

ORIGINAL MANUSCRIPT

Optimal experimental design for cytogenetic dose-response calibration curves

Manuel Higuera^{a,c}, Adam Howes^{b,c}, and Jesús López Fidalgo^d

^aDepartamento de Matemáticas y Computación, Universidad de La Rioja; ^bDepartment of Mathematics, Imperial College London; ^cBasque Center for Applied Mathematics; ^dInstituto Cultura y Sociedad, Universidad de Navarra

ARTICLE HISTORY

Compiled January 28, 2020

ABSTRACT

Purpose: To introduce optimal experimental design techniques in the cytogenetic biological dosimetry practice. This includes the development of a new optimality criterion for the calibration of radiation doses.

Materials and Methods: The most typical optimal design criterion and the one developed in this research are introduced and applied in an example from the literature. In another example from the literature, a simulation study has been performed to compare the standard error of the dose estimation using different experimental designs. An RStudio project and a GitHub project have been developed to reproduce these results.

Results: In the paper it is observed that the application of optimal experimental design techniques can reduce the standard error of biodosimetric dose estimations.

Conclusions: Optimal experimental design techniques jointly with practitioners requirements may be applied. This practice would not involve an additional laboratory work.

KEYWORDS

Biodosimetry; I -optimality; calibration.

CONTACT Manuel Higuera. Email: manuel.higuera@unirioja.es

1. Introduction

Ionising radiation (IR) may be absorbed by humans, leading to negative health consequences. After a radiation event, absorbed dose estimates allow for appropriate clinical actions to be made. IRs produce damage at a cellular level in the form of chromosomal aberrations, which may be used as biomarkers of the absorbed dose. To study the effect of IR, calibration dose-response curves are built.

These curves are based on the irradiation of n *in vitro* blood samples which simulate homogeneous whole body exposures. Each sample is irradiated with a dose d_i ($i = 1 \dots n$). After the irradiation, m_i blood cells are analysed for sample i and y_{ij} chromosomal aberrations are scored for each cell ($j = 1, \dots, m_i$). For a low linear energy transfer (LET) whole body irradiation, it is assumed that the number of chromosomal aberrations follow a Poisson distribution whose intensity is a quadratic function of the absorbed dose, *i.e.*

$$Y \sim \text{Pois}(C + \alpha D + \beta D^2), \quad (1)$$

where Y represents the number of aberrations, D is the absorbed dose and $\{C, \alpha, \beta\}$ is the calibration parameter set. These parameters are calculated by maximum likelihood estimation. It is important to remark that in this Poisson model the link function is the identity, instead of the usual logarithmic link. This assumption is supported by the biological process of production of IR induced chromosomal aberrations, Hall and Giaccia (2012), and statistically in Oliveira et al. (2016). For high LET exposures it is assumed there is no quadratic dose effect, *i.e.* $\beta = 0$, Hall and Giaccia (2012).

The experimental design of these curves is based on the International Atomic Energy Agency (IAEA) suggestions. IAEA (2011) states that at least 10 doses should be used in the range $(0, 5]$ Gy, plus a control sample, and at least 4 of them in $(0, 1]$ Gy. The sample size, *i.e.* the number of scored cells at each dose, should aim to detect 100 chromosomal aberrations at each dose, but for the lower doses it is suggested to score a number of blood cells in the range $[3000, 5000]$. It is also suggested to reduce the variance of the linear term, *i.e.* α . Table 1 shows the experiment calibration data used

to construct the dicentric dose-response curve in Barquinero et al. (1995).

In this example, the fitted calibration parameters and their standard errors, estimated by maximum likelihood, are the following:

$$\begin{aligned}\hat{C} &= 0.00128, & SE(\hat{C}) &= 0.00047; \\ \hat{\alpha} &= 0.02103, & SE(\hat{\alpha}) &= 0.00516; \\ \hat{\beta} &= 0.06307, & SE(\hat{\beta}) &= 0.00401.\end{aligned}\tag{2}$$

The dose estimation leads to an inverse regression problem. The irradiated dose cannot be modeled as a regression function of the frequency of chromosomal aberrations, note the doses are fixed values selected by the experimenter. Given a sample with a total of S dicentrics in N blood cells, the dose estimation is the positive value \hat{D} which solves the equation

$$\hat{C} + \hat{\alpha}\hat{D} + \hat{\beta}\hat{D}^2 = S/N \Rightarrow \hat{D} = \frac{-\hat{\alpha} + \sqrt{\hat{\alpha}^2 - 4\hat{\beta}(\hat{C} - S/N)}}{2\hat{\beta}}.\tag{3}$$

This equation is based on the fact that the Poisson maximum likelihood estimator for the mean is the sample mean.

The purpose of this work is to provide experimental designs which reduce variance of the dose estimation in the design space (here the range of doses). Next section (Section 2) introduces the optimal experimental design field and presents a new criteria for calibration models.

2. Optimal experimental design

Let Y be a dependent variable, in our problem the number of dicentrics per blood cell, whose probability mass function is $f(Y|D, \theta)$, here the Poisson probability mass function. D is the explanatory variable, the absorbed dose here, and $\theta = \{\hat{C}, \hat{\alpha}, \hat{\beta}\}$ is the parameter set. Given a prediction function for the explanatory variable, in our case the estimation of the absorbed dose $\eta(N, \theta) = \hat{D}$ (see Equation 3), the aim is to find optimal (in some sense) experimental designs for the explanatory variable.

The sample size N is the number of blood cells analyzed. Two observed samples sizes will be used: $N = 500$, the number of cells recommended to analyze by the IAEA, and $N = 50$, the number of cells recommended to analyze in emergency situations for fast triage radio-protection response (see Barquinero et al. (2017)).

An *exact experimental design of size n* consists of a collection of points d_i , $i = 1, \dots, n$, in a given compact *design space*, \mathcal{X} . Some of these points may be repeated and a probability measure can be defined assigning to each different point the proportion of times it appears in the design. This leads to the idea of extending the definition of experimental design to any probability measure (*approximate design*). From the optimal experimental design point of view we can restrict the search to discrete designs of the type

$$\xi = \left\{ \begin{array}{cccc} d_1 & d_2 & \dots & d_n \\ w_1 & w_2 & \dots & w_n \end{array} \right\},$$

where d_i , $i = 1, \dots, n$ are the support points and $\xi(d_i) = w_i$ is the proportion of experiments to be made, here the number of scored cells, at point d_i . Thus, $w_i \geq 0$ and $\sum_{i=1}^n w_i = 1$.

The Fisher information matrix (FIM) of a design ξ is given by

$$M(\xi, \theta) = \sum_{d \in \mathcal{X}} I(d, \theta) \xi(d), \quad (4)$$

where

$$I(d, \theta) = \mathbb{E} \left[\frac{\partial \log f(Y|d, \theta)}{\partial \theta} \frac{\partial \log f(Y|d, \theta)}{\partial \theta^T} \right]$$

is the FIM at a particular point d . It is evaluated at some nominal value of θ . The nominal value usually represents the best guess for the parameters vector θ at the beginning of the experiment.

It can be proved that the inverse of this matrix is asymptotically proportional to the covariance matrix of the parameter estimators. An optimality design criterion,

$\Phi[M(\xi, \theta)]$, aims to minimize the covariance matrix in some sense and therefore the inverse of the information matrix. For simplicity $\Phi(\xi)$ will be used instead of $\Phi[M(\xi, \theta)]$. The main optimality design criterion is the D -optimal, which maximizes the FIM determinant, *i.e.* $\Phi_D(\xi) = \det M(\xi, \theta)$. In this paper a modification of the I -optimality criterion is considered, jointly to the D - and c -optimality ones.

The I -optimality criterion minimizes the average of the predictions of the response variable, that is the function

$$\Phi_I(\xi, N) = \int_{\mathcal{X}} \frac{\partial \eta(N, \theta)}{\partial \theta^T} M^-(\xi, \theta) \frac{\partial \eta(N, \theta)}{\partial \theta^T} dY,$$

where the superscript $-$ stands for the generalized inverse class of the matrix. In this paper the interest is not in estimating Y , but the absorbed dose, \hat{D} , explicitly known. Therefore, the modified I -criterion, say I_M -optimality requires the expected value of $\eta(N, \theta)$ with respect S , *i.e.*

$$\Phi_{I_M}(\xi) = \int_{\mathcal{X}} \text{E} \left[\frac{\partial \eta(N, \theta)}{\partial \theta^T} \right] M^-(\xi, \theta) \text{E} \left[\frac{\partial \eta(N, \theta)}{\partial \theta^T} \right] dD. \quad (5)$$

The c -optimality criterion is used to estimate a linear combination of the parameters, say $c^T \theta$, and it is defined by $\Phi_c(\xi) = c^T M^-(\xi, \theta) c$. Although the generalized inverse is unique only for nonsingular matrices the value of $c^T M(\xi, \theta)^- c$ is invariant for any member of the generalized inverse class if and only if $c^T \theta$ is estimable with the design ξ . IAEA (2011) states that a significant effort should be made to reduce the statistical uncertainty associated with the linear dose effect coefficient of yield, α . Consequently the c -optimality criterion is applied here for $c = (0, 1, 0)$, and will be referred as c_α -optimality criterion.

The goodness of a design is measured by its efficiency, defined by

$$\text{eff}_\Phi(\xi) = \frac{\Phi(\xi^*)}{\Phi(\xi)}.$$

The efficiency can be multiplied by 100 and be reported in percentage. If the function has a homogeneity property there is a practical statistical interpretation. Thus, if the

efficiency of a design is 50% this means that the design needs to double the total number of observations to perform as well as the optimal design.

In order to check whether a particular design is optimal or not there is a celebrated General Equivalence Theorem (GET). See Kiefer and Wolfowitz (1960), or Whittle (1973) for a more general version. This theorem is valid for approximate designs and convex criteria. It is quite useful also for building efficient algorithms for computing optimal designs. If the criterion function, Φ , is differentiable the Equivalence Theorem has a friendly and much more useful version. Let $\psi(D, \xi)$ be the Frechet directional derivative in the direction of a one-point design at D ,

$$\psi(D, \xi) = \lim_{\varepsilon \rightarrow 0^+} \frac{\Phi[(1 - \varepsilon)M(\xi, \theta) + \varepsilon I(D, \alpha)] - \Phi[M(\xi, \theta)]}{\varepsilon}.$$

This function is frequently called the sensitivity function. The GET states that under some conditions of the criterion function, $\psi(D, \xi)$ achieves its minimum value, zero, at the support points of the optimal design.

This theorem provides also an easy-to-build bound for the Φ -efficiency of a design, ξ ,

$$\text{eff}_{\Phi}(\xi) \geq 1 + \frac{\min_D \psi(D, \xi)}{\Phi(\xi)}.$$

For c -optimality the Elfving's graphic method (Elfving (1952)) can be used to construct the optimal design and this will not be needed.

More details on the theory of optimal experimental designs may be found, e.g., in Pazman (1986), Fedorov and Hackl (1997) and Atkinson et al. (2007).

In next section (Section 3), these criteria are applied in a literature experiment to compare efficiencies. López-Fidalgo and Amo-Salas (2019) considered a similar calibration problem with the essential difference of considering the error in the dose with its corresponding probability distribution.

3. Application example

Applying the Barquinero et al. (1995) fitted coefficients as nominal values (see Equation 2), the I_M -optimal local design for $N = 500$ and for $N = 50$ are shown in Table 2. It is also displayed the D -optimal and the c_α -optimal designs. The design space for the absorbed doses is $[0, 5]$ Gy.

The sensitivity functions of the optimal designs, are positive over all the design space, Figure 1, and consequently the GET states they are locally optimal.

As the three optimal designs obtained here are extremal with only 3 different doses, sequences are optimized, analogously to López-Fidalgo and Wong (2002), to produce I_M -suboptimal designs which look similar to the IAEA directions (IAEA (2011)). For this aim, sequences are proposed for both doses and weights to be optimized. For the doses, the following series is proposed for n points

$$\{D_1 = 0, D_2 = r/4, D_i = 5r^{n-i} \mid 5^{\frac{1}{4-n}} < r < 1, i = 3, \dots, n\},$$

where constrains are set for fixing a control point and the upper bound dose (5 Gy), for having at least one point in the range (0, 0.25) Gy, and the rest of the points greater than 0.25 and lower than 5. Then r is optimized. For their respective weights, the sequence

$$W_i = \frac{f(D_i)}{\sum_{j=1}^n f(D_j)}$$

is defined, where

$$f(x) = Ax^2 + Bx + 1, 0 < -\frac{B}{2A} < 5, f(0) > 0, f(5) > 0, f\left(-\frac{B}{2A}\right) > 0.$$

These constraints are to ensure the function is positive in the space design \mathcal{X} and the critical point is inside \mathcal{X} . A and B are optimized, the independent term is fixed to 1 in this quadratic shape to avoid multiple equivalent solutions by proportionality.

Table 2 also shows two I_M -suboptimal designs, one for $N = 50$ and another for

$N = 500$. Table 3 shows the efficiencies with respect the designs analyzed here.

All the results in this section can be reproduced by the RStudio project (RStudio Team (2016)) code available under request to the corresponding author.

3.1. *Sensitivity analysis*

Changes on the true calibration coefficients are done to check the efficiency of the optimal designs (modifying the nominal values) respect the original optimal design. Tables 4 and 5 show the efficiencies of different modified optimal designs with respect the $I_M^{N=50}$ - and $I_M^{N=500}$ -optimal designs. The nominal values used to compute the design remain fixed while true values of the parameters are consider in a neighborhood of them. The neighborhood was built using the standard errors of the estimators of the parameter in the example of reference.

In Tables 4 and 5 the first column indicates the modification done to the original nominal values. The second column in both tables describes the optimal design for the indicated modification. The third column is the efficiency of the “modified” optimal design with respect the optimal one for the original nominal values.

The original nominal values in this example are the maximum likelihood estimators of the Barquinero et al. (1995) dicentric assay experiment. These values are stated in Equation 2. This means that the calibration curve of the experiment is $0.00128 + 0.02103D + 0.06307D^2$. The first row of Table 4 is represents the modification of subtracting the standard error of \hat{C} to the nominal value of the intercept, *i.e.* the calibration curve results $0.00081 + 0.02103D + 0.06307D^2$. For these nominal values the $I_M^{N=50}$ -optimal design suggest to use the 15.93% of the cells as control, to irradiate the 45.85% to 0.95 Gy, and the 38.22% to 5 Gy. The $I_M^{N=50}$ -efficiency of this design in the scenario of the original nominal values is 94.20%.

4. Distribution for design space

When applying I_M -criterion, a distribution for the doses in the space design may be considered, to focus the reduction of the variance of the dose estimation to a determinate range. Note that in Expression 5 the reduction of the variance of the dose

estimation is distributed uniformly in all the dose range.

To consider the I_M -criterion with a different distribution with probability function $f(D)$, this criterion function Expression 5 may be modified as

$$\Phi_{I_M}(\xi) = \int_{\mathcal{X}} f(D) \mathbb{E} \left[\frac{\partial \eta(N, \theta)}{\partial \theta^T} \right] M^-(\xi, \theta) \mathbb{E} \left[\frac{\partial \eta(N, \theta)}{\partial \theta^T} \right] dD.$$

In Section 5, as part of the comparative study, some probability functions $f(D)$ are checked, but in this study the authors are focused on the equally distributed space design. Practitioners interested in constructing cytogenetic dose-response curves with the techniques described in this work, will have the choice to select the distribution for the space design they want to consider, *e.g.* they may be interested in more reduced statistical uncertainty dose estimations for the 0-1 Gy range.

5. Comparative study

A comparative study is performed in order to check the variance of the dose estimations for different experimental designs. Both parametric and non-parametric simulations are performed. They will be described later in detail. The experimental designs considered are the classical, $I_M^{N=500}$ -optimal, $I_M^{N=50}$ -optimal, c_α -optimal and D-optimal. In the non-parametric simulations, since the support points d_i must be contained within the available data, quasi-optimal versions of the above designs are used instead. Additionally, the simulations are performed at a range $\{0.125, 0.25, 0.5, 0.75, 0.875, 1\}$ of different observation irradiation fractions, where irradiation fraction 1 corresponds to full-body exposure.

The number of cells simulated in the observation is 500 and in the calibration is 20000. For each dose, experimental design and irradiation fraction, 10000 dose estimates are simulated. The standard deviations of these simulations are shown in Tables 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 and 17 together with Figures 2 and 3. In these tables, the minimum standard deviation at each scenario is highlighted in **bold**.

The scripts which reproduce the results of these simulations are available on <https://github.com/athowes/biodose>.

5.1. Simulation details

Firstly, some preliminary steps are taken:

- Using the Barquinero *et al.* 1997 Table 1 data, nominal values for the calibration parameters are fit by maximum likelihood estimation resulting in the values

$$\hat{C} = 9.054 \cdot 10^{-4}, \quad \hat{\alpha} = 3.431 \cdot 10^{-2}, \quad \hat{\beta} = 5.702 \cdot 10^{-2}$$

and the covariance matrix

$$\hat{\Sigma} = \begin{pmatrix} 0.009726 & -0.038362 & 0.012924 \\ -0.038362 & 2.351729 & -1.012357 \\ 0.012924 & -1.012357 & 0.885871 \end{pmatrix} \cdot 10^{-5}.$$

- For each approximate experimental design ξ , the exact experimental design

$$\xi^* = \left\{ \begin{array}{cccc} d_1 & d_2 & \dots & d_n \\ m_1 & m_2 & \dots & m_n \end{array} \right\}$$

is generated where the number of cells m_i at each design point d_i sum to the total calibration sample size. In this case, that is to say $\sum_{i=1}^n m_i = 20,000$. This is done using rounding if necessary.

5.1.1. Observed full-body exposure simulations

The number of chromosome aberrations y_z is simulated in each cell z in both the exact experimental design ξ^* and the observed sample. As previously noted, there are two methods for simulating y_z . The steps of the parametric simulation for cell z at dose d , with full-body exposure $f = 1$, are:

S1 Generate (C, α, β) from $\mathcal{N}((\hat{C}, \hat{\alpha}, \hat{\beta}), \hat{\Sigma})$

S2 Calculate $\lambda = C + \alpha \cdot d + \beta \cdot d^2$

S3 Generate y_z from $\text{Pois}(\lambda)$

Alternatively, in the non-parametric simulations y_z is determined by re-sampling with

replacement from the Barquinero *et al.* 1997 Barquinero et al. (1997) Table 1 data corresponding to dose d .

In both methods the calibration parameters are then refitted by maximum likelihood estimation using the simulated calibration data. The absorbed dose estimation \hat{D} may then be calculated using the simulated observed sample, of size 500, and the refitted calibration parameters via Equation 3. In the case that there is either no solution or only negative solutions to Equation 3 then \hat{D} is set to zero.

5.1.2. Observed partial-body exposure simulations

For observed partial-body exposure simulations, the calibration data is simulated as above. As before the calibration parameters are refitted by maximum likelihood estimation. However, for the cells in the observed sample the simulations also incorporate the fraction f of the body which has been irradiated.

In the case of partial-body exposure the number of chromosome aberrations is believed to follow a zero-inflated Poisson distribution, with additional parameter ω , the proportion of extra zeroes (see Higuera et al. (2016)). The value of ω may be estimated using Dolphin's method (IAEA (2011))

$$f = \frac{(1 - \omega) \exp(d/d_0)}{\omega + (1 - \omega) \exp(d/d_0)}, \quad (6)$$

where d_0 is the 37% cell survival dose. From experimental evidence d_0 is believed to be between 2.7 and 3.5 Gy. As such, the steps of the parametric partial body simulation for each cell z in the observed sample are:

- S1 Generate d_0 from $\mathcal{U}(2.7, 3.5)$
- S2 Calculate $\omega = \left(\frac{f}{1-f} \exp(-d/d_0) + 1 \right)^{-1}$ (rearranging Equation 6)
- S3 Generate (C, α, β) from $\mathcal{N}((\hat{C}, \hat{\alpha}, \hat{\beta}), \hat{\Sigma})$
- S4 Calculate $\lambda = C + \alpha \cdot d + \beta \cdot d^2$
- S5 Generate y_z from $\text{ZIP}(\omega, \lambda)$

Similarly to the whole-body exposure simulations, the non-parametric partial-body simulations of the cells in the observed sample are made by re-sampling with replace-

ment, this time from the Barquinero et al. (1997). Table 2 data corresponding to dose d and irradiated fraction f .

For both parametric and non-parametric simulations a zero-inflated Poisson distribution is fitted to the observation data to give maximum likelihood estimate $\hat{\lambda}$. This is used in place of the sample mean in Equation 3 and again combined with the simulated calibration parameters to give the dose estimate \hat{D} . As before, in the case that there is either no solution or only negative solutions then \hat{D} is set to zero.

Table 18 summarizes the standard deviations of dose estimations using the classical design divided by the standard deviations of dose estimations using the optimal design in the same situation, *i.e.* assuming the same dose, the same irradiated fraction body and the same type of simulation. These ratios are grouped in four dose ranges. The mean and the 95% most probably interval (MPI) are shown for each group.

6. Final remarks

A new optimal experimental design criterion is proposed here to minimize the dose estimation error for classical cytogenetic biodosimetry calibration curves. The results displayed here show how the I_M -efficiency can be increased significantly defining non-extremal alternative designs which are inside the IAEA directions (IAEA (2011)). Approximate designs are proposed balancing both the minimization of the dose estimation variance and the experimenters suggestions.

The application of these results allow to define new experimental designs taking into account the results in previous experiments. For instance, the alternative design for irradiating 18261 blood cells (the same number of cells scored in the Barquinero et al. (1995) experiment) based on the 11 points $I_M^{N=500}$ -suboptimal design at Table 3 would be

$$\left\{ \begin{array}{l} \text{Dose (Gy):} \quad 0.00 \quad 0.19 \quad 0.48 \quad 0.64 \quad 0.86 \quad 1.15 \quad 1.54 \quad 2.07 \quad 2.78 \quad 3.73 \quad 5.00 \\ \text{Scored cells:} \quad 2064 \quad 1815 \quad 1484 \quad 1325 \quad 1144 \quad 958 \quad 812 \quad 800 \quad 1117 \quad 2143 \quad 4599 \end{array} \right\}.$$

In Table 18 it is observed that the reduction of the standard error of dose estimations in optimal designs with respect the classical design increases at higher doses. In fact, at

lowest doses (≤ 0.25 Gy) the classical design seems to produce lower standard errors. Here it is necessary to consider that only the I_M -optimal design is focused on reducing the standard error of dose estimations, and as the simulated samples are of size 500, only the $I_M^{N=500}$ -optimal design is indicated to reduce the average of the standard error of whole-body irradiations in the design space (0-5 Gy). However, in general, application of OED techniques reduces the standard error of dose estimations. For instance, the standard error for doses higher than 2 Gy is reduced in average a 15%. For more specific targets, *e.g.* to reduce the standard error at doses lower than 1 Gy for partially irradiated samples, a new design can be generated by giving more weight to the 0-1 Gy interval in the design space, assuming a partially irradiated sample is given.

Suboptimal designs are proposed here in order to come to a compromise between OED techniques and experimenters' experience. This implies to deteriorate the theoretical optimization, but it is important to remark that in practice an optimal design is as optimal as true its model hypothesis. Consequently, experimenters can distrust an extremal design with only three points. In this particular practice, the spread of dose points can respond to the worry of biodosimetrists in a dose range values which they consider critical, for instance at doses lower than 1 Gy.

This research proposes the application of OED techniques for designing cytogenetic dose-response curves. A new optimality criterion is proposed to minimize the calibrated estimation of dose. In Section 3 designs are calculated and analyzed for the Barquinero et al. (1995) dicentric assay experiment, based on this new criterion and in other well-known criteria. The results of this section are reproducible by the R code scripts available (under request to the corresponding author) in the form a RStudio project (RStudio Team (2016)). In Section 5, a comparative study via simulations is performed. These simulations compare the error of the dose estimations of different designs using the Barquinero et al. (1997) dicentrics plus rings data. The GitHub project *biodose* (<https://github.com/athowes/biodose>) reproduce the results of this comparative study.

Acknowledgements

This work was supported the Basque Government under Grant BERC 360 2014-2017; the Spanish Ministry of Economy and Competitiveness MINECO and FEDER under Grant BCAM Severo Ochoa excellence accreditation SEV-2017-0718; BIOSTATNET under 2014 and 2016 calls for *Programa de estancias BIOSTATNET jóvenes investigadores*; the Spanish Research Agency and fondos FEDER under Grants MTM2013-47879-C2-1-P and MTM2016-80539-C2-R1.

Disclosure statement

The authors report no conflicts of interest.

References

- Ainsbury EA, Higuera M, Puig P, Einbeck J, Samaga D, Barquinero JF, *et al.* 2017. Uncertainty of fast biological radiation dose assessment for emergency response scenarios. *Int J Radiat Biol.* 93:127–135.
- Atkinson AC, Donev A, Tobias R. 2007. *Optimum Experimental Designs, with SAS.* Oxford University Press.
- Barquinero JF, Barrios L, Caballín MR, Miró R, Ribas M, Subias A, Egozcue J. 1995. Establishment and validation of a dose-effect curve for gamma-rays by cytogenetic analysis. *Mutat Res.* 326:65–69.
- Barquinero JF, Barrios L, Caballín MR, Miró R, Ribas M, Subias A, Egozcue J. 1997. Biological dosimetry in simulated in vitro partial irradiations. *Int J Radiat Biol.* 71:435–440.
- Elfving G. 1952. Optimum allocation in linear regression theory. *Ann Math Stat.* 23:255–262.
- Fedorov VV, Hackl P. 1997. *Model-Oriented Design of Experiments.* Springer: New York.
- Hall EJ, Giaccia AJ. 2012. *Radiobiology for the radiologist, 7th edition.* Lippincott Williams & Wilkins: Philadelphia.
- Higuera M, Puig P, Ainsbury EA, Vinnikov VA, Rothkamm K. 2016. A new Bayesian model applied to cytogenetic partial body irradiation estimation. *Radiat Prot Dosim.* 168:330–336.
- IAEA. 2011. *Cytogenetic Dosimetry: Applications in Preparedness for and Response to Radiation Emergencies.* International Atomic Energy Agency: Vienna.

- Kiefer J, Wolfowitz J. 1960. The Equivalence of two Extremum Problems. *Canad J Math.* 12:363–366.
- López-Fidalgo J, Wong WK. 2002. Design for the Michaelis-Menten model. *J Theor Biol.* 215:1–11.
- López-Fidalgo J, Amo-Salas M. 2019. Optimal dose calibration in radiotherapy. Submitted.
- Oliveira M, Einbeck J, Higuera M, Ainsbury EA, Rothkamm K, Puig P. 2016. Zero-inflated regression models for radiation-induced chromosome aberration data: A comparative study. *Biom J.* 58:259–279.
- Pazman A. 1986. *Foundations of Optimum Experimental Design.* Springer Netherlands.
- RStudio Team. 2016. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL <http://www.rstudio.com/>.
- Whittle P. 1973. Some General Points in the Theory of Optimal Experimental Design. *J Roy Statist Soc Ser B.* 35:123–130.

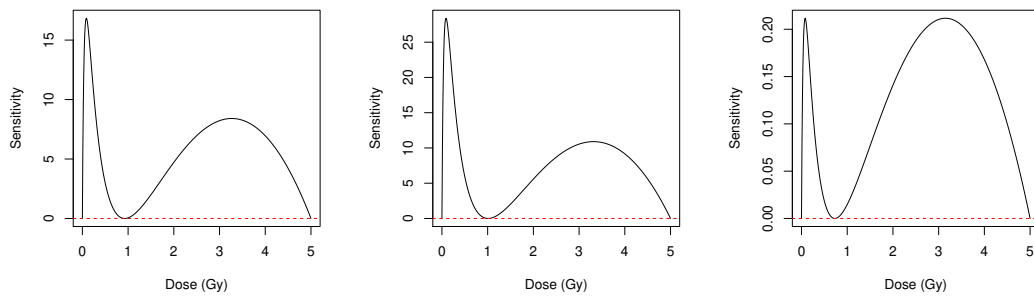


Figure 1. From left to right, sensitivity functions of the $I_M^{N=500}$ -, $I_M^{N=50}$ - and c_α -optimal designs for the Barquinero et al. (1995) dicentric assay.

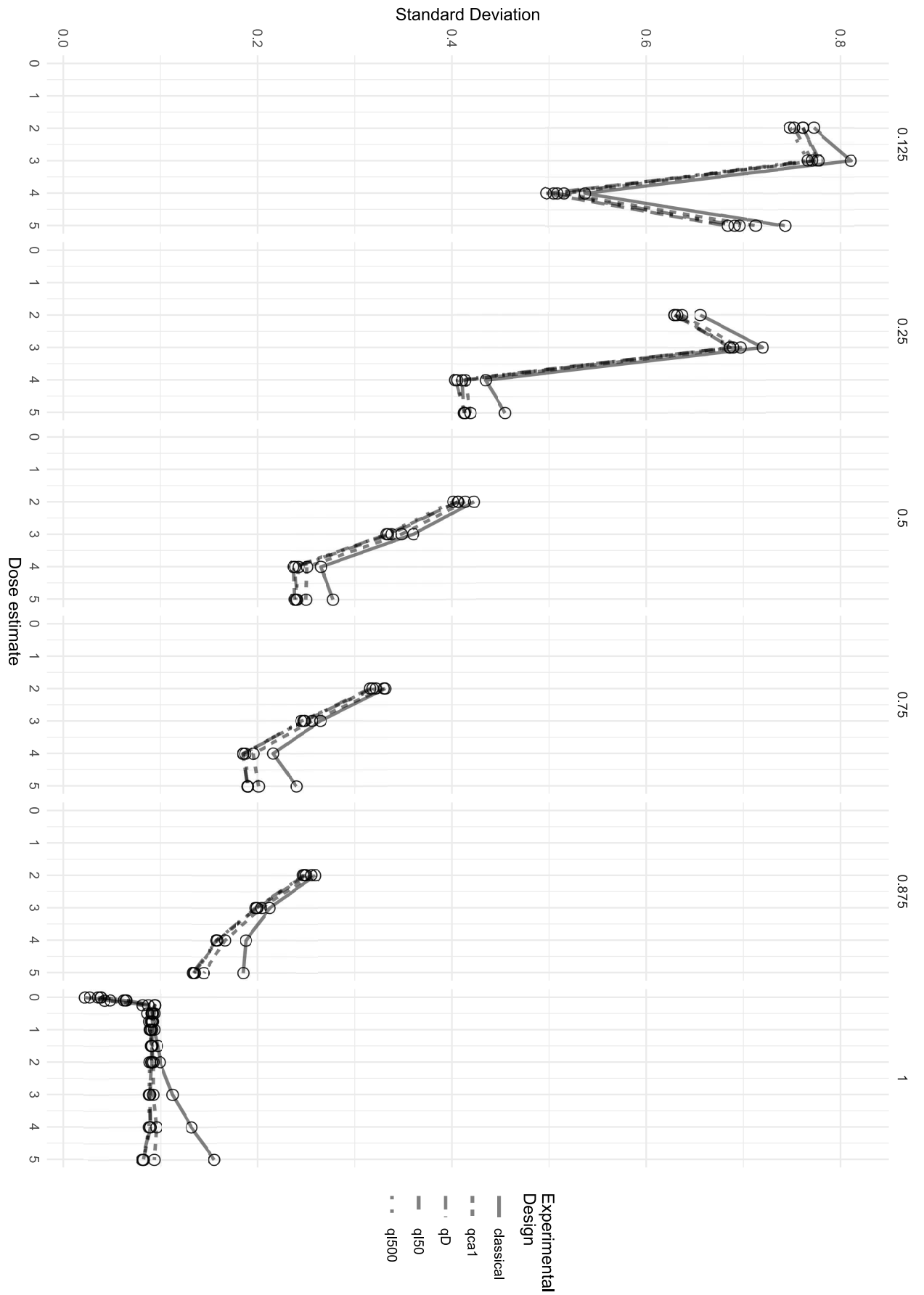


Figure 2. Non-parametric.

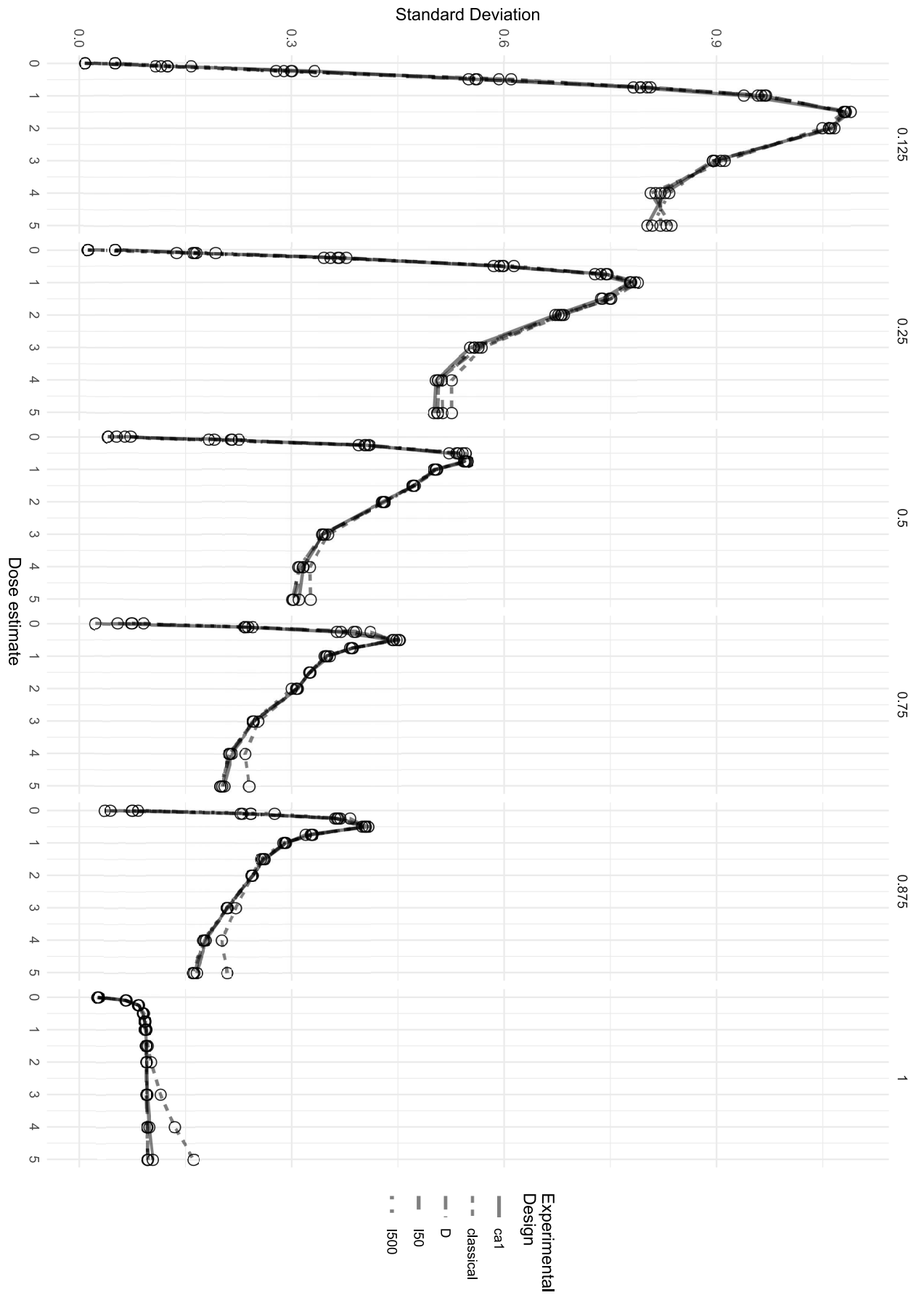


Figure 3. Standard deviation of simulated dose estimations; 10,000 simulations via parametric bootstrap, plot faceted by irradiated fraction

Table 1. Dicentric distribution within cells, cells analysed m , total number of dicentrics detected s , and sample means s/m for each irradiated dose.

Dose (Gy)	Number of dicentrics						m	s	s/m
	0	1	2	3	4	5			
0.00	4992	8					5000	8	0.002
0.10	4988	14					5002	14	0.003
0.25	1987	20	1				2008	22	0.011
0.50	1947	55					2002	55	0.027
0.75	1736	92	4				1832	100	0.050
1.00	1064	99	5				1168	109	0.093
1.50	474	76	12				562	100	0.178
2.00	251	62	16	3			332	103	0.310
3.00	104	72	15	2			193	108	0.560
4.00	35	41	21	4	2		103	103	1.000
5.00	11	19	11	9	6	3	59	107	1.814

Table 2. First column at each design represents the weight (%) at 0 Gy. All designs have a 5 Gy point whose weight is omitted.

Name	Design (Gy and %)									
Classical	27.38	0.1	0.25	0.5	0.75	1	1.5	2	3	4
$I_M^{N=500}$ -optimal	12.08	27.39	11.00	10.96	10.03	6.40	3.08	1.82	1.06	0.56
$I_M^{N=50}$ -optimal	24.89					0.93				
c_α -optimal	12.22					49.62				
D -optimal	33.33					1.01				
$I_M^{N=50}$ -suboptimal, $n = 10$	20.95		0.19	0.69	0.91	1.21	1.61	2.13	2.83	3.76
$I_M^{N=500}$ -suboptimal, $n = 11$	11.29		17.80	10.75	8.15	5.27	2.46	0.51	1.08	7.40
		0.19	0.48	0.64	0.86	1.15	1.54	2.07	2.78	3.73
		9.94	8.12	7.26	6.27	5.25	4.45	4.38	6.12	11.74

Table 3. Designs efficiencies (%).

Name	$I_M^{N=500}$ -eff	$I_M^{N=50}$ -eff	c_α -eff	D -eff
Classical	8.26	10.43	43.57	37.85
$I_M^{N=500}$ -optimal	100.00	86.33	72.65	84.79
$I_M^{N=50}$ -optimal	91.73	100.00	62.08	96.74
c_α -optimal	49.46	50.62	100.00	62.85
D -optimal	81.40	93.15	52.51	100.00
$I_M^{N=50}$ -suboptimal, $n = 10$	82.43	87.18	52.91	87.12
$I_M^{N=500}$ -suboptimal, $n = 11$	88.53	78.02	54.82	75.84

Table 4. Sensitivity analysis for $I_M^{N=50}$ -criterion (%). First column at each design represents the weight (%) at 0 Gy. All designs have a 5 Gy point whose weight is omitted.

Modification	Design (Gy and %)	$I_M^{N=50}$ -eff
$-SE(\hat{C})$	15.93 0.95 45.85	94.20
$+SE(\hat{C})$	42.96 0.77 49.07	50.76
$-0.003D$	43.21 0.78 46.56	60.25
$+SE(\hat{\alpha})D$	15.72 1.02 46.55	94.04
$-SE(\hat{\beta})D^2$	23.22 1.02 42.40	99.84
$+SE(\hat{\beta})D^2$	26.81 1.00 40.94	99.80
$+0.00047 - 0.000015D$	42.65 0.76 49.71	49.17
$-SE(\hat{C}) + SE(\hat{\beta})D^2$	16.55 0.93 45.45	95.02

Table 5. Sensitivity analysis for $I_M^{N=500}$ -criterion (%). First column at each design represents the weight (%) at 0 Gy. All designs have a 5 Gy point whose weight is omitted.

Modification	Design (Gy and %)	$I_M^{N=50}$ -eff
$-SE(\hat{C})$	8.79	0.90 51.32
$+SE(\hat{C})$	39.77	0.87 46.32
$-0.003D$	33.77	0.92 43.40
$+SE(\hat{\alpha})D$	9.09	0.97 51.52
$-SE(\hat{\beta})D^2$	11.54	0.94 50.00
$+SE(\hat{\beta})D^2$	12.76	0.91 49.23
$+0.00047 - 0.000015D$	42.27	0.79 48.93
$-SE(\hat{C}) - SE(\hat{\beta})D^2$	9.00	0.89 51.11

Table 6. Non-parametrically simulated standard deviations for whole body irradiations.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0274	0.0487	0.0877	0.0921	0.0917	0.0935	0.0970	0.1000	0.1128	0.1321	0.1555
$qI_{M=500}^N$	0.0389	0.0653	0.0940	0.0911	0.0905	0.0896	0.0916	0.0891	0.0885	0.0884	0.0830
$qI_{M=50}^N$	0.0387	0.0646	0.0941	0.0913	0.0921	0.0901	0.0905	0.0913	0.0892	0.0896	0.0818
qc_α	0.0226	0.0426	0.0822	0.0869	0.0888	0.0893	0.0907	0.0926	0.0925	0.0958	0.0936
qD	0.0361	0.0628	0.0942	0.0935	0.0915	0.0913	0.0913	0.0907	0.0892	0.0898	0.0828

Table 7. Parametrically simulated standard deviations for whole body irradiations.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0256	0.0661	0.0837	0.0901	0.0929	0.0952	0.0971	0.1017	0.1154	0.1351	0.1619
$I_{N=500\text{-op}}^N$	0.0279	0.0669	0.0845	0.0904	0.0925	0.0946	0.0951	0.0954	0.0955	0.0958	0.0966
$I_{M=50\text{-op}}^M$	0.0274	0.0669	0.0829	0.0909	0.0924	0.0927	0.0940	0.0946	0.0947	0.0959	0.0974
$c_{\alpha\text{-op}}$	0.0268	0.0661	0.0834	0.0893	0.0927	0.0939	0.0953	0.0955	0.0966	0.0993	0.1042
D-op	0.0261	0.0656	0.0837	0.0914	0.0940	0.0952	0.0953	0.0957	0.0953	0.0966	0.0969

Table 8. Non-parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.875.

Design	Irradiated dose (Gy)			
	2	3	4	5
Classical	0.2595	0.2127	0.1884	0.1857
$qI_{N=500}^M$	0.2470	0.1992	0.1575	0.1356
$qI_{M=50}^N$	0.2483	0.1979	0.1577	0.1336
qC_α	0.2554	0.2039	0.1666	0.1446
qD	0.2498	0.1996	0.1588	0.1346

Table 9. Non-parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.75.

Design	Irradiated dose (Gy)			
	2	3	4	5
Classical	0.3315	0.2649	0.2163	0.2402
$qI_{N=500}^M$	0.3219	0.2453	0.1876	0.1905
$qI_{M=50}^N$	0.3161	0.2476	0.1855	0.1900
qC_α	0.3300	0.2563	0.1958	0.2011
qD	0.3185	0.2485	0.1853	0.1900

Table 10. Parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.875.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0441	0.2280	0.3642	0.3992	0.3284	0.2914	0.2610	0.2447	0.2208	0.2015	0.2091
$I_{N=500\text{-op}}^N$	0.0751	0.2760	0.3687	0.4053	0.3292	0.2880	0.2565	0.2431	0.2096	0.1749	0.1607
$I_{M=50\text{-op}}^M$	0.0831	0.2308	0.3823	0.4040	0.3300	0.2900	0.2621	0.2439	0.2083	0.1764	0.1609
$c_{\alpha\text{-op}}$	0.0363	0.2424	0.3659	0.4084	0.3201	0.2927	0.2615	0.2450	0.2084	0.1791	0.1664
$D\text{-op}$	0.0753	0.2422	0.3617	0.4025	0.3271	0.2911	0.2590	0.2448	0.2096	0.1771	0.1629

Table 11. Parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.75.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0910	0.2349	0.3697	0.4484	0.3855	0.3490	0.3264	0.3073	0.2531	0.2346	0.2401
$I_{N=500\text{-op}}^N$	0.0740	0.2447	0.3905	0.4524	0.3849	0.3506	0.3261	0.3007	0.2459	0.2116	0.1997
$I_{M=50\text{-op}}^M$	0.0747	0.2386	0.4109	0.4425	0.3846	0.3470	0.3253	0.3068	0.2451	0.2134	0.1996
$c_{\alpha\text{-op}}$	0.0230	0.2332	0.3641	0.4444	0.3860	0.3491	0.3250	0.3090	0.2464	0.2157	0.2047
$D\text{-op}$	0.0545	0.2344	0.3876	0.4517	0.3824	0.3540	0.3257	0.3076	0.2464	0.2117	0.2020

Table 12. Non-parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.5.

Design	Irradiated dose (Gy)			
	2	3	4	5
Classical	0.4228	0.3604	0.2654	0.2776
$qI_{N=500}^N$	0.4016	0.3327	0.2424	0.2397
$qI_{M=50}^M$	0.4061	0.3378	0.2387	0.2407
qC_α	0.4137	0.3486	0.2508	0.2498
qD	0.4067	0.3342	0.2370	0.2383

Table 13. Non-parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.25.

Design	Irradiated dose (Gy)			
	2	3	4	5
Classical	0.6558	0.7198	0.4350	0.4547
$qI_M^{N=500}$	0.6293	0.6896	0.4056	0.4129
$qI_M^{N=50}$	0.6296	0.6864	0.4038	0.4138
qC_α	0.6370	0.6971	0.4140	0.4192
qD	0.6321	0.6863	0.4106	0.4125

Table 14. Parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.5.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0529	0.2255	0.4025	0.5331	0.5473	0.5041	0.4723	0.4306	0.3517	0.3257	0.3268
$I_{N=500\text{-op}}^N$	0.0409	0.1833	0.4082	0.5456	0.5456	0.5016	0.4712	0.4293	0.3470	0.3097	0.3031
$I_{M=50\text{-op}}^M$	0.0406	0.2158	0.4101	0.5358	0.5488	0.5040	0.4715	0.4312	0.3447	0.3162	0.3031
$c_{\alpha\text{-op}}$	0.0648	0.2145	0.4045	0.5225	0.5434	0.5018	0.4717	0.4279	0.3441	0.3167	0.3105
D-op	0.0727	0.1908	0.3943	0.5414	0.5441	0.5051	0.4739	0.4319	0.3428	0.3117	0.3014

Table 15. Parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.25.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0121	0.1376	0.3550	0.5981	0.7441	0.7849	0.7485	0.6842	0.5684	0.5261	0.5263
$I_{N=500\text{-op}}^N$	0.0126	0.1632	0.3461	0.6142	0.7362	0.7789	0.7366	0.6800	0.5580	0.5072	0.5068
$I_{M=50\text{-op}}^M$	0.0126	0.1928	0.3675	0.5996	0.7457	0.7889	0.7506	0.6766	0.5576	0.5123	0.5133
$c_{\alpha\text{-op}}$	0.0507	0.1656	0.3654	0.5851	0.7433	0.7775	0.7387	0.6724	0.5519	0.5043	0.5017
$D\text{-op}$	0.0512	0.1608	0.3771	0.5936	0.7279	0.7769	0.7506	0.6815	0.5642	0.5074	0.5062

Table 16. Non-parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.125.

Design	Irradiated dose (Gy)			
	2	3	4	5
Classical	0.7723	0.8104	0.5375	0.7432
$qI_{N=500}^N$	0.7482	0.7661	0.5054	0.6911
$qI_{M=50}^M$	0.7523	0.7703	0.5092	0.6960
qC_α	0.7614	0.7759	0.5161	0.7126
qD	0.7610	0.7780	0.4974	0.6842

Table 17. Parametrically simulated standard deviations for partial body irradiations, irradiated fraction 0.125.

Design	Irradiated dose (Gy)										
	0	0.1	0.25	0.5	0.75	1	1.5	2	3	4	5
Classical	0.0085	0.1159	0.2893	0.6103	0.7931	0.9636	1.0806	1.0493	0.9119	0.8077	0.8363
$I_{N=500\text{-op}}^N$	0.0511	0.1251	0.2780	0.5623	0.7918	0.9683	1.0817	1.0587	0.8956	0.8333	0.8098
$I_{M=50\text{-op}}^M$	0.0508	0.1081	0.3012	0.5498	0.8024	0.9585	1.0892	1.0586	0.9066	0.8144	0.8298
$c_{\alpha\text{-op}}$	0.0086	0.1250	0.3325	0.5597	0.7820	0.9389	1.0831	1.0611	0.8981	0.8272	0.8025
$D\text{-op}$	0.0084	0.1583	0.2997	0.5925	0.8073	0.9701	1.0794	1.0662	0.8951	0.8212	0.8213

Table 18. Ratio of standard deviations of dose estimations grouped in dose ranges to compare the classical design against the optimized ones.

Dose range (Gy)	Mean	95% MPI
[0, 0.25]	0.96	(0.17, 1.29)
(0.25, 1]	1.01	(0.97, 1.06)
(1, 2]	1.02	(0.98, 1.08)
(2, 5]	1.15	(0.99, 1.68)