

OPEN ACCESS

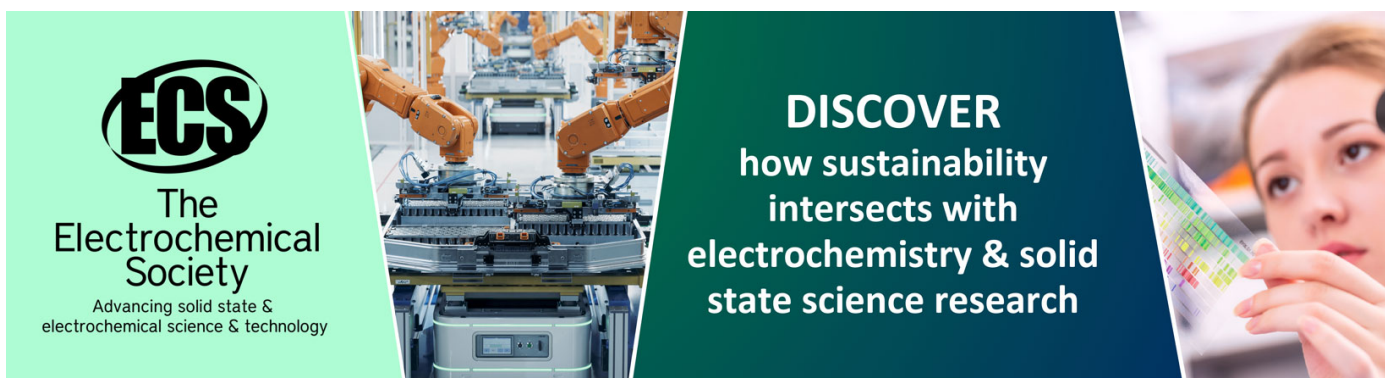
PlanJury: probabilistic plan evaluation revisited

To cite this article: M Witte *et al* 2014 *J. Phys.: Conf. Ser.* **489** 012062

View the [article online](#) for updates and enhancements.

You may also like

- [Deep learning based MLC aperture and monitor unit prediction as a warm start for breast VMAT optimisation](#)
L Vandewinckele, T Reynders, C Weltens et al.
- [Threshold-driven optimization for reference-based auto-planning](#)
Troy Long, Mingli Chen, Steve Jiang et al.
- [Use of knowledge based DVH predictions to enhance automated re-planning strategies in head and neck adaptive radiotherapy](#)
Elisabetta Cagni, Andrea Botti, Agnese Chendi et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

PlanJury: probabilistic plan evaluation revisited

M Witte, J-J Sonke and M van Herk

The Netherlands Cancer Institute, Amsterdam, The Netherlands

E-mail: m.witte@nki.nl

Abstract. *Purpose:* Over a decade ago, the ‘Van Herk margin recipe paper’ introduced plan evaluation through DVH statistics based on population distributions of systematic and random errors. We extended this work for structures with correlated uncertainties (e.g. lymph nodes or parotid glands), and considered treatment plans containing multiple (overlapping) dose distributions (e.g. conventional lymph node and hypo-fractionated tumor doses) for which different image guidance protocols may lead to correlated errors.

Methods: A command-line software tool ‘PlanJury’ was developed which reads 3D dose and structure data exported from a treatment planning system. Uncertainties are specified by standard deviations and correlation coefficients. Parameters control the DVH statistics to be computed: e.g. the probability of reaching a DVH constraint, or the dose absorbed at given confidence in a (combined) volume. Code was written in C++ and parallelized using OpenMP. Testing geometries were constructed using idealized spherical volumes and dose distributions.

Results: Negligible stochastic noise could be attained within two minutes computation time for a single target. The confidence to properly cover both of two targets was 90% for two synchronously moving targets, but decreased by 7% if the targets moved independently. For two partially covered organs at risk the confidence of at least one organ below the mean dose threshold was 40% for synchronous motion, 36% for uncorrelated motion, but only 20% for either of the organs separately. Two abutting dose distributions ensuring 91% confidence of proper target dose for correlated motions led to 28% lower confidence for uncorrelated motions as relative displacements between the doses resulted in cold spots near the target.

Conclusions: Probabilistic plan evaluation can efficiently be performed for complicated treatment planning situations, thus providing important plan quality information unavailable in conventional PTV based evaluations.

1. Introduction

Compensation for geometric uncertainties when planning external beam radiation treatment is traditionally accomplished through the application of margins, expanding the Clinical Target Volume (CTV) to a Planning Target Volume (PTV), or an Organ At Risk (OAR) to a Planning organ-at-Risk Volume (PRV). These surrogate volumes are then used both for plan optimization and plan evaluation purposes. However, such surrogate volumes offer only limited information about the dose that will likely be absorbed in the various tissues. Probability based plan optimization techniques which do not use PTV margins are under development [1]. Furthermore, Monte Carlo plan evaluation techniques based on population statistics of geometric uncertainties have been proposed [2–4]. Recently, ICRU report 83 stated that confidence intervals on absorbed dose should be considered part of Level 2 (*state-of-the-art techniques*) or Level 3 (*optional research-and-development*) reporting whenever possible [5].

Monte Carlo plan evaluation relies on the estimation of dose absorbed in tissues for many



possible geometrical instances. This process is much simplified if the 3D dose distribution is assumed not to change as a result of the geometrical variations, usually valid for deep seated tumors in homogeneous tissue. Then, geometrical instances are approximated as displacements of delineated volumes with respect to the dose distribution. This assumption also implies that patient setup errors and organ motion have an identical effect and can be grouped together.

Uncertainties of shape invariant Regions Of Interest (ROIs) have previously [3] been described using Standard Deviations (SDs) of systematic and random translations and rotations. In this work we extend this description to the case of multiple ROIs [6] and/or multiple dose distributions [7], for which correlated uncertainties may exist. For example, patient alignment based on bony structures could result in correlated systematic errors in multiple tumors, while correction based on a variable tumor position could induce correlated random errors in lymph node positions. Like the SDs, such correlations should be measured from patient series for each imaging and set-up protocol.

2. Methods

2.1. PlanJury

A standalone command-line software tool ‘PlanJury’ was developed for the Microsoft Windows, Oracle Solaris and GNU/Linux operating systems. PlanJury is configured from a MS Windows ini-style file, and will read 3D planned dose and structure data from files exported using the scripting capability of Philips Pinnacle, or from an in-house developed file format. Structures may be read in as contour sets, CT masks or surface triangulations.

A cloud of random points is generated representing a specified ROI. Systematic translational and optional rotational errors are sampled from Gaussian distributions, and for each systematic error N_{frac} random errors. Every volume point is transformed using a systematic error and each of the associated random errors consecutively, dose values are acquired at the transformed locations using trilinear interpolation, and accumulated dose is computed with and/or without α/β correction. The accumulated dose values are summarized in a dose-volume histogram per systematic instance. An efficient computational structure was established by first preparing all necessary transformation matrices, which can then be used by all random volume points. The loop over all ($\approx 10^4$) random points is parallelized using an OpenMP pragma in the C++ language, using a single atomic (i.e. non-parallel) statement to accumulate the DVH values.

From the sampled DVHs per systematic instance, any number of dosimetric parameters may be extracted. For example, by specifying the lines `GTV.DVH.1.Dose = 95%` and `GTV.DVH.1.Volume = 99%` in the ini-file, PlanJury will return the probability of acquiring at least 95% of dose prescription to 99% of the GTV. The lines `GTV.DVH.2.Prob = 90%` and `GTV.DVH.2.Volume = 99%` could be added to also return the dose which is absorbed at 90% confidence in 99% of the GTV. Similarly, confidence levels and/or expectation values of EUD, TCP and NTCP can be computed.

2.1.1. Multiple ROIs PlanJury may combine multiple ROIs into a compound volume. Error standard deviations can be set per sub-volume, and correlation coefficients for similar errors (e.g. systematic translations in the left-right direction) between sub-volumes can be specified. Gaussian errors generated using the Box-Muller transform are converted into correlated samples using a linear transformation based on the spectral decomposition of the covariance matrices (e.g. http://en.wikipedia.org/wiki/Multivariate_normal_distribution).

The way in which dose in the compound volume needs to be scored depends on the clinical relation between the sub-volumes. Two modes of operation are available to work with a compound volume:

Combined. Transformations on the random dots are performed per sub-volume, and dots are combined into a single DVH per systematic instance. This mode can e.g. be used to compute

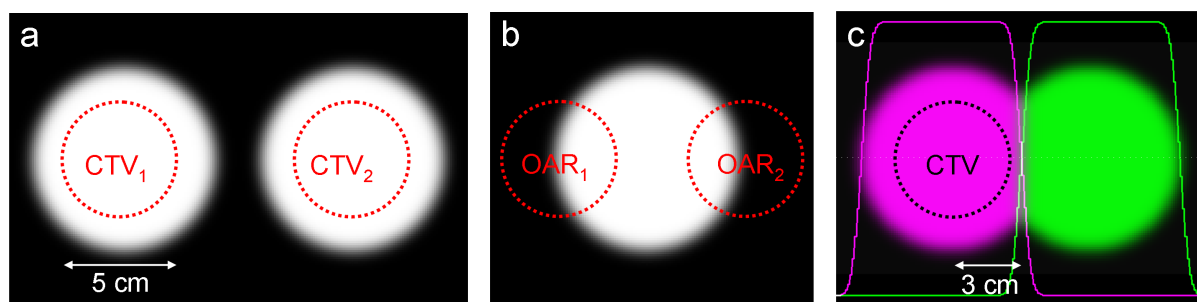


Figure 1. Testing geometries for plan evaluation computations. *a)* Two spherical targets each lie within a perfectly conforming dose distribution with Gaussian penumbra. *b)* Two organs at risk are partially covered with high dose. *c)* Two spherical dose distributions were intersected by a plane and joined to form a continuous dose distribution. Relative displacements between the two doses may result in cold spots near a CTV inside one of the composing distributions. Dose profiles through the center line are indicated.

the minimum dose to either of a number of delineated targets.

Separately. The sampling process results in DVHs per systematic instance for every sub-volume separately. When extracting the dosimetric parameters, a qualifier (e.g. Minimum) specifies which DVH to select for any given systematic instance. This mode can e.g. be used to handle the situation in which one out of two (or more) functioning organs ensures uncomplicated outcome, and one is interested in the maximum dose to the remaining organ (receiving lowest dose).

2.1.2. Multiple dose distributions Treatment plans containing multiple dose distributions can be handled by defining a number of sets. Each of these sets contain a number of dose distributions to be delivered on the same days. Accumulation is always performed without α/β correction for doses within a set, whereas accumulation over sets (i.e. different days) may include fractionation effects. Systematic and random SDs can be specified per dose distribution, and correlation coefficients between doses can be specified for similar random errors within a set, and for systematic errors both within and between sets. (It is assumed random errors may correlate during a day but not from one day to another.) Multiple dose distributions and multiple targets may be combined.

2.2. Testing geometries

Testing geometries were prepared based on a $\varnothing 5$ cm spherical target (modeled as a geodesic sphere) and perfectly conforming dose with a Gaussian penumbra ($SD_{pen} = 0.32$ cm). A 1 mm dose grid resolution was used, and the sphere was sampled with a density of 250 random points per cc. Dose prescription was set at 66 Gy in 24 fractions. Systematic and random errors were assumed to be $SD = 0.3$ cm, and the margin between target and the 95% isodose was fine-tuned such that PlanJury would estimate a 90.0% probability of receiving 95% dose prescription to 99.9% of the spherical volume. Geometries were created for the case of two targets, two OARs, and two overlapping dose distributions (fig. 1a–c).

3. Results

Computations were performed on a workstation with dual Intel Xeon X5650 hexa-core processors running at 2.67GHz and hyper-threading enabled, thus capable of computing 24 threads simultaneously. For a given number of systematic error samples, stochastic residuals were estimated by repeating the determination of each result based on the

Table 1. Confidence levels on minimum DVH parameters for the combined CTV₁ and CTV₂ depicted in fig. 1a. Computed variables are printed boldface, variables in regular typeface were fixed. Stochastic noise intervals are quoted between square brackets. Three correlation coefficients for CTV₁ and CTV₂ were chosen between fully synchronous (= 1) and fully independent (= 0), and apply to systematic and random translations in all directions.

Correlation coefficients		Dose (%) (at least)	Probability (%)	Volume (%) (at least)
1	<i>D</i>	95.0 [94.9, 95.0]	90	99.9
	<i>P</i>	95	89.9 [89.8, 90.0]	99.9
	<i>V</i>	95	90	99.9 [99.9, 99.9]
0.5	<i>D</i>	93.5 [93.4, 93.5]	90	99.9
	<i>P</i>	95	84.3 [84.2, 84.5]	99.9
	<i>V</i>	95	90	99.6 [99.6, 99.6]
0	<i>D</i>	93.2 [93.2, 93.3]	90	99.9
	<i>P</i>	95	83.2 [83.1, 83.3]	99.9
	<i>V</i>	95	90	99.6 [99.6, 99.6]

first and last half of the samples only. The resulting intervals will be quoted between square brackets. Computation times for the probability of receiving at least 95% dose in 99.9% of a single CTV were recorded for various numbers of systematic samples:
5000 samples took 11.6 seconds with resulting $P = 89.4\%$ [88.5%, 90.3%];
10000 samples took 22.6 seconds with resulting $P = 90.1\%$ [89.8%, 90.4%];
50000 samples took 111.5 seconds with resulting $P = 90.0\%$ [90.0%, 90.0%].
Subsequent computations were done using 50000 systematic instances.

3.1. Two targets

DVH results for the case of two CTVs inside their respective high dose regions (fig. 1a) are listed in table 1. For true minimum dose computations (volume = 100%) the results for fully correlated motion would be expected to be identical to the case of a single CTV. However, the computation of true minimum dose in the sampled CTV would depend on a single random point, resulting in considerable stochastic fluctuations. As a compromise, near-minimum dose computations for 99.9% of the volume were performed. For the near-minimum dose the computed probability of $P = 89.9\%$ still corresponds well with the fine-tuned single CTV probability of $P = 90.0\%$. For fully uncorrelated motion the confidence is less at $P = 83.2\%$

3.2. Two OARs

Results for the two OARs partially covered by high dose (fig. 1b) are listed in table 2. It is seen that the probability of receiving no higher than 20 Gy to whichever OAR happens to receive the lowest mean dose is highest for fully correlated motion and decreases with the correlation coefficients, but is always higher than the corresponding probability in either of the OARs separately (results for OAR₂ are the same as for OAR₁ and are omitted).

3.3. Two dose distributions

In the case of two abutting dose distributions (fig. 1c) the DVH analysis for fully correlated motion is similar to the case of a perfectly conformal dose distribution (table 3); the slight increase to $P = 90.8\%$ can be contributed to the effectively larger margin behind the joining

Table 2. Confidence levels on maximum mean ROI dose for the two OARs depicted in fig. 1b. Results for OAR_{Min} are based on whichever OAR receives lowest mean dose for each systematic instance. Correlation coefficients apply to systematic and random translations in all directions.

ROI	Correlation coefficients	Mean dose (Gy) (at most)	Probability (%)
OAR _{Min}	1	<i>D</i> 24.0 [24.0, 24.0]	95
		<i>P</i> 20	40.0 [39.7, 40.3]
	0.5	<i>D</i> 26.4 [26.4, 26.4]	95
		<i>P</i> 20	38.9 [38.8, 39.1]
	0	<i>D</i> 28.2 [28.2, 28.3]	95
		<i>P</i> 20	36.2 [36.1, 36.2]
OAR ₁	–	<i>D</i> 33.3 [33.3, 33.3]	95
		<i>P</i> 20	19.9 [19.7, 20.0]

Table 3. Confidence levels on minimum DVH parameters for a CTV irradiated using two dose distributions (fig. 1c). The correlation coefficients govern the relative systematic and random shifts between the doses in all directions.

Correlation coefficients	Dose (%) (at least)	Probability (%)	Volume (%) (at least)
1	<i>D</i> 95.3 [95.2, 95.3]	90	99.9
	<i>P</i> 95	90.8 [90.7, 90.9]	99.9
	<i>V</i> 95	90	99.9 [99.9, 99.9]
0.5	<i>D</i> 84.9 [84.8, 85.0]	90	99.9
	<i>P</i> 95	67.3 [67.1, 67.5]	99.9
	<i>V</i> 95	90	97.3 [97.3, 97.4]
0	<i>D</i> 79.4 [79.4, 79.5]	90	99.9
	<i>P</i> 95	61.7 [61.6, 61.8]	99.9
	<i>V</i> 95	90	96.6 [96.6, 96.7]

interface. For uncorrelated uncertainties cold spots can occur near the joining interface, and as a result the probability of delivering adequate dose to the CTV decreases to 61.7%.

4. Discussion

By concentrating on a single spherical target with an idealized perfectly conformal dose distribution, the original work [3] on DVH sampling could lead to the (in)famous margin recipe $M = 2.5\Sigma + 0.7\sigma$. In the current work we have extended this sampling approach and considered more complicated clinical circumstances, involving multiple ROIs or multiple dose distributions. To describe such situations, correlations between the errors which occur need to be addressed. As there may be many different ways in which complications could present, we did not strive for parametrized simplifications (recipes). Instead, we developed a software tool which may be clinically deployed to perform probability based plan evaluation under various clinically challenging circumstances. By applying this tool to idealized test situations we have connected our results to those published before, and plotted out the ways the results may change for various

challenging clinical scenarios.

The rationale to combine volumes when computations are performed for two or more targets is that the clonogenic cells present in all target sub-volumes should receive adequate dose to ensure global tumor control. Likewise, we have shown how to handle two OARs with a common physiological function, and for which one of the organs can suffice to ensure uncomplicated treatment outcome. On the other hand, we did not consider correlations between dose deposition to e.g. tumors and OARs, or to OARs with independent functionality. Coherent motion patterns and patient setup errors can certainly result in correlated dose deposition in such volumes. However, correlations between the associated clinical effects are typically not relevant. The reason is that dose-effects are assumed to be local, and the clinical effects are assumed not to influence one another. For this reason, clinical trials are usually not separately analyzed for toxicity in patient subgroups with and without treatment failure, nor for failure depending on the occurrence of toxicity.

For treatment sites involving tissue inhomogeneities and/or air-tissue interfaces (such as lung and head-and-neck) the 3D dose distribution may itself be sensitive to the occurrence of geometrical variations. Especially for charged particle therapies this sensitivity may be high. Then, Monte Carlo plan evaluation should re-evaluate the dose as geometries are sampled, and distinguish rigid patient setup errors from organ motion, which should be described as patient deformations. Population based methods to generate patient geometries using principal component analysis have been investigated [8, 9], but full 3D dose reconstruction for the numerous sampled geometries may be impractical.

While the tools described here enable advanced treatment plan evaluation, they do not directly suggest how to improve a plan which does not meet clinical requirements. A solution would be to incorporate the techniques involving multiple ROIs and/or dose distributions also in plan optimization software [1], a major effort which was not yet endeavored.

5. Conclusions

An efficient and flexible tool for probabilistic plan evaluation was developed, able to handle challenging clinical planning circumstances. PlanJury can provide important plan quality information unavailable in conventional PTV or PRV based evaluations, thus aiding the creation of treatment plans with an improved balance between the probability of proper target dose on the one hand and the risk of treatment induced complications on the other.

References

- [1] Bohoslavsky R, Witte M, Janssen T and van Herk M 2013 *Phys. Med. Biol.* **58** 3563–3580
- [2] Killoran J, Kooy H, Gladstone D, Welte F and Beard C 1997 *Int. J. Radiat. Oncol. Biol. Phys.* **37** 213–221
- [3] van Herk M, Remeijer P, Rasch C and Lebesque J V 2000 *Int. J. Radiat. Oncol. Biol. Phys.* **47** 1121–1135
- [4] Baum C, Alber M, Birkner M and Nusslin F 2004 *Phys. Med. Biol.* **49** 5475–5488
- [5] ICRU 2010 *ICRU Report 83*
- [6] Schaake E, Belderbos J, Buikhuisen W, Rossi M, Burgers J and Sonke J J 2012 *Radiother. Oncol.* **105** 150–154
- [7] Franssen G, Peulen H, Glaser L, Belderbos J, Sonke J, Rossi M, Heres Diddens F and Damen E 2013 *Radiother. Oncol.* **106S** 109
- [8] Budiarto E, Keijzer M, Storchi P, Hoogeman M, Bondar L, Mutanga T, de Boer H and Heemink A 2011 *Phys. Med. Biol.* **56** 1045–1061
- [9] Sohn M, Sobotta B and Alber M 2012 *Phys. Med. Biol.* **57** 3693–3709