Journal of Physics: Conference Series **489** (2014) 012090 doi:10.1088/1742-6596/489/1/012090

Derivation and representation of dose-volume response from large clinical trial data sets: an example from the RADAR prostate radiotherapy trial

MA Ebert[1,2], K Foo[3,4], A Haworth[5,6], SL Gulliford[7], R Kearvall[1], A Kennedy[1], S Richardson[1], M Krawiec[1], N Stewart[1], DJ Joseph[1,8] and JW Denham[9]

[1] Department of Radiation Oncology, Sir Charles Gairdner Hospital, Perth, Austral
[2] School of Physics, University of Western Australia, Perth, Australia
[3] Radiation Oncology, Sydney Cancer Centre, Sydney, Australia
[4] University of Sydney, Sydney, Australia
[5] Department of Physical Sciences, Peter MacCallum Cancer Centre, Melbourne, Australia
[6] University of Melbourne, Melbourne, Australia
[7] Joint Department of Physics, Institute of Cancer Research and Royal Marsden National Health Service Foundation Trust, Sutton, UK
[8] School of Surgery, University of Western Australia, Perth, Australia
[9] University of Newcastle, Newcastle, Australia

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

Abstract. Large multicentre radiotherapy trials incorporating assessment of multiple outcomes at multiple timepoints can generate extensive datasets. We have investigated graphical techniques for presentation of this data and the associated underlying dose-volume response information, necessary for guiding statistical analyses and translating outcomes to future patient treatments. A relational database was used to archive reviewed plan data for patients accrued to the TROG 03.04 RADAR trial. Viewing software was used to clean and enhance the data. Scripts were developed to export arbitrary dose-histogram data which was combined with clinical toxicity data with a median follow-up of 72 months. Graphical representations of dose-volume response developed include prevalence atlasing, univariate logistic regression and dose-volume-point odds ratios, and continuous cut-point derivation via ROC analysis. These representations indicate variable association of toxicities across structures and time-points.

1. Introduction

We have been investigating the relationship between dose-volume data and clinical outcome for the rectum for patients accrued to the TROG 03.04 'RADAR' prostate radiotherapy trial. In this context, we have developed a series of representations of the data that can provide graphical and statistically informative representations of those relationships. The results are graphical displays to demonstrate the intricate relationships between dose, volume, anatomical definition, specific toxicity and toxicity grade.

2. Description

2.1. Trial data collection

The RADAR trial focused on the influence of duration of androgen deprivation on prostate cancer patients and included an external beam radiotherapy (EBRT) dose-escalation component (two-phase treatment to 66, 70, 74 or 78 Gy) and optional HDR brachytherapy. Of 1071 participants accrued from 24 contributing centres, 813 had EBRT alone and of these, 754 complete planning data sets were retrieved from the participating centres' planning systems and archived. Descriptions of the technical aspects of the RADAR protocol, assessed toxicities and principal outcomes have been given elsewhere [1, 2].

A software platform ('SWAN') was developed to facilitate review and processing of participant planning data [3]. SWAN allows visual review and manipulation of most plan objects, automated review of multiple protocol items, reporting back to submitting centres and archive of plan objects into a relational database. SWAN was used to review and archive all data presented here. The RADAR protocol required delineation of rectum from 1 cm below the GTV to where the rectum turns horizontally into the sigmoid colon. This definition was poorly adhered to and most rectums were manually corrected or re-delineated by our team. Rectal contours were also extended inferiorly to the base of the ischial tuberosities to create an 'anorectum' structure.

2.2. Data processing and extraction

Histogram calculation, export and link to clinical outcomes. Java scripts were developed to apply SWAN's independent dose-volume histogram (DVH) and dose-surface histogram (DSH) algorithms to archived data. The DVH algorithm calculates the exact area of dose-voxels encompassed by joined delineation points for a structure. DVH and DSH data were automatically calculated at 1 cGy resolution and exported for all archived patient plans for rectum and anorectum. A script was also developed to separate out the lowest 3 cm of the anorectum, defining an 'anal canal', for which DVH/DSH data was exported. Prior to histogram calculation, the multiphase RT data were combined to form an equivalent dose in 2 Gy fractions (EQD_2) based on α/β = 3 Gy. Dose-volume data was linked (via trial ID) to all patient outcomes data, exported from the clinical database in November 2012. Median follow-up time was 72 months (maximum 108 months). The data were collated as an array (one element per patient) of structures in Matlab 2012b (Mathworks, Natick MA), which was subsequently used to code analysis. Peak grade across the 'late' follow-up period ≥ 12 months was determined for each toxicity.

2.3. Derivation and representation of dose-volume response

2.3.1. Prevalence – DVH atlasing. The QUANTEC report on normal tissue effects in radiotherapy detailed the advantages of completeness of presentation of data derived from clinical studies [4]. A dose-volume atlas demonstrates, for a given anatomical structure the number of patients whose DVH/DSH passes through a discrete element ('pixel') of dose (or EQD_2)/volume space, and the associated proportion of those patients experiencing an 'event' (a specific toxicity grade definition at a specific time-point or longitudinally across timepoints), and can thus be used to present dose-outcomes data to facilitate meta-analyses. This technique was initially investigated by Jackson et al [5], who previously demonstrated a variety of characteristics of dose-volume data that can be related to patient rectal toxicities [6]. The 'atlas of complication incidence' has since been used for visual association of parameters with trial outcomes [7-9]. Our implementation utilises pixels of dimensions 5 Gy EQD_2 and 5% relative volume/surface area, each coloured according to the fraction of events/patients with a histogram that passes through that pixel. Text values are added to each pixel, being the number of

events over the number of histograms passing through the pixel. Alternative representations can be made based on various descriptive statistics [5] or alternative transformations of dose [7].

Figure 1 illustrates atlases for rectal bleeding incidence of grade ≥ 2, as recorded at trial time-point 36 months post-randomisation, and as summarised as the peak grade each patient experiences over the entire follow-up period. With the values tabulated in the plot, a cumulative incidence score as presented by Jackson et al [5] can alternatively be derived.

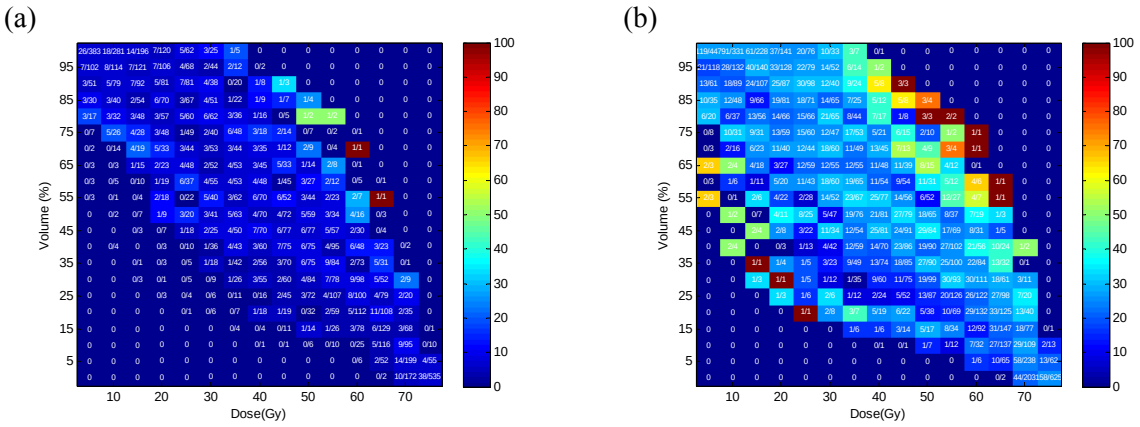

Figure 1. Atlases of prevalence of rectal bleeding of grade ≥ 2 excluding patients with baseline values > 0, at (a) the 36 month time-points, and (b) peak value across whole late (≥ 12 months post radiotherapy) follow-up period. Dose/volume intervals are 5 Gy/5%. Text values indicate event rates in each atlas 'pixel', with the ratio displayed as a colour-scaled percentage, for all patients with follow-up at that time-point and who were not excluded due to baseline toxicity.

2.3.2. Dose-effect – univariate analysis and odds ratios. As pointed out by Jackson *et al* [5], summarising data via an atlas puts it in an excellent form for regression analysis. As a means of illustrating the impact of dose-volume across an atlas, we have combined univariate logistic regression at each EQD_2 interval (arbitrarily selecting 1 Gy intervals), yielding an odds-ratio (OR) for toxicity incidence per % volume increase at each interval, with determination of an OR at each individual dose-volume point for toxicity incidence when the population is dichotomised above and below each volume point at each EQD_2 interval. An alternative atlas summarising this OR-distribution can be generated with OR on a colour scale. The dichotomous OR is considered significant if the lower 95% confidence interval (CI) is greater than 1.0. This can be indicated (in this case, a grey outline) for pixels where the OR is significant. Samples, again for grade > 1 rectal bleeding, are shown in Figure 2. The OR calculation routine developed by Cardillo [10] was used.

2.3.3. Volumetric cutpoints. The ability of specific volumes to discriminate events at each EQD_2 interval can be used to establish volumetric cut-points using the receiver operator characteristic (ROC) curve, derived using volume as the independent predictor [11-13], and assessed using the area under the curve (AUC). A significance level for a cutpoint can be estimated by re-sampling methods (see on-line resources supplementary to this publication). Sample distributions of resulting cutpoints as a function of EQD_2 from the RADAR trial data are shown in Figure 2 for peak late grade of rectal bleeding for grade ≥ 2. Given expected co-linearity of derived cut-points, p-values have been adjusted for multiple testing according to the Holm-Bonferoni step-down method [14]. Derivation of optimal cutpoints using a maximally-selected test statistic has also previously been demonstrated by Buettner et al [15].

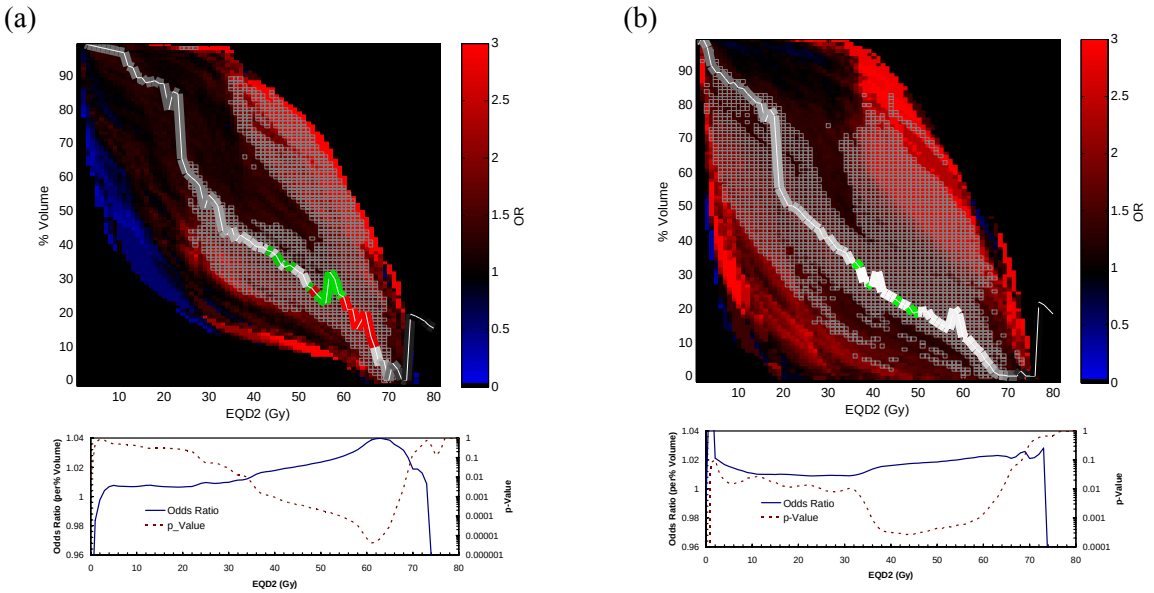

Figure 2. Results of univariate logistic regression, dichotomous OR and optimal cutpoint for rectal bleeding, peak grade across whole late follow-up period of ≥ 2. The bottom plot in each figure shows the OR from logistic regression at each EQD_2 interval plotted with the corresponding p-value. The top plot overlays the ROC-derived cutpoint as a function of dose over the dichotomous OR. The OR is shown as a colour scale at each discrete 1 Gy/1% 'pixel' (as shown), with a grey outline when the lower 95% CI is greater than 1.0. The cutpoint derived from the ROC Youden index is shown as a thick line which is colour-scaled according to the ROC AUC or multiple-testing corrected p-value - from black (AUC ≤ 0.5) to white (AUC at its maximum), and then green for $0.05 < p \leq 0.01$, and red for $p \leq 0.01$. (a) For DVH of whole anorectum, and (b) for DVH of anal-canal.

3. Discussion

Multiple factors can influence the observed associations between dose-volume (or dose-surface area) points derived by these mapping techniques, including the distribution of DVH/DSH data at each EQD_2 value. The absence of significant areas of association when randomly-generated data is used (see supplementary online resources) suggests however, that the significant regions observed on the maps and cutpoint derivations are in fact the result of significant associations between dose-volume points and resulting patient toxicities. The issue then remains of what use can be made of that information.

A significant cutpoint establishes a volumetric index significantly separating patients with and without a specific resulting toxicity. By observing the patterns of this significance, as a function of dose, histogram type and anatomical structure, we can drive hypotheses of the underlying response mechanisms and set constraints for histogram-based optimisation of future patient treatment. By comparing ROC-derived cutpoints with the underlying dichotomous OR values (ie., is shown in the plots of Figure 2) we also establish the precision with which a cutpoint is defined. In Figure 2 for example, we see a wide region of significant dichotomous OR values suggesting the resulting cutpoint values are not well defined, with the potential to take on large confidence intervals. Validation of the cutpoints could be undertaken by comparing them with values derived using alternative techniques, and investigation of such techniques is currently underway.

4. Conclusion

DVH atlasing continues to be established as a standard for the presentation of dose-response data and is ideally suited to the analysis presented above. The variation of significance of association between volume and outcome is apparent when comparing graphical representations spanning

anatomical structures (as shown for anorectum and anal canal above) and event definitions. Such representations are now driving derivation of univariate regression and proportional hazards assessment for presentation in more comprehensive publications.

Acknowledgements
We acknowledge funding from the Cancer Council Western Australia (John Nott Fellowship 2012), Cancer Australia and the Diagnostics and Technology Branch of the Australian Government Department of Health and Ageing (grant 501106), the National Health and Medical Research Council (grants 300705, 455521, 1006447), Abbott Laboratories, Novartis Pharmaceuticals. We gratefully acknowledge the support of participating RADAR centers, C Wilcox, the Trans-Tasman Radiation Oncology Group, and past SWAN programmers – J Blight, S Price, B Hooton and E van der Wath.

References
[1] Kearvell R, Haworth A, Ebert MA, Murray J, Hooton B, Richardson S, et al. *JMIRO*. 2013;In press.
[2] Denham JW, Wilcox C, Lamb DS, Spry NA, Duchesne G, Atkinson C, et al. *Radioth Oncol*. 2012;105:184-92.
[3] Ebert MA, Haworth A, Kearvell R, Hooton B, Coleman R, Spry NA, et al. *Radioth Oncol*. 2008;86:200-10.
[4] Jackson A, Marks LB, Bentzen SM, Eisbruch A, Yorke ED, Ten Haken RK, et al. *Int J Radiat Oncol Biol Phys*. 2010;76:S155-S60.
[5] Jackson A, Yorke ED, Rosenzweig KE. *Sem Radiat Oncol*. 2006;16:260-8.
[6] Jackson A, Skwarchuk MW, Zelefsky MJ, Cowen DM, Venkatraman ES, Levegrun S, et al. *Int J Radiat Oncol Biol Phys*. 2001;49:685-98.
[7] Liu F, Yorke ED, Belderbos JSA, Borst GR, Rosenzweig KE, Lebesque JV, et al. *Int J Radiat Oncol Biol Phys*. 2013;85:182-9.
[8] Mutter RW, Liu F, Abreu A, Yorke E, Jackson A, Rosenzweig KE. *Int J Radiat Oncol Biol Phys*. 2012;82:1783-90.
[9] Gulliford SL, Miah AB, Brennan S, McQuaid D, Clark CH, Partridge M, et al. *Radioth Oncol*. 2012;104:205-12.
[10] Cardillo G. Odds: compute odds and risk ratio on a 2x2 matrix. http://www.mathworks.com/matlabcentral/fileexchange/15347; 2007.
[11] Gulliford SL, Partridge M, Sydes MR, Andreyev J, Dearnaley DP. *Radioth Oncol*. 2010;94:241-7.
[12] van der Laan HP, van den Bergh A, Schilstra C, Vlasman R, Meertens H, Langendijk JA. *Int J Radiat Oncol Biol Phys*. 2008;70:1138-45.
[13] Kim B, Park HC, Oh D, Shin EH, Ahn YC, Kim J, et al. *Radiat Oncol J*. 2012;30:43-8.
[14] Westfall PH, Young SS. Resampling-BasedMultiple Testing: Examples and Methods for p-Value Adjustment. New york: Wiley; 1993.
[15] Buettner F, Gulliford SL, Webb S, Sydes MR, Dearnaley DP, Partridge M. *Phys Med Biol*. 2009;54:6535-48.