-sigma). Sagittal and transverse planes of Gafchromic film (EBT or EBT2, Ashland) intersect the primary target. The planar dose is normalized to the TLD dose in the primary target (4).

Measured doses are compared with those calculated by the treatment planning system. Regions of interest corresponding to the volume of the TLD are compared with the TLD dose; planar dose distributions are compared with the film doses. For the phantom irradiation to “pass,” the dose measured by each of the 6 TLDs must agree with the planning system within $\pm 7\%$. Planar dose distributions must also agree. Prior to 2012, this was based on a 4 mm distance-to-agreement criterion in the region between the primary target and the organ at risk. Since 2012, 85% of pixels must pass a 7%/4-mm gamma criterion on each film. Although the institutions are given planning constraints, acceptability of the phantom irradiation does not hinge on meeting these constraints, but solely on the agreement between planned and measured doses.

The head and neck phantom has been irradiated more than 1000 times. Although these date back to 2001, most have occurred since 2008. Irradiation of this phantom has included a diverse distribution of delivery techniques, treatment planning systems, and linear accelerators (5).

Institutional IMRT QA

When developing the phantom plan, the institution is instructed to also conduct and submit IMRT QA of the plan according to their individual clinical practice.

For this study, institutions’ IMRT QA records were retrospectively reviewed for the following: whether the institution claimed the plan passed or failed IMRT QA, the IMRT QA device used, criteria for acceptance, agreement between their measurement and calculation, and whether absolute or relative dosimetry was used. Records were excluded if no viable IMRT QA information was submitted. Additionally, 19 plans (typically from early phantom results) were excluded because the monitor units may have been altered between the IMRT QA results and the phantom results and the size of the change, if any, could not be confirmed. A total of 855 phantom irradiations remained and were used for initial analysis.

Analysis

Truth tables were constructed to compare the pass/fail results of IMRT QA and the IROC Houston phantom; these tables identified the sensitivity and specificity of IMRT QA relative to the phantom. Sensitivity is how often IMRT QA indicates a plan fails when the phantom indicates a failing plan; specificity is how often IMRT QA indicates a plan passes when the phantom indicates a passing plan.

Because many hospitals submitted their IMRT QA results without explicitly stating whether they considered the results to be passing, we constructed an initial truth table assuming that all submitted cases passed IMRT QA unless the institution explicitly stated otherwise ($n = 5$); we assumed that institutions were unlikely to irradiate the phantom with a plan they believed to be flawed. However, in several cases, the IMRT QA results were poor even when the institution claimed that the results passed their institutional criteria. Therefore, we also evaluated institutions' IMRT QA results on the basis of interpretation by IROC Houston, rather than the declared pass/fail status from the institution. Based on common clinical practice, we declared plans to have failed IMRT QA if all reported absolute point dose measurement(s) from a point dosimeter disagreed by $>3\%$, composite planar dose analysis had $<90\%$ of pixels passing a $3\%/3\text{-mm}$ (or looser) gamma criterion, or field-by-field planar analysis had multiple fields with $<90\%$ of pixels passing this same gamma criterion. Plans were considered to have passed IMRT QA with absolute dose agreement of better than 3% or more than 90% of pixels passing a $3\%/3\text{-mm}$ (or tighter) gamma criterion. Plans were excluded if they did not include sufficient details or did not use gamma criteria that fit into these categories. A second truth table was constructed with the remaining 745 plans.

The analysis for these 745 plans was also subdivided by device and analysis technique when there were a sufficient number of cases. Truth tables were constructed for IMRT QA conducted with both an ion chamber and a planar detector ($n = 91$), an ion chamber alone (including those already counted when used with a planar detector; $n = 325$), film ($n = 71$), and a MapCheck device ($n = 322$). Truth tables were also constructed for planar dosimeters used in absolute dose mode ($n = 295$) and relative dose mode ($n = 97$). Uncertainty in the sensitivity and specificity was calculated using a normal approximation of the variance and the binomial nature of the underlying sensitivity (or specificity) data. Equation 1 shows the standard deviation where p is the sensitivity (or specificity) and n is the number of samples used to calculate sensitivity (or specificity):

$$\sigma = \sqrt{\frac{p(1-p)}{n}} \quad (1)$$

For plans in which an ion chamber was used to measure absolute dose, we compared the percent dose difference between the ion chamber and treatment plan with the percent dose difference between the TLD in the phantom and the treatment plan (averaged over all TLDs). Similarly, for planar results, we compared the percent of pixels passing gamma in the IMRT QA device using a $3\%/3\text{-mm}$ criterion (averaged over all fields for field-by-field analysis) with the percent of pixels passing gamma in the phantom films (averaged over both planes). Even though gamma criteria was not used to define pass versus fail in the phantoms until 2012, this calculation was done starting in 2008, and all gamma results were included in this analysis. Regression analysis was performed for both comparisons.

Finally, receiver operator characteristic (ROC) curves were constructed for the 3 most common detectors (ion chamber, film, and MapCheck), to compare the performance of these devices while allowing the threshold to vary (i.e., not limited to a 3% dose difference

threshold for the ion chamber and a 90% threshold of pixels passing gamma). Planar analyses were limited to solely those done with a 3%/3-mm criterion for consistency (52 film results and 286 MapCheck results). Analysis was done with R, using an alpha of 0.05. There were insufficient samples with other devices to perform similar analysis.

Results

Of the 855 phantom irradiations and IMRT QA results initially analyzed, 122 (14%) failed the phantom, whereas 5 (0.6%) were declared by the institution to have failed IMRT QA (Figure 2a). Correspondingly, the IMRT QA results showed a sensitivity of 2% ($\pm 1\%$ standard deviation), indicating that they overwhelmingly failed to detect a plan that would fail the phantom. Specificity was $99.6 \pm 0.2\%$, indicating that IMRT QA almost perfectly predicted plans that would pass the phantom; this largely reflects that essentially all plans passed IMRT QA (Table 1).

When IROC Houston interpreted whether a plan had failed the institution's IMRT QA, many more plans were described as failing (Figure 2b). Seventy six plans (10%) failed IMRT QA, whereas 103 (14%) failed the phantom. Despite a more similar number of failing plans in this analysis, the sensitivity of IMRT QA remained poor ($18 \pm 4\%$): the plans that failed IMRT QA rarely corresponded to the plans that failed the phantom (Table 1).

The performance of specific IMRT QA dosimeters is also shown in Table 1. The ion chamber and planar device combination showed the highest sensitivity ($54 \pm 14\%$), although this is still relatively poor. The ion chamber alone or film alone showed poorer sensitivities and the MapCheck device showed the poorest sensitivity; plans that failed the RPC phantom were almost never identified as problematic by MapCheck. Although the 95% confidence intervals of the sensitivity of most devices overlapped, MapCheck performed significantly more poorly than the ion chamber + planar device. In general, the sensitivity of all devices was very low. In contrast, the specificity of all devices was relatively high; for all devices, IMRT QA declared most plans to be acceptable.

When planar/array devices were used in absolute versus relative dose mode (Table 1), sensitivity appeared, surprisingly, to be higher for relative mode ($21 \pm 9\%$ vs $3 \pm 3\%$). However, this difference was not statistically significant ($p=0.06$, Fisher exact test).

For IMRT QA results using an ion chamber, the percent dose difference for IMRT QA compared with the percent dose difference for the phantom TLD is shown in Figure 3. Although the TLD and ion chamber measurement points are not exactly the same location (or even in the same geometry), they are both designed to determine absolute dose agreement in the high-dose region and a correlation is expected. Regression analysis did show a significant positive relationship between them ($P = 0.002$); however, the slope of the line, rather than being around unity, was 0.1, and the R^2 value was only 0.03, indicating a very weak relationship.

Similar analysis was done for the percent of pixels passing gamma between planar IMRT QA and the phantom film (Figure 4). Although the analysis was in different planes, a relationship would still be expected. A significant positive relationship was present ($P <$

0.001), but with a very small slope (0.06) and an R^2 of 0.05, indicating a very weak relationship between planar IMRT QA and phantom results.

ROC curves are shown in Figure 5 for the ion chamber, film, and MapCheck devices, including the areas under the curve (AUCs). The AUCs for the 3 devices ranged between 0.606 and 0.699 and were not significantly different (p-values >0.42), and represent poor ability of these dosimeters to predict phantom results. The low AUC indicates that, in the aggregate manner these devices were used in this study, no threshold will provide both high sensitivity and high specificity.

Discussion

IMRT QA results were not predictive of IROC Houston phantom results. While IMRT QA must offer good specificity to be efficient and thereby avoid trouble-shooting acceptable plans, good sensitivity is more important; if the plan is unacceptable, IMRT QA should indicate this. Our sensitivity analyses showed that IMRT QA was a poor tool for predicting failing the phantom.

IMRT QA declared a much higher percentage of plans to be acceptable than the phantom did. Only 0.6–10% of plans failed IMRT QA, depending on whether the institution or RPC interpretation was used. These values straddle the range of failing rates reported in the literature, 1–3% (9–11), but are all less than the 14% fail rate of the IROC Houston phantom. There are several reasons for this discrepancy. First, several forms of common IMRT QA have recently been shown to not detect known large errors (12-). Second, when IMRT QA fails clinically, it is common to repeat measurements or take new measurements until a passing measurement is obtained (9,13), a process that is not possible with the phantom. Third, the phantom is an end-to-end test, including a more challenging alignment process and increased risk of setup error (which is not detectable and leads to a failing phantom result). Although setup errors are a concern in patient care, they are not expected to agree between IMRT QA and phantom irradiations, which would inappropriately lower the sensitivity. However, this issue actually contributed relatively little to the disparity in failing rates (sensitivity) in our study. Comparable to previous results (3) of the IROC Houston phantom, of the 103 phantom failures in our study, 72 were because of a dose difference in the TLD, 11 were because of disagreement in the planar results, and 20 were because of both. In only 15 cases was a setup error identified as part of the cause of failure. Excluding these 15 cases, the sensitivity (for all IMRT QA devices) changed minimally and not significantly from 18% to 20%.

More importantly than the difference in failing rates between IMRT QA and the IROC Houston phantom is the very poor overlap between the plans that failed each. Figure 2 shows very dramatically that very few of the plans that failed IMRT QA failed the phantom, and very few of the plans that failed the phantom failed IMRT QA. Across all devices and analysis methods, the number of false negatives and false positives were much larger than the number of true negatives. This poor agreement, combined with the poor agreement shown in Figures 3 and 4, is disconcerting. Although the details of the process are different, both IMRT QA and IROC Houston phantoms are designed to verify the accurate delivery of

IMRT. However, for all conditions evaluated in this study, these two methods showed highly disparate results. Other recent evaluations of the IMRT QA process have also found poor sensitivity of some forms of IMRT QA to detect plan errors. Although the mechanisms are not always clear, several concerns with common QA methods have been raised, such as the limitations of using a single ion chamber point to assess an entire plan (9,14), the inclusion of many low dose points in an array measurement that invariably pass a global dose difference criterion and therefore inflate the percent of pixels passing gamma (14,16), use of a lax gamma criteria (14–16), use of a single gamma metric to summarize an entire plan or failure to review gamma distributions carefully (16), and dosimeter deficiencies (14,15). This work supports those findings and highlights the need for further study of IMRT QA to optimize the QA process.

The ROC analysis (Figure 5) showed that, over a range of thresholds for dose differences and percent of pixels passing gamma, the ion chamber, film, and MapCheck detectors were all equally, and poorly, capable of predicting the phantom results. However, the thresholds commonly used clinically (3% dose difference, 90% of pixels passing gamma) are relatively loose, leading to the very poor sensitivities shown in Table 1. To achieve 50% sensitivity, a dose agreement threshold of 2.0% for the ion chamber is required or a threshold of 97% of pixels passing gamma for MapCheck is required (3%/3 mm). The tradeoff for this improved sensitivity is low specificity at these thresholds (70% for the ion chamber and 67% for MapCheck); that is, if high sensitivity were achieved, the QA device would have a relatively high likelihood of reporting a failing plan when it would have passed the phantom. In general, the low sensitivity observed in these devices at clinical thresholds is consistent with previous findings comparing different IMRT QA techniques (14–16). These results emphasize that the physics community should work towards identifying QA processes that offer large areas under the ROC curve. This will then allow for an optimal threshold to be identified.

Our findings have implications for multi-institutional clinical trials. With the goal of consistent dose delivery, the IROC Houston phantom has a conceptual advantage because all irradiation results are processed, analysed, and evaluated in a consistent manner. These phantoms allow for consistent criteria to be established. In contrast, between different institutions, IMRT QA comprises many different devices, delivery techniques, and analyses, with no well-established criteria for assessment (13). These conceptual concerns, combined with the results of this work that show substantial disagreement in terms of results, indicate that use of in-house patient-specific IMRT QA for plan verification for clinical trials is currently unrealistic.

Conclusions

In-house patient-specific IMRT QA failed to detect unacceptable plan delivery as measured by the IROC Houston head and neck phantom. This was true for all IMRT QA devices and analysis techniques evaluated. These findings indicate that IMRT QA is not a reasonable substitute for a phantom irradiation in the context of clinical trials. More generally, our findings highlight inconsistencies in QA results and suggest the need for improving the

patient-specific IMRT QA process. Further study is needed to optimize this process, including guidelines to accomplish this goal.

Acknowledgments

This work was funded in part by Public Health Service Grant CA10953 (NCI) and Cancer Center Support Grant P30CA016672.

References

1. Das IJ, Cheng CW, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems. *J Natl Cancer Inst.* 2008; 100:300–307. [PubMed: 18314476]
2. Peters LF, O'Sullivan B, Giralt J, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02. 02. *J Clin Oncol.* 2010; 28:2996–3001. [PubMed: 20479390]
3. Bekelman J, Deye J, Vikram B, et al. Redesigning radiotherapy quality assurance: opportunities to develop an efficient, evidence-based system to support clinical trials – report of the National Cancer Institute work group on radiotherapy quality assurance. *Int J Radiat Oncol Biol Phys.* 2012; 83(3): 782–790. [PubMed: 22425219]
4. Molineu A, Followill DS, Balter P, 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.* 2005; 63:577–583. [PubMed: 16168849]
5. Molineu A, Hernandez N, Nguyen T, Ibbott G, Followill DS. Credentialing results from IMRT irradiations of an anthropomorphic head and neck phantom. *Med Phys.* 2013; 40(2):022101–022101-8. [PubMed: 23387762]
6. Ibbott GS, Molineu A, Followill DS. Independent evaluations of IMRT through the use of an anthropomorphic phantom. *Technol Cancer Res Treat.* 2006; 5:481–487. [PubMed: 16981790]
7. Kirby T, Hanson WF, Johnston DA. Uncertainty analysis of absorbed dose calculations from thermoluminescence dosimeters. *Med Phys.* 1992; 19:1427–1433. [PubMed: 1461205]
8. Scarboro SB, Followill DS, Howell RM, Kry SF. Variations in photon energy spectra of a 6 MV beam and their impact on TLD response. *Med Phys.* 2011; 38(5):2619–2628. [PubMed: 21776799]
9. Dong L, Antolak J, Salehpour M, et al. Patient-specific point dose measurement for IMRT monitor unit verification. *Int J Radiat Oncol Biol Phys.* 2003; 56:867–877. [PubMed: 12788197]
10. Fenoglietto P, Laliberte B, Ailleres N, et al. Eight years of IMRT quality assurance with ionization chambers and film dosimetry experience of the montpellier comprehensive cancer center. *Radiat Oncol.* 2011; 6:1–11. [PubMed: 21208415]
11. Pulliam K, Followill DS, Court L, et al. A review of more than 13000 patient-specific IMRT QA results. *J Appl Clin Med Phys.* In press.
12. Kruse JJ. On the insensitivity of single field planar dosimetry to IMRT inaccuracies. *Med Phys.* 2010; 37:2516–2524. [PubMed: 20632563]
13. Nelms BE, Zhen H, Tome WA. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys.* 2011; 38:1037–1044. [PubMed: 21452741]
14. McKenzie EM, Balter P, Stingo F, Jones J, Followill DS, Kry SF. Toward optimizing patient-specific IMRT QA techniques in the accurate detection of dosimetrically acceptable and unacceptable patient plans. *Med Phys.* In review.
15. Carrasco P, Jornet N, Latorre A, Eudaldo T, Guiz A, Ribas M. 3D DVH-based metric analysis versus per-beam planar analysis in IMRT pretreatment verification. *Med Phys.* 2012; 39:5040–5049. [PubMed: 22894429]
16. Nelms BE, Chan MF, Jarry G, et al. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys.* 2013; 40:111722. [PubMed: 24320430]

17. Nelms BE, Simon JA. A survey on planar IMRT QA analysis. *J Appl Clin Med Phys*. 2007; 8(3): 76–90. [PubMed: 17592457]

Summary

We compared institutional patient-specific IMRT QA results with IROC Houston phantom results. Though both tools are designed to test the accuracy of IMRT plan delivery, we found that no IMRT QA device could reasonably predict whether a plan would fail the IROC Houston phantom. This indicates that IMRT QA is not a suitable replacement for an independent credentialing phantom and raises concerns about the level of consistency among IMRT evaluation methods.

Figure 1. IROC Houston's head and neck phantom. The imaging/dosimetry insert is disassembled, showing the 2 targets in brown and a small avoidance structure in olive.

		RPC				RPC	
		Fail	Pass			Fail	Pass
Inst QA	Fail	2	3	Inst QA	Fail	19	57
	Pass	120	730		Pass	84	585

(a)
(b)

Figure 2.

Truth table for institutional IMRT QA results versus IROC Houston phantom results for head and neck phantom plans. (a) All plans were assumed to pass institutional IMRT QA unless the institution explicitly stated otherwise. (b) Institutional IMRT QA results were determined by IROC Houston using the common criteria of 3% dose or >90% of pixels passing 3%/3 mm gamma.

Figure 3.

Percent differences between dose measurements and treatment planning system calculations for institutional IMRT QA compared with the TLD in the IROC Houston phantom. The linear trend line should ideally have a slope of 1, but instead is nearly flat.

Figure 4.

Percent of pixels passing gamma for institutional IMRT QA compared with the IROC Houston phantom films. The linear trend line should ideally have a slope of 1, but instead is nearly flat.

Figure 5.

ROC curves for the ion chamber (a), film (b), and MapCheck (c), devices, indicating the sensitivity and specificity of the devices across thresholds (dose difference for the ion chamber and percent of pixels passing 3%/3 mm gamma for film and MapCheck). The area under the curve (AUC) is shown for each device.

Table 1

Sensitivity and specificity (including standard deviation) of institutional intensity-modulated radiotherapy (IMRT) quality assurance (QA) results compared with IROC Houston phantom results. “All results” includes all IMRT QA devices and analysis techniques, whereas “device” and “mode” list results for IMRT QA conducted with particular devices or using a particular analysis method (for planar devices).

Results	No. of results	Sensitivity (standard deviation), %	Specificity (standard deviation), %
All results			
Institution claim	855	2 (1)	99.6 (0.2)
Evaluated by IROC H	745	18 (4)	91 (1)
Device			
Ion chamber + planar	91	54 (14)	79 (5)
Ion chamber	325	25 (6)	90 (2)
Film	71	33 (16)	82 (5)
MapCheck	322	14 (5)	94 (2)
Mode			
Absolute	295	3 (3)	94 (1)
Relative	97	21 (9)	91 (3)