Comparison of DVH data from multiple radiotherapy treatment planning systems

2010 Phys. Med. Biol. 55 N337

(http://iopscience.iop.org/0031-9155/55/11/N04)

Phys. Med. Biol. **55** (2010) N337–N346 doi:10.1088/0031-9155/55/11/N04

NOTE

Comparison of DVH data from multiple radiotherapy treatment planning systems

M A Ebert[1,2,8], A Haworth[3,4], R Kearvell[1], B Hooton[1], B Hug[5], N A Spry[1,6], S A Bydder[1,7] and D J Joseph[1,7]

[1] Department of Radiation Oncology, Sir Charles Gairdner Hospital, Perth, Western Australia, Australia
[2] School of Physics, University of Western Australia, Perth, Western Australia, Australia
[3] Department of Physical Sciences, Peter McCallum Cancer Centre, Melbourne, Victoria, Australia
[4] School of Applied Sciences, RMIT University, Melbourne, Victoria, Australia
[5] Medical Imaging and Applied Physics, Curtin University of Technology, Bentley, Western Australia, Australia
[6] School of Medicine and Pharmacology, University of Western Australia, Perth, Western Australia, Australia
[7] School of Surgery and Pathology, University of Western Australia, Perth, Western Australia, Australia

E-mail: Martin.Ebert@health.wa.gov.au

Received 1 December 2009, in final form 1 April 2010
Published 12 May 2010
Online at stacks.iop.org/PMB/55/N337

Abstract

This study examined the variation of dose–volume histogram (DVH) data sourced from multiple radiotherapy treatment planning systems (TPSs). Treatment plan exports were obtained from 33 Australian and New Zealand centres during a dosimetry study. Plan information, including DVH data, was exported from the TPS at each centre and reviewed in a digital review system (SWAN). The review system was then used to produce an independent calculation of DVH information for each delineated structure. The relationships between DVHs extracted from each TPS and independently calculated were examined, particularly in terms of the influence of CT scan slice and pixel widths, the resolution of dose calculation grids and the TPS manufacturer. Calculation of total volume and DVH data was consistent between SWAN and each TPS, with the small discrepancies found tending to increase with decreasing structure size. This was significantly influenced by the TPS model used to derive the data. For target structures covered with relatively uniform dose distributions, there was a significant difference between the minimum dose in each TPS-exported DVH and that calculated independently.

(Some figures in this article are in colour only in the electronic version)

[8] Author to whom any correspondence should be addressed.

1. Introduction

Dose–volume histograms (DVHs) are widely used in radiotherapy to summarize complex three-dimensional (3D) dose distributions for quantitative treatment plan analysis (Chen 1988). Typically during multicentre clinical trials, data will be sourced from multiple treatment planning systems (TPSs). There is a temptation to assume that collated DVH data are accurate and consistent across the TPSs used to derive it. However, there are multiple sources of uncertainty that can lead to differences in calculated, reported and/or exported DVH data. The use of digital imaging and computation in treatment planning demands that discrete data be used which can lead to variations in image, dose calculation and DVH calculation resolution (Chung *et al* 2006, Panitsa *et al* 1998). The interpolation and calculation algorithms used in any individual TPS may also contribute to variations in actual DVH values.

In order to determine whether or not independent DVH calculation on collected trial data is necessary, the present study used clinically relevant data sets obtained from commercial TPSs in current common use to quantify the variation that may be expected when collecting DVH data. It was also aimed at determining if specific parameters of the treatment planning and data export processes introduced any systematic variation into exported DVH data.

2. Method and materials

2.1. Data source

Exported treatment plans were available from a multicentre dosimetric study employing an anthropomorphic phantom (Ebert *et al* 2008a). The pelvic phantom, incorporating variable-density organs which could be distinguished on computed tomography (CT) images, was treated according to prostate and/or rectal protocols at multiple centres across Australia and New Zealand. At each of 33 centres, the phantom was CT scanned and treatment plans developed (including delineation of organs). 36 treatment plans were collected across the centres, exported from the TPS at each centre in either the format of the Radiation Therapy Oncology Group (RTOG format) (Baxter *et al* 1982), and/or the radiotherapy extension of the Digital Imaging and Communication in Medicine standard (DICOM-RT format) (NEMA 2001). The representation of TPS manufacturers in the data is shown in table 1. The methods for volume and DVH calculation used by the TPS models investigated here are summarized in table 2.

2.2. Data review and independent DVH calculation

The DVHs in exported plans were reviewed using the SWAN system[9] (Ebert *et al* 2008b). SWAN is a non-commercial software system developed for clinical trial treatment plan review and data analysis, facilitating access to the DVH data that were exported from each TPS. SWAN allows an independent calculation of DVHs for all structures defined in the export using a procedure based on a previously reported method (Straube *et al* 2005). 3D elements describing delineated structures are assumed to have width in the transverse imaging plane equal to the sum of the half-distance to adjacent axial image slices. The image set volume is defined by a grid of voxels sampled at a resolution specified by the user. The centre of each voxel is examined and if found to lie within a structure, is included in the volume for that structure. The dose value at the centre of the cube is interpolated in 3D from the dose grid. Although inherently inefficient, this is a robust method, suitable for concave and bifurcated

[9] Results presented here are relevant to SWAN version 2.0.3.

Table 1. Treatment planning systems (TPSs) from which planning data were collected.

Manufacturer	Model	Versions	Number in study	CT slice thickness (mm)	CT pixel widths (mm)	Dose grid voxel width (mm)	DVH dose resolution (cGy)
Elekta (Stockholm, Sweden)	XiO	4.1 to 4.33.02	9	2.5–3.0	0.80–0.89	2.5–3.0	10–15
Nucletron (Veenendaal, The Netherlands)	Plato	2.6.3	2	3.0	0.98	1.95	6–11
	Theraplan Plus[a]	2.7.4 3.8	1	3.0	0.94	5.0	70
Philips (Guildford, Surrey, UK)	Pinnacle	6.2b to 8.0 d	13	2.5–5.0	0.78–1.25	3.5–4.0	5–15
Varian (Palo Alto, CA)	Eclipse	7.1.67 to 7.3.10	8	2.5–3.0	0.86–0.98	2.5	1–15

[a] Previously manufactured by Theratronics.

Table 2. Volume and DVH calculation algorithms used in the investigated TPSs.

TPS	DVH calculation	Reference
XiO	Regular sampling on a user-defined grid with interpolation from the dose grid. User-defined bin size	Product manuals and manufacturer representative
Pinnacle	Regular sampling on a user-defined dose grid. User-defined bin size	Product manuals, manufacturer representative and (Bedford *et al* 2003)
Theraplan plus	Regular sampling on a user-defined dose grid. User-defined bin size	Product manuals provided by manufacturer and (Mangili *et al* 2004)
Eclipse	Regular sampling on a user-defined grid based on interpolated structure outline. User-defined bin size	Product manuals provided by manufacturer
Plato	Random sampling on a user-defined dose grid. User-defined bin size	Manufacturer representative
SWAN	Regular sampling on a user-defined grid with tri-linear interpolation from the dose grid. User-defined bin size	

structures. It is not necessarily correct to suggest that the DVH data from any TPS-export or that calculated in SWAN are more 'accurate' since there is scope for ambiguity in the interpretation of a treatment plan on the original TPS or as represented in a plan export. In the context of this study, SWAN simply provides independent DVH calculation to act as a reference for comparison of DVHs from different TPSs.

The DVHs stored in TPS plan exports were validated against hardcopies from contributing centres by visual inspection for any gross differences. For each of the 36 treatment plans, DVHs for the delineated structures were calculated in SWAN using a spatial sampling resolution which exceeded the minimum pixel size in exported image sets (<0.7 mm). Both the TPS and SWAN DVH values for each structure for each treatment plan were exported to text files for analysis. DVHs for a total of 227 structures, ranging in volume from approximately

2 cm^3 to 1500 cm^3, were obtained from the 36 treatment plans, for comparison against the 227 corresponding DVHs calculated with SWAN.

2.3. Analysis methods

To provide a measure of agreement for each exported DVH relative to the independent calculation in SWAN, a comparison method was developed similar to gamma analysis used for spatial comparison of dose distributions (Low *et al* 1998). The gamma analysis method is commonly used to provide a combination of direct comparison of dose differences and distance-to-agreement between measured and calculated dose distributions. For application to DVH comparison, every point in a TPS-derived DVH is compared to every point in the corresponding SWAN-derived DVH to see if points fall within acceptable limits of each other. Acceptance criteria levels (volume difference on the vertical axis or dose-to-agreement on the horizontal axis) were selected to compare each cumulative DVH point. The DVH of highest resolution was selected as the reference. A gamma value is obtained for each DVH point i (with dose d_i and volume v_i) with a volume-difference criterion of $\Delta V_R\%$ (as a percentage of the total structure volume, V_{Tot}) and dose-to-agreement criterion of $\Delta D_R\%$ (as a percentage of the maximum DVH dose, $D_{\max}$) via

$$\gamma_i = \min\{\Gamma[(d_i, v_i), (d_r, v_r)]\} \; \forall \; \{r = 1..P\}, \tag{1}$$

where there are P bins (each rth bin having absolute dose d_r and volume v_r) in the reference DVH, and where

$$\Gamma[(d_i, v_i), (d_r, v_r)] = \left[\left(\frac{100 \cdot (v_r - v_i)}{\Delta V_R \cdot V_{\text{Tot}}}\right)^2 + \left(\frac{100 \cdot (d_r - d)_i}{\Delta D_R \cdot D_{\max}}\right)^2\right]^{1/2}. \tag{2}$$

A value of $\gamma_i < 1$ indicates agreement for the DVH bin i. Figure 1 shows sample DVHs with γ value variations across all dose bins and an indication of the percentage failing to meet $\gamma_i < 1$.

The measures used to evaluate agreement between exported DVHs and the independent calculation were as follows.

- The variation in percentage of DVHs meeting the γ criteria for all of their dose bins (examined for multiple values of ΔV_R and ΔD_R).
- Percentage agreement in absolute total volume (relative to SWAN) for each DVH. Structure volume is obtained from the cumulative DVH value at zero dose, allowing comparison of the TPS and SWAN volume estimates for the 227 delineated structures for which complete data were available.
- Percentage failing gamma (γ) criteria—for each DVH, this is the percentage of bins failing the γ criteria (i.e. with $\gamma_i < 1$ for specific values of ΔV_R and ΔD_R).

For structures covered by relatively uniform dose distributions, the γ evaluation would tend to underestimate the influence of DVH differences, since typically only the small number of DVH points near the uniform dose value will yield γ values indicating discrepancy. This would be particularly relevant for target volumes (i.e. CTV, GTV, PTV (ICRU 1999)), where in the context of models based on a purely parallel structure, dose response may be sensitive to the minimum dose (Sanchez-Nieto and Nahum 1999). As such, for target volumes, the differences in the minimum, mean and maximum dose extracted from exported DVHs and SWAN-calculated DVHs were also separately examined.

Multivariate regression with backward elimination of independent variables (TPS manufacturer, CT slice thickness, CT pixel size, dose grid resolution, DVH dose bin size)

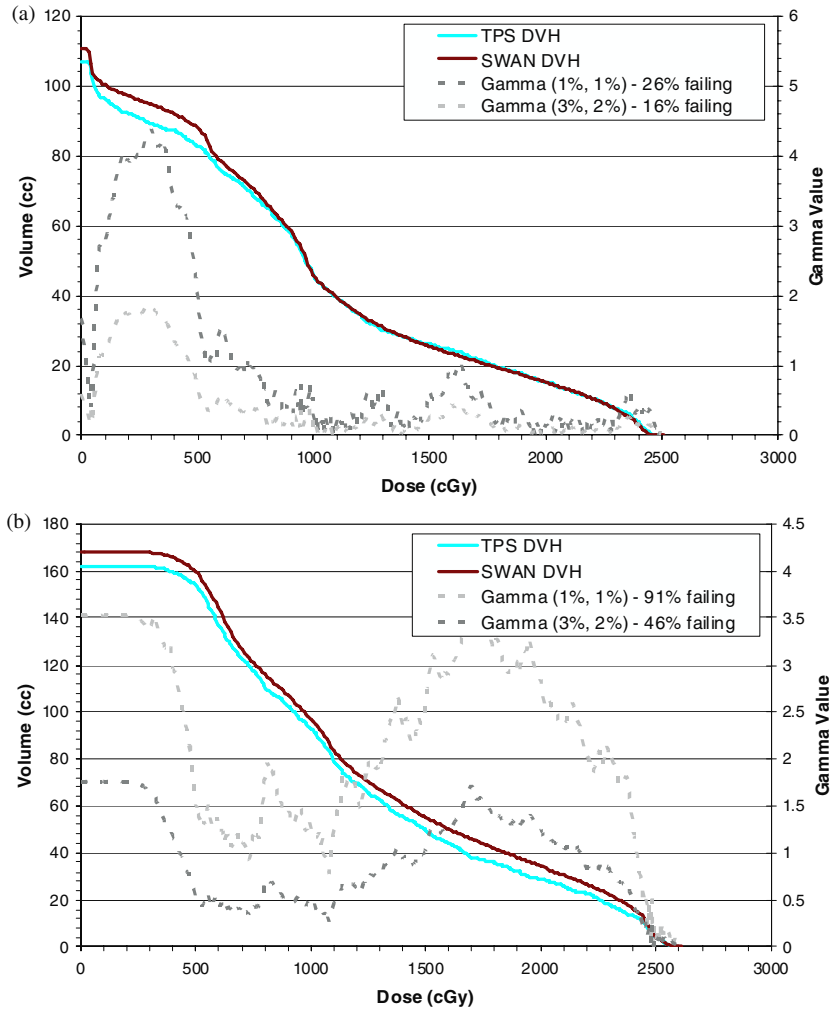

Figure 1. Examples of DVHs from both the TPS and SWAN, from the same plan, for the same structure, together with the resulting γ value distributions for two combinations of acceptance criteria (ΔV_R, ΔD_R) as shown.

was undertaken with SPSS 17.0 (SPSS Inc., Chicago, IL) to determine the factors influencing volume differences and gamma criteria comparison. Interaction between all terms was also investigated. Values of $\Delta V_R = 1\%$ and $\Delta D_R = 1\%$ were found to provide a sensitive indication of DVH mismatch across the data sets examined (whilst maintaining ΔD_R above the DVH dose resolution) and were used in the regression analysis.

3. Results

Table 3 indicates the overall agreement between exported DVHs and DVHs independently calculated in SWAN for a variety of γ test agreement criteria. The comparison shows the percentage of the 227 DVHs where at least 95% of DVH bins had $\gamma < 1$.

Table 3. Relationship of TPS-exported DVHs to independently calculated DVHs across all structures from all examined TPS exports on the basis of γ value for variation in agreement criteria. The values indicate, at each combination of (ΔV_R, ΔD_R), the percentage of all DVHs with at least 95% of $\gamma_i < 1$.

ΔD ΔV	1%	2%	5%	10%
1%	78%	82%	85%	88%
2%	88%	89%	93%	94%
5%	96%	99%	99%	99%
10%	100%	100%	100%	100%

In figure 2(a), the influence of absolute structure volume on SWAN/TPS agreement is shown. Multiple regression analysis indicated that TPS manufacturer was the only significant predictor for percentage volume difference ($p = 0.011$). As an example, the percentage of DVH samples failing the γ criteria over more than 10% of points (at $\Delta V_R = 1\%$ and $\Delta D_R = 1\%$) was Xio—0%; Plato—29%; Pinnacle—2%; Theraplan Plus—14%; and Eclipse—13%. The percentage of points failing gamma criteria (figure 2(b)) was dependent on dose grid resolution ($p < 0.001$; univariate regression coefficient -15% mm^{-1}), TPS manufacturer ($p < 0.02$) and CT pixel size ($p < 0.03$; univariate regression coefficient +57% mm^{-1}), although there was significant correlation between the variables TPS manufacturer and dose grid resolution, and TPS manufacturer and CT pixel size.

Comparison of differences in the minimum, mean and maximum dose between the TPS-exported DVHs and independently calculated DVHs for the 50 target volumes across the 25 prostate treatment plans indicated excellent agreement between mean and maximum doses, as shown in figure 3. However, there were significant differences ($p = 0.0001$) between values for the minimum dose with a mean difference of -0.9% and the absolute standard deviation in this difference of 1.5%.

4. Discussion

This study demonstrated that the calculation of total volume and DVH data was consistent between SWAN and each TPS, with the small discrepancies found tending to increase with decreasing structure size. This was significantly influenced by the TPS model used to derive the data. The variety of commercial TPS models examined in this study (see table 1) reflects the situation in radiation oncology departments across Australia and New Zealand. The advantage of utilizing data from a dosimetric intercomparison with a single phantom was the comparison of plan exports based on a single geometry together with well-defined planning protocols. The data were also collected under controlled conditions (all study visits for collection of this data were performed in person by two members from a pool of four investigators). This enabled extraction of data using the review system, consistent with that obtained at the source TPS. The disadvantage was the limited number of resulting data sets. As such, especially for the TPSs represented by only one or two samples, the general applicability of these results for each specific model of TPS should not be assumed.

From table 3 it is apparent that satisfaction of the γ evaluation is more sensitive to volume criteria than dose criteria. A principal reason for this is the presence of target volumes in the sample group where relatively uniform dose delivery led to largely rectangular cumulative DVHs. For these, a difference (between the export and independent calculation) in percentage

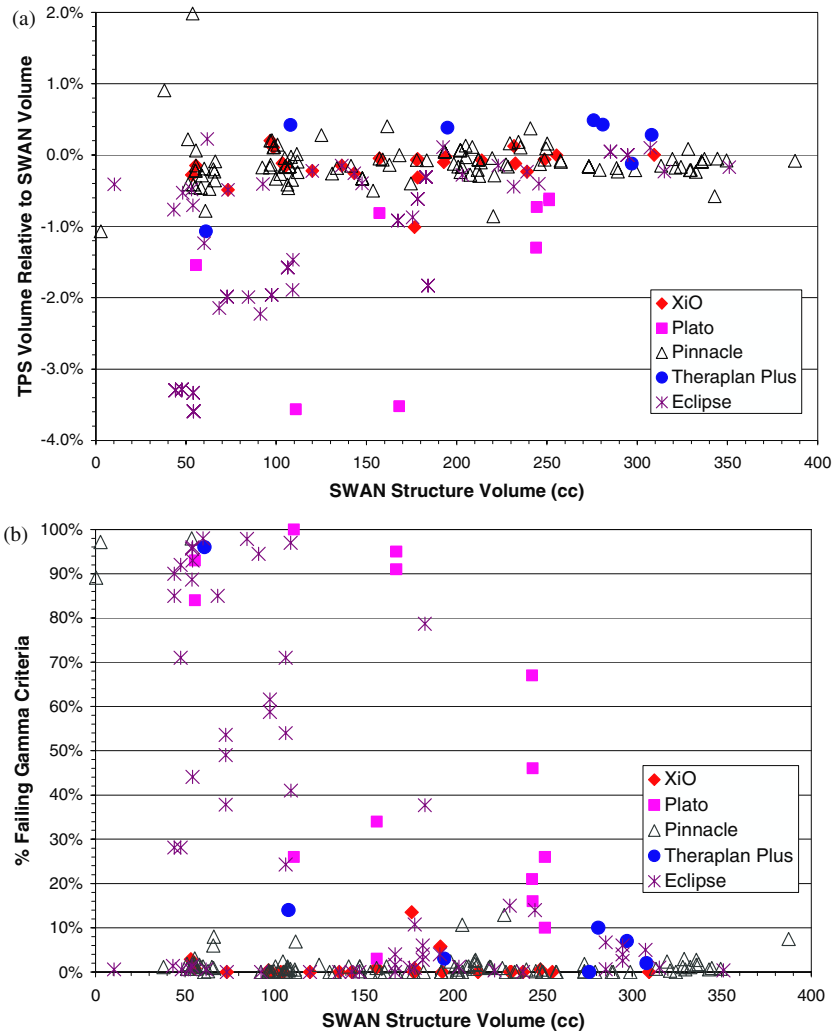

Figure 2. Comparison of (a) volume and (b) DVH calculation (γ criteria for $\Delta V_R = 1\%$ and $\Delta D_R = 1\%$) for all structures according to the structure volume and planning system model. (Values for larger volumes not shown although relative volume differences were $< 0.1\%$ and % failing γ criteria were $< 5\%$.)

total volume for a structure outside the ΔV_R criteria would lead to almost 100% failure in γ evaluation.

DVH information exported by each TPS is seen to vary relative to that generated by the SWAN review system as volumes become smaller (particularly below approximately 250 cm^3), with a dependence on the TPS of origin as shown in figure 2. By using the SWAN-calculated DVHs as reference, we have been able to quantify not necessarily the absolute accuracy of DVH calculation, but the degree of consistency between DVH calculations from multiple commercial TPSs. The similarity of the DVH calculation algorithms used by SWAN, XiO and Pinnacle (see table 2) is consistent with the agreement between these systems as shown in figure 2. The influence of the TPS model and export parameters on DVH calculation suggests that independent DVH calculation is useful for providing consistency and reproducibility

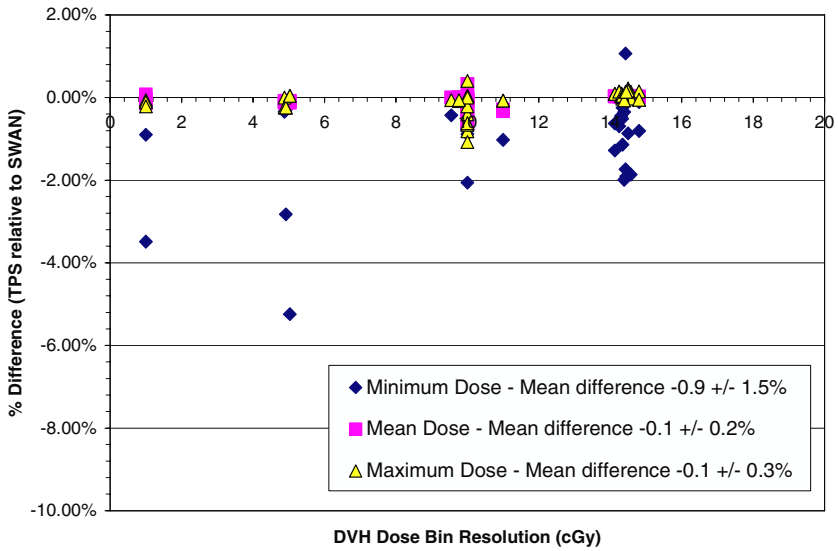

Figure 3. Differences in the minimum, mean and maximum dose for TPS-exported DVHs and SWAN-calculated DVHs for 50 different target volumes extracted from 25 separate prostate treatment plans (TPS relative to SWAN). The values are plotted against the dose resolution of the TPS DVH.

across collected data sets, reducing variability and bias in dose–volume quantities. This is of particular relevance in the processing of DVH data from multicentre clinical trials.

It should be noted that the percentage volume differences highlighted in figure 2(a) represent relatively small discrepancies in calculated volume. If all 227 structures examined are considered as equivalent spheres (i.e. spheres having the same volume as each structure), then the SWAN and TPS equivalent spheres agree to within a radius of 0.5 mm (less than 2% difference in effective radius between SWAN and each TPS).

The DVH analysis results indicate a dependence on export parameters, although correlations of variables (in particular, TPS model and CT characteristics) inhibit identification of specific trends. For example, DVH calculation (as evaluated with gamma analysis) would be expected to agree more consistently as the dose grid size decreases. A negative regression coefficient was obtained, however, indicating a contradictory trend (DVH agreement *increasing* with *increasing* grid size), and the most likely explanation for this is the limited range of grid sizes sampled in this study and the correlation of grid size with TPS manufacturer. In table 1 for example, it is seen that all results for the Plato TPS were obtained at the highest dose grid resolution. As opposed to other systems (see table 2), Plato uses random sampling for DVH calculation which, on a limited number of points, can cause discrepancy (Lu and Chin 1993). Similarly, greater agreement would be expected with decreasing CT slice and pixel sizes, in contrast to what was observed in this study.

Niemierko and Goitein (1990) have shown that random sampling can be used with much greater efficiency than regular sampling in a way which avoids spatial correlation issues, which have been examined previously (van t' Veld and Bruinvis 1995). For iterative inverse methods, where computational efficiency is mandatory, small random fluctuations in parameter estimates can be tolerated and Niemierko and Goitein (1990) have shown that the effects are more than tolerable when translated to model-based cost functions such as complication probabilities.

The precision of DVH calculation is inherently limited by the resolution of the dose and delineation grids used in TPSs. In theory, as grid resolution decreases (i.e. the voxel size increases), errors due to partial-volume effects on a regular grid are proportional to the square of each grid dimension (Shidong *et al* 1997), with this relationship occasionally complicated by spatial correlation of pseudo-regular structures with the sampling grid (Lu and Chin 1993, van t' Veld and Bruinvis 1995).

It is interesting to note that in figure 3 almost all TPS-exported DVHs gave target volume minimum doses that were greater than that independently calculated in SWAN. For such volumes, the minimum dose is typically located at the periphery of the outlined structure. As such, the determination of the minimum dose will be dependent on factors such as whether or not the TPS automatically extends volumes an extra half image-slice for inclusion in DVH calculation (Ackerly *et al* 2003), the size and coincidence of the dose grid with image planes, and the interpolation methods used at the periphery of the structure. The DVH calculation in SWAN 2.0.3 does include half image-slice regions outside the superior and inferior limits of each volume. Information obtained from manufacturers regarding the same effect in their TPSs was not consistently clear.

The reliability of TPS-derived DVH data collected in multicentre trials depends on the accuracy required in relating DVH constraints (or model parameters) to outcomes. Uncertainty in derivation of dose–volume parameters, related to measurement of treatment outcomes, adds to the multitude of uncertainties associated with collection and analysis of trial data (Deasy *et al* 2002, Pettersen *et al* 2008). The influence of geometric and dosimetric uncertainties should be considered relative to the criteria used for DVH comparison here—1% dose and 1% volume. Uncertainties in level I dosimetry (i.e. accelerator output) are of this order (Bentzen *et al* 2000, Ebert *et al* 2008c) and the addition of patient geometries and shifts from reference conditions increase dosimetric uncertainty. Patient motion and anatomical changes, inter-observer variability in delineation and potentially vague demarcation of tumours and/or their functional distributions lead to uncertainties that far exceed 1% in volume (Hamilton and Ebert 2005), suggesting that the agreement between DVHs shown in this study is acceptable within the noise level that is inherent in collected trial data. From table 3, we can expect volume and dose calculation within 5% and 2% respectively for DVH data collected under the range of conditions tested in this study (as detailed in table 1).

Acknowledgments

This work was supported by the Cancer Foundation of Western Australia. This research received funding from Cancer Australia and the Diagnostics and Technology Branch of the Australian Government Department of Health and Ageing. We are grateful to the Trans-Tasman Radiation Oncology Group, the 'Elvis' dosimetry study team and all radiation oncology centres that supplied TPS data.

References

Ackerly T, Andrews J, Ball D, Guerrieri M, Healy B and Williams I 2003 Discrepancies in volume calculations between different radiotherapy treatment planning systems *Aust. Phys. Eng. Sci. Med.* **26** 91–3

Baxter B S, Hitchner L E and Maguire G Q Jr 1982 *A Standard Format for Digital Image Exchange* (Madison, WI: AAPM)

Bedford J L, Childs P J, Nordmark Hansen V, Mosleh-Shirazi M A, Verhaegen F and Warrington A P 2003 Commissioning and quality assurance of the Pinnacle(3) radiotherapy treatment planning system for external beam photons *Br. J. Radiol.* **76** 163–76

Bentzen S M, Bernier J, Davis J B, Horiot J C, Garavaglia G, Chavaudra J, Johansson K A and Bolla M 2000 Clinical impact of dosimetry quality assurance programmes assessed by radiobiological modelling of data from the thermoluminescent dosimetry study of the European Organization for Research and Treatment of Cancer *Eur. J. Cancer* **36** 615–20

Chen G T Y 1988 Dose volume histograms in treatment planning *Int. J. Radiat. Oncol. Biol. Phys.* **14** 1319–20

Chung H T, Jin H, Palta J, Suh T S and Kim S 2006 Dose variations with varying calculation grid size in head and neck IMRT *Phys. Med. Biol.* **51** 4841–56

Deasy J O, Niemierko A, Herbert D, Yan D, Jackson A, Ten Haken R K, Langer M and Sapareto S 2002 Methodological issues in radiation dose–volume outcome analyses: summary of a joint AAPM/NIH workshop *Med. Phys* **29** 2109–27

Ebert M A, Harrison K, Cornes D, Howlett S, Joseph D J, Kron T, Hamilton C S and Denham J W 2008a A comprehensive Australasian multi-centre dosimetric intercomparison—issues, logistics and recommendations *J. Med. Imaging Radiat. Oncol.* **86** 200–10

Ebert M A, Haworth A, Kearvell R, Hooton B, Coleman R, Spry N A, Bydder S and Joseph D J 2008b Detailed review and analysis of complex radiotherapy clinical trial planning data: evaluation and initial experience with the SWAN software system *Radiother. Oncol.* **86** 200–10

Ebert M A, Howlett S J, Harrison K, Cornes D, Hamilton C S and Denham J W 2008c Linear-accelerator x-ray output: a multicentre chamber-based intercomparison study in Australia and New Zealand *Aust. Phys. Eng. Sci. Med.* **31** 255–66

Hamilton C S and Ebert M A 2005 Volumetric uncertainty in radiotherapy *Clin. Oncol.* **17** 456–64

ICRU 1999 *Prescribing, Recording and Reporting Photon Beam Therapy* (Bethesda, MD: International Commission on Radiological Units)

Low D A, Harms W B, Mutic S and Purdy J A 1998 A technique for the quantitative evaluation of dose distributions *Med. Phys.* **25** 656–61

Lu X Q and Chin L M 1993 Sampling techniques for the evaluation of treatment plans *Med. Phys.* **20** 151–61

Mangili P, Stea L, Cattani F, Lappi S, Giglioli F, Calamia E, Ziglio F, Martinelli R and Longobardi B 2004 Comparative study of permanent interstitial prostate brachytherapy post-implant evaluation among seven Italian institutes *Radiother. Oncol.* **71** 13–21

NEMA 2001 *Digital Imaging and Communications in Medicine (DICOM) Standard* (Washington, DC: National Electrical Manufacturers Association)

Niemierko A and Goitein M 1990 Random sampling for evaluating treatment plans *Med. Phys.* **17** 753–62

Panitsa E, Rosenwald J C and Kappas C 1998 Quality control of dose volume histogram computation characteristics of 3D treatment planning systems *Phys. Med. Biol.* **43** 2807–16

Pettersen M N, Aird E and Olsen D R 2008 Quality assurance of dosimetry and the impact on sample size in randomized clinical trials *Radiother. Oncol.* **86** 195–9

Sanchez-Nieto B and Nahum A E 1999 The delta-TCP concept: a clinically useful measure of tumor control probability *Int. J. Radiat. Oncol. Biol. Phys.* **44** 369–80

Shidong L, Arthur B, Yong L and George T Y C 1997 Analysis of the dose–surface histogram and dose–wall histogram for the rectum and bladder *Med. Phys.* **24** 1107–16

Straube W, Mathews J, Bosch W R and Purdy J A 2005 DVH analysis: consequences for quality assurance of multi-institutional clinical trials *Med. Phys.* **32** 2021–2

van t' Veld A A and Bruinvis I A 1995 Influence of shape on the accuracy of grid-based volume computations *Med. Phys.* **22** 1377–85