



HAL
open science

Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations

Martin Andrew Ebert, Sarah Gulliford, Oscar Acosta, Renaud de Crevoisier, Todd Mcnutt, Wilma D Heemsbergen, Marnix G Witte, Giuseppe Palma, Tiziana Rancati, Claudio Fiorino

► **To cite this version:**

Martin Andrew Ebert, Sarah Gulliford, Oscar Acosta, Renaud de Crevoisier, Todd Mcnutt, et al.. Spatial descriptions of radiotherapy dose: normal tissue complication models and statistical associations. *Physics in Medicine and Biology*, 2021, 66 (12), pp.12TR01. <10.1088/1361-6560/ac0681>. <hal-03246458>

HAL Id: hal-03246458

<https://hal.science/hal-03246458v1>

Submitted on 21 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

PHYSICS IN MEDICINE AND BIOLOGY**TOPICAL REVIEW****Spatial descriptions of radiotherapy dose: Normal tissue complication models and statistical associations**

- Martin A Ebert^{1,2,3}
- Sarah Gulliford^{4,5}
- Oscar Acosta⁶
- Renaud de Crevoisier^{6a}
- Todd McNutt⁷
- Wilma D Heemsbergen⁸
- Marnix Witte⁹
- Giuseppe Palma¹⁰
- Tiziana Rancati¹¹
- Claudio Fiorino¹²

¹ School of Physics, Mathematics and Computing, University of Western Australia, Crawley, Western Australia

² Department of Radiation Oncology, Sir Charles Gairdner Hospital, Nedlands, Western Australia

³ 5D Clinics, Claremont, Western Australia

⁴ Department of Radiotherapy Physics, University College Hospitals London, United Kingdom

⁵ Department of Medical Physics and Bioengineering, University College London, United Kingdom

⁶ Univ Rennes, CLCC Eugène Marquis, INSERM, LTSI - UMR 1099, F-35000 Rennes, France

⁷ Johns Hopkins University, Baltimore, Maryland, USA

⁸ Department of Radiotherapy, Erasmus MC Cancer Institute, the Netherlands

⁹ The Netherlands Cancer Institute, Amsterdam, The Netherlands

¹⁰ Institute of Biostructures and Bioimaging, National Research Council, Napoli, Italy

¹¹ Prostate Cancer Program, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

¹² Medical Physics, San Raffaele Scientific Institute, Milano, Italy

Corresponding author:

Dr Martin A Ebert

1
2
3 Radiation Oncology
4

5
6 Sir Charles Gairdner Hospital
7

8 Hospital Ave
9

10
11 Nedlands, Western Australia 6009
12

13
14 Australia
15

16
17 Tel: +61 423 976 746
18

19
20 Email: Martin.Ebert@health.wa.gov.au
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Accepted Manuscript

1		
2		
3	1	INTRODUCTION.....2
4		
5	2	STRATEGIES FOR CHARACTERISING DOSE DISTRIBUTIONS3
6		
7	2.1	DOSE-VOLUME APPROACHES 3
8		
9	2.1.1	<i>The advantages of dose-volume approaches..... 3</i>
10	2.1.2	<i>The disadvantages of dose-volume approaches..... 4</i>
11		
12	2.2	APPROACHES THAT PRESERVE SPATIAL INFORMATION..... 5
13		
14	2.2.1	<i>1D precision dose-volume approaches 7</i>
15	2.2.2	<i>2D surface mapping..... 8</i>
16	2.2.3	<i>3D feature extraction 9</i>
17	2.2.4	<i>3D volume mapping..... 9</i>
18		
19		
20	3	PRACTICAL CONSIDERATIONS.....10
21		
22	3.1	REQUIRED TECHNICAL DATA..... 10
23		
24	3.1.1	<i>Imaging..... 10</i>
25	3.1.2	<i>Structures..... 11</i>
26	3.1.3	<i>Dose..... 11</i>
27	3.1.4	<i>Treatment description..... 12</i>
28		
29	3.2	OUTCOME DATA..... 13
30	3.3	DATA SOURCES..... 14
31		
32		
33	4	STATISTICAL AND MODELLING CONSIDERATIONS.....15
34		
35	4.1	FEATURE SELECTION..... 16
36		
37	4.1.1	<i>Candidate dosimetric features and collinearity..... 16</i>
38	4.1.2	<i>Feature reduction..... 16</i>
39	4.1.3	<i>Co-variates..... 17</i>
40	4.1.4	<i>Models and Algorithms..... 17</i>
41	4.1.5	<i>Voxel-wise models..... 18</i>
42	4.1.6	<i>Significance..... 20</i>
43		
44	4.2	PERFORMANCE, VALIDITY AND REPORTING..... 21
45		
46	4.2.1	<i>Model performance..... 21</i>
47	4.2.2	<i>Model validity..... 21</i>
48	4.2.3	<i>Model reporting..... 21</i>
49		
50		
51		
52	5	REVIEW OF METHODS – SPATIAL DOSE ASSOCIATIONS WITH COMPLICATIONS AND
53		APPLICATIONS TO NTCP CALCULATION22
54		
55	5.1	USE OF HISTOGRAM-BASED FEATURES 26
56		
57	5.1.1	<i>Description..... 26</i>
58	5.1.2	<i>Examples..... 27</i>
59		
60		

1		
2		
3	5.2	VOXEL-WISE ASSESSMENT 28
4		
5	5.2.1	<i>Description</i> 28
6	5.2.2	<i>Examples</i> 29
7		
8		2D dose-surface outcome mapping 29
9		3D voxel-wise outcome mapping 29
10		NTCP from voxel-wise methods 31
11		
12	5.3	SPATIAL PARAMETERISATION OF DOSE DISTRIBUTIONS 32
13		
14	5.3.1	<i>Description</i> 32
15	5.3.2	<i>Examples</i> 32
16		
17		Parameterisation of 2D dose 32
18		Parameterisation of 3D dose 33
19		Supervised broad spatial descriptors 34
20		Unsupervised broad spatial descriptors 35
21		
22	5.4	SPATIAL CLUSTERING 36
23		
24	5.4.1	<i>Description</i> 36
25	5.4.2	<i>Examples</i> 36
26		
27	6	ONGOING ENDEAVOURS 37
28		
29	6.1	MODEL DEVELOPMENT AND VALIDATION 37
30	6.2	MODEL GENERALISATION AND EXTENSION 38
31	6.3	INCLUDING INTRA AND INTER-FRACTION CHANGES 38
32	6.4	POTENTIAL APPLICATIONS OF ARTIFICIAL INTELLIGENCE 39
33	6.5	UNDERSTANDING PATHOPHYSIOLOGY 40
34	6.6	MODEL APPLICATION 40
35		
36		
37		
38		
39	7	CONCLUSION 41
40		
41	8	ACKNOWLEDGEMENTS 42
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		

1 **Abstract**

2 For decades, dose-volume information for segmented anatomy has provided the essential data for
3 correlating radiotherapy dosimetry with treatment-induced complications. Dose-volume information
4 has formed the basis for modelling those associations via normal tissue complication (NTCP) models
5 and for driving treatment planning. Limitations to this approach have been identified. Many studies
6 have emerged demonstrating that the incorporation of information describing the spatial nature of
7 the dose distribution, and potentially its correlation with anatomy, can provide more robust
8 associations with toxicity and seed more general NTCP models. Such approaches are culminating in
9 the application of computationally intensive processes such as machine learning and the application
10 of neural networks. The opportunities these approaches have for individualising treatment,
11 predicting toxicity and expanding the solution space for radiation therapy are substantial and have
12 clearly widespread and disruptive potential. Impediments to reaching that potential include issues
13 associated with data collection, model generalisation and validation.

14 This review examines the role of spatial models of complication and summarises relevant published
15 studies. Sources of data for these studies, appropriate statistical methodology frameworks for
16 processing spatial dose information and extracting relevant features are described. Spatial
17 complication modelling is consolidated as a pathway to guiding future developments towards
18 effective, complication-free radiotherapy treatment.

19 **Keywords:** radiotherapy, complications, modelling, dosimetry

20 **Word count:** ~15,500 (~22,000 with bibliography)

21

Accepted Manuscript

1 Introduction

In radiotherapy, the risk of treatment-induced toxicity is the limiting factor for dose escalation in pursuit of an increase in local control. The prediction of radio-induced side-effects guides the physician and the patient between treatment alternatives and enables treatment optimisation by integrating predictive models within computerised planning.

Radio-induced toxicity is classically linked to the dose-volume relationship, patient clinical parameters (such as medical history and adjuvant treatments) and intrinsic radiosensitivity. With steady increases in computational capabilities and increased efforts to gather and analyse relevant data (Deasy *et al.*, 2010), exploiting information from more available data with integrative approaches is now feasible.

The dose-volume toxicity relationship has been widely investigated. In 2010, the Quantitative Analysis of Normal Tissue Effects in the Clinic (QUANTEC) review summarized the three-dimensional dose/volume/outcome data to update and refine the related normal tissue tolerance guidelines (Marks *et al.*, 2010), initially provided by Emami *et al.* (1991). Dose-volume histogram (DVH) based normal tissue complication probability (NTCP) models attempt to condense the dose-volume information into a number that expresses the risk of a certain toxicity. Most NTCP models are phenomenological and have the advantage of being characterized by few parameters (typically ≤ 3). Different approaches have been historically developed to model NTCP, with the Lyman–Kutcher–Burman (LKB) model being one of the first and most commonly employed (Lyman, 1985). Even if prediction of toxicity and treatment plan evaluation with the NTCP-based models is still common practice, these kinds of models present limitations reducing their prediction capability. DVHs reduce the 3D (or even 4D) dose distribution within an organ to a unidimensional and discrete representation of the dose-volume relationship, inhibiting the ability of models to account for the actual underlying complexity.

Spatial NTCP models have sought to geometrically represent the 3D dose distribution. This allows information on the pattern as well as the amount of dose to be characterised. Recent spatial NTCP models have sought to geometrically represent the 3D dose distribution in a single coordinate system via a spatial normalisation for a joint analysis of dose at the lowest sampling scale (pixel and voxel levels, referred to from here as “pixel-wise” in 2D and “voxel-wise” in 3D) (e.g., (Marcello *et al.*, 2020a; Mylona *et al.*, 2020b; Palma *et al.*, 2020a; Jiang *et al.*, 2019; Palma *et al.*, 2019b)). These low spatial-scale methodologies have allowed the unravelling of the local dose-effect relationship across a population at each single voxel in a common coordinate system in different organs. Models

1
2
3 54 can also be created by defining and analysing spatial features of the 3D dose distribution (e.g.
4 55 Buettner *et al.* (2012b)) or abstractions such as the dose surface map (DSM) (e.g. Heemsbergen *et al.*
5 56 (2020)). These spatial methods, and others described below, have been pursued to improve
6 57 prediction and classification. Such models may also facilitate identification of the underlying
7 58 aetiology of radio-induced injury and be used to improve patient-specific treatment planning. They
8 59 are likely to reduce toxicity (Drean *et al.*, 2016b; Lafond *et al.*, 2020), and may one day inform or
9 60 help validate *in silico* models of treatment toxicity (e.g. Cicchetti *et al.*, 2020)).

16 61 The goal of this review is to describe these recent spatial dose-effect investigations and NTCP
17 62 models and provide some guidance around their development.

21 63 **2 Strategies for characterising dose distributions**

24 64 **2.1 Dose-volume approaches**

26 65 The concept of the dose-volume relationship of a defined region of interest became commonplace
27 66 when both 3D dose computation and 3D segmentation (“contouring”) of regions became practical.
28 67 The cumulative DVH synthesises the dose vs volume relationship as a function representing the
29 68 percentage of volume that receives at least a certain dose.

34 69 *2.1.1 The advantages of dose-volume approaches*

35 70 The primary advantage of the dose-volume approach is linked to the wealth of knowledge obtained
36 71 through prior studies of radiation and the resounding clinical success of such approaches. Today’s
37 72 radiation therapy is driven by dose-volume constraints based on the results of published studies and
38 73 meta-analyses. So much so, that today’s dose distributions contain little information outside the
39 74 bounds of these dose volume parameters, as they are controlled for in clinical practice.

45 75 Dose-volume metrics are easily understood and are based on the natural compartmentalisation of
46 76 the body into organs. Reporting of them can be reduced to a table of numeric entries representing
47 77 the quality of the complex 3D treatment plan. To even further simplify their presentations, software
48 78 applications have reduced them to colour codes to indicate alerts when a plan may violate one of
49 79 the treatment goals. In busy clinics, this facilitates rapid evaluation. Dose-volume metrics are also
50 80 convenient when defining the goals for optimization in inverse treatment planning.

56 81 Radiobiological models have been developed to bridge the gap between the physical dose-based
57 82 objectives to drive treatment planning and the clinical dose goals reflecting the toxicity risks. Those
58 83 commonly studied, such as NTCP, tumour control probability (TCP) and the complication-free

1
2
3 84 tumour control probability (P+) (Källman *et al.*, 1992), have typically been designed to operate on
4
5 85 DVH information.
6
7

8 86 2.1.2 *The disadvantages of dose-volume approaches*

9 87 Fundamentally the DVH assumes that every sub-volume of the region is of equal importance to the
10 88 function of that tissue and is equally sensitive to radiation dose. Realistically, the segmented regions
11 89 in radiotherapy are typically bulk anatomy and do not reflect the microstructure of anatomy that
12 90 may be impacted by radiation. Therefore, the DVH may be too coarse of a feature to adequately
13 91 model the impact radiation may have on the anatomy.
14
15
16
17
18

19 92 The assumption that each element of tissue is equally important to the NTCP function and equally
20 93 sensitive to radiation dose is simply not true for many anatomical regions typically segmented in
21 94 treatment planning. For example, the parotid glands consist of acinar cells producing saliva and a
22 95 ductal region that carries the saliva to the oral cavity. Similarly, a kidney is made up of several cell
23 96 types and structures. In other cases, such as the oesophagus and rectum, organ structure consists of
24 97 a mucosal layer surrounded by muscle tissue. These structures may have different risks when the
25 98 dose is high to the entire circumference of the structure versus when it has the same volume of dose
26 99 oriented longitudinally along the structure. Understanding the true causal relationships between
27 100 radiation dose and normal tissue dysfunction is limited with dose volume metrics that are naive to
28 101 the detailed components of the anatomy.
29
30
31
32
33
34
35

36 102 The spinal cord has a complex spatial arrangement of functional sub-units (FSUs - compartments
37 103 that accomplish part of an organ's function), and subsequently a complex inter-relationship with
38 104 overall organ function. Precise pre-clinical experiments performed by Bijl *et al.* (2003) identified
39 105 large variations in dose-volume based predictors of paralysis in rats when the spatial patterns of
40 106 irradiation were changed. Conventionally, simple maximum cord dose has been used to predict
41 107 subsequent complications (Kirkpatrick *et al.*, 2010). However, the inhomogeneity in irradiation now
42 108 afforded with stereotactic spinal radiosurgery exceeds the predictive capability of dose-volume
43 109 analysis (Medin and Boike, 2011). Similarly, models of lung complication had focused on their
44 110 parallel-like nature and mean lung dose (MLD) had conventionally been used as a principal predictor.
45 111 However, evidence for more localised dose-response in humans emerged nearly two decades ago
46 112 (Seppenwoolde *et al.*, 2004), following extensive animal experiments (as well reviewed by Voshart *et*
47 113 *al.* (2021)).
48
49
50
51
52
53
54
55
56
57

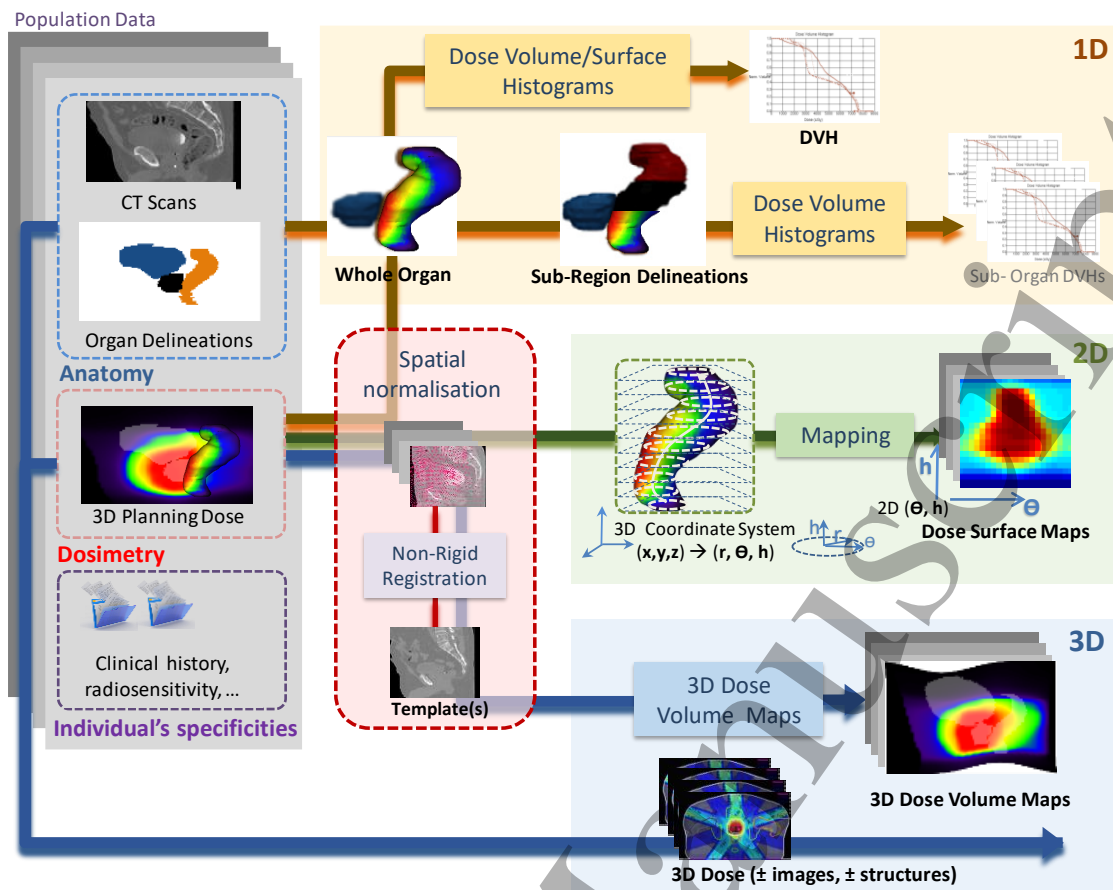
58 114 Further, analysis has traditionally been limited to dose-volume metrics for single organs. Many
59 115 human functions involve multiple components of anatomy. Swallowing, for example involves many
60

1
2
3 116 muscles in the head and neck region. These muscles may be able to compensate for one another and
4
5 117 the impact of a dose pattern across the set of muscles and its impact on swallowing can be quite
6
7 118 complicated. In many cases, a significant portion of the anatomy is not contoured at all, and the
8
9 119 dose-volume metrics can only be computed for contoured regions. Contouring with high spatial
10
11 120 detail in routine workflow remains burdensome. Similarly, in a shift of spatial focus for dose-toxicity
12
13 121 association, the impact on lung toxicity from cardiac irradiation has been identified (Tucker *et al.*,
14
15 122 2014; van Luijk *et al.*, 2005).

16 123 Multiple spatial dose distributions (an essentially infinite number) will yield the same or similar DVH.
17
18 124 Dependence on a dose-volume approach requires an assumption that all those distributions will lead
19
20 125 to the same toxicity – the problem of *degeneracy*. Conversely, dose-volume derived NTCP models
21
22 126 from studies involving specific irradiation techniques will have been derived with minimal variation
23
24 127 in DVH between patients. Extrapolation of DVH and NTCP metrics beyond the specific context in
25
26 128 which they were derived is known to be dangerous. Due to this limitation, as well as many other
27
28 129 sources of variations between cohorts, DVH-based complication models derived for one treatment
29
30 130 approach tend not to be applicable to alternative irradiation strategies in the same sites (Troeller *et*
31
32 131 *al.*, 2015).

32 132 **2.2 Approaches that preserve spatial information**

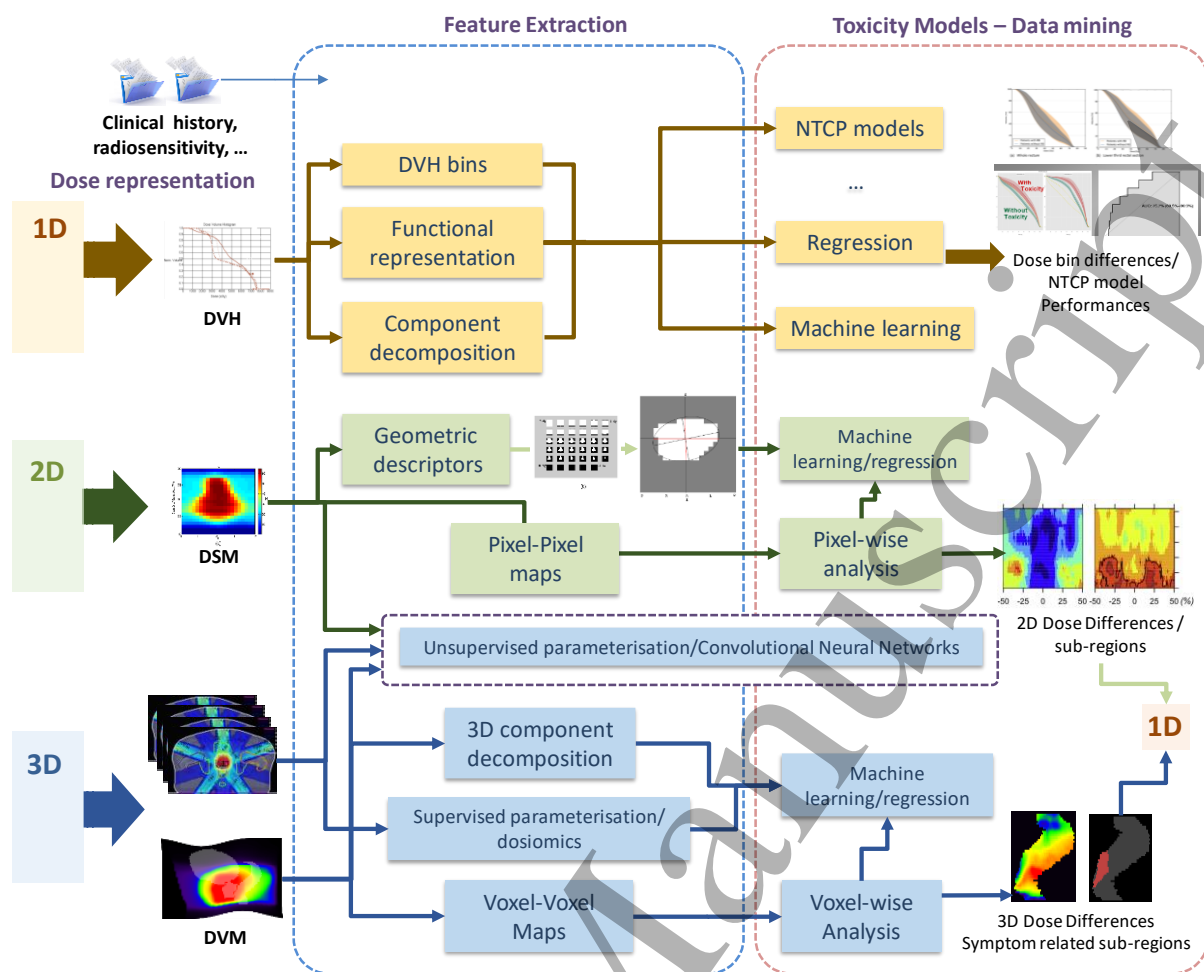
33 133 To overcome the limitation of whole organ DVHs, recent approaches have investigated the existence
34
35 134 of spatial signatures of dose distributions across dimensionalities and at diverse spatial scales. Here
36
37 135 we describe the processing of treatment planning data (Figure 1) required to achieve extraction of
38
39 136 features describing spatial distributions at the various spatial scales and development of subsequent
40
41 137 toxicity models (Figure 2). Practical applications of these features and models are described in
42
43 138 Section 5.



139

140 Figure 1. Processing workflows for preparing data for toxicity modelling across dimensionalities. Orange, path for
 141 histogram development (1D data); green, path for 2D dose surface maps (2D data); and blue, path for 3D dose volume
 142 maps (3D data). Some data sources and processes may not be used in all approaches, and these are indicated with
 143 dashed borders.

Accepted Manuscript



144

145 **Figure 2.** Data flow in the extraction of dosimetric features and construction of toxicity models. Features extracted from
 146 1D, 2D or 3D data are exploited following different strategies, leading to different kinds of predictive models (NTCP,
 147 machine learning or general regression). 2D DSM and 3D DVM models may require the entire population dose to be
 148 mapped to a single coordinate system before being analysed.

149 2.2.1 1D precision dose-volume approaches

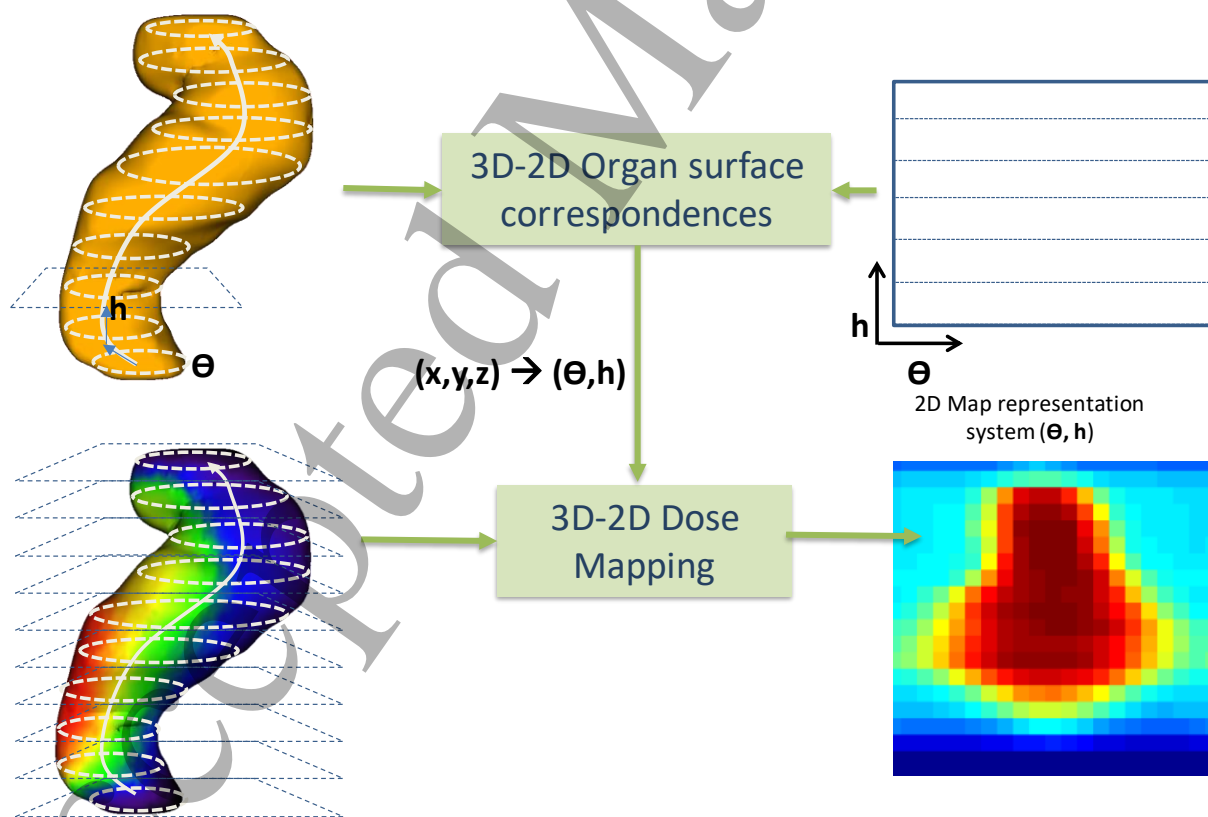
150 The simplest approach is to identify a more precise sub-region of the organ where dosimetry and
 151 DVH metrics are most correlated with outcome. Improvements in NTCP models, and evidence of
 152 correlations between local dose and side-effects, have been provided by undertaking DVH analysis
 153 (or analysis with related histogram information) at spatial scales below the organ level. Partitioning
 154 the organs for computing sub-region DVHs for example has demonstrated a sub-anatomical
 155 dependence for specific toxicities (Ebert *et al.*, 2015b; Heemsbergen *et al.*, 2005; Peeters *et al.*,
 156 2006b; Stenmark *et al.*, 2014). The question that may arise is whether the organ partitions are
 157 anatomically-equivalent across individuals allowing DVH comparisons. If sub-region partitions
 158 between patients are generated following the same geometrical criteria, then they can refer to the
 159 same anatomico-physiological regions. The identification of correlative regions can be derived

60

160 manually (e.g. Gulliford *et al.* (2017)), or by identifying clusters of correlated pixels and voxels in 2D
 161 and 3D representations (e.g. Drean *et al.* (2016b)). DVH-based features of those sub-regions can be
 162 used to validate their association with complications.

163 2.2.2 2D surface mapping

164 Spatial considerations on the distribution of dose to an organ surface can be achieved with dose
 165 surface mapping (DSM). DSMs present a virtual unfolded planar representation of the dose
 166 distribution across an organ wall. Such mapping has been implemented following different strategies
 167 (Sanchez-Nieto *et al.*, 2001; Hoogeman *et al.*, 2004; Munbodh *et al.*, 2008; Tucker *et al.*, 2006b;
 168 Witztum *et al.*, 2016). A 2D image is constructed via parametric mapping from the 3D coordinate
 169 system of the organ wall. The general idea is depicted with a rectal DSM in Figure 3, where a direct
 170 relationship exists between the 3D cylindrical coordinates and the 2D (θ, h) space. Thus, each pixel in
 171 2D corresponds to a portion of the organ wall where the dose is mapped and propagated by
 172 interpolation. Once constructed, dose surface maps can be used to undertake “pixel-wise” analysis
 173 (Yahya *et al.*, 2017), or parameterised using geometric descriptors such as lateral and longitudinal
 174 extent (Buettner *et al.*, 2009b) or texture features (Chen *et al.*, 2018).



175
 176 **Figure 3: Construction of a dose surface map (DSM) for the rectum by establishment of a direct relationship between the**
 177 **3D coordinate system and a planar $O(\theta, h)$ space.**

1
2
3 178 Crucial aspects in this construction are the definition of the origin (i.e. 0,0) and the resolution and
4 size of the 2D images. If the rectum was the organ to be studied (e.g. (Buettner *et al.*, 2009a;
5 179 Moulton *et al.*, 2017)) a cylindrical coordinate system for building the DSM has been used. In
6 180 Buettner *et al.* (2009b) the contour was thus cut at the posterior-most position on each CT-slice and
7 181 unwrapped to a map of 21x21 pixels. Witzum *et al.* (2016) raised some of the issues concerning
8 182 tortuous structures. They developed a raytracing approach to create dose surface maps for the
9 183 duodenum accounting for the bend in the structure, following an inner path.
10 184

11
12
13
14
15
16 185 In other hollow organs such as the bladder a similar slice-based methodology has been applied. In
17 186 works from Palorini *et al.* (2016a) and Yahya *et al.* (2017), 1 mm-resolution DSMs were generated
18 187 (cranial-caudal direction), by virtually cutting bladder contours at the points intersecting the sagittal
19 188 plane passing through its centre-of-mass. Because of the large inter-individual bladder variability
20 189 some issues arise when having large and small bladders to map together for population analysis or
21 190 where some parts of the bladder are not equally mapped. In Mylona *et al.* (2020a) this was
22 191 addressed with an anisotropic vertical interpolation to the smallest bladder, aligned at the bladder
23 192 base.

24 193 2.2.3 3D feature extraction

25 194 It is feasible to reduce the complex 3D voxel-level dose information to a smaller number of features
26 195 via an appropriate spatial parameterisation. One such approach is to describe the spatial distribution
27 196 within an organ via 3D moments (Buettner *et al.*, 2012b; Dean *et al.*, 2016). Alternatively, borrowing
28 197 from the world of imaging analytics, supervised descriptions can be obtained via spatial texture
29 198 features (“dosiomics” (Liang *et al.*, 2019; Rossi *et al.*, 2018), “dosomics” (Placidi *et al.*, 2020) or
30 199 “radiomorphology” (Jiang *et al.*, 2019)), or unsupervised learning can be employed via neural
31 200 networks.

32 201 2.2.4 3D volume mapping

33 202 At a fine scale, dose-outcome correlations can be investigated at the voxel level. For voxel-wise
34 203 comparisons to be meaningful, anatomical correspondence across the individuals must be ensured.
35 204 This pre-processing step is referred to as “spatial normalisation”, whose goal is to define geometrical
36 205 transformations aimed at registering and resampling inter-individual anatomies and doses into a
37 206 common coordinate system as depicted in Figure 1 (e.g. (Monti *et al.*, 2020; Acosta *et al.*, 2013;
38 207 Rigaud *et al.*, 2019; Acosta and De Crevoisier, 2019)). This 3D-3D dose mapping to a common
39 208 coordinate system to create a dose-volume map (DVM) remains challenging. Such mapping may be
40 209 obtained via a parametric representation of the anatomy in a spherical or cylindrical coordinate
41 210

210 system as in Chen *et al.* (2013). It may be more precisely computed through existing non-rigid
211 (deformable) registration methods (McWilliam *et al.*, 2017; Monti *et al.*, 2018; Marcello *et al.*,
212 2020a), or tailored to a particular anatomy as proposed for the rectum in Drean *et al.* (2016a), or for
213 the bladder in Mylona *et al.* (2019) using spatial descriptors. Depending on the investigated
214 anatomical site, organ-driven registration methods may be more precise than the ones based on
215 intensity levels. This is the case in Acosta *et al.* (2013), Drean *et al.* (2016a) and Mylona *et al.* (2019)
216 where anatomical mapping based on 3D structural models of the considered organs were proposed.
217 These approaches require, nevertheless, a precedent segmentation of some of the considered
218 structures such as the urethra (Acosta *et al.*, 2017). However, when inter-individual registration is to
219 new patients without identified structures or is to be structure-agnostic, image information alone
220 must be used (McWilliam *et al.*, 2017; Monti *et al.*, 2018; Abravan *et al.*, 2020).

221 The 3D spatial normalisation approach can also be used to align anatomy for derivation of DSMs,
222 especially in the case of pixel-wise analysis, or for the purpose of sub-region identification.

223 **3 Practical Considerations**

224 The development of spatial response models places specific demands on the nature of technical data
225 collected for their construction. When interpreting, utilising or publishing a spatial complication
226 model, factors impacting the underlying technical data should be considered and appropriately
227 reported and accommodated (see section 4.2.3). The relevance and quality of patient outcome data
228 is of similar or even greater importance for the derivation of useful models. Additional data types
229 can constitute modifying and stratifying co-variables, such as patient demographics and co-
230 morbidities, disease staging, treatment characteristics (techniques, timing, adjuvant treatments),
231 pathologic and genetic information.

232 **3.1 Required technical data**

233 Due to the computational nature of spatial models, it is assumed that required data will be available
234 in digital form which could be arbitrary in-house, native or proprietary formats, or more generally in
235 prescribed formats such as Digital Imaging and Communications in Medicine (DICOM) (NEMA),
236 Neuroimaging Informatics Technology Initiative (NiftI) (NiftI, 2020) or Nearly Raw Raster Data
237 (NRRD) (SourceForge, 2020). The three principal technical data ingredients for model development
238 are briefly described below.

239 **3.1.1 Imaging**

240 Anatomical imaging typically provides the reference space for derived models, guides definition of
241 segmented structures and facilitates intra- and inter-individual registration. The robustness of spatial
242 models can depend significantly on the sensitivity and specificity of imaging, particularly through
243 influence on the definition of structures (e.g. (Roach *et al.*, 2019)).

244 3.1.2 Structures

245 Many of the processes for characterising spatial dose distributions presented in Sections 2.2 and 5
246 operate on information related to anatomical and functional structures. The definition of such
247 structures can be made manually by observers at the time of patient treatment planning or manually
248 through retrospective review of collated data. Alternatively, autosegmentation routines utilising
249 anatomical atlases (Kennedy *et al.*, 2019) or artificial intelligence approaches (Fu *et al.*, 2020) can be
250 used. Structure segmentation can represent a significant source of uncertainty in the derivation and
251 application of models, with multiple contributing factors:

- 252 • Geometric variability: The location and extent of structures will depend on multiple factors
253 relating to image quality, image sensitivity and specificity, inter-observer variability (e.g.
254 (Roach *et al.*, 2019)), organ deformation and motion (e.g. (Palorini *et al.*, 2016a)), errors and
255 limitations in image registration, bias propagated from atlas definitions or neural network
256 learning environments or selection of a patient template (see Section 2.2) (Acosta *et al.*,
257 2010).
- 258 • Structure definition: A common source of undesired variability, particularly when pooling
259 data sources or during validation, is variable definition of anatomical structures (e.g.
260 (Nitsche *et al.*, 2017)). Models need to operate on like-definitions. Variability and ambiguity
261 can be reduced through the use of consensus definitions, reviews of definitions such as
262 within the QUANTEC reports (Bentzen *et al.*, 2010), or published standards (Wright *et al.*,
263 2019).
- 264 • Structure naming: Structure naming can often be problematic for scripting model
265 development, particularly when data comes from multiple institutions. This can be
266 ameliorated through use of naming conventions (e.g. (Mayo *et al.*, 2018; Santanam *et al.*,
267 2012)) or ontologies (Phillips *et al.*, 2020). Note that spatial models may utilise or give rise to
268 non-standard structures (such as predictive clusters identified in DSMs and DVMs).

269 3.1.3 Dose

1
2
3 270 As indicated in Figure 1 and Figure 2, access to multi-dimensional descriptions of dose distributions,
4
5 271 or features derived from them, represents a common minimum level of required technical data.
6
7 272 When deriving and applying spatial models, several aspects of these data should be considered:

- 8
9 273 • Accuracy: Although dose distributions are frequently available based on planned or intended
10
11 274 treatments, correct models will be based on dose distributions which have been verified or
12
13 275 accumulated as delivered (e.g. (Shelley *et al.*, 2017; Jaffray *et al.*, 2010)). Accuracy should
14
15 276 ideally have been assessed independently, such as via participation in credentialing exercises
16
17 277 (e.g., (Ebert *et al.*, 2011; Molineu *et al.*, 2013; Weber *et al.*, 2014)). Deformations of dose,
18
19 278 due perhaps to the intra-individual accumulation process (Tilly *et al.*, 2013) or inter-
20
21 279 individual co-registration (see Section 2.2.3) will impact on the accuracy of dose
22
23 280 representation.
- 24
25 281 • Precision: Spatial resolution in the description of dose can impact the ability to precisely
26
27 282 represent the underlying response effects. The resolution of dose calculation has previously
28
29 283 been shown to impact even dose-volume based models (Ebert *et al.*, 2010; Kim *et al.*, 2018).
30
31 284 Variation in resolution can have a moderate impact on dosimetric texture features (Placidi *et*
32
33 285 *al.*, 2020). With an increasing need to develop models for precision stereotactic treatments,
34
35 286 precise descriptions of steep dose gradients across spatially-limited structures are required
36
37 287 (e.g. (Ryu *et al.*, 2007; Hrycushko *et al.*, 2019; Gale *et al.*, 2017; Kim *et al.*, 2014)).
- 38
39 288 • Completeness: Dose calculations are often limited in extent relative to potentially-involved
40
41 289 anatomy, such as when based on cone beam CT data obtained with accelerator-mounted
42
43 290 imaging systems. This can inhibit spatial models, particularly those relating to low-doses
44
45 291 over extensive regions of anatomy.
- 46
47 292 • Temporal features: Dose fractionation, inter- and intra-fraction dose temporal patterns can
48
49 293 impact complication incidence (Dörr, 2015). Changes in response due to variable dose-per-
50
51 294 fraction, either between voxels or due to variable treatment phases, may need to be
52
53 295 incorporated into the model. Such variations may also be accounted for using equieffective
54
55 296 dose estimates (Bentzen *et al.*, 2012), noting that this leads to spatial discontinuities where
56
57 297 parameters vary between tissues. The complexity of temporal dose effects increases
58
59 298 significantly when intra-treatment variations due to organ motion or the pharmacokinetics
60
61 299 of radionuclide deliveries are considered.

300 3.1.4 Treatment description

1
2
3 301 Treatment factors, such as patient set-up at imaging and treatment, patient preparation, the use of
4 302 immobilisation and fixation devices, may be co-variables of importance to the specificity of a model.
5
6 303 This information is often not captured in DICOM fields or through oncology information systems.
7
8

9 304 **3.2 Outcome Data**

10 305 Outcome information, providing the known output for a model (the “endpoint” or “event
11 306 incidence”), comes in diverse forms. For complication outcome, we are typically concerned with
12 307 organ-specific symptoms of radiation injury which may manifest over months or years. These can be
13 308 graded at discrete (ordinal) levels using standardised clinician- or patient-reported instruments such
14 309 as provided by the Common Terminology Criteria for Adverse Events (Trotti *et al.*, 2003) developed
15 310 by the United States (US) National Cancer Institute (NCI), instruments developed in-house or by
16 311 various international collaboratives. The trend is towards the use of patient-reported complications
17 312 for outcome. This is because the severity of symptoms are often under-reported by clinicians (Xiao
18 313 *et al.*, 2013), and follows recognition of the importance of focussing on symptoms with the most
19 314 impact on patients’ quality of life. Although definitions can vary, complications are typically graded
20 315 according to indicative symptoms and required interventions (GX – Grade X):
21
22
23
24
25
26
27
28
29

- 30 316 • G0 – symptoms are absent
- 31 317 • G1 – the complication is mild and no interventions are required
- 32 318 • G2 – the complication is moderate and some local intervention might be required
- 33 319 • G3 – the complication is severe and intervention is required, though is not life-threatening
- 34 320 • G4 – the complication is life-threatening and major intervention is required
- 35 321 • G5 – the complication has caused death

36 322 Whilst some models can utilise continuous outcomes, for NTCP models it is common to convert
37 323 measures to a binary endpoint classification. These may be either determined at fixed time-points
38 324 following treatment, as incidence at any time during follow-up, or the time-to-event incidence if
39 325 temporal features can be incorporated in the model. The definition, interpretation, collection and
40 326 application of complication outcome measures are notorious sources of uncertainty in outcome
41 327 modelling. Multiple factors should be kept in mind related to model accuracy and generalisability:
42
43
44
45
46
47
48
49
50
51

- 52 328 • Specificity of the included patient cohort.
 - 53 329 • The relevance of an outcome to patient quality-of-life.
 - 54 330 • Variations in scoring mechanisms and criteria.
 - 55 331 • Variations in follow-up time or time between measurements.
- 56
57
58
59
60

- 1
2
3 332 • The identification and influence of comorbidities, concurrent treatments or pre-existing
4 333 morbidity.
- 5
6 334 • The influence of social and/or technical factors on measures.
- 7
8 335 • The nature of the data source, as discussed below.
- 9
10

11 336 3.3 Data sources

12
13 337 When considering sources of data for spatial complication models, we can consider the ability of
14 338 those sources to meet specific criteria for development of generalizable, robust and powerful
15 339 models. A source should provide large volumes of high quality, well-curated data for patients with
16 340 diverse characteristics and treated with diverse techniques (noting that data diversity can lead to
17 341 unexpectedly biased results, as discussed in Section 6.1). The sub-optimal performance of many
18 342 radiotherapy outcomes models can largely be blamed on the paucity and lack of diversity of
19 343 available data (Luo *et al.*, 2020).

20
21
22
23
24
25
26 344 Table 1 lists specific criteria, provides some examples of sources and attempts to describe, via
27 345 generalisations, how likely each source is to meet the criteria. In Table 1, *quality* infers the
28 346 completeness, accuracy and consistency of technical and outcome data. *Diversity* relates to the
29 347 variability in studied populations, radiotherapy technique and overall patient treatment, including
30 348 trial vs non-trial contexts (Chen *et al.*, 2016; Krauss, 2018). Diversity also pertains to inter-individual
31 349 variations in spatially-localised dose (note also the implications of diversity for model
32 350 generalizability, as discussed in Section 6.2).

33
34
35
36
37
38
39 351 Some points to note in relation to Table 1:

- 40
41 352 - **Single-institution studies** enable ready access to appropriate high-quality data though with
42 353 minimal variability and typically only small patient numbers. Collated data is rarely made
43 354 available outside the institution.
- 44
45
46 355 - **Multi-centre clinical trials** often employ rigorous data collation. However, such trials will
47 356 rarely be statistically powered specifically for the purpose of spatial response modelling and
48 357 so the sample size may be insufficient. Software systems developed over the last couple of
49 358 decades, both in-house and commercially, have facilitated quality assessment of technical
50 359 data by multicentre trials groups (e.g. (Ebert *et al.*, 2010; Deasy *et al.*, 2003; La Macchia *et al.*,
51 360 2012; Meroni *et al.*, 2019; Roelofs *et al.*, 2014; Deasy and Adita, 2013; Purdy, 2008;
52 361 Purdy *et al.*, 1998)). Although the quality of clinical trial data can be advantageous, variations
53 362 from trial conditions in the clinic, including participant selection, can bias model predictions
54 363 relative to non-trial practice (Ohri *et al.*, 2013).

- 364 - **Data pooling and publication.** International policies are trending towards data availability
 365 and interoperability (e.g. (Hayman *et al.*, 2019; Taichman *et al.*, 2017)). In Table 1 we
 366 distinguish “public” pooling and publication, such as provided by the Cancer Imaging Archive
 367 (www.cancerimagingarchive.net, (Clark *et al.*, 2013)), from “private” pooling, such as might
 368 be achieved via manufacturer-led knowledge base collaboratives and user-communities.
 369 Both public and private data pools have the potential for development of large cohorts with
 370 data variability, though data quality may be ambiguous if not well documented.
- 371 - **Federated data access** can enable accessing large patient cohorts spanning multiple
 372 repositories, including clinical systems at individual treatment centres. Ethical and socio-
 373 political issues can be minimised if model parameters can be estimated for data at each site,
 374 before being combined centrally (Deist *et al.*, 2017). Although no published evidence was
 375 found that spatial complication models have been derived through this approach, the
 376 potential for validation of developed models is significant.

377

378 **Table 1: Potential sources of data for spatial models and their ability to meet desirable criteria for forming**
 379 **statistically-powerful, generalisable models that meet current standards for validation and translation.**

Source	Technical data quality	Outcome data quality	Variability/ diversity	Sample size	FAIR	Facilitates validation
Single institution studies	High	High	Low	Low	Low	Low
Multicentre clinical trials	High	High	Medium	Medium	High	Medium
Public data pooling and publication	Variable	Variable	Medium	Medium	High	High
Private data pooling	Variable	Low	Medium	High	Medium	Variable
Federated data access	Low	Low	High	High	Variable	High

380 ^a Including dosimetric accuracy.

381 ^b A data source will have a high ability to satisfy this criterion if it meets the FAIR principles
 382 (Wilkinson *et al.*, 2016) defined by the FORCE 11 (Future of Research Communications and e-
 383 Scholarship) community, of data being findable, accessible, interoperable and re-usable.

384 ^c e.g., Treatment planning system manufacturer-facilitated knowledge base consortia.

385 4 Statistical and Modelling Considerations

386

387

388

389

390

391

392

386 A central aim of using spatial dose descriptors to model dose-complication is to reduce the impact of
387 degeneracy relative to dose-volume approaches. It is important that the process utilised maintains
388 the principles associated with robust, unambiguous statistical analysis and interpretation. Here we
389 summarise such relevant considerations.

390 **4.1 Feature Selection**

391 *4.1.1 Candidate dosimetric features and collinearity*

392 An important step in developing phenomenological NTCP models (van der Schaaf *et al.*, 2015) is to
393 start off with a list of potential prognostic factors based on the literature and underlying
394 radiobiological assumptions (e.g. assumed α/β ratio). This can reduce the number of false positive
395 findings and guide the feature reduction process (Palma *et al.*, 2020a; Heinze *et al.*, 2018). The
396 inhomogeneous physical dose distribution can be aggregated into dose features (Figure 2) that
397 represent the biological dose received and are predictive for the toxicity endpoint of interest. The
398 result may be just a small number of features as derived from a spatial parameterisation. However,
399 hundreds to thousands of dose features can be retrieved from a spatial voxel-wise 3D dose
400 distribution, even though the sample size may be quite limited, and collinearity is likely. Candidate
401 prognostic factors selected from a group of correlated variables are typically those that have the
402 highest predictive power at univariate analysis compared to the correlated variables that are a priori
403 excluded. A general rule of thumb is that correlation between candidate variables for a multivariable
404 model should be below ≈ 0.7 (El Naqa *et al.*, 2009; Schaake *et al.*, 2016).

405 *4.1.2 Feature reduction*

406 The generally accepted rule of thumb is that regression models should be used with a minimum of
407 10 “events per variable” EPV (Peduzzi *et al.*, 1996). This rule has been criticized as being too strict -
408 Vittinghoff and McCulloch (2007) instead recommend a minimum of 7 EPV. After pre-processing the
409 dataset to a candidate list of features considered for modelling, a variable selection algorithm must
410 be chosen (Heinze *et al.*, 2018; Steyerberg and Vergouwe, 2014). Valid approaches to reduce the
411 number of features (and clinical co-variates) to the most predictive in a multivariate model are: 1)
412 select variables for the final multivariable model based on their univariate model estimates, using a
413 p value threshold; 2) backward and forward selection tools like Wald, Likelihood Ratio and
414 conditional regression methods; and 3) the LASSO method (least absolute shrinkage and selection
415 operator) which is a logistic regression analysis with a penalty for the magnitude of the regression
416 coefficients to prevent overfitting (Tibshirani, 1996; Buettner *et al.*, 2011; Gabryś *et al.*, 2018).
417 Consideration can be given to reduction of features through use of their principal components

1
2
3 418 (e.g.(Chen *et al.*, 2011)). Additionally, feature selection can be combined with the method to
4
5 419 determine association with outcome through algorithms such as random forest, and through the
6
7 420 stability of features in associative models derived from sampled sub-sets of the full data (i.e.
8
9 421 “bootstrapping”). Adequate feature reduction is vital to ensuring the ability for a model to be
10
11 422 generalised. An excellent overview of techniques is provided in Guyon and Elisseeff (2003).

12 13 423 *4.1.3 Co-variates*

14 424 The inclusion of clinical factors in NTCP models may improve the predictive power of the model
15
16 425 considerably (Defraene *et al.*, 2012; Morimoto *et al.*, 2019; Rancati *et al.*, 2011; Dean *et al.*, 2017;
17
18 426 Palma *et al.*, 2020b). A preselection of all treatment- patient- and tumour- related factors by an
19
20 427 educated guess is needed to avoid false positive results. For this purpose, a literature search is
21
22 428 recommended to define candidate clinical factors to be considered subsequently in model building
23
24 429 (Steyerberg and Vergouwe, 2014).

25 26 430 *4.1.4 Models and Algorithms*

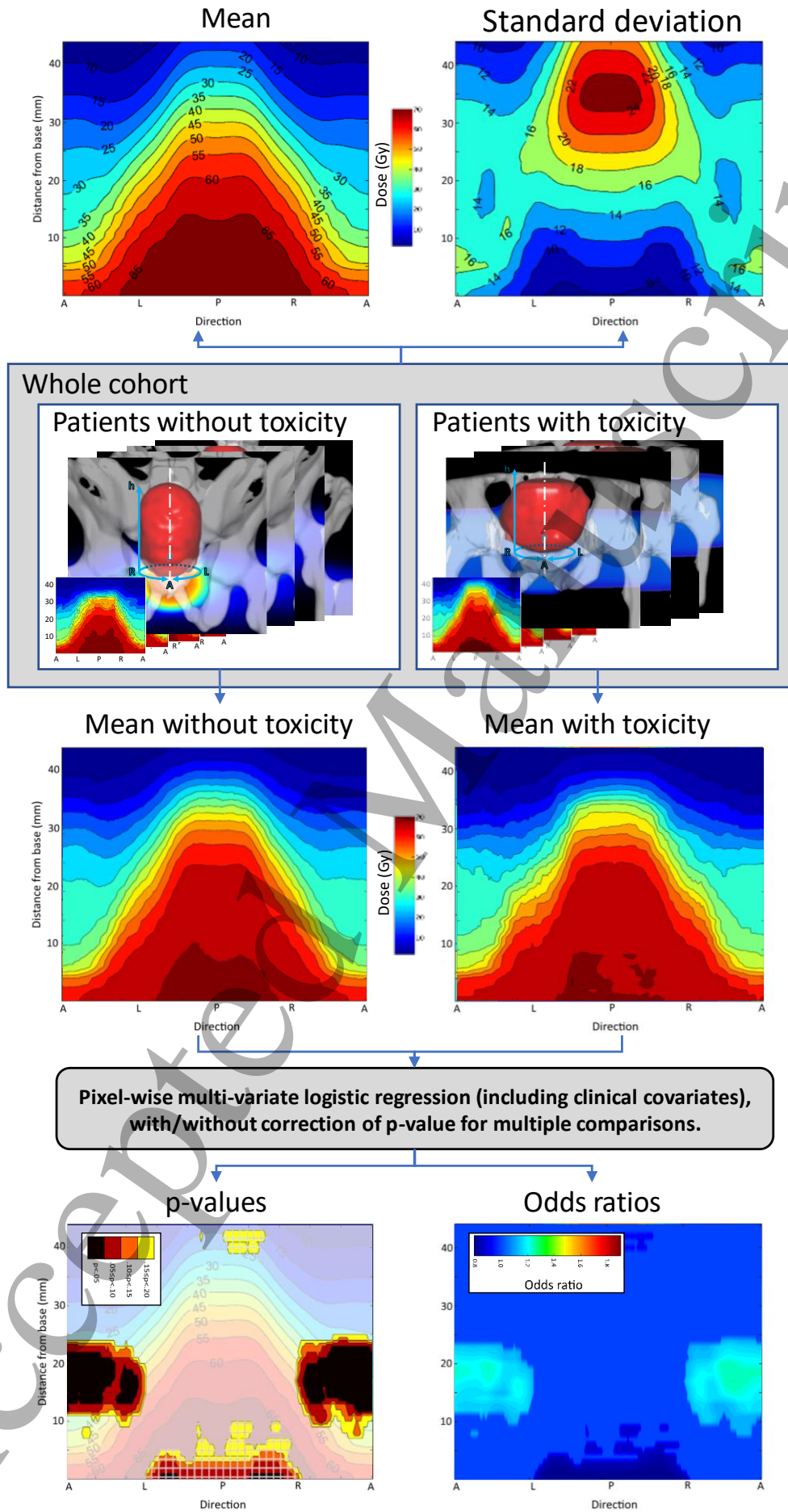
27 431 To parametrize the dose-dependence of an organ at risk, typically a sigmoid-shaped function is
28
29 432 fitted, like the LKB model, the Relative Seriality (RS) model, and the general logistic regression model
30
31 433 (Trott *et al.*, 2012). It has been demonstrated that the general applied logistic regression technique
32
33 434 produces very similar dose-effect curves as the LKB and RS model (Defraene *et al.*, 2012). A
34
35 435 prerequisite is that the type and pattern of toxicity (i.e. the dependent variable) has to be translated
36
37 436 and captured in a ‘present (1)/not present (0)’ for logistic regression modelling.

38
39 437 As an alternative in the current information age, data mining and machine learning approaches for
40
41 438 toxicity prediction research are increasingly applied (Robertson *et al.*, 2015; Beasley *et al.*, 2018;
42
43 439 Gabryś *et al.*, 2018; Luo *et al.*, 2020; Dean *et al.*, 2018; Palma *et al.*, 2019a). Commonalities and
44
45 440 differences between the more conventional methods of model-based statistical inference and the
46
47 441 rapidly progressing field of data driven machine learning have given rise to an active debate (c.f. the
48
49 442 field of imaging in neuroscience (Bzdok, 2017)). It has been shown that machine learning approaches
50
51 443 do not, by default, result in better predictions (Yahya *et al.*, 2016; Dean *et al.*, 2018). Unsupervised
52
53 444 machine learning approaches aim to produce accurate predictions for unseen data based on a large
54
55 445 body of training data, and do not depend on tractable relations between variables, which can limit
56
57 446 sensible extrapolation of the associated models. Conventional regression, on the other hand, may
58
59 447 reveal the specific dependence of a given variable on several independent variables within a data
60
448 set. From this comes the opportunity to extrapolate beyond the initial model fitting, beyond the
449 initial conditions under which data were acquired, by adaptation.

1
2
3 450 Selection of the appropriate statistical test(s) depends on the nature of the predicted outcome. If
4
5 451 time to event is considered important, parameters of a proportional hazards model may be
6
7 452 inspected (provided proportionality of the hazard is valid), or e.g. accelerated failure time models
8
9 453 may be employed (Bradburn *et al.*, 2003). On the other hand, when fixed time point differences or
10
11 454 incidence over multiple time points are considered sufficiently descriptive, parametric t-tests or
12
13 455 nonparametric signed-rank tests can be performed (Lumley *et al.*, 2002). Rather than to seek
14
15 456 rejection of a null-hypothesis, Bayesian analysis may provide a more informative description of
16
17 457 observed differences (Kruschke, 2013).

18 458 *4.1.5 Voxel-wise models*

19 459 Although conventional statistics can be applied at a pixel-wise or voxel-wise level, a comparison of
20
21 460 the aggregated data dichotomised by endpoint is a commonly used approach. Detailed descriptions
22
23 461 and formalisms of the process for voxel-wise analysis for toxicity studies are provided by Acosta and
24
25 462 De Crevoisier (2019) and Palma *et al.* (2020a). The idea of identifying local dose-response patterns
26
27 463 by voxel-wise analysis based on two-sample tests was derived from neuroimaging studies where the
28
29 464 aim is to discover voxel-wise changes due to a specific disease (Ashburner and Friston, 2000;
30
31 465 Whitwell, 2009). When comparing DSMS/DVMs, the null-hypothesis is that there is no difference
32
33 466 between the dose distributions of the patients with and without toxicity, which can be tested either
34
35 467 using parametric (e.g. Student's T-test) or nonparametric tests (e.g. the Mann-Whitney U test or the
36
37 468 Wilcoxon rank-sum test). In both cases, a map of p-values can be filled in voxel by voxel, pinpointing
38
39 469 where are the significant differences between the groups of patients.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



1
2
3 471 **Figure 4: Example of voxel (pixel)-wise assessment of a dose-complication relationship for a change in international**
4 472 **prostate symptom score (IPSS) of ≥ 10 in a prostate radiotherapy trial cohort (Yahya *et al.*, 2017). Bladder DSMs have**
5 473 **been derived by cylindrical mapping using 200 equally-spaced radial samples at each interpolated 1 mm slice vertically**
6 474 **from the bladder base (up to 45 mm for all patients). Pixel-wise logistic regression incorporating clinical factors leads to**
7 475 **the (uncorrected) p-value and odds ratio maps shown, including demonstration of a confounding protective effect of**
8 476 **dose at the bladder base.**

11
12 477 Logistic regression, LKB modelling or Logit dose-response modelling are possible alternative
13 478 approaches to studying local dose-response effects at the voxel-level (as illustrated via an example
14 479 for DSMs in Figure 4). For each voxel in the DSM/DVM, the relationship between the dose and the
15 480 incidence of a selected toxicity endpoint is calculated. When the actuarial incidence of the side-
16 481 effect is considered, Cox regression constitutes a suitable choice (Marcello *et al.*, 2020a; Marcello *et*
17 482 *al.*, 2020b). This analysis produces a map of best-fit parameters, constant and b-coefficient for dose
18 483 for the logistic regression, TD50 (uniform dose corresponding to 50% complication probability) and
19 484 slope at TD50 for LKB and Logit models, $H_0(t)$ and β -coefficient for dose when Cox is used. This kind
20 485 of analysis allows identifying regions where the dose-response is steeper vs areas where it is
21 486 shallow, thus providing a hypothesis for treatment optimization on selected sub-regions.

22
23
24
25
26
27
28
29
30 487 Clinical risk factors can be included, with the inclusion of multiple b-coefficients/ β -coefficients in
31 488 logistic and Cox regressions and with the addition of dose modifying factors in LKB and Logit models
32 489 (Peeters *et al.*, 2006a). Of note, in this case, a map of effect sizes for the clinical risk factors is
33 490 produced, with a variation of effect sizes at the voxel level. Discussion is still open on the meaning of
34 491 these variations, with the possibility of a clinical factor to be a protective factor in some voxels and a
35 492 risk factor in others. A possible alternative way to include clinical risk factors is to use local dose-
36 493 based modelling to determine areas with different dose-response curves and apply an adjustment
37 494 for clinical risk factors at a sub-region level or a patient level.

45 495 *4.1.6 Significance*

46 496 From a modelling perspective a large variation over the population provides the best opportunity to
47 497 derive a high-quality dose-effect model (see Table 1). Techniques that result in high rates of toxicity
48 498 do not necessarily exhibit a large variation over patients. When deriving statistics at the voxel-level,
49 499 dose deposited by external beams gives rise to correlations between dose variables. Establishing
50 500 significance based on per-voxel null hypothesis testing (see Section 5.2) severely suffers from
51 501 multiple testing issues – the likelihood of incorrect rejection of that hypothesis. Methods based on
52 502 estimated false discovery rate (FDR) have been proposed, which have been shown to hold under
53 503 positive dependencies (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001; Storey, 2002).
54
55
56
57
58
59
60 504 Permutation methods can be used to establish significance based on test statistics aggregated over

1
2
3 505 the individual voxels (Manly, 1997; Chen *et al.*, 2013; Palma *et al.*, 2020a): a pertinent global
4 506 threshold of the single-voxel test statistic is derived, leading to the selection of voxels that exceed
5 507 that value. Reporting the adjusted map beyond arbitrary thresholds for significance (such as the
6 508 commonly-used $p = 0.05$) might be suggested, allowing readers to make a more informed conclusion
7 509 by also considering the trends and spatial patterns of association, rather than focusing on specific
8 510 highly significant voxels (Palorini *et al.*, 2016b).

14 511 **4.2 Performance, validity and reporting**

15 512 Three main purposes of statistical models can be identified: 1) predictive/prognostic models,
16 513 focussing on event prediction; 2) explanatory models explaining difference in outcome via
17 514 explanatory variables, focussing on (causal) relationships and the magnitude of effects; and 3)
18 515 descriptive models with the main purpose to capture accurately the association between the
19 516 dependent variable and the independent variables, which may focus on both elements of
20 517 prediction, relationships and magnitude of effects (Shmueli, 2010).

27 518 *4.2.1 Model performance*

28 519 NTCP models are descriptive models, describing the relationship between biological dose, clinical
29 520 cofactors, and toxicity risks. To evaluate discriminative (predictive) power, the performance of the
30 521 model is commonly reported through the area under the receiver operating characteristic curve
31 522 (AUC) which is a measure that combines the specificity and sensitivity in one number (Dean *et al.*,
32 523 2018; Men *et al.*, 2019). In case of a large imbalance in the data, the F-score based on precision-
33 524 recall could additionally be considered (Saito and Rehmsmeier, 2015).

40 525 *4.2.2 Model validity*

41 526 The internal validity of a prediction model concerns the reproducibility of the underlying data. To
42 527 avoid overfitting and unstable models, preferred methods for internal validation are cross-validation
43 528 and bootstrap resampling techniques (Heinze *et al.*, 2018; Steyerberg and Vergouwe, 2014; Xu *et al.*,
44 529 2012). For the external validation of the model, concerning the generalizability of the results to
45 530 other similar patient populations outside the database and outside the institution, independent
46 531 validation datasets are required (Bentzen *et al.*, 2010). A relevant example is provided by Mylona *et al.*
47 532 (2020b), where dosimetry for sub-regions in the bladder was found to be more predictive of
48 533 complications than that for the whole organ, as validated in an external cohort.

56 534 *4.2.3 Model reporting*

1
2
3 535 It is recommended to report at least the following characteristics of a developed phenomenological
4 536 (data-driven) NTCP model (Jackson *et al.*, 2010; Collins *et al.*, 2015): study population, received
5
6 537 treatment, definition and measurement of predicted outcome, dose-volume information of full
7
8 538 organs and relevant sub-volumes, basic statistical data on incidence of toxicity including number of
9
10 539 subjects and number of events, complication rates associated with constraints, available follow-up
11
12 540 time, statistical motivation of sample size, handling of missing data, numerical range and median of
13
14 541 the dosimetric variables of interest, model parameter estimates and their standard errors, applied
15
16 542 feature selection method (model building algorithm), candidate variable list, applied validation
17
18 543 methods, goodness-of-fit and discriminative power of the final model. For spatial models, it is also
19
20 544 recommended to report dose-grid resolution and dose calculation algorithm (Placidi *et al.*, 2020),
21
22 545 and a definition for the algorithms used in extraction of features (e.g. (Zwanenburg *et al.*, 2020)). A
23
24 546 checklist for transparent reporting is available through the TRIPOD initiative (Collins *et al.*, 2015).

25 547 **5 Review of methods – spatial dose associations with complications** 26 27 548 **and applications to NTCP calculation**

28
29 549 Section 2 defined, in general terms, approaches that may be used, in various dimensions, to
30
31 550 represent dose information in ways that retain spatial information from which features may be
32
33 551 extracted. Section 3 detailed where the data may be obtained from to inform those processes, and
34
35 552 for describing the complication outcomes with which the features will be correlated, using the
36
37 553 statistical processes described in Section 4. We can now review publications which attempt to
38
39 554 combine these to derive NTCP models and for examining associations of spatial dose information
40
41 555 with complication incidence.

42
43 556 Evidence of improved predictive capabilities with models which are inclusive of spatial information
44
45 557 have been emerging from analysis of isolated data sets over the last 10 – 15 years. Table 2 provides
46
47 558 a summary of some previously published analyses where a comparison has been made between
48
49 559 histogram-based toxicity models and those incorporating various forms of spatial dose information.
50
51
52
53
54
55
56
57
58
59
60

560

561 **Table 2: A selection of published studies comparing histogram-based toxicity modelling to models incorporating spatial information. Note that some studies incorporated multiple**
 562 **approaches to spatial feature extraction.**

Reference	Tumour site	Evaluated region	Toxicity endpoint	Spatial method	Comparison	Impact on complication prediction
Heemsbergen <i>et al.</i> (2005)	Prostate	Rectum	Various	Pixel-wise, DSM sub-regions	Total rectum DSH vs sub-regions alone	Several specific toxicities only associated with spatially-localised dose
Peeters <i>et al.</i> (2005)	Prostate	Rectum	Acute rectal \geq G2	Dose-length parameters	Total rectum DVH and DSH vs addition of spatial features	Most significant DVH and dose-length parameter both improved final model
Peeters <i>et al.</i> (2006b)	Prostate	Anorectum	Various	DVHs for sub-regions (rectum, anus)	Total anorectum DVH vs sub-regions	Specific toxicities better predicted by sub-region dosimetry
Acosta <i>et al.</i> (2013)	Prostate	Rectum	Rectal bleeding	Voxel-wise dose, DVM	Rectum DVH vs voxel-wise	Rectal bleeding only correlates with identified local dose levels, not with total rectum DVH.
Drean <i>et al.</i> (2016b)	Prostate	Rectum	Rectal bleeding	Voxel-wise and manually identified sub-region	Rectum DVH vs different sub-regions	DVH-based inferior-anterior hemi rectum (voxel-wise identified sub-region) performed best.
Casares-Magaz <i>et al.</i> (2017)	Prostate	Rectum	Various	Pixel-wise DSM	Rectal DVH and DSH vs pixel-wise	For all endpoints DSM-based parameters showed better AUCs (mean 0.64) than the best DSH/DSH-based parameters (mean 0.61)
Rossi <i>et al.</i> (2018)	Prostate	Rectum	Rectal bleeding Faecal leakage	3D texture features	Rectal DVH vs addition of spatial features	Bleeding - AUC increased 0.68 to 0.72; leakage - AUC increased from 0.68 to 0.75
Buettner <i>et al.</i> (2009a)	Prostate	Rectum	Rectal bleeding	CNN on DSM	Rectal DSH vs addition of spatial features	AUC increased from 0.59 to 0.64
Buettner <i>et al.</i> (2011)	Prostate	Rectum	Various	Parameterised DSM	NTCP based on rectal vs addition of spatial features	AUC increased from 0.59 to 0.63 – 0.67
Zhen <i>et al.</i> (2017)	Cervix	Rectum	General toxicity	DSM texture features and CNN (with transfer learning)	Peak dose-indices vs texture features vs CNN	AUC 0.47-0.58 (dose-indices), 0.70 (texture features), 0.89 (CNN)
Wilkins <i>et al.</i> (2020)	Prostate	Anorectum	Various	Parameterised DSM; manual sub-regions	Rectal DVH vs sub-region DVH vs DSM	DSM-based parameters did not improve prediction compared to DVH-based parameters;

				(rectum, anal-canal)	features	sub-region dosimetry not identified as more predictive
Heemsbergen <i>et al.</i> (2010)	Prostate	Bladder	Urinary obstruction	DVM, voxel-wise (specific local dose points)	DSH-based total bladder vs addition of local point dose trigone	Both DVH point and local trigone dose point added to final model
Improta <i>et al.</i> (2016)	Prostate	Bladder	IPSS toxicity score	Pixel-wise DSM	Bladder DSH vs addition of spatial features	AUC increased from 0.58-0.71 to 0.66-0.77
Palorini <i>et al.</i> (2016b)	Prostate	Bladder	Acute urinary symptoms	Pixel-wise and parameterised DSM	Bladder DSH vs parameterised DSM	For all endpoints DSM-based parameters showed better AUCs than the best DSH-based parameters
Rossi <i>et al.</i> (2018)	Prostate	Bladder	Nocturia Incontinence	3D texture features	Bladder DVH vs addition of texture features	Nocturia - AUC increased from 0.63 to 0.67; Incontinence - AUC increased from 0.71 to 0.73
Mylona <i>et al.</i> (2019)	Prostate	Bladder, urethra	Acute and late urinary symptoms	Sub-regions derived from voxel-wise DVM analysis	Bladder DVH vs sub-regions DVHs	AUC improvements in both acute and late toxicity in several sub-regions including the urethra (AUCs ≥ 0.62)
Beasley <i>et al.</i> (2018)	H&N	Head region	Trismus	Voxel-wise DVM	Organ vs sub-region DVH	Identified voxel cluster most significant
Buettner <i>et al.</i> (2012b)	H&N	Salivary glands	Xerostomia	Parameterised 3D organ dose distribution	Mean dose vs 3D moments	AUC increased from ~ 0.7 to > 0.8
Gabryś <i>et al.</i> (2018)	H&N	Parotid glands	Xerostomia	Parameterised 3D dose	Mean dose and parotid DVH vs addition of multiple spatial features	AUC increased from < 0.6 to $0.68 - 0.78$ for dose-gradient features
Men <i>et al.</i> (2019)	H&N	Glands	Xerostomia	3D dose CNN and CT images	Combinations of basic dose-volume metrics, clinical parameters, and CNN based on images, structures and dose	AUC increased from 0.56 for dose metrics alone to 0.84 with all CNN information
Monti <i>et al.</i> (2017)	H&N	Neck region	Dysphagia	Voxel-wise DVM	Sub-region mean dose and multi-organ DVH (Alterio <i>et al.</i> , 2017)	AUC confirmed between multi-organ vs voxel-wise analysis (~ 0.8)
Dean <i>et al.</i> (2018)	H&N	Pharyngeal mucosa	Dysphagia	3D spatial parameterisation	Organ DVH vs spatial features	DVH features as predictive as spatial features (AUC $\sim 0.71-0.82$) and maintained on external

						validation
Dean <i>et al.</i> (2016)	H&N	Approximated oral mucosa	Acute mucositis	3D moments	Organ DVH vs addition of spatial features	No improvement
Dean <i>et al.</i> (2017)	H&N	Oral cavity	Acute mucositis	Sub-region definition (mucosal surface)	Organ DVH vs sub-region	No improvement
Palma <i>et al.</i> (2016)	Thorax	Lung	Lung fibrosis	Voxel-wise DVM identified sub-regions	Whole-lung mean dose vs sub-region-based mean dose	AUC increased from 0.60 to 0.75
Palma <i>et al.</i> (2019a)	Thorax	Lung	Lung fibrosis	Voxel-wise DVM	LKB vs 3D model (PACE)	AUC increased from 0.66 to 0.85
Lee <i>et al.</i> (2020)	Lung	Oesophagus	Weight loss	3D texture features	Combinations of DVH and radiomic/dosiomic features	Substantial increases in AUC though addition of spatial features
Liang <i>et al.</i> (2020)	Lung	Lung	Pneumonitis	3D dose texture features and CNN	DVH vs NTCP vs dosiomics vs 3D CNN	AUC increased from 0.676 (DVH) to 0.744 (NTCP) to 0.782 (dosiomics) to 0.842 (CNN)

563

564 Figure 5 illustrates the progression from dose-volume to spatial models of varying complexity, using
 565 the relationship of pelvic radiotherapy dose to gastrointestinal complications as an example.
 566 References describing studies are provided, grouped according to the complexity of anatomical
 567 information used and by the spatial dose features used in the investigations. Many of the cited
 568 studies are discussed in more detail below.

		Anatomical information					
		2D	3D				
		2D rectum BEV	3D rectum	Manual sub-regions	Broader pelvic anatomy	Statistical sub-regions	Structure agnostic
Primary feature definitions	1D (Histograms)	Estimated DVH	Emami et al 1991				
		Calculated DVH		Fiorino et al 2002; Gulliford et al 2004; Rancati et al 2004; Sohn et al 2007; Michalski et al 2010; Tomatis et al 2012; Ospina et al 2014; Fargeas et al 2018	Peeters et al 2006b; Buettner et al 2012; Stenmark et al 2014; Ebert et al 2015b; Gulliford et al 2017; Wilkins et al 2020	Smeenk et al 2012; Shaake et al 2016	Drean et al 2016b; Mylona et al 2020b
		DSH/DWH/zDVH		Cheng and Das 1999; Meijer et al 1999; Fiorino et al 2003	Buettner et al 2012; Kim et al 2014; Ebert et al 2015b; Wilkins et al 2020		
	2D (Dose-surface maps)	2D pixel-wise		Wortel et al 2015; Onjukka et al 2019			
		Parameterised		Heemsbergen et al 2005; Munbodh et al 2008; Buettner et al 2009b; Buettner et al 2011; Moulton et al 2017; Shelley et al 2017; Vanneste et al 2018; Henderson et al 2018; Casares-Magaz et al 2017, 2019	Buettner et al 2012; Wilkins et al 2020		
		Cluster model		Tucker et al 2006b			
		Supervised		Casares-Magaz et al 2017; Chen et al 2018			
		Unsupervised		Buettner et al 2009a;			
	3D (Dose-volume maps)	3D voxel-wise		Acosta et al 2013; Fargeas et al 2013;			Drean et al 2013; Marcello 2020; Ospina et al 2013; Chen et al 2011;
		Parameterised					Chen et al 2011;
		Supervised		Rossi et al 2018			
		Unsupervised		Zhen et al 2017			Coloigner et al 2015

569

570 Figure 5: Illustration of the variety and evolution of methods for incorporation of dosimetric features into dose-
 571 complication association studies and NTCP models in the context of gastro-intestinal toxicity. References are provided as
 572 examples for studies involving various combinations of anatomical information and dosimetric feature extraction and
 573 are not exhaustive. Studies can be further broken down according to the model used for association with complication
 574 (see e.g., Acosta and De Crevoisier (2019)). (BEV – beam’s eye view).

575 **5.1 Use of histogram-based features**

576 **5.1.1 Description**



1
2
3 577 The degeneracy of the spatial dose distribution into the associated DVH of a structure may be
4
5 578 moderated if the dose distribution can be correlated with more specific descriptions of the
6
7 579 underlying functional structures themselves. This can be achieved, for example, by breaking a given
8
9 580 structure down spatially into more precise or component sub-structures according to some
10
11 581 anatomical or statistical criterion (as described in Section 2.2.1). The DVH characteristics of each sub-
12
13 582 structure can be considered independently. Analysis of more specific structures can also reveal that
14
15 583 dose to the originally-hypothesised structure of interest may be less correlated with complication
16
17 584 than alternative adjacent structures. It is also possible to utilise additional spatial information
18
19 585 regarding the structure (such as medical imaging scans) to modify the basic DVH information being
20
21 586 used as input to a dose-volume based NTCP model.

21 587 5.1.2 Examples

22
23 588 A first class of models is based on the assumption that the organs can be thought of as organized in
24
25 589 functional sub-units (FSUs). If the density of FSUs $f(\vec{r})$ is not homogeneous throughout the
26
27 590 considered structure Ω , a more informative version of the DVH would be weighted by the
28
29 591 corresponding $f(\vec{r})$ yielding $fDVH(D_0)$, defined as:

$$31 \quad 592 \quad fDVH(D_0) = \frac{\int_{\Omega} f(\vec{r})H[D(\vec{r}) - D_0]d\vec{r}}{\int_{\Omega} f(\vec{r})d\vec{r}}$$

32
33
34
35 593 where $H(\cdot)$ is the Heaviside step function (Lu *et al.*, 1997). Though DVH-based NTCP models would
36
37 594 be better recast on fDVH, it has been recognized that the derivation of the detailed underlying
38
39 595 arrangement of FSUs in most anatomical sites still requires dedicated studies from techniques such
40
41 596 as functional imaging (e.g. (Arslan *et al.*, 2018; Lee and Park, 2020)).

42
43 597 For lung, a low-cost variation on the fDVH concept is represented by the dose-mass histogram
44
45 598 (DMH), in which the mass density (easily estimated from the planning CT) is considered as a
46
47 599 surrogate of FSU density. As expected, the DMH results to be independent of breathing phase
48
49 600 (Nioutsikou *et al.*, 2005; Cella *et al.*, 2015). Interestingly, however, a study on the risk of
50
51 601 postoperative pulmonary complications among oesophageal cancer patients found no evidence of
52
53 602 significant benefits from the substitution of DVHs with DMHs within the NTCP model (Tucker *et al.*,
54
55 603 2006a).

56 604 Similarly, for hollow organs such as the rectum, the absence of FSUs within the wall content led to
57
58 605 the development of the dose-wall histogram (DWH). DWHs represent the DVH of the organ wall only
59
60 606 based on the segmented outer organ contour (Meijer *et al.*, 1999). The dose-surface histogram

607 (DSH) lies, instead, on the histogram of the dose delivered to a representative surface of the organ.
608 Two main approaches have been proposed for the DSH computation: one based on the interpolation
609 of the dose on the organ surface (Lu *et al.*, 1995), and one normalizing the DVH of the organ wall by
610 the shell depth in the limit of vanishing thickness (Palma and Cella, 2019). There is often a strong
611 correlation between the various histogram types (Fiorino *et al.*, 2003; Carillo *et al.*, 2012; Hoogeman
612 *et al.*, 2005). An exception is when the irradiation technique delivers a dose gradient that is steep
613 relative to the organ size, such as found by Kim *et al* for prostate cancer patients treated with
614 stereotactic radiotherapy (Kim *et al.*, 2014).

615 A first hybrid approach for including a notion of spatial dose distribution within a histogram
616 framework is the zDVH (Cheng and Das, 1999), which expresses the volume receiving a given dose at
617 a given cranio-caudal position in the form of a 2D histogram.

618 An effective approach based on pathophysiological knowledge of the toxicity aetiology consists in
619 splitting a heterogeneous district into component substructures to achieve better DVH-response
620 predictions. This approach has been made for the anorectum (Peeters *et al.*, 2006b; Ebert *et al.*,
621 2015a) and the bladder trigone (Ghadjar *et al.*, 2014; Henderson *et al.*, 2018). Outcome associations
622 have also been undertaken over broader spatial ranges of anatomy than conventionally
623 hypothesised. For rectal toxicity in pelvic radiotherapy for example, although the gastrointestinal
624 tract is usually targeted for derivation of associations, alternative structures can provide stronger
625 associations with specific toxicities. Smeenk *et al* showed that incontinence was more strongly
626 associated to dose to the pelvic floor muscles (Smeenk *et al.*, 2012), whilst Gulliford *et al* discovered
627 the importance of dose to the peri-rectal fat space for control-like symptoms (Gulliford *et al.*, 2017).

628 The emergence of voxel-wise toxicity analyses in radiation oncology has fostered a data-driven
629 evolution of this approach. This is aimed at defining, on a statistical basis, the relevant anatomical
630 substructures involved in the development of radiation induced morbidity and from which
631 histogram-based features can be extracted. This approach is described in Sections 5.2 and 5.4.

632 **5.2 Voxel-wise assessment**

633 *5.2.1 Description*

634 In contrast to analyses based on known or hypothesised FSUs as in Section 5.1, the use of voxel-wise
635 methods points to an “agnostic”/bottom-up approach. Once the DSMs/DVMs in a cohort are
636 spatially registered to a common coordinate system (see Section 2.2 for relevant details) in a way
637 that they can be compared voxel-wise, the regions which are significantly associated to the

particular (toxicity) outcome are identified by statistical inference. Different approaches can be used, as described in Section 4.1.5. In general, the final goal of voxel-wise analysis is to identify regions driving the clinical manifestation of radio-induced side effects, i.e. to find clusters of voxels where the dose is significantly different in patients with/without toxicity (see also Section 5.4). The resulting organ sub-regions do not consider any prior anatomical or functional division. They can provide information to make inferences on the differential radio-sensibility of some organs or the simultaneous implication of different structures on some radio-induced toxicities.

Voxel-wise assessment does not by default generate an NTCP model. DVHs in the regions that were highlighted as statistically associated with the selected outcome should be considered to derive NTCP models following a classical dose-response analysis. Alternatively, a total complication risk can be formed from aggregation of risks determined at the voxel level.

5.2.2 Examples

2D dose-surface outcome mapping

2D DSMs (Section 2.2.2) are usually generated from an anatomical structure and restricted to the surface of this structure. This choice produces results which can be easily translated into organ sub-regions to be spared. Historically, the first analyses of DSMs in the radiotherapy field were related to hollow organs whose geometry could be easily associated with a cylinder, such as the oesophagus (Chen *et al.*, 2013; Dankers *et al.*, 2017) and the rectum (Casares-Magaz *et al.*, 2019; Munbodh *et al.*, 2008; Onjukka *et al.*, 2019; Sanchez-Nieto *et al.*, 2001; Tucker *et al.*, 2006b; Wortel *et al.*, 2015). Although pixel-wise assessment can be made to derive patterns of response, significant progress has been made by parameterising the DSMs, reducing the number of features and providing parameters for NTCP models, as discussed in Section 5.3.

Pixel-wise studies have related DSMs for the bladder with a number of early and late urinary endpoints (Palorini *et al.*, 2016b; Mylona *et al.*, 2020a; Palorini *et al.*, 2016a; Yahya *et al.*, 2017; Improta *et al.*, 2016). Recently a method for the calculation of DSM for the heart was implemented by using a modified cylindrical coordinate system (McWilliam *et al.*, 2020). DSMs of the heart were analysed to infer possible local dose effect for survival after lung cancer radiotherapy. The rationale for considering heart DSMs rather than DVMs resides in the location on the surface of some clinically relevant sub-regions, such as the coronary arteries, the electrical conduction system and the myocardium.

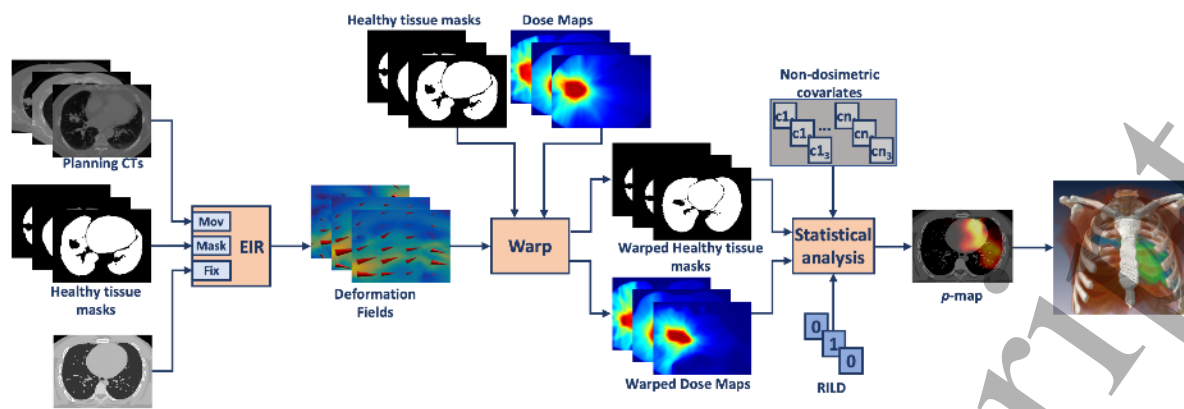
3D voxel-wise outcome mapping

1
2
3 669 DVMs (Section 2.2.4) can be generated either starting from an anatomical structure and restricted to
4
5 670 its volume or independently from any structure. The second choice has the power to embrace a
6
7 671 totally agnostic approach regarding which organs/tissues are involved in radio-induced toxicity and
8
9 672 entails the possibility of highlighting the interaction between different organs and FSUs. Notably,
10
11 673 special care should be taken in order to counteract the possibility of finding significant areas which
12
13 674 offer no feasible anatomical explanation and which could lead to inappropriate organ-sparing
14
15 675 objectives in treatment planning.

16 676 Organ-based DVMs were considered in the literature for the analysis of both rectal (Acosta *et al.*,
17
18 677 2013; Drean *et al.*, 2016b; Mylona *et al.*, 2020b; Shelley *et al.*, 2017; Marcello *et al.*, 2020b) and
19
20 678 urinary (Mylona *et al.*, 2019; Mylona *et al.*, 2020a; Mylona *et al.*, 2020b; Marcello *et al.*, 2020a)
21
22 679 toxicity. These kinds of analysis heavily build upon robust co-registration methods, which become
23
24 680 even more critical when organs highly prone to organ motion and variable filling are considered.

25 681 The first published example of the use of quasi-organ-agnostic DVMs was from Heemsbergen *et al.*
26
27 682 (2010) investigating urinary toxicity. In this case, the 3D reconstruction started with the definition of
28
29 683 the outer surface of the prostate and with the identification of the spatial coordinates of the
30
31 684 prostate centre of mass. After that, for every patient, a spherical surface was considered, extending
32
33 685 6 cm from the prostate. Every voxel inside this region was identified through polar coordinates
34
35 686 (distance from the prostate centre of mass and two angles identifying the vector connecting the
36
37 687 single voxel to the prostate centre of mass) and the absorbed dose in each voxel was calculated by
38
39 688 trilinear interpolation of the nearest dose points of the individual dose grid.

40 689 Regression coefficients associated to each voxel specifically in the salivary glands have been used to
41
42 690 shed light on the regional radio-sensitivity of the glands (Jiang *et al.*, 2019). Other studies considered
43
44 691 DVMs without any restriction to specific contoured organs for investigation of local dose effects in
45
46 692 the thoracic/head and neck region (Beasley *et al.*, 2018; McWilliam *et al.*, 2017; Monti *et al.*, 2017;
47
48 693 Palma *et al.*, 2016; Palma *et al.*, 2019d; Palma *et al.*, 2019c; Green *et al.*, 2020), with interest in the
49
50 694 association of dose pattern with lung toxicity, acute dysphagia, trismus and survival.
51
52
53
54
55
56
57
58
59
60



695

696 **Figure 6: Flowchart for 3D voxel-wise analysis of dosimetric association with lung toxicity following stereotactic body**
 697 **lung radiotherapy, reproduced from Palma *et al.* (2019c). Structures and dose distributions are spatially normalised to a**
 698 **common coordinate system (CCS) via course alignment of structures and CT elastic image registration (EIR). Voxel-wise**
 699 **significance maps are then derived by calculating a test statistic at the voxel level and adjusting for multiple comparisons**
 700 **via a permutation algorithm.**

701 ***NTCP from voxel-wise methods***

702 As already pointed out, although voxel-wise analysis can identify important organ sub-regions, it
 703 does not provide an NTCP.

704 NTCP can be derived by classical dose-response modelling on the specific sub-regions, either
 705 including the whole DVH as calculated in the specific identified areas or choosing some particular
 706 DVH cut-points. Examples of this kind of approach can be found in (Buettner *et al.*, 2009b;
 707 Heemsbergen *et al.*, 2010; Mylona *et al.*, 2019; Onjukka *et al.*, 2019; Palma *et al.*, 2016; Casares-
 708 Magaz *et al.*, 2017). In Drean *et al.* (2016b), parameters for the LKB model were computed within
 709 the voxel-wise derived sub-region.

710 More sophisticated and more global approaches were also developed, taking information from
 711 voxel-wise analysis directly into account. Vinogradskiy *et al.* (2012) proposed a modified LKB model
 712 where the lung dose in each voxel was weighted using a user-defined spatial weighting matrix which
 713 could be derived by a previous voxel-wise analysis. Jiang *et al.* (2019) demonstrated prediction of
 714 xerostomia induced by the irradiation of the salivary glands in head-and-neck cancer patients using a
 715 ridge logistic regression model directly dealing with the local dose delivered to each voxel of the
 716 organ at risk. The framework was naturally able to include non-dosimetric predictors in the NTCP
 717 model.

718 Palma *et al.* (2019a) established a new formalism, called PACE (Probabilistic Atlas for normal tissue
 719 Complication Estimation in radiation therapy), which incorporates regional dose information coming

60

1
2
3 720 from regression performed at the voxel-wise with clinical risk factors. PACE builds upon the LKB
4 721 model and substitutes the generalized equivalent uniform dose (EUD) with a weighted combination
5 722 of risks as calculated by regression at the voxel level, using confidence intervals for predicted risks as
6 723 weights (thus giving more weight to more certain predictions).
7
8
9

10 724 **5.3 Spatial parameterisation of dose distributions**

11 725 *5.3.1 Description*

12
13
14 726 Attempts to reduce the number of features, reduce collinearity and generalise models from voxel-
15 727 wise analyses can be made by parameterising the dose distribution. For analyses restricted to
16 728 specific organs, this will typically involve functional parameterisation of DSMs and organ-constrained
17 729 DVMs, with or without registration to a template geometry. The resulting parameters can then
18 730 become co-variates in regression models or supervised machine learning models. More widespread
19 731 dose distributions can be parameterised using techniques borrowed from imaging analytics –
20 732 namely, the supervised derivation of specific feature classes (“dosimics” (Liang *et al.*, 2019)).
21 733 Unsupervised classification of outcome based on the dose distribution can also be attempted with
22 734 convolutional neural networks, with or without the inclusion of anatomical and functional imaging
23 735 information.
24
25
26
27
28
29
30
31
32
33

34 736 *5.3.2 Examples*

35 737 ***Parameterisation of 2D dose***

36
37 738 The process of derivation of parameters for geometric descriptors from a DSM is illustrated in Figure
38 739 2, particularly in the context of investigating rectal complications due to prostate radiotherapy.
39 740 Concentric isodoses on the rectal wall from prostate radiotherapy can be thresholded systematically
40 741 at different doses and fitted with an ellipse. Parameterised geometrical features can then be
41 742 calculated (Buettner *et al.*, 2009b).
42
43
44
45
46
47

48 743 Previous studies on rectal toxicity following prostate radiotherapy indicate that spatial dose metrics
49 744 such as lateral extent of dose around the circumference of the rectum, longitudinal extent and
50 745 eccentricity derived from rectal dose surface maps (DSM), are related to toxicities including rectal
51 746 bleeding and loose stools (Buettner *et al.*, 2009b; Moulton *et al.*, 2017). Interestingly, a recent test of
52 747 this approach failed to demonstrate any improvement over DVH-based prediction of rectal toxicity
53 748 (Wilkins *et al.*, 2020). This result may be confounded by the differences between planned and
54 749 delivered dose distributions (see Section 3.1.3), with Shelley *et al.* (2017) finding parameters derived
55
56
57
58
59
60

1
2
3 750 from DSMs for accumulated dose being more predictive than those from planned dose, as also
4
5 751 found by Casares-Magaz *et al.* (2019) at the pixel-level.

6
7
8 752 Casares-Magaz *et al.* (2017) demonstrated that parameterised DSMs performed slightly better than
9
10 753 DSHs when predicting rectal toxicity and produced results for more endpoints by quantifying the
11
12 754 dose when a DSM was subdivided to a 3x3 matrix. Vanneste *et al.* (2018) used DSMs to evaluate the
13
14 755 effect of hydrogel rectal spacers on dose to the rectum for prostate radiotherapy.

15
16 756 Although most published data relating to parameterised rectal DSMs is from patients who received
17
18 757 prostate radiotherapy, Chen *et al.* (2018) detail the use of DSM to relate the dose from both external
19
20 758 beam and brachytherapy for a cohort of cervix patients. The two dose distributions were non-rigidly
21
22 759 registered, and a rectal DSM created from the summed dose distribution. Both volumetric and
23
24 760 texture metrics were calculated, and principal component analysis used to provide inputs to a
25
26 761 support vector machine-based model. Area and texture parameters were found to be important and
27
28 762 to have an improved AUC compared to the standard Groupe Européen de Curiethérapie/European
29
30 763 Society for Radiotherapy (GEC-ESTRO) model.

31 764 ***Parameterisation of 3D dose***

32
33 765 For a solid structure such as the parotid it is possible to define metrics to quantify the relative 3D
34
35 766 spatial distribution of dose to the whole organ. Buettner *et al.* (2012b) used 3D spatial invariant
36
37 767 moments to characterise the morphology of the dose distribution to the parotid in terms of centre
38
39 768 of mass, spread and skewness. They showed that minimising the dose to cranial and lateral regions
40
41 769 of the parotid gland would decrease the incidence of xerostomia. The model containing spatial
42
43 770 metrics had a significantly-improved performance compared to the standard predictive of model of
44
45 771 mean dose. 3D moments were also used in a comparison of the conventional oral mucosa outline
46
47 772 and a novel segmentation to predict acute mucositis (Dean *et al.*, 2016). Dose distributions to the
48
49 773 two organ-at-risk (OAR) structures were calculated and used as inputs to both penalised logistic
50
51 774 regression and random forest models. In this example, using the novel segmentation and spatial
52
53 775 metrics did not improve model performance compared to a model built on fractional dose-volume
54
55 776 data for the conventional structure. (Dean *et al.*, 2018) studied acute dysphagia using moments and
56
57 777 dose-volume-length and -circumference data for the pharyngeal mucosa. They demonstrated that
58
59 778 although the length and circumference that received over 1 Gy per fraction were shown to be
60
779 important, a penalised logistic regression NTCP model trained purely on dose-volume data
780 performed equally well on internal validation and was superior when applied to an external
781 validation cohort.

782 ***Supervised broad spatial descriptors***

783 The papers described so far have used a variety of bespoke methods to parameterise the spatial
784 distribution of dose. However, synergy with the field of radiomics allows spatial dose distributions
785 to be characterised by a vast array of standardised metrics (Zwanenburg *et al.*, 2020). Here the 3D
786 radiotherapy dose distribution can be characterised in full with or without reference to structure
787 segmentation. Radiomic features from patient images can be integrated to derive models predictive
788 of complication (Talamonti *et al.*, 2019).

789 One study which assessed this concept and compared predictions to previous work is described by
790 Gabryś *et al.* (2018). This study compared models to predict xerostomia starting with a standard
791 model based on mean dose and parotid-specific spatial metrics described above (Buettner *et al.*,
792 2012b). Spatial descriptors were extended for the parotid gland to consider entropy along with
793 dosiomic descriptors of DVH shape and general dosiomic features describing the gradient of the
794 entire 3D dose distribution. The manuscript describes comparisons of many models including
795 conventional statistical and machine learning approaches. Additionally, feature selection and class
796 balance approaches were compared. Overall the strongest features identified were parotid gland
797 volume eccentricity and the spread of the contralateral parotid dose distribution. The contralateral
798 dose gradient of the 3D dose distribution (right to left) was also identified on univariate analysis but
799 did not feature strongly in the final multivariate analysis. Similarly, Lee *et al.* (2020) informed
800 machine learning algorithms with combinations of dose-volume, radiomics and dosiomics features,
801 together with clinical co-variates. Resulting predictive models of weight loss in lung cancer
802 radiotherapy with greater accuracy than models based on dose-volume and clinical co-variates
803 alone, though still with a relatively modest AUC of 0.71.

804 Liang *et al.* (2019) compared conventional dosimetric models with models derived from dosiomic
805 features to predict radiation pneumonitis. It was demonstrated that a multivariate NTCP model
806 including the Grey Level Co-occurrence Matrix (GLCM) contrast and Grey Level Run Length Matrix
807 (GLRLM) (which has similarities to lateral and longitudinal extent described previously)
808 outperformed models based on dose-volume parameters and conventional NTCP model parameters.
809 Rossi *et al.* (2018) included texture analysis features in a study to predict genitourinary and
810 gastrointestinal toxicity following prostate radiotherapy. The 3D texture analysis features for the
811 rectum and bladder were derived from standard radiomics and used alongside non-treatment
812 related features (such as age, staging and comorbidities) and DVH-based metrics to build
813 multivariate logistic regression NTCP models. It was demonstrated that for gastrointestinal
814 endpoints inclusion of texture features improved the AUC compared to models containing only

1
2
3 815 clinical and DVH-based features. Results for genitourinary toxicity were generally not improved by
4
5 816 any dosimetric features.
6
7

8 817 ***Unsupervised broad spatial descriptors***

9 818 An alternative to utilising crafted dosimetric descriptors of broad distributions is to apply neural
10 819 networks. For example, Buettner *et al.* (2009a) used DSMs derived by the rectum unfolding as input
11
12 820 for a rectal bleeding model based on locally-connected neural networks able to account for the
13
14 821 topology of the dose distribution. The higher performance achieved by such models, compared to
15
16 822 the more traditional fully-connected conventional neural nets based on DSHs, suggested that
17
18 823 morphological aspects of the dose distributions play a relevant role in the development of radiation
19
20 824 induced morbidity. Zhen *et al.* (2017) utilised a convolutional neural network (CNN) to distinguish
21
22 825 rectal DSMs indicative of toxicity, incorporating transfer learning to compensate limited patient
23
24 826 data.

25 827 CNNs can be used to extract unspecified higher-level features of 3D dose distributions which can
26
27 828 directly classify the distributions as likely to lead to complications, and studies have begun to emerge
28
29 829 demonstrating this with varying combinations of ancillary information. Ibragimov et al (Ibragimov *et*
30
31 830 *al.*, 2018; Ibragimov *et al.*, 2019) utilised CNNs incorporating 3D dose information, supplemented
32
33 831 with transfer learning from previous abdominal imaging, for hepatobiliary toxicity prediction
34
35 832 following stereotactic liver radiotherapy. Incorporating the CNNs with more conventional features
36
37 833 including dose-volume parameters, dose-fractionation and clinical co-variables increased the model
38
39 834 predictions (increase in AUC from 0.79 to 0.85). In a strategy which preferences identifying likely
40
41 835 toxicity (i.e., minimising false negatives), the CNN approach halved the number of false positive
42
43 836 predictions relative to DVH-based prediction. Ibragimov et al were able to extend this approach to a
44
45 837 structure-agnostic spatial assessment to map anatomical regions where dose variations associate
46
47 838 with toxicity. This revealed regions associated with the hepatobiliary tract and liver as specific focus
48
49 839 regions to guide dose planning (Ibragimov *et al.*, 2020).

50 840 In a progression from the dosimetrics approach, Liang *et al.* (2020) utilised CNNs incorporating the 3D
51
52 841 dose distribution for predicting pneumonitis following volumetric-modulated radiotherapy. A
53
54 842 superior prediction (AUC 0.842) was achieved relative to regression models incorporating
55
56 843 dosimetric, NTCP and dosimetrics features (AUC < 0.782). Class activation maps were used to identify
57
58 844 lung regions associated with increased or reduced high-grade toxicity.

59 845 In head and neck cancer radiotherapy, Men *et al.* (2019) used CNNs which incorporated one or more
60 846 of the planning CT images, planned 3D dose and segmented anatomy, for prediction of grade ≥ 2

1
2
3 847 xerostomia, and compared prediction against regression models incorporating dose with or without
4 848 clinical co-variates. The CNNs provided greater accuracy ($AUC < 0.84$), compared to the regression
5 849 models, for all combinations of 3D information except for when 3D dose was removed.
6
7
8

9 850 **5.4 Spatial clustering**

11 851 *5.4.1 Description*

12 852 Thames *et al.* (2004) proposed that hot spots distributed as small areas throughout an organ at risk
13 853 are likely to cause a different response than if the highest dose covers one contiguous region. This
14 854 difference would translate through to a difference in toxicity prediction using NTCP models which
15 855 describe the clustering of damage to FSUs. This concept of spatial dose clusters forms bridges
16 856 between voxel-wise assessment, definition of sub-regions and spatial analysis based on spatial
17 857 parameterisation. The cluster models highlight the relevance of including both the number and the
18 858 spatial location of radiation-sterilized FSUs in a comprehensive NTCP model (Thames *et al.*, 2004). In
19 859 a general sense, these models suggest that a volume receiving at least a given dose value is more
20 860 likely associated with a radiation-induced toxicity if it corresponds to a connected spatial cluster
21 861 rather than if spatially scattered (Deasy and El Naqa, 2008).
22
23
24
25
26
27
28
29
30
31

32 862 *5.4.2 Examples*

33 863 Tucker et al demonstrated a practical application of the method described by Thames using rectal
34 864 DSMs (Tucker *et al.*, 2006b). Nine case-control pairs with very similar absolute DSH but with and
35 865 without grade 2 rectal bleeding were used to fit a local-effect cluster model. The logistic function
36 866 describing the probability of damage for each voxel in each DSM had 2 unknown parameters. The
37 867 model was fitted to maximise the relationship between maximum cluster size (considering 2-
38 868 connectivity) between the cases and controls. Although the cohort was very small, the authors were
39 869 able to find parameter values which separated cases from controls and inferred that dose
40 870 distributions in the region of 30 Gy were important for the prediction of rectal bleeding.
41
42
43
44
45
46
47

48 871 Chao et al (Chao *et al.*, 2020; Chao *et al.*, 2018) also developed spatial cluster metrics based on the
49 872 method proposed by Thames et al. They demonstrated that maximum cluster size for the superior 5
50 873 cm of the oesophagus was not related to conventional dose-volume and NTCP metrics and inferred
51 874 that spatial distributions were not represented by conventional dose metrics. They applied a cluster-
52 875 based approach to model xerostomia (Chao *et al.*, 2019). The metrics included mean cluster size and
53 876 largest cluster size normalised to the volume of the gland which were incorporated into LKB models.
54 877 Although no conventional (DVH-based) LKB model was derived from the cohort a comparison was
55
56
57
58
59
60

1
2
3 878 made with published models utilising mean dose ($n=1$). TD50 was higher for the thresholded cluster
4
5 879 model at just below 40 Gy compared to 26 Gy from the literature.
6
7

8 880 The concept of spatial cluster models can be expanded using percolation theory, which has origins in
9
10 881 statistical physics and considers how clusters are connected. Originally proposed at a similar time to
11
12 882 Thames' (Thames *et al.*, 2004) work on cluster models, Myers and Niemierko (2004) presented the
13
14 883 use of percolation theory for predicting NTCP from clusters. Gale *et al.* (2017) describe how the
15
16 884 concept can be applied to geometric arrangements of FSUs to predict toxicity for both parallel and
17
18 885 serial organs.

19 886 Several studies considered the clusters of organ voxel L_p^- whose dose-toxicity association exceeded
20
21 887 some statistical significance threshold p . They showed that the mean dose in such clusters could be
22
23 888 a more powerful predictor of toxicity than traditional metrics associated to the organ considered as
24
25 889 a whole structure. Hence, an NTCP model can be proficiently trained as a logistic regression of the
26
27 890 patients' outcomes as a function of simple dose metrics in the cluster L_p^- propagated from the
28
29 891 common coordinate system of the voxel-wise analysis (see Section 5.2) to each individual native
30
31 892 space. In this way, sub-regions have been identified in different locations such as the lungs (Palma *et*
32
33 893 *al.*, 2016), the heart (McWilliam *et al.*, 2017), head and neck (Monti *et al.*, 2017), the rectum (Acosta
34
35 894 *et al.*, 2013; Drean *et al.*, 2016b) and the bladder (Mylona *et al.*, 2019).

35 895 **6 Ongoing Endeavours**

36 37 38 896 **6.1 Model development and validation**

39 897 As for other approaches to radiotherapy complication modelling, a major issue is represented by the
40
41 898 quantity and quality of data available to researchers. Relative to DVH-based models, spatial methods
42
43 899 require more comprehensive data (see Section 3). Despite the abundance of relevant data
44
45 900 generated continuously around the world and the technical capability to collect it, and despite
46
47 901 decades of pleas (e.g. (Deasy *et al.*, 2010)), remarkably little data has become available to progress
48
49 902 this type of analysis. Based on legislative constraints (i.e., ownership, privacy and patient consent
50
51 903 needs) it is likely, at least in the next few years, that data will prevalently come from clinical trials
52
53 904 where their recovery, storage and access are already planned.

54
55 905 The implementation of innovative trials including large cohorts of clinical data (Baumann *et al.*,
56
57 906 2016) could rapidly change the landscape. Such trials could multiply the opportunities for developing
58
59 907 models, provide opportunities for validating models, and enable the merging of different large
60
908 cohorts to increase feature diversity. A specific issue may concern the possibility of introducing

909 unpredictable biases if pooling together cohorts of patients treated, for instance, at different dose
910 levels with largely different spatial locations of the high-dose volumes. Uncontrolled voxel-wise
911 comparisons could lead to “false” spatial effects due, for instance, to the higher incidence or
912 prevalence of side effects in cohorts delivering systematically higher doses and/or treating larger
913 volumes. Ideally, the availability of large cohorts should be accompanied by a proper grouping of
914 patients to make the different patient groups comparable.

915 **6.2 Model generalisation and extension**

916 Apart from the critical issues related to generalizability of NTCP models such as technical, temporal
917 or geographical variabilities (van der Schaaf *et al.*, 2015), a few specific points deserve discussion.

918 The interplay between the spatial patterns of a certain modality/technique and the inter-individual
919 variability is a challenging issue: well driven studies may help in quantifying the real impact of a
920 modality with respect to another. The generalizability of models across different modalities need
921 high-quality studies and extensive validation. One confounding problem is that the patterns of dose
922 delivered today already reflect the existing knowledge based on dose volume metrics. As these
923 models mature, there is the potential for radiobiological predictions that consider the spatial pattern
924 of dose that can drive the optimization of treatment plans towards more favourable dose patterns
925 beyond that of the traditional dose-volume metrics.

926 Another important field of investigation regarding model generalisation is represented by the
927 challenge of combined therapies. Data from studies testing radiotherapy-only vs combined therapy
928 (for instance chemotherapy, immunotherapy) could help in assessing spatial dosimetry correlations
929 specifically linked to the action, for instance, of a drug and making possible local dose corrections
930 incorporating its effect. Similarly, highly non-conventional dose and dose-rate distributions, such as
931 from ultra-fast irradiation (Esplen *et al.*, 2020) or molecular radiotherapy (Stokke *et al.*, 2017) will
932 offer new data sources with which to generalise derived models.

933 **6.3 Including intra and inter-fraction changes**

934 In many situations, both intra and inter-fraction anatomical and geometrical changes may have a
935 significant impact in modifying the delivered dose with respect to the planned one. In particular, the
936 prevalence of systematic over random changes may potentially blur (or even hide) the correlation
937 with toxicities; consequently, investigations quantifying these effects are needed. As an example,
938 the impact of variable bladder filling on bladder DSM can be assessed from daily cone-beam CT
939 imaging: one recent study showed a relatively small impact of variable filling on bladder DSM during
940 image-guided radiotherapy of prostate cancer (Palorini *et al.*, 2016a). A statistical approach based on

1
2
3 941 Gaussian-like variations of local doses likely works in several situations but is expected to fail in
4
5 942 others, such as when the phenomenon itself is prevalently non-Gaussian. Shelley *et al.* (2017)
6
7 943 demonstrated superiority in toxicity prediction from rectal DSMs formed from estimated delivered
8
9 944 rather than planned dose.

10
11 945 When the toxicity rate is small (say, <10%), those few patients with large systematic changes
12
13 946 resulting in a relevant increase of dose to proximal OARs may jeopardise results. Greater efforts may
14
15 947 especially be expected in trying to incorporate individually-assessed anatomical modifications in
16
17 948 stereotactic body radiation therapy (SBRT) (Magallon-Baro *et al.*, 2019), looking to the 3D dose-of-
18
19 949 the-day and/or to the accumulated dose instead of the planned dose distribution. SBRT is also prone
20
21 950 to be associated with larger effects due to both the reduced margins and the high dose per fraction,
22
23 951 dealing with an enhanced impact on critical regions even with small anatomical/geometrical
24
25 952 changes. Relevant effects due to systematic deviations between the planned and the delivered dose
26
27 953 may occur even in unexpected situations and the availability of in-room imaging information is of
28
29 954 paramount importance to identify them. The recently reported correlation between shift toward the
30
31 955 heart of field isocentre during delivery and poorer survival in lung cancer patients treated with SBRT
32
33 956 is a highly paradigmatic example (Johnson-Hart *et al.*, 2018).

34
35 957 Similarly, intra-fraction changes are known to significantly affect the delivered dose in specific sites.
36
37 958 Breathing-induced motion can be highly anisotropic and variable between patients in the different
38
39 959 thoracic and abdominal areas. Although, to our knowledge, no studies have reported on the impact
40
41 960 of intra-fraction motion on spatial models, more relevant research in this area is needed.

42 961 **6.4 Potential applications of artificial intelligence**

43 962 The rise of deep learning approaches for image segmentation, pattern recognition and patient
44
45 963 classification adds many opportunities to extend this field (El Naqa and Das, 2020). Ready access to
46
47 964 advanced deep learning tools is making this kind of analysis more popular (with examples given in
48
49 965 Section 5.3). A merit of these methods is the opportunity to consider features mostly “hidden” to
50
51 966 the human eye and to find complex correlations in a multi-layer approach. On the other hand, this
52
53 967 same merit may also constitute a disadvantage from the point of view of interpretability of the
54
55 968 results and consequent confidence in clinically applying them; in fact, any attempt to maintain some
56
57 969 causality to explain any correlation is largely lost. A major issue regarding artificial intelligence
58
59 970 models is their intrinsically higher difficulty in being validated. Valdes and Interian (2018) provide a
60
971 timely summary of the potential for mis-interpretation in such complex approaches. Keeping the
972 models as simple and interpretable as possible should be considered valuable: the benefit of the

1
2
3 973 addition of deep learning based spatial signatures should always be demonstrated and quantified in
4
5 974 validation cohorts.
6
7

8 975 **6.5 Understanding pathophysiology**

9
10 976 An intriguing and relevant field of investigation related to NTCP models based on 3D/2D similarity
11 977 comparisons concerns the meaning of the resulting regions whose dose differences are predictive of
12 978 toxicity. As already underlined, the information resulting from these analyses cannot be
13 979 automatically associated to a specific cause, being intrinsically a phenomenological finding (i.e.:
14 980 simply reflecting some statistical correlation). Moreover, the assessment of specific
15 981 volumes/surfaces within the body/OARs apparently more “sensitive” to radiation can be biased by
16 982 unknown factors or just due to geometrical or technical issues. Any hypothetical causality has to be
17 983 considered as a strength of such models, in case the found results are consistent with known
18 984 physiological processes/functionalities. As examples, identification of the bladder trigone as a
19 985 structure likely to be highly sensitive (Rancati *et al.*, 2017; Henderson *et al.*, 2018; Yahya *et al.*, 2017)
20 986 is consistent with the involvement of the trigone in the physiology of urination, and the physiological
21 987 connection between the heart and lungs (Ghobadi *et al.*, 2012) adds validity to correlation of heart
22 988 dose with lung toxicity (Palma *et al.*, 2019c; Palma *et al.*, 2019d). Any hypothesis generated by such
23 989 models would deserve to be tested in pre-clinical and clinical studies. Animal models may be well
24 990 used to verify the existence of spatial effects. Conversely, pre-clinical research may first explain
25 991 specific patterns of toxicity that may be confirmed later by studies dealing with dose similarity
26 992 comparison. An interesting example is the evidence of spatial dosimetry effects within parotids
27 993 impacting xerostomia, due to the sparing (or not) of stem cells contained in the ductal region. Such
28 994 observations have been reported in animal experiments (van Luijk *et al.*, 2015) and confirmed by a
29 995 3D dose comparison investigation on data from a large patient cohort treated for head-neck cancer
30 996 (Jiang *et al.*, 2019).
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45

46 997 **6.6 Model application**

47 998 Although examples of practical applications of NTCP models incorporating spatial dosimetric
48 999 features are rare, it is likely that a few of the most robust results will increasingly influence planning
50 1000 optimization. When a causal relationship between a spatial effect and the pattern of the
51 1001 corresponding side effects is apparent, changes may be easily implemented in clinical practice. Two
52 1002 examples are the previously-cited cases of the bladder trigone for prostate cancer and the ductal
53 1003 region of the parotid glands. The latter, cited above as originating in pre-clinical studies, is being
54 1004 assessed within a clinical trial (van Luijk *et al.*, 2015), which is probably the first example of a trial
55
56
57
58
59
60

1005 specifically looking to the possibility to exploit information regarding the spatial dose distribution
1006 within an OAR to reduce toxicity.

1007 A likely progression will be the incorporation of spatial models into tools to evaluate the planned 3D
1008 dose distribution and for generating NTCP and risk estimates. This could be accomplished, for
1009 instance, within clinical trials or as an additional tool for plan quality assurance, in complement with
1010 conventional DVH-based EUD/NTCP estimates. The propagation of identified sensitive sub-regions to
1011 an individual would facilitate toxicity-minimised planning, without the need to modify current
1012 optimisation methods (Acosta and De Crevoisier, 2019). This has been demonstrated by Lafond *et al.*
1013 (2020). A subsequent natural extension would be the possibility to implement these models directly
1014 into the optimization engine. However, the general adoption of spatial models is greatly inhibited by
1015 the prior evolution of the planning process and optimisation engines in the context of dose-volume
1016 approaches. For spatial models that cannot be formulated via dose-volume terminology, research
1017 planning systems are required to enable inclusion of the relevant predicted complication models in
1018 optimisation constraints and objectives or via scripting capabilities of commercial planning systems
1019 (e.g. (Voutilainen, 2016)). With the growth of artificial intelligence based planning systems, there is
1020 considerable scope for building automated planning algorithms that directly incorporate spatial
1021 models to augment or replace dose-volume based optimisation (Wang *et al.*, 2019).

1022 Intriguingly, for models which are agnostic to segmented structures, plan optimisation could in
1023 principle be feasible without the incorporation of dose-volume data for OARs. This would permit a
1024 segmentation-free plan optimization. In the same direction, this kind of approach could also find
1025 applications in overall treatment optimization, directly considering patient outcome as the goal and
1026 incorporating possible “systemic” effects due to the irradiation of multiple organs and to the
1027 interaction with the immune system (for instance through the implicit consideration of the
1028 incidental irradiation of nodes and of the vascular system). Similarly, one could hypothesize
1029 applications in combined treatments to include the effect of modifying agents at the voxel-level, and
1030 to “virtual human” simulation in the optimisation of patient-specific treatments.

1031 **7 Conclusion**

1032 The field reviewed in in this paper is still in its infancy. However, models which consider the spatial
1033 characteristics of radiotherapy dose will permit the expansion, or at least fine-graining, of the
1034 solution space for radiotherapy treatment planning and improving the prediction of treatment
1035 complications. The potential for large-scale relevant applications in treatment personalization, plan

1
2
3 1036 optimization and evaluation are more than promising. Rapid developments and extensive
4 applications are expected in the coming years.
5 1037
6
7

8 1038 **8 Acknowledgements**

9
10 1039 ME acknowledges funding support from the National Health and Medical Research Council (NHMRC
11 grant 1077788). TR was partially supported by the Fondazione Italo Monzino. OA and RD
12 1040 acknowledge partial funding from a French government grant (through the CominLabs excellence
13 1041 laboratory and managed by the National Research Agency in the “Investing for the Future” program,
14 1042 under reference ANR-10-LABX-07-01). SG is supported by a Cancer Research UK Centres Network
15 1043 Accelerator Award Grant (A21993) to the ART-NET Consortium.
16 1044
17
18
19
20

21 1045

22
23
24 1046
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

References

- 1
2
3 1047
4 1048
5 1049 Abravan A, Faivre-Finn C, Kennedy J, McWilliam A and van Herk M 2020 Radiotherapy-Related
6 1050 Lymphopenia Affects Overall Survival in Patients With Lung Cancer *Journal of Thoracic*
7 1051 *Oncology* **15** 1624-35
8 1052 Acosta O and De Crevoisier R 2019 *Modelling Radiotherapy Side Effects - Practical Applications for*
9 1053 *Planning Optimisation*, ed T Rancati and C Fiorino (Boca Raton: CRC Press) pp 415-40
10 1054 Acosta O, Dowling J, Cazoulat G, Simon A, Salvado O, de Crevoisier R and Haigron P 2010 *Prostate*
11 1055 *Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention: International*
12 1056 *Workshop, Held in Conjunction with MICCAI 2010, Beijing, China, September 24, 2010.*
13 1057 *Proceedings*, ed A Madabhushi, et al. (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 42-
14 1058 51
15 1059 Acosta O, Drean G, Ospina J D, Simon A, Haigron P, Lafond C and de Crevoisier R 2013 Voxel-based
16 1060 population analysis for correlating local dose and rectal toxicity in prostate cancer
17 1061 radiotherapy *Phys Med Biol* **58** 2581-95
18 1062 Acosta O, Mylona E, Le Dain M, Voisin C, Lizee T, Rigaud B, Lafond C, Gnep K and de Crevoisier R
19 1063 2017 Multi-atlas-based segmentation of prostatic urethra from planning CT imaging to
20 1064 quantify dose distribution in prostate cancer radiotherapy *Radioth Oncol* **125** 492-9
21 1065 Alterio D, Gerardi M A, Cella L, Spoto R, Zurlo V, Sabbatini A, Fodor C, D'Avino V, Conson M, Valoriani
22 1066 F, Ciardo D, Pacelli R, Ferrari A, Maisonneuve P, Preda L, Bruschini R, Cossu Rocca M, Rondi
23 1067 E, Colangione S, Palma G, Dicuonzo S, Orecchia R, Sanguineti G and Jereczek-Fossa B A 2017
24 1068 Radiation-induced acute dysphagia : Prospective observational study on 42 head and neck
25 1069 cancer patients *Strahlenther Onkol* **193** 971-81
26 1070 Arslan S, Ktena S I, Makropoulos A, Robinson E C, Rueckert D and Parisot S 2018 Human brain
27 1071 mapping: A systematic comparison of parcellation methods for the human cerebral cortex
28 1072 *NeuroImage* **170** 5-30
29 1073 Ashburner J and Friston K J 2000 Voxel-Based Morphometry—The Methods *NeuroImage* **11** 805-21
30 1074 Baumann M, Krause M, Overgaard J, Debus J, Bentzen S M, Daartz J, Richter C, Zips D and Bortfeld T
31 1075 2016 Radiation oncology in the era of precision medicine *Nat Rev Cancer* **16** 234-49
32 1076 Beasley W, Thor M, McWilliam A, Green A, Mackay R, Slevin N, Olsson C, Pettersson N, Finizia C,
33 1077 Estilo C, Riaz N, Lee N Y, Deasy J O and van Herk M 2018 Image-based Data Mining to Probe
34 1078 Dosimetric Correlates of Radiation-induced Trismus *Int J Radiat Oncol Biol Phys* **102** 1330-8
35 1079 Benjamini Y and Hochberg Y 1995 Controlling the False Discovery Rate: A Practical and Powerful
36 1080 Approach to Multiple Testing *J Royal Stat Soc Ser B* **57** 289-300
37 1081 Benjamini Y and Yekutieli D 2001 The Control of the False Discovery Rate in Multiple Testing under
38 1082 Dependency *Annal Stat* **29** 1165-88
39 1083 Bentzen S M, Constine L S, Deasy J O, Eisbruch A, Jackson A, Marks L B, Ten Haken R K and Yorke E D
40 1084 2010 Quantitative Analyses of Normal Tissue Effects in the Clinic (QUANTEC): an
41 1085 introduction to the scientific issues *Int J Radiat Oncol Biol Phys* **76** S3-S9
42 1086 Bentzen S M, Dorr W, Gahbauer R, Howell R W, Joiner M C, Jones B, Jones D T, van der Kogel A J,
43 1087 Wambersie A and Whitmore G 2012 Bioeffect modeling and equieffective dose concepts in
44 1088 radiation oncology—terminology, quantities and units *Radioth Oncol* **105** 266-8
45 1089 Bijl H P, van Luijk P, Coppes R P, Schippers J M, Konings A W and van der Kogel A J 2003 Unexpected
46 1090 changes of rat cervical spinal cord tolerance caused by inhomogeneous dose distributions *Int*
47 1091 *J Radiat Oncol Biol Phys* **57** 274-81
48 1092 Bradburn M J, Clark T G, Love S B and Altman D G 2003 Survival analysis part II: multivariate data
49 1093 analysis—an introduction to concepts and methods *Br J Cancer* **89** 431-6
50 1094 Buettner F, Gulliford S L, Webb S and Partridge M 2009a Using dose-surface maps to predict
51 1095 radiation-induced rectal bleeding: a neural network approach *Phys Med Biol* **54** 5139
52 1096 Buettner F, Gulliford S L, Webb S and Partridge M 2011 Modeling late rectal toxicities based on a
53 1097 parameterized representation of the 3D dose distribution *Phys Med Biol* **56** 2103-18
54
55
56
57
58
59
60

- 1
2
3 1098 Buettner F, Gulliford S L, Webb S, Sydes M R, Dearnaley D P and Partridge M 2009b Assessing
4 1099 correlations between the spatial distribution of the dose to the rectal wall and late rectal
5 1100 toxicity after prostate radiotherapy: an analysis of data from the MRC RT01 trial (ISRCTN
6 1101 47772397) *Phys Med Biol* **54** 6535-48
- 8 1102 Buettner F, Gulliford S L, Webb S, Sydes M R, Dearnaley D P and Partridge M 2012a The dose-
9 1103 response of the anal sphincter region--an analysis of data from the MRC RT01 trial *Radioth*
10 1104 *Oncol* **103** 347-52
- 11 1105 Buettner F, Miah A B, Gulliford S L, Hall E, Harrington K J, Webb S, Partridge M and Nutting C M
12 1106 2012b Novel approaches to improve the therapeutic index of head and neck radiotherapy:
13 1107 an analysis of data from the PARSPORT randomised phase III trial *Radioth Oncol* **103** 82-7
- 14 1108 Bzdok D 2017 Classical Statistics and Statistical Learning in Imaging Neuroscience *Front Neurosci* **11**
15 1109 543
- 17 1110 Carillo V, Cozzarini C, Chietera A, Perna L, Gianolini S, Maggio A, Botti A, Rancati T, Valdagni R and
18 1111 Fiorino C 2012 Correlation between surrogates of bladder dosimetry and dose-volume
19 1112 histograms of the bladder wall defined on MRI in prostate cancer radiotherapy *Radioth*
20 1113 *Oncol* **105** 180-3
- 21 1114 Casares-Magaz O, Bülow S, Pettersson N J, Moiseenko V, Pedersen J, Thor M, Einck J, Hopper A,
22 1115 Knopp R and Muren L P 2019 High accumulated doses to the inferior rectum are associated
23 1116 with late gastro-intestinal toxicity in a case-control study of prostate cancer patients treated
24 1117 with radiotherapy *Acta oncologica* **58** 1543-6
- 26 1118 Casares-Magaz O, Muren L P, Moiseenko V, Petersen S E, Pettersson N J, Høyer M, Deasy J O and
27 1119 Thor M 2017 Spatial rectal dose/volume metrics predict patient-reported gastro-intestinal
28 1120 symptoms after radiotherapy for prostate cancer *Acta oncológica* **56** 1507-13
- 29 1121 Cella L, D'Avino V, Palma G, Conson M, Liuzzi R, Picardi M, Pressello M C, Boboc G I, Battistini R,
30 1122 Donato V and Pacelli R 2015 Modeling the risk of radiation-induced lung fibrosis: Irradiated
31 1123 heart tissue is as important as irradiated lung *Radioth Oncol* **117** 36-43
- 32 1124 Chao M, Wei J, Lo Y-C and Peñagaricano J A 2020 Dose cluster model parameterization of the parotid
33 1125 gland in irradiation of head and neck cancer *Phys Eng Sci Med* **43** 143-53
- 35 1126 Chao M, Wei J, Lo Y C and Penagaricano J A 2019 Percolation Based Cluster Models Fully
36 1127 Incorporating Spatial Dose Distribution in Assessment of Parotid Gland Radiation Induced
37 1128 Complication in Head and Neck Cancer Treatment *Int J Radiat Oncol Biol Phys* **105** S169-S70
- 38 1129 Chao M, Wei J, Narayanasamy G, Yuan Y, Lo Y-C and Peñagaricano J A 2018 Three-dimensional
39 1130 cluster formation and structure in heterogeneous dose distribution of intensity modulated
40 1131 radiation therapy *Radioth Oncol* **127** 197-205
- 41 1132 Chen B, Acosta O, Kachenoura A, Ospina Arango J, Dréan G, Simon A, Bellanger J-J, Haigron P and
42 1133 Crevoisier R 2011 Spatial Characterization and Classification of Rectal Bleeding in Prostate
43 1134 Cancer Radiotherapy with a Voxel-Based Principal Components Analysis Model for 3D Dose
44 1135 Distribution. In: *14th Prostate Cancer Imaging. Image Analysis and Image-Guided*
45 1136 *Interventions - International Workshop, Held in Conjunction with MICCAI · MICCAI 2011*, ed A
46 1137 Madabhushi, *et al.* pp 60-9
- 48 1138 Chen C, Witte M, Heemsbergen W and Herk M v 2013 Multiple comparisons permutation test for
49 1139 image based data mining in radiotherapy *Radiat Oncol* **8** 293-
- 50 1140 Chen J, Chen H, Zhong Z, Wang Z, Hrycushko B, Zhou L, Jiang S, Albuquerque K, Gu X and Zhen X 2018
51 1141 Investigating rectal toxicity associated dosimetric features with deformable accumulated
52 1142 rectal surface dose maps for cervical cancer radiotherapy *Radiat Oncol* **13** 125-
- 54 1143 Chen Y W, Mahal B A, Muralidhar V, Nezoslosky M, Beard C J, Den R B, Feng F Y, Hoffman K E, Martin
55 1144 N E, Orto P F and Nguyen P L 2016 Association Between Treatment at a High-Volume Facility
56 1145 and Improved Survival for Radiation-Treated Men With High-Risk Prostate Cancer *Int J*
57 1146 *Radiat Oncol Biol Phys* **94** 683-90
- 58 1147 Cheng C W and Das I J 1999 Treatment plan evaluation using dose-volume histogram (DVH) and
59 1148 spatial dose-volume histogram (zDVH) *Int J Radiat Oncol Biol Phys* **43** 1143-50
- 60

- 1
2
3 1149 Cicchetti A, Laurino F, Possenti L, Rancati T and Zunino P 2020 In silico model of the early effects of
4 1150 radiation therapy on the microcirculation and the surrounding tissues *Phys Med* **73** 125-34
5 1151 Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M,
6 1152 Tarbox L and Prior F 2013 The Cancer Imaging Archive (TCIA): maintaining and operating a
7 1153 public information repository *J Digit Imaging* **26** 1045-57
8 1154 Collins G S, Reitsma J B, Altman D G and Moons K G M 2015 Transparent reporting of a multivariable
9 1155 prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement *BMC*
10 1156 *Medicine* **13** 1
11 1157 Coloigner J, Fargeas A, Kachenoura A, Wang L, Dréan G, Lafond C, Senhadji L, Crevoisier R d, Acosta O
12 1158 and Albera L 2015 A Novel Classification Method for Prediction of Rectal Bleeding in Prostate
13 1159 Cancer Radiotherapy Based on a Semi-Nonnegative ICA of 3D Planned Dose Distributions
14 1160 *IEEE J Biomed Health Inform* **19** 1168-77
15 1161 Dankers F, Wijsman R, Troost E G, Monshouwer R, Bussink J and Hoffmann A L 2017 Esophageal wall
16 1162 dose-surface maps do not improve the predictive performance of a multivariable NTCP
17 1163 model for acute esophageal toxicity in advanced stage NSCLC patients treated with intensity-
18 1164 modulated (chemo-)radiotherapy *Phys Med Biol* **62** 3668-81
19 1165 Dean J, Wong K, Gay H, Welsh L, Jones A-B, Schick U, Oh J H, Apte A, Newbold K, Bhide S, Harrington
20 1166 K, Deasy J, Nutting C and Gulliford S 2018 Incorporating spatial dose metrics in machine
21 1167 learning-based normal tissue complication probability (NTCP) models of severe acute
22 1168 dysphagia resulting from head and neck radiotherapy *Clin Transl Radiat Oncol* **8** 27-39
23 1169 Dean J A, Welsh L C, Wong K H, Aleksic A, Dunne E, Islam M R, Patel A, Patel P, Petkar I, Phillips I,
24 1170 Sham J, Schick U, Newbold K L, Bhide S A, Harrington K J, Nutting C M and Gulliford S L 2017
25 1171 Normal Tissue Complication Probability (NTCP) Modelling of Severe Acute Mucositis using a
26 1172 Novel Oral Mucosal Surface Organ at Risk *Clin Oncol (R Coll Radiol)* **29** 263-73
27 1173 Dean J A, Wong K H, Welsh L C, Jones A B, Schick U, Newbold K L, Bhide S A, Harrington K J, Nutting C
28 1174 M and Gulliford S L 2016 Normal tissue complication probability (NTCP) modelling using
29 1175 spatial dose metrics and machine learning methods for severe acute oral mucositis resulting
30 1176 from head and neck radiotherapy *Radioth Oncol* **120** 21-7
31 1177 Deasy J and Adita A 2013 *Informatics in Radiation Oncology*, ed G Starkschall and R A Siochi (Bosa
32 1178 Roca: CRC Press Inc)
33 1179 Deasy J O, Bentzen S M, Jackson A, Ten Haken R K, Yorke E D, Constine L S, Sharma A and Marks L B
34 1180 2010 Improving normal tissue complication probability models: the need to adopt a "data-
35 1181 pooling" culture *Int J Radiat Oncol, Biol, Phys* **76** S151-S4
36 1182 Deasy J O, Blanco A I and Clark V H 2003 CERR: a computational environment for radiotherapy
37 1183 research *Med Phys* **30** 979-85
38 1184 Deasy J O and El Naqa I 2008 Image-based modeling of normal tissue complication probability for
39 1185 radiation therapy *Cancer Treat Res* **139** 215-56
40 1186 Defraene G, Van den Bergh L, Al-Mamgani A, Haustermans K, Heemsbergen W, Van den Heuvel F
41 1187 and Lebesque J V 2012 The Benefits of Including Clinical Factors in Rectal Normal Tissue
42 1188 Complication Probability Modeling After Radiotherapy for Prostate Cancer *Int J Radiat Oncol,*
43 1189 *Biol, Phys* **82** 1233-42
44 1190 Deist T M, Jochems A, van Soest J, Nalbantov G, Oberije C, Walsh S, Eble M, Bulens P, Coucke P, Dries
45 1191 W, Dekker A and Lambin P 2017 Infrastructure and distributed learning methodology for
46 1192 privacy-preserving multi-centric rapid learning health care: euroCAT *Clin Transl Radiat Oncol*
47 1193 **4** 24-31
48 1194 Dörr W 2015 Radiobiology of tissue reactions *Annals of the ICRP* **44** 58-68
49 1195 Drean G, Acosta O, Lafond C, Simon A, de Crevoisier R and Haigron P 2016a Interindividual
50 1196 registration and dose mapping for voxelwise population analysis of rectal toxicity in prostate
51 1197 cancer radiotherapy *Med Phys* **43** 2721-30
52
53
54
55
56
57
58
59
60

- 1
2
3 1198 Drean G, Acosta O, Ospina J D, Fargeas A, Lafond C, Correge G, Lagrange J L, Crehange G, Simon A,
4 1199 Haigron P and de Crevoisier R 2016b Identification of a rectal subregion highly predictive of
5 1200 rectal bleeding in prostate cancer IMRT *Radioth Oncol* **119** 388-97
- 6 1201 Dréan G, Acosta O, Ospina J D, Voisin C, Rigaud B, Simon A, Haigron P and de Crevoisier R 2013 How
7 1202 to identify rectal sub-regions likely involved in rectal bleeding in prostate cancer
8 1203 radiotherapy. In: *IX International Seminar on Medical Information Processing and Analysis*,
9 1204 ed SPIE (Mexico DF: SPIE) p 9
- 10 1205 Ebert M A, Bulsara M, Haworth A, Kearvell R, Richardson S, Kennedy A, Spry N A, Bydder S A, Joseph
11 1206 D J and Denham J W 2015a Technical quality assurance during the TROG 03.04 RADAR
12 1207 prostate radiotherapy trial: Are the results reflected in observed toxicity rates? *J Med*
13 1208 *Imaging Radiat Oncol* **59** 99-108
- 14 1209 Ebert M A, Foo K, Haworth A, Gulliford S L, Kennedy A, Joseph D J and Denham J W 2015b
15 1210 Gastrointestinal Dose-Histogram Effects in the Context of Dose-Volume Constrained
16 1211 Prostate Radiation Therapy: Analysis of Data From the RADAR Prostate Radiation Therapy
17 1212 Trial *Int J Radiat Oncol Biol Phys* **91** 595-603
- 18 1213 Ebert M A, Harrison K M, Howlett S J, Cornes D, Bulsara M, Hamilton C S, Kron T, Joseph D J and
19 1214 Denham J W 2011 Dosimetric intercomparison for multicenter clinical trials using a patient-
20 1215 based anatomic pelvic phantom *Med Phys* **38** 5167-75
- 21 1216 Ebert M A, Haworth A, Kearvell R, Hooton B, Hug B, Spry N A, Bydder S A and Joseph D J 2010
22 1217 Comparison of DVH data from multiple radiotherapy treatment planning systems *Phys Med*
23 1218 *Biol* **55** N337-N46
- 24 1219 El Naqa I, Bradley J D, Lindsay P E, Hope A J and Deasy J O 2009 Predicting radiotherapy outcomes
25 1220 using statistical learning techniques *Phys Med Biol* **54** S9-s30
- 26 1221 El Naqa I and Das S 2020 The role of machine and deep learning in modern medical physics *Med Phys*
27 1222 **47** e125-e6
- 28 1223 Emami B, Lyman J, Brown A, Coia L, Goitein M, Munzenrider J E, Shank B, Solin L J and Wesson M
29 1224 1991 Tolerance of normal tissue to therapeutic irradiation *Int J Radiat Oncol, Biol, Phys* **21**
30 1225 109-22
- 31 1226 Esplen N M, Mendonca M S and Bazalova-Carter M 2020 Physics and biology of ultrahigh dose-rate
32 1227 (FLASH) radiotherapy: a topical review *Phys Med Biol*
- 33 1228 Fargeas A, Acosta O, Ospina Arrango J D, Ferhat A, Costet N, Albera L, Azria D, Fenoglio P,
34 1229 Créhange G, Beckendorf V, Hatt M, Kachenoura A and de Crevoisier R 2018 Independent
35 1230 component analysis for rectal bleeding prediction following prostate cancer radiotherapy
36 1231 *Radioth Oncol* **126** 263-9
- 37 1232 Fargeas A, Kachenoura A, Acosta O, Albera L, Drean G and De Crevoisier R 2013 Feature extraction
38 1233 and classification for rectal bleeding in prostate cancer radiotherapy: A PCA based method
39 1234 *IRBM* **34** 296-9
- 40 1235 Fiorino C, Gianolini S and Nahum A E 2003 A cylindrical model of the rectum: comparing dose-
41 1236 volume, dose-surface and dose-wall histograms in the radiotherapy of prostate cancer *Phys*
42 1237 *Med Biol* **48** 2603-16
- 43 1238 Fiorino C, Vavassori V, Sanguineti G, Bianchi C, Cattaneo G M, Piazzolla A and Cozzarini C 2002
44 1239 Rectum contouring variability in patients treated for prostate cancer: impact on rectum
45 1240 dose-volume histograms and normal tissue complication probability *Radioth Oncol* **63** 249-
46 1241 55
- 47 1242 Fu Y, Lei Y, Wang T, Curran W J, Liu T and Yang X 2020 Deep learning in medical image registration: a
48 1243 review *Phys Med Biol* **65** 20TR01
- 49 1244 Gabryś H S, Buettner F, Sterzing F, Hauswald H and Bangert M 2018 Design and Selection of Machine
50 1245 Learning Methods Using Radiomics and Dosiomics for Normal Tissue Complication
51 1246 Probability Modeling of Xerostomia *Front Oncol* **8** 35
- 52
53
54
55
56
57
58
59
60

- 1
2
3 1247 Gale N, House M and Ebert M A 2017 Using percolation networks to incorporate spatial-dose
4 1248 information for assessment of complication probability in radiotherapy *Australas Phys Eng*
5 1249 *Sci Med* **40** 869-80
- 6
7 1250 Ghadjar P, Zelefsky M J, Spratt D E, Munck af Rosenschöld P, Oh J H, Hunt M, Kollmeier M,
8 1251 Happersett L, Yorke E, Deasy J O and Jackson A 2014 Impact of Dose to the Bladder Trigone
9 1252 on Long-Term Urinary Function After High-Dose Intensity Modulated Radiation Therapy for
10 1253 Localized Prostate Cancer *Int J Radiat Oncol Biol Phys* **88** 339-44
- 11 1254 Ghobadi G, van der Veen S, Bartelds B, de Boer R A, Dickinson M G, de Jong J R, Faber H,
12 1255 Niemantsverdriet M, Brandenburg S, Berger R M, Langendijk J A, Coppes R P and van Luijk P
13 1256 2012 Physiological interaction of heart and lung in thoracic irradiation *Int J Radiat Oncol Biol*
14 1257 *Phys* **84** e639-46
- 15 1258 Green A, Vasquez Osorio E, Aznar M C, McWilliam A and van Herk M 2020 Image Based Data Mining
16 1259 Using Per-voxel Cox Regression *Front Oncol* **10** 1178
- 17 1260 Gulliford S L, Ghose S, Ebert M A, Kennedy A, Dowling J, Mitra J, Joseph D J and Denham J W 2017
18 1261 Radiotherapy dose-distribution to the perirectal fat space (PRS) is related to gastrointestinal
19 1262 control-related complications *Clin Transl Radiat Oncol* **7** 62-70
- 20 1263 Gulliford S L, Webb S, Rowbottom C G, Corne D W and Dearnaley D P 2004 Use of artificial neural
21 1264 networks to predict biological outcomes for patients receiving radical radiotherapy of the
22 1265 prostate *Radioth Oncol* **71** 3-12
- 23 1266 Guyon I and Elisseeff A 2003 An introduction to variable and feature selection *J. Mach. Learn. Res.* **3**
24 1267 1157-82
- 25 1268 Hayman J A, Dekker A, Feng M, Keole S R, McNutt T R, Machtay M, Martin N E, Mayo C S, Pawlicki T,
26 1269 Smith B D, Kudner R, Dawes S and Yu J B 2019 Minimum Data Elements for Radiation
27 1270 Oncology: An American Society for Radiation Oncology Consensus Paper *Pract Radiat Oncol*
28 1271 **9** 395-401
- 29 1272 Heemsbergen W D, Al-Mamgani A, Witte M G, van Herk M, Pos F J and Lebesque J V 2010 Urinary
30 1273 Obstruction in Prostate Cancer Patients From the Dutch Trial (68 Gy vs. 78 Gy): Relationships
31 1274 with Local Dose, Acute Effects, and Baseline Characteristics *Int J Radiat Oncol Biol Phys* **78**
32 1275 19-25
- 33 1276 Heemsbergen W D, Hoogeman M S, Hart G A M, Lebesque J V and Koper P C M 2005 Gastrointestinal
34 1277 toxicity and its relation to dose distributions in the anorectal region of prostate cancer
35 1278 patients treated with radiotherapy *Int J Radiat Oncol Biol Phys* **61** 1011-8
- 36 1279 Heemsbergen W D, Incrocci L, Pos F J, Heijmen B J M and Witte M G 2020 Local Dose Effects for Late
37 1280 Gastrointestinal Toxicity After Hypofractionated and Conventionally Fractionated Modern
38 1281 Radiotherapy for Prostate Cancer in the HYPRO Trial *Front Oncol* **10**
- 39 1282 Heinze G, Wallisch C and Dunkler D 2018 Variable selection - A review and recommendations for the
40 1283 practicing statistician *Biom J* **60** 431-49
- 41 1284 Henderson D R, Murray J R, Gulliford S L, Tree A C, Harrington K J and Van As N J 2018 An
42 1285 Investigation of Dosimetric Correlates of Acute Toxicity in Prostate Stereotactic Body
43 1286 Radiotherapy: Dose to Urinary Trigone is Associated with Acute Urinary Toxicity *Clin Oncol (R*
44 1287 *Coll Radiol)* **30** 539-47
- 45 1288 Hoogeman M S, Peeters S T, de Bois J and Lebesque J V 2005 Absolute and relative dose-surface and
46 1289 dose-volume histograms of the bladder: which one is the most representative for the actual
47 1290 treatment? *Phys Med Biol* **50** 3589-97
- 48 1291 Hoogeman M S, van Herk M, de Bois J, Muller-Timmermans P, Koper P C M and Lebesque J V 2004
49 1292 Quantification of local rectal wall displacements by virtual rectum unfolding *Radioth Oncol*
50 1293 **70** 21-30
- 51 1294 Hrycushko B, van der Kogel A J, Phillips L, Folkert M R, Sayre J W, Vernino S, Hassan-Rezaeian N,
52 1295 Foster R D, Yamada Y, Timmerman R and Medin P M 2019 Spinal Nerve Tolerance to Single-
53 1296 Session Stereotactic Ablative Radiation Therapy *Int J Radiat Oncol Biol Phys* **104** 845-51
- 54
55
56
57
58
59
60

- 1
2
3 1297 Ibragimov B, Toesca D, Chang D, Yuan Y, Koong A and Xing L 2018 Development of deep neural
4 1298 network for individualized hepatobiliary toxicity prediction after liver SBRT *Med Phys* **45**
5 1299 4763-74
6
7 1300 Ibragimov B, Toesca D A S, Chang D T, Yuan Y, Koong A C, Xing L and Vogelius I R 2020 Deep learning
8 1301 for identification of critical regions associated with toxicities after liver stereotactic body
9 1302 radiation therapy *Med Phys* **47** 3721-31
10 1303 Ibragimov B, Toesca D A S, Yuan Y, Koong A C, Chang D T and Xing L 2019 Neural Networks for Deep
11 1304 Radiotherapy Dose Analysis and Prediction of Liver SBRT Outcomes *IEEE J Biomed Health*
12 1305 *Inform* **23** 1821-33
13 1306 Improta I, Palorini F, Cozzarini C, Rancati T, Avuzzi B, Franco P, Degli Esposti C, Del Mastro E, Girelli G,
14 1307 Iotti C, Vavassori V, Valdagni R and Fiorino C 2016 Bladder spatial-dose descriptors correlate
15 1308 with acute urinary toxicity after radiation therapy for prostate cancer *Phys Med* **32** 1681-9
16 1309 Jackson A, Marks L B, Bentzen S M, Eisbruch A, Yorke E D, Ten Haken R K, Constine L S and Deasy J O
17 1310 2010 The lessons of QUANTEC: recommendations for reporting and gathering data on dose-
18 1311 volume dependencies of treatment outcome *Int J Radiat Oncol, Biol, Phys* **76** S155-S60
19 1312 Jaffray D A, Lindsay P E, Brock K K, Deasy J O and Tomé W A 2010 Accurate accumulation of dose for
20 1313 improved understanding of radiation effects in normal tissue *Int J Radiat Oncol, Biol, Phys* **76**
21 1314 S135-S9
22
23 1315 Jiang W, Lakshminarayanan P, Hui X, Han P, Cheng Z, Bowers M, Shpitser I, Siddiqui S, Taylor R H,
24 1316 Quon H and McNutt T 2019 Machine Learning Methods Uncover Radiomorphologic Dose
25 1317 Patterns in Salivary Glands that Predict Xerostomia in Patients with Head and Neck Cancer
26 1318 *Adv Radiat Oncol* **4** 401-12
27
28 1319 Johnson-Hart C N, Price G J, Faivre-Finn C, Aznar M C and van Herk M 2018 Residual Setup Errors
29 1320 Towards the Heart After Image Guidance Linked With Poorer Survival in Lung Cancer
30 1321 Patients: Do We Need Stricter IGRT Protocols? *Int J Radiat Oncol Biol Phys* **102** 434-42
31 1322 Källman P, Lind B K and Brahme A 1992 An algorithm for maximizing the probability of complication-
32 1323 free tumour control in radiation therapy *Phys Med Biol* **37** 871-90
33 1324 Kennedy A, Dowling J, Greer P B, Holloway L, Jameson M G, Roach D, Ghose S, Rivest-Henault D,
34 1325 Marcello M and Ebert M A 2019 Similarity clustering-based atlas selection for pelvic CT
35 1326 image segmentation *Med Phys* **46** 2243-50
36
37 1327 Kim D W, Cho L C, Straka C, Christie A, Lotan Y, Pistenmaa D, Kavanagh B D, Nanda A, Kueplian P,
38 1328 Brindle J, Cooley S, Perkins A, Raben D, Xie X J and Timmerman R D 2014 Predictors of rectal
39 1329 tolerance observed in a dose-escalated phase 1-2 trial of stereotactic body radiation therapy
40 1330 for prostate cancer *Int J Radiat Oncol Biol Phys* **89** 509-17
41 1331 Kim K-H, Chung J-B, Suh T S, Kang S-W, Kang S-H, Eom K-Y, Song C, Kim I-A and Kim J-S 2018
42 1332 Dosimetric and radiobiological comparison in different dose calculation grid sizes between
43 1333 Acuros XB and anisotropic analytical algorithm for prostate VMAT *PLOS ONE* **13** e0207232
44 1334 Kirkpatrick J P, van der Kogel A J and Schultheiss T E 2010 Radiation Dose–Volume Effects in
45 1335 the Spinal Cord *Int J Radiat Oncol, Biol, Phys* **76** S42-S9
46
47 1336 Krauss A 2018 Why all randomised controlled trials produce biased results *Annals of Medicine* **50**
48 1337 312-22
49 1338 Kruschke J K 2013 Bayesian estimation supersedes the t test *J Exp Psychol Gen* **142** 573-603
50 1339 La Macchia M, Fellin F, Amichetti M, Cianchetti M, Gianolini S, Paola V, Lomax A J and Widesott L
51 1340 2012 Systematic evaluation of three different commercial software solutions for automatic
52 1341 segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer *Radiat*
53 1342 *Oncol* **7** 160-
54
55 1343 Lafond C, Barateau A, N'Guessan J, Perichon N, Delaby N, Simon A, Haigrón P, Mylona E, Acosta O
56 1344 and de Crevoisier R 2020 Planning With Patient-Specific Rectal Sub-Region Constraints
57 1345 Decreases Probability of Toxicity in Prostate Cancer Radiotherapy *Front Oncol* **10** 1597
58
59
60

- 1
2
3 1346 Lee S H, Han P, Hales R, Voong K R, Noro K, Sugiyama S, Haller J W, McNutt T and Lee J 2020 Multi-
4 1347 view radiomics and dosiomics analysis with machine learning for predicting acute-phase
5 1348 weight loss in lung cancer patients treated with radiotherapy *Phys Med Biol*
6
7 1349 Lee S J and Park H J 2020 Single photon emission computed tomography (SPECT) or positron
8 1350 emission tomography (PET) imaging for radiotherapy planning in patients with lung cancer: a
9 1351 meta-analysis *Scientific Reports* **10** 14864
10 1352 Liang B, Tian Y, Chen X, Yan H, Yan L, Zhang T, Zhou Z, Wang L and Dai J 2020 Prediction of Radiation
11 1353 Pneumonitis With Dose Distribution: A Convolutional Neural Network (CNN) Based Model
12 1354 *Front Oncol* **9**
13 1355 Liang B, Yan H, Tian Y, Chen X, Yan L, Zhang T, Zhou Z, Wang L and Dai J 2019 Dosiomics: Extracting
14 1356 3D Spatial Features From Dose Distribution to Predict Incidence of Radiation Pneumonitis
15 1357 *Front Oncol* **9**
16
17 1358 Lu Y, Li S, Spelbring D, Song P, Vijayakumar S, Pelizzari C and Chen G T Y 1995 Dose-surface
18 1359 histograms as treatment planning tool for prostate conformal therapy *Med Phys* **22** 279-84
19 1360 Lu Y, Spelbring D R and Chen G T Y 1997 Functional dose - volume histograms for functionally
20 1361 heterogeneous normal organs *Phys Med Biol* **42** 345-56
21 1362 Lumley T, Diehr P, Emerson S and Chen L 2002 The importance of the normality assumption in large
22 1363 public health data sets *Annu Rev Public Health* **23** 151-69
23 1364 Luo Y, Chen S and Valdes G 2020 Machine learning for radiation outcome modeling and prediction
24 1365 *Med Phys* **47** e178-e84
25
26 1366 Lyman J T 1985 Complication Probability as Assessed from Dose-Volume Histograms *Rad Res* **8** 13-9
27 1367 Magallon-Baro A, Loi M, Milder M T W, Granton P V, Zolnay A G, Nuytens J J and Hoogeman M S
28 1368 2019 Modeling daily changes in organ-at-risk anatomy in a cohort of pancreatic cancer
29 1369 patients *Radioth Oncol* **134** 127-34
30 1370 Manly B F J 1997 *Randomization, Bootstrap and Monte Carlo Methods in Biology* (London: Chapman
31 1371 and Hall)
32 1372 Marcello M, Denham J W, Kennedy A, Haworth A, Steigler A, Greer P, Holloway L, Dowling J,
33 1373 Jameson M, Roach D, Joseph D J, Gulliford S L, Dearnaley D P, Sydes M R, Hall E and Ebert M
34 1374 A 2020a Increased dose to organs in urinary tract associates with measures of genitourinary
35 1375 toxicity in pooled voxel-based analysis of 3 randomized Phase III trials *Front Oncol* **10**
36 1376 Marcello M, Denham J W, Kennedy A, Haworth A, Steigler A, Greer P B, Holloway L C, Dowling J A,
37 1377 Jameson M G, Roach D, Joseph D J, Gulliford S L, Dearnaley D P, Sydes M R, Hall E and Ebert
38 1378 M A 2020b Relationships between rectal and perirectal doses and rectal bleeding or
39 1379 tenesmus in pooled voxel-based analysis of 3 randomised phase III trials *Radiotherapy and*
40 1380 *Oncology* **150** 281-92
41
42 1381 Marks L B, Yorke E D, Jackson A, Ten Haken R K, Constine L S, Eisbruch A, Bentzen S M, Nam J and
43 1382 Deasy J O 2010 Use of normal tissue complication probability models in the clinic *Int J Radiat*
44 1383 *Oncol, Biol, Phys* **76** S10-S9
45
46 1384 Mayo C S, Moran J M, Bosch W, Xiao Y, McNutt T, Popple R, Michalski J, Feng M, Marks L B, Fuller C
47 1385 D, Yorke E, Palta J, Gabriel P E, Molineu A, Matuszak M M, Covington E, Masi K, Richardson S
48 1386 L, Ritter T, Morgas T, Flampouri S, Santanam L, Moore J A, Purdie T G, Miller R C, Hurkmans
49 1387 C, Adams J, Jackie Wu Q R, Fox C J, Siochi R A, Brown N L, Verbakel W, Archambault Y,
50 1388 Chmura S J, Dekker A L, Eagle D G, Fitzgerald T J, Hong T, Kapoor R, Lansing B, Jolly S,
51 1389 Napolitano M E, Percy J, Rose M S, Siddiqui S, Schadt C, Simon W E, Straube W L, St James S
52 1390 T, Ulin K, Yom S S and Yock T I 2018 American Association of Physicists in Medicine Task
53 1391 Group 263: Standardizing Nomenclatures in Radiation Oncology *Int J Radiat Oncol Biol Phys*
54 1392 **100** 1057-66
55
56 1393 McWilliam A, Dootson C, Graham L, Banfill K, Abravan A and van Herk M 2020 Dose surface maps of
57 1394 the heart can identify regions associated with worse survival for lung cancer patients treated
58 1395 with radiotherapy *Phys Imag Radiat Oncol* **15** 46-51
59
60

- 1
2
3 1396 McWilliam A, Kennedy J, Hodgson C, Vasquez Osorio E, Faivre-Finn C and van Herk M 2017 Radiation
4 1397 dose to heart base linked with poorer survival in lung cancer patients *Eur J Can* **85** 106-13
5 1398 Medin P M and Boike T P 2011 Spinal cord tolerance in the age of spinal radiosurgery: lessons from
6 1399 preclinical studies *Int J Radiat Oncol, Biol, Phys* **79** 1302-9
7 1400 Meijer G J, van den Brink M, Hoogeman M S, Meinders J and Lebesque J V 1999 Dose-wall
8 1401 histograms and normalized dose-surface histograms for the rectum: a new method to
9 1402 analyze the dose distribution over the rectum in conformal radiotherapy *Int J Radiat Oncol*
10 1403 *Biol Phys* **45** 1073-80
11 1404 Men K, Geng H, Zhong H, Fan Y, Lin A and Xiao Y 2019 A Deep Learning Model for Predicting
12 1405 Xerostomia Due to Radiation Therapy for Head and Neck Squamous Cell Carcinoma in the
13 1406 RTOG 0522 Clinical Trial *Int J Radiat Oncol Biol Phys* **105** 440-7
14 1407 Meroni S, Cavatorta C, Barra S, Cavagnetto F, Scarzello G, Scaggion A, Pecori E, Diletto B, Alessandro
15 1408 O, Massimino M, Gianolini S, Pignoli E and Gandola L 2019 A dedicated cloud system for real-
16 1409 time upfront quality assurance in pediatric radiation therapy *Strahlenther Onkol* **195** 843-50
17 1410 Michalski J M, Gay H, Jackson A, Tucker S L and Deasy J O 2010 Radiation dose-volume effects in
18 1411 radiation-induced rectal injury *Int J Radiat Oncol, Biol, Phys* **76** S123-S9
19 1412 Molineu A, Hernandez N, Nguyen T, Ibbott G and Followill D 2013 Credentialing results from IMRT
20 1413 irradiations of an anthropomorphic head and neck phantom *Med Phys* **40** 022101
21 1414 Monti S, Pacelli R, Cella L and Palma G 2018 Inter-patient image registration algorithms to
22 1415 disentangle regional dose bioeffects *Scientific Reports* **8** 4915
23 1416 Monti S, Paganelli C, Buizza G, Preda L, Valvo F, Baroni G, Palma G and Cella L 2020 A novel
24 1417 framework for spatial normalization of dose distributions in voxel-based analyses of brain
25 1418 irradiation outcomes *Phys Med* **69** 164-9
26 1419 Monti S, Palma G, D'Avino V, Gerardi M, Marvaso G, Ciardo D, Pacelli R, Jereczek-Fossa B A, Alterio D
27 1420 and Cella L 2017 Voxel-based analysis unveils regional dose differences associated with
28 1421 radiation-induced morbidity in head and neck cancer patients *Scientific Reports* **7** 7220
29 1422 Morimoto M, Bijl H P, A V D S, Xu C J, Steenbakkens R, Chouvalova O, Yoshioka Y, Teshima T and
30 1423 Langendijk J A 2019 Development of Normal Tissue Complication Probability Model for
31 1424 Trismus in Head and Neck Cancer Patients Treated With Radiotherapy: The Role of
32 1425 Dosimetric and Clinical Factors *Anticancer Res* **39** 6787-98
33 1426 Moulton C R, House M J, Lye V, Tang C I, Krawiec M, Joseph D J, Denham J W and Ebert M A 2017
34 1427 Spatial features of dose-surface maps from deformably-registered plans correlate with late
35 1428 gastrointestinal complications *Phys Med Biol* **62** 4118
36 1429 Munbodh R, Jackson A, Bauer J, Schmidtlein C R and Zelefsky M J 2008 Dosimetric and anatomic
37 1430 indicators of late rectal toxicity after high-dose intensity modulated radiation therapy for
38 1431 prostate cancer *Med Phys* **35** 2137-50
39 1432 Murphy K, Ginneken B v, Reinhardt J M, Kabus S, Ding K, Deng X, Cao K, Du K, Christensen G E, Garcia
40 1433 V, Vercauteren T, Ayache N, Commowick O, Malandain G, Glocker B, Paragios N, Navab N,
41 1434 Gorbunova V, Sporring J, Bruijne M d, Han X, Heinrich M P, Schnabel J A, Jenkinson M,
42 1435 Lorenz C, Modat M, McClelland J R, Ourselin S, Muenzing S E A, Viergever M A, Nigris D D,
43 1436 Collins D L, Arbel T, Peroni M, Li R, Sharp G C, Schmidt-Richberg A, Ehrhardt J, Werner R,
44 1437 Smeets D, Loeckx D, Song G, Tustison N, Avants B, Gee J C, Staring M, Klein S, Stoel B C,
45 1438 Urschler M, Werlberger M, Vandemeulebroucke J, Rit S, Sarrut D and Pluim J P W 2011
46 1439 Evaluation of Registration Methods on Thoracic CT: The EMPIRE10 Challenge *IEEE Trans Med*
47 1440 *Imaging* **30** 1901-20
48 1441 Myers C and Niemierko A 2004 Percolation-based cluster models of dose-volume effects *Int J Radiat*
49 1442 *Oncol Biol Phys* **60** S157
50 1443 Mylona E, Acosta O, Lizee T, Lafond C, Crehange G, Magné N, Chiavassa S, Supiot S, Arango Ospina J
51 1444 D, Campillo-Gimenez B, Castelli J and de Crevoisier R 2019 Voxel-based analysis for
52 1445 identification of urethro-vesical subregions predicting urinary toxicity after prostate cancer
53 1446 radiotherapy *Int J Radiat Oncol Biol Phys* **104** 343-54
54
55
56
57
58
59
60

- 1
2
3 1447 Mylona E, Cicchetti A, Rancati T, Palorini F, Fiorino C, Supiot S, Magne N, Crehange G, Valdagni R,
4 1448 Acosta O and de Crevoisier R 2020a Local dose analysis to predict acute and late urinary
5 1449 toxicities after prostate cancer radiotherapy: Assessment of cohort and method effects
6 1450 *Radioth Oncol* **147** 40-9
7
8 1451 Mylona E, Ebert M, Kennedy A, Joseph D, Denham J, Steigler A, Supiot S, Acosta O and de Crevoisier
9 1452 R 2020b Rectal and Urethro-vesical Subregions for Toxicity Prediction After Prostate Cancer
10 1453 Radiotherapy: validation of voxel-based models in an independent population *Int J Radiat*
11 1454 *Oncol Biol Phys*
12 1455 NEMA Digital Imaging and Communications in Medicine (DICOM) Standard. (Rosslyn, VA, USA:
13 1456 National Electrical Manufacturers Association)
14 1457 NIfTI 2020 Neuroimaging Informatics Technology Initiative. Neuroimaging Informatics Technology
15 1458 Initiative)
16 1459 Nioutsikou E, Webb S, Panakis N, Bortfeld T and Oelfke U 2005 Reconsidering the definition of a
17 1460 dose-volume histogram *Phys Med Biol* **50** L17-9
18 1461 Nitsche M, Brannath W, Brückner M, Wagner D, Kaltenborn A, Temme N and Hermann R M 2017
19 1462 Comparison of different contouring definitions of the rectum as organ at risk (OAR) and
20 1463 dose-volume parameters predicting rectal inflammation in radiotherapy of prostate cancer:
21 1464 which definition to use? *The British journal of radiology* **90** 20160370-
22 1465 Ohri N, Shen X, Dicker A P, Doyle L A, Harrison A S and Showalter T N 2013 Radiotherapy protocol
23 1466 deviations and clinical outcomes: A meta-analysis of cooperative group clinical trials *J Natl*
24 1467 *Cancer Inst* **105** 387-93
25 1468 Onjukka E, Fiorino C, Cicchetti A, Palorini F, Improta I, Gagliardi G, Cozzarini C, Degli Esposti C,
26 1469 Gabriele P, Valdagni R and Rancati T 2019 Patterns in ano-rectal dose maps and the risk of
27 1470 late toxicity after prostate IMRT *Acta oncologica* **58** 1757-64
28 1471 Ospina J D, Zhu J, Chira C, Bossi A, Delobel J B, Beckendorf V, Dubray B, Lagrange J L, Correa J C,
29 1472 Simon A, Acosta O and de Crevoisier R 2014 Random forests to predict rectal toxicity
30 1473 following prostate cancer radiation therapy *Int J Radiat Oncol Biol Phys* **89** 1024-31
31 1474 Palma G and Cella L 2019 A new formalism of Dose Surface Histograms for robust modeling of skin
32 1475 toxicity in radiation therapy *Phys Med* **59** 75-8
33 1476 Palma G, Monti S, Buonanno A, Pacelli R and Cella L 2019a PACE: A Probabilistic Atlas for Normal
34 1477 Tissue Complication Estimation in Radiation Oncology *Front Oncol* **9** 130-
35 1478 Palma G, Monti S and Cella L 2020a Voxel-based analysis in radiation oncology: A methodological
36 1479 cookbook *Phys Med* **69** 192-204
37 1480 Palma G, Monti S, Conson M, Pacelli R and Cella L 2019b Normal tissue complication probability
38 1481 (NTCP) models for modern radiation therapy *Semin Oncol* **46** 210-8
39 1482 Palma G, Monti S, D'Avino V, Conson M, Liuzzi R, Pressello M C, Donato V, Deasy J O, Quarantelli M,
40 1483 Pacelli R and Cella L 2016 A Voxel-Based Approach to Explore Local Dose Differences
41 1484 Associated With Radiation-Induced Lung Damage *Int J Radiat Oncol Biol Phys* **96** 127-33
42 1485 Palma G, Monti S, Thor M, Rimner A, Deasy J O and Cella L 2019c Spatial signature of dose patterns
43 1486 associated with acute radiation-induced lung damage in lung cancer patients treated with
44 1487 stereotactic body radiation therapy *Phys Med Biol* **64** 155006
45 1488 Palma G, Monti S, Xu T, Scifoni E, Yang P, Hahn S M, Durante M, Mohan R, Liao Z and Cella L 2019d
46 1489 Spatial dose patterns associated with radiation pneumonitis in a randomized trial comparing
47 1490 intensity-modulated photon therapy with passive scattering proton therapy for locally
48 1491 advanced non-small cell lung cancer *Int J Radiat Oncol Biol Phys* **104** 1124-32
49 1492 Palma G, Taffelli A, Fellin F, D'Avino V, Scartoni D, Tommasino F, Scifoni E, Durante M, Amichetti M,
50 1493 Schwarz M, Amelio D and Cella L 2020b Modelling the risk of radiation induced alopecia in
51 1494 brain tumor patients treated with scanned proton beams *Radioth Oncol* **144** 127-34
52 1495 Palorini F, Botti A, Carillo V, Gianolini S, Improta I, Iotti C, Rancati T, Cozzarini C and Fiorino C 2016a
53 1496 Bladder dose-surface maps and urinary toxicity: Robustness with respect to motion in
54 1497 assessing local dose effects *Phys Med* **32** 506-11
55
56
57
58
59
60

- 1
2
3 1498 Palorini F, Cozzarini C, Gianolini S, Botti A, Carillo V, Iotti C, Rancati T, Valdagni R and Fiorino C 2016b
4 1499 First application of a pixel-wise analysis on bladder dose surface maps in prostate cancer
5 1500 radiotherapy *Radioth Oncol* **119** 123-8
6
7 1501 Peduzzi P, Concato J, Kemper E, Holford T R and Feinstein A R 1996 A simulation study of the number
8 1502 of events per variable in logistic regression analysis *J Clin Epidem* **49** 1373-9
9 1503 Peeters S T, Hoogeman M S, Heemsbergen W D, Hart A A, Koper P C and Lebesque J V 2006a Rectal
10 1504 bleeding, fecal incontinence, and high stool frequency after conformal radiotherapy for
11 1505 prostate cancer: normal tissue complication probability modeling *Int J Radiat Oncol Biol Phys*
12 1506 **66** 11-9
13 1507 Peeters S T, Hoogeman M S, Heemsbergen W D, Slot A, Tabak H, Koper P C and Lebesque J V 2005
14 1508 Volume and hormonal effects for acute side effects of rectum and bladder during conformal
15 1509 radiotherapy for prostate cancer *Int J Radiat Oncol Biol Phys* **63** 1142-52
16 1510 Peeters S T H, Lebesque J V, Heemsbergen W D, van Putten W L J, Slot A, Dielwart M F H and Koper P
17 1511 C M 2006b Localized volume effects for late rectal and anal toxicity after radiotherapy for
18 1512 prostate cancer *Int. J. Radiat. Oncol. Biol. Phys.* **64** 1151-61
19 1513 Phillips M H, Serra L M, Dekker A, Ghosh P, Luk S M H, Kalet A and Mayo C 2020 Ontologies in
20 1514 radiation oncology *Phys Med* **72** 103-13
21 1515 Placidi L, Lenkowicz J, Cusumano D, Boldrini L, Dinapoli N and Valentini V 2020 Stability of dosomics
22 1516 features extraction on grid resolution and algorithm for radiotherapy dose calculation *Phys*
23 1517 *Med* **77** 30-5
24 1518 Purdy J A 2008 Quality assurance issues in conducting multi-institutional advanced technology
25 1519 clinical trials *Int J Radiat Oncol, Biol, Phys* **71** S66-70
26 1520 Purdy J A, Harms W B, Michalski J and Bosch W R 1998 Initial experience with quality assurance of
27 1521 multi-institutional 3D radiotherapy clinical trials. A brief report *Strahlentherapie und*
28 1522 *Onkologie* **174 Suppl 2** 40-2
29 1523 Rancati T, Fiorino C, Fellin G, Vavassori V, Cagna E, Casanova Borca V, Girelli G, Menegotti L, Monti A
30 1524 F, Tortoreto F, Delle Canne S and Valdagni R 2011 Inclusion of clinical risk factors into NTCP
31 1525 modelling of late rectal toxicity after high dose radiotherapy for prostate cancer *Radioth*
32 1526 *Oncol* **100** 124-30
33 1527 Rancati T, Fiorino C, Gagliardi G, Cattaneo G M, Sanguineti G, Borca V C, Cozzarini C, Fellin G,
34 1528 Foppiano F, Girelli G, Menegotti L, Piazzolla A, Vavassori V and Valdagni R 2004 Fitting late
35 1529 rectal bleeding data using different NTCP models: results from an Italian multi-centric study
36 1530 (AIROPROS0101) *Radioth Oncol* **73** 21-32
37 1531 Rancati T, Palorini F, Cozzarini C, Fiorino C and Valdagni R 2017 Understanding urinary toxicity after
38 1532 radiotherapy for prostate cancer: first steps forward *Tumori* **103** 395-404
39 1533 Rigaud B, Simon A, Castelli J, Lafond C, Acosta O, Haigron P, Cazoulat G and de Crevoisier R 2019
40 1534 Deformable image registration for radiation therapy: principle, methods, applications and
41 1535 evaluation *Acta oncologica* **58** 1225-37
42 1536 Roach D, Holloway L C, Jameson M G, Dowling J A, Kennedy A, Greer P B, Krawiec M, Rai R, Denham
43 1537 J, De Leon J, Lim K, Berry M E, White R T, Bydder S A, Tan H T, Croker J D, McGrath A,
44 1538 Matthews J, Smeenk R J and Ebert M A 2019 Multi-observer contouring of male pelvic
45 1539 anatomy: Highly variable agreement across conventional and emerging structures of interest
46 1540 *J Med Imaging Radiat Oncol* **63** 264-71
47 1541 Robertson S P, Quon H, Kiess A P, Moore J A, Yang W, Cheng Z, Afonso S, Allen M, Richardson M,
48 1542 Choflet A, Sharabi A and McNutt T R 2015 A data-mining framework for large scale analysis
49 1543 of dose-outcome relationships in a database of irradiated head and neck cancer patients
50 1544 *Med Phys* **42** 4329-37
51 1545 Roelofs E, Dekker A, Meldolesi E, van Stiphout R G P M, Valentini V and Lambin P 2014 International
52 1546 data-sharing for radiotherapy research: an open-source based infrastructure for multicentric
53 1547 clinical data mining *Radioth Oncol* **110** 370-4
54
55
56
57
58
59
60

- 1
2
3 1548 Rossi L, Bijman R, Schillemans W, Aluwini S, Cavedon C, Witte M, Incrocci L and Heijmen B 2018
4 1549 Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in
5 1550 radiotherapy *Radioth Oncol* **129** 548-53
6
7 1551 Ryu S, Jin J Y, Jin R, Rock J, Ajlouni M, Movsas B, Rosenblum M and Kim J H 2007 Partial volume
8 1552 tolerance of the spinal cord and complications of single-dose radiosurgery *Cancer* **109** 628-
9 1553 36
10 1554 Saito T and Rehmsmeier M 2015 The precision-recall plot is more informative than the ROC plot
11 1555 when evaluating binary classifiers on imbalanced datasets *PLoS One* **10** e0118432
12 1556 Sanchez-Nieto B, Fenwick J, Nahum A and Dearnaley D P 2001 Biological dose surface maps:
13 1557 evaluation of 3D dose data for tubular organs *Radioth Oncol* **61** S52
14 1558 Santanam L, Hurkmans C, Mutic S, van Vliet-Vroegindewij C, Brame S, Straube W, Galvin J,
15 1559 Tripuraneni P, Michalski J and Bosch W 2012 Standardizing naming conventions in radiation
16 1560 oncology *Int J Radiat Oncol Biol Phys* **83** 1344-9
17 1561 Schaake W, van der Schaaf A, van Dijk L V, Bongaerts A H, van den Bergh A C and Langendijk J A 2016
18 1562 Normal tissue complication probability (NTCP) models for late rectal bleeding, stool
19 1563 frequency and fecal incontinence after radiotherapy in prostate cancer patients *Radioth*
20 1564 *Oncol* **119** 381-7
21
22 1565 Seppenwoolde Y, De Jaeger K, Boersma L J, Belderbos J S A and Lebesque J V 2004 Regional
23 1566 differences in lung radiosensitivity after radiotherapy for non-small-cell lung cancer *Int J*
24 1567 *Radiat Oncol Biol Phys* **60** 748-58
25 1568 Shelley L E A, Scaife J E, Romanchikova M, Harrison K, Forman J R, Bates A M, Noble D J, Jena R,
26 1569 Parker M A, Sutcliffe M P F, Thomas S J and Burnet N G 2017 Delivered dose can be a better
27 1570 predictor of rectal toxicity than planned dose in prostate radiotherapy *Radioth Oncol* **123**
28 1571 466-71
29 1572 Shmueli G 2010 To Explain or to Predict? *Statist. Sci.* **25** 289-310
30 1573 Smeenk R J, Hoffmann A L, Hopman W P, van Lin E N and Kaanders J H 2012 Dose-effect
31 1574 relationships for individual pelvic floor muscles and anorectal complaints after prostate
32 1575 radiotherapy *Int J Radiat Oncol Biol Phys* **83** 636-44
33 1576 Söhn M, Alber M and Yan D 2007 Principal Component Analysis-Based Pattern Analysis of Dose-
34 1577 Volume Histograms and Influence on Rectal Toxicity *Int J Radiat Oncol Biol Phys* **69** 230-9
35 1578 SourceForge 2020 NRRD. SourceForge)
36 1579 Stenmark M H, Conlon A S C, Johnson S, Daignault S, Litzenberg D, Marsh R, Ritter T, Vance S, Kazzi
37 1580 N, Feng F Y, Sandler H, Sanda M G and Hamstra D A 2014 Dose to the inferior rectum is
38 1581 strongly associated with patient reported bowel quality of life after radiation therapy for
39 1582 prostate cancer *Radioth Oncol* **110** 291-7
40 1583 Steyerberg E W and Vergouwe Y 2014 Towards better clinical prediction models: seven steps for
41 1584 development and an ABCD for validation *Eur Heart J* **35** 1925-31
42 1585 Stokke C, Gabiña P M, Solný P, Ciccone F, Sandström M, Gleisner K S, Chiesa C, Spezi E, Paphiti M,
43 1586 Konijnenberg M, Aldridge M, Tipping J, Wissmeyer M, Brans B, Bacher K, Kobe C and Flux G
44 1587 2017 Dosimetry-based treatment planning for molecular radiotherapy: a summary of the
45 1588 2017 report from the Internal Dosimetry Task Force *EJNMMI Physics* **4** 27
46 1589 Storey J D 2002 A Direct Approach to False Discovery Rates *J Royal Stat Soc B* **64** 479-98
47 1590 Taichman D B, Sahni P, Pinborg A, Peiperl L, Laine C, James A, Hong S-T, Haileamlak A, Gollogly L,
48 1591 Godlee F, Frizelle F A, Florenzano F, Drazen J M, Bauchner H, Baethge C and Backus J 2017
49 1592 Data Sharing Statements for Clinical Trials — A Requirement of the International Committee
50 1593 of Medical Journal Editors *NEJM* **376** 2277-9
51 1594 Talamonti C, Piffer S, Greto D, Mangoni M, Ciccarone A, Dicarolo P, Fantacci M E, Fusi F, Oliva P,
52 1595 Palumbo L, Favre C, Livi L, Pallotta S and Retico A 2019 Radiomic and Dosiomic Profiling of
53 1596 Paediatric Medulloblastoma Tumours Treated with Intensity Modulated Radiation Therapy.
54 1597 (Cham: Springer International Publishing) pp 56-64
55
56
57
58
59
60

- 1
2
3 1598 Thames H D, Zhang M, Tucker S L, Liu H H, Dong L and Mohan R 2004 Cluster models of dose-volume
4 1599 effects *Int J Radiat Oncol Biol Phys* **59** 1491-504
- 5 1600 Tibshirani R 1996 Regression Shrinkage and Selection via the Lasso *J Royal Stat Soc Ser B* **58** 267-88
- 6 1601 Tilly D, Tilly N and Ahnesjö A 2013 Dose mapping sensitivity to deformable registration uncertainties
7 1602 in fractionated radiotherapy - applied to prostate proton treatments *BMC Med Phys* **13** 2-
- 8 1603 Tomatis S, Rancati T, Fiorino C, Vavassori V, Fellin G, Cagna E, Mauro F, Girelli G, Monti A and
9 1604 Baccolini M 2012 Late rectal bleeding after 3D-CRT for prostate cancer: development of a
10 1605 neural-network-based predictive model *Phys Med Biol* **57** 1399
- 11 1606 Troeller A, Yan D, Marina O, Schulze D, Alber M, Parodi K, Belka C and Söhn M 2015 Comparison and
12 1607 limitations of DVH-based NTCP models derived from 3D-CRT and IMRT data for prediction of
13 1608 gastrointestinal toxicities in prostate cancer patients by using propensity score matched pair
14 1609 analysis *Int J Radiat Oncol Biol Phys* **91** 435-43
- 15 1610 Trott K-R, Doerr W, Facoetti A, Hopewell J, Langendijk J, van Luijk P, Ottolenghi A and Smyth V 2012
16 1611 Biological mechanisms of normal tissue damage: Importance for the design of NTCP models
17 1612 *Radioth Oncol* **105** 79-85
- 18 1613 Trotti A, Colevas A D, Setser A, Rusch V, Jaques D, Budach V, Langer C, Murphy B, Cumberlin R,
19 1614 Coleman C N and Rubin P 2003 CTCAE v3.0: development of a comprehensive grading
20 1615 system for the adverse effects of cancer treatment *Semin Radiat Oncol* **13** 176-81
- 21 1616 Tucker S L, Liao Z, Dinh J, Bian S X, Mohan R, Martel M K and Grosshans D R 2014 Is there an impact
22 1617 of heart exposure on the incidence of radiation pneumonitis? Analysis of data from a large
23 1618 clinical cohort *Acta oncologica* **53** 590-6
- 24 1619 Tucker S L, Liu H H, Wang S, Wei X, Liao Z, Komaki R, Cox J D and Mohan R 2006a Dose-volume
25 1620 modeling of the risk of postoperative pulmonary complications among esophageal cancer
26 1621 patients treated with concurrent chemoradiotherapy followed by surgery *Int J Radiat Oncol*
27 1622 *Biol Phys* **66** 754-61
- 28 1623 Tucker S L, Zhang M, Dong L, Mohan R, Kuban D and Thames H D 2006b Cluster model analysis of
29 1624 late rectal bleeding after IMRT of prostate cancer: A case-control study *Int J Radiat Oncol*
30 1625 *Biol Phys* **64** 1255-64
- 31 1626 Valdes G and Interian Y 2018 Comment on 'Deep convolutional neural network with transfer
32 1627 learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study'
33 1628 *Phys Med Biol* **63** 068001
- 34 1629 van der Schaaf A, Langendijk J A, Fiorino C and Rancati T 2015 Embracing Phenomenological
35 1630 Approaches to Normal Tissue Complication Probability Modeling: A Question of Method *Int J*
36 1631 *Radiat Oncol, Biol, Phys* **91** 468-71
- 37 1632 van Luijk P, Novakova-Jiresova A, Faber H, Schippers J M, Kampinga H H, Meertens H and Coppes R P
38 1633 2005 Radiation damage to the heart enhances early radiation-induced lung function loss
39 1634 *Cancer Res* **65** 6509-11
- 40 1635 van Luijk P, Pringle S, Deasy J O, Moiseenko V V, Faber H, Hovan A, Baanstra M, van der Laan H P,
41 1636 Kierkels R G J, van der Schaaf A, Witjes M J, Schippers J M, Brandenburg S, Langendijk J A,
42 1637 Wu J and Coppes R P 2015 Sparing the region of the salivary gland containing stem cells
43 1638 preserves saliva production after radiotherapy for head and neck cancer *Sci Transl Med* **7**
44 1639 305ra147-305ra147
- 45 1640 Vanneste B G L, Buettner F, Pinkawa M, Lambin P and Hoffmann A L 2018 Ano-rectal wall dose-
46 1641 surface maps localize the dosimetric benefit of hydrogel rectum spacers in prostate cancer
47 1642 radiotherapy *Clin Transl Radiat Oncol* **14** 17-24
- 48 1643 Vinogradskiy Y, Tucker S L, Liao Z and Martel M K 2012 A novel method to incorporate the spatial
49 1644 location of the lung dose distribution into predictive radiation pneumonitis modeling *Int J*
50 1645 *Radiat Oncol Biol Phys* **82** 1549-55
- 51 1646 Vittinghoff E and McCulloch C E 2007 Relaxing the rule of ten events per variable in logistic and Cox
52 1647 regression *Am J Epidemiol* **165** 710-8
- 53
54
55
56
57
58
59
60

- 1
2
3 1648 Voshart D C, Wiedemann J, van Luijk P and Barazzuol L 2021 Regional Responses in Radiation-
4 1649 Induced Normal Tissue Damage *Cancers (Basel)* **13** 367
5 1650 Voutilainen A 2016 Spatial Objectives in Radiation Therapy Treatment Planning. In: *School of Science:*
6 1651 *Aalto University)*
7 1652 Wang C, Zhu X, Hong J C and Zheng D 2019 Artificial Intelligence in Radiotherapy Treatment
8 1653 Planning: Present and Future *Technology in Cancer Research & Treatment* **18**
9 1654 1533033819873922
10 1655 Weber D C, Vallet V, Molineu A, Melidis C, Teglas V, Naudy S, Moeckli R, Followill D S and Hurkmans
11 1656 C W 2014 IMRT credentialing for prospective trials using institutional virtual phantoms:
12 1657 results of a joint European Organization for the Research and Treatment of Cancer and
13 1658 Radiological Physics Center project *Radiat Oncol* **9** 123
14 1659 Whitwell J L 2009 Voxel-based morphometry: an automated technique for assessing structural
15 1660 changes in the brain *J Neurosci* **29** 9661-4
16 1661 Wilkins A, Naismith O, Brand D, Fernandez K, Hall E, Dearnaley D and Gulliford S 2020 Derivation of
17 1662 Dose/Volume Constraints for the Anorectum from Clinician- and Patient-Reported Outcomes
18 1663 in the CHHiP Trial of Radiation Therapy Fractionation *Int J Radiat Oncol, Biol, Phys* **106** 928-
19 1664 38
20 1665 Wilkinson M D, Dumontier M, Aalbersberg I J, Appleton G, Axton M, Baak A, Blomberg N, Boiten J W,
21 1666 da Silva Santos L B, Bourne P E, Bouwman J, Brookes A J, Clark T, Crosas M, Dillo I, Dumon O,
22 1667 Edmunds S, Evelo C T, Finkers R, Gonzalez-Beltran A, Gray A J, Groth P, Goble C, Grethe J S,
23 1668 Heringa J, t Hoen P A, Hooft R, Kuhn T, Kok R, Kok J, Lusher S J, Martone M E, Mons A, Packer
24 1669 A L, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S A, Schultes E, Sengstag T,
25 1670 Slater T, Strawn G, Swertz M A, Thompson M, van der Lei J, van Mulligen E, Velterop J,
26 1671 Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J and Mons B 2016 The FAIR Guiding
27 1672 Principles for scientific data management and stewardship *Sci Data* **3** 160018
28 1673 Witztum A, George B, Warren S, Partridge M and Hawkins M A 2016 Unwrapping 3D complex hollow
29 1674 organs for spatial dose surface analysis *Med Phys* **43** 6009
30 1675 Wortel R C, Witte M G, van der Heide U A, Pos F J, Lebesque J V, van Herk M, Incrocci L and
31 1676 Heemsbergen W D 2015 Dose-surface maps identifying local dose-effects for acute
32 1677 gastrointestinal toxicity after radiotherapy for prostate cancer *Radiat Oncol* **117** 515-20
33 1678 Wright J L, Yom S S, Awan M J, Dawes S, Fischer-Valuck B, Kudner R, Mailhot Vega R and Rodrigues G
34 1679 2019 Standardizing Normal Tissue Contouring for Radiation Therapy Treatment Planning: An
35 1680 ASTRO Consensus Paper *Pract Radiat Oncol* **9** 65-72
36 1681 Xiao C, Polomano R and Bruner D W 2013 Comparison between patient-reported and clinician-
37 1682 observed symptoms in oncology *Cancer Nurs* **36** E1-e16
38 1683 Xu C J, van der Schaaf A, Van't Veld A A, Langendijk J A and Schilstra C 2012 Statistical validation of
39 1684 normal tissue complication probability models *Int J Radiat Oncol Biol Phys* **84** e123-9
40 1685 Yahya N, Ebert M A, Bulsara M, House M J, Kennedy A, Joseph D J and Denham J W 2016 Statistical-
41 1686 learning strategies generate only modestly performing predictive models for urinary
42 1687 symptoms following external beam radiotherapy of the prostate: A comparison of
43 1688 conventional and machine-learning methods *Med Phys* **43** 2040-52
44 1689 Yahya N, Ebert M A, House M J, Kennedy A, Matthews J, Joseph D J and Denham J W 2017 Modeling
45 1690 Urinary Dysfunction After External Beam Radiation Therapy of the Prostate Using Bladder
46 1691 Dose-Surface Maps: Evidence of Spatially Variable Response of the Bladder Surface *Int J*
47 1692 *Radiat Oncol Biol Phys* **97** 420-6
48 1693 Zhen X, Chen J, Zhong Z, Hrycushko B, Zhou L, Jiang S, Albuquerque K and Gu X 2017 Deep
49 1694 convolutional neural network with transfer learning for rectum toxicity prediction in cervical
50 1695 cancer radiotherapy: a feasibility study *Phys Med Biol* **62** 8246-63
51 1696 Zwanenburg A, Vallières M, Abdalah M A, Aerts H J W L, Andrearczyk V, Apte A, Ashrafinia S, Bakas S,
52 1697 Beukinga R J, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook G J R, Davatzikos C,
53 1698 Depeursinge A, Desseroit M-C, Dinapoli N, Dinh C V, Echegaray S, Naqa I E, Fedorov A Y,
54
55
56
57
58
59
60

1
2
3 1699 Gatta R, Gillies R J, Goh V, Götz M, Guckenberger M, Ha S M, Hatt M, Isensee F, Lambin P,
4 1700 Leger S, Leijenaar R T H, Lenkowicz J, Lippert F, Losnegård A, Maier-Hein K H, Morin O,
5 1701 Müller H, Napel S, Nioche C, Orhac F, Pati S, Pfaehler E A G, Rahmim A, Rao A U K, Scherer J,
6 1702 Siddique M M, Sijtsema N M, Fernandez J S, Spezi E, Steenbakkens R J H M, Tanadini-Lang S,
7 1703 Thorwarth D, Troost E G C, Upadhaya T, Valentini V, Dijk L V v, Griethuysen J v, Velden F H P
8 1704 v, Whybra P, Richter C and Löck S 2020 The Image Biomarker Standardization Initiative:
9 1705 Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping
10 1706 *Radiology* **295** 328-38

11
12
13 1707
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Accepted Manuscript