Eurographics Conference on Visualization (EuroVis) 2018
J. Heer, H. Leitte, and T. Ropinski
(Guest Editors)

Volume 37 (2018), Number 3

Bladder Runner: Visual Analytics for the Exploration of RT-Induced Bladder Toxicity in a Cohort Study

R.G. Raidou[1], O. Casares-Magaz[2], A. Amirkhanov[1], V. Moiseenko[3], L.P. Muren[2], J.P. Einck[3], A. Vilanova[4], M.E. Gröller[1,5]

[1]TU Wien, Austria, [2]Department of Medical Physics, Århus University Hospital, Denmark
[3]Department of Radiation Medicine and Applied Sciences, UC San Diego, United States
[4]Delft University of Technology, The Netherlands, [5]VRVis Research Center, Austria

Abstract

We present the Bladder Runner*, a novel tool to enable detailed visual exploration and analysis of the impact of bladder shape variation on the accuracy of dose delivery, during the course of prostate cancer radiotherapy (RT). Our tool enables the investigation of individual patients and cohorts through the entire treatment process, and it can give indications of RT-induced complications for the patient. In prostate cancer RT treatment, despite the design of an initial plan prior to dose administration, bladder toxicity remains very common. The main reason is that the dose is delivered in multiple fractions over a period of weeks, during which, the anatomical variation of the bladder – due to differences in urinary filling – causes deviations between planned and delivered doses. Clinical researchers want to correlate bladder shape variations to dose deviations and toxicity risk through cohort studies, to understand which specific bladder shape characteristics are more prone to side effects. This is currently done with Dose-Volume Histograms (DVHs), which provide limited, qualitative insight. The effect of bladder variation on dose delivery and the resulting toxicity cannot be currently examined with the DVHs. To address this need, we designed and implemented the* Bladder Runner*, which incorporates visualization strategies in a highly interactive environment with multiple linked views. Individual patients can be explored and analyzed through the entire treatment period, while inter-patient and temporal exploration, analysis and comparison are also supported. We demonstrate the applicability of our presented tool with a usage scenario, employing a dataset of 29 patients followed through the course of the treatment, across 13 time points. We conducted an evaluation with three clinical researchers working on the investigation of RT-induced bladder toxicity. All participants agreed that* Bladder Runner *provides better understanding and new opportunities for the exploration and analysis of the involved cohort data.*

CCS Concepts

•***Human-centered computing*** → ***Visual analytics;*** •***Applied computing*** → ***Life and medical sciences;***

1. Introduction

Radiotherapy (RT) is one of the most common therapeutical options for prostate cancer, being administered to 60% of all patients at some stage of the treatment [DJFB05]. The aim of RT is to deliver a high dose of radiation to the tumor, while sparing the surrounding normal tissues. Prior to the administration of the RT dose, an optimized treatment plan is performed, which takes into account the locations of the tumor and the healthy surrounding organs. The prescribed dose is not administered all at once, but in multiple fractions over a period of several weeks. During this time, the organs adjacent to the prostate gland – and especially the bladder – exhibit high day-to-day anatomical variations, due to differences in urinary filling and other inherent anatomical characteristics [CMMH*17]. The anatomical variation of the bladder can cause deviations between planned and delivered doses, with severe side effects for the patient [ZLH*08].

Several clinical studies were conducted in the past, to evaluate the impact of the bladder shape variation on the dose distribution during RT treatment [CMMH*17,MLK*07,ZLH*08]. These previous studies were limited to the use of *Dose-Volume Histograms* (DVHs). A DVH relates dose to tissue volume in RT planning and summarizes dose distributions in simple 2D plots, as shown in Figure 1. Here, three different patients receive the same relative RT dose (60%), but different bladder volumes (5%, 10% and 12%) are affected. Essentially, only a small amount (5%) of the bladder volume is affected by a high RT dose in Patient 1 (blue), compared to the other two patients. Although DVHs are being used as an evaluation tool for RT, they do not include any anatomical information. They do not show where a particular dose is administered, do not reflect any shape changes and consider only organ volumes.

The evaluation of RT-induced toxicity is usually conducted in

cohorts of patients for outcome analysis, i.e., to evaluate the robustness of specific RT treatment strategies [CMMH*17]. There is a hypothesis that certain subgroups of patients may be at a higher risk of bladder toxicity than others. For example, patients with highly variable bladder shapes may require more attention during treatment planning or even different RT strategies, to minimize genitourinary side effects. So far, attempts to produce DVH-based guidelines for the probability of bladder toxicity have been unsuccessful or inconclusive. One possible explanation is that daily variations of the bladder shape have an impact on the actual delivered dose, but a DVH-based analysis cannot represent it. Currently, the exploration and analysis of this data is challenging due to a lack of corresponding tools, and adequate visualization strategies need to be designed.

The *contribution* of our work is the design and implementation of a tool, the *Bladder Runner*, which allows clinical researchers to visually explore and analyze the impact of bladder shape variations on the dose distribution during the course of RT treatment. The investigation of RT-induced bladder toxicity in *Bladder Runner* can be conducted for individual patients and in a cohort. To our knowledge, there is no other tool to serve this purpose. Our approach incorporates two *components*:

(C1) Individual patient exploration and analysis through the entire course of RT treatment:
- **(C1-a)** For the quantification and exploration of bladder shape variation.
- **(C1-b)** For the exploration and analysis of the impact of bladder shape variation on the dose distribution.
- **(C1-c)** For the exploration of toxicity risk due to bladder shape variation, with respect to the initial RT plan.

(C2) Cohort exploration and analysis through the entire course of RT treatment:
- **(C2-a)** For the inter-patient quantification, exploration, and analysis of bladder shape variation.
- **(C2-b)** For the inter-patient exploration and analysis of the impact of bladder shape variation on the dose distribution.

2. Clinical Background and Cohort Data Description

RT research is aiming at designing more effective and better targeted treatments with less side effects for the patients [DJFB05]. The goal of RT is not only to treat the tumor effectively, but also to preserve the healthy tissues and organs around it, minimizing as much as possible the risk of patient complications. For this work, we focus on External Beam RT (EBRT), where radiation is administered with external sources at different angles around the patient body. However, the approaches described in this paper could be adapted and generalized to cover also Brachytherapy procedures, where radiation is administered with sources positioned inside the area requiring treatment.

In prostate cancer RT treatment, genitourinary toxicity, i.e., the manifestation of bladder-related acute and late complications

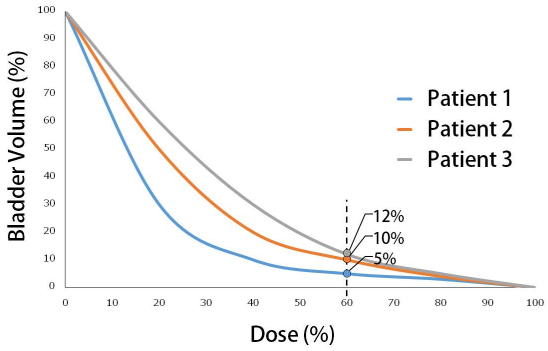

Figure 1: *Example of a Dose-Volume Histogram (DVH). High relative RT doses (60%) affect a small amount of the bladder volume (5%) of Patient 1 (blue), in comparison to the other two patients.*

due to involuntary over-radiation of the bladder, remains common [ZLH*08]. Current means of RT planning exploration and assessment are limited to Dose-Volume Histograms (DVHs), which – as described in Section 1 – do not offer insight with respect to the specific anatomic location of a dose delivery and induced toxicity.

In addition to that, RT is delivered in a fractionated manner, where protocols requiring patients to be treated in 8-9 consecutive weeks are common. Bladder tissue preservation, though, is evaluated only once, when the RT plan is designed. During the course of the treatment, the bladder tissue can be flexible and variable in shape and anatomical characteristics, mainly due to differences in urinary filling. This variation of shape and anatomical characteristics of the organ can have an impact on the accurate administration of the dose fractions with unwanted side effects for the patient [CMMH*17].

In current clinical practice, the most up-to-date RT protocols use image-guidance. The initial RT plan is performed and afterwards, just before every dose administration, new CT scans are acquired. These are used to mainly realign – if necessary – the estimated position of the prostate tumor with the initial plan. Bladder shape variation is assessed solely qualitatively. In essence, priority is always given on the tumor position, and only gross variations of the bladder, e.g., an empty bladder instead of a full, will prompt changes to the initial plan. Otherwise, the initial plan is not adapted.

In contrast to clinical practice, clinical research shows increasing interest in trying to model the bladder shape variation and in attempting to correlate it to deviations of dose distributions (between planned and administered doses) and bladder toxicity [CMMH*17]. To do so, retrospective cohort studies are conducted, where a large number of patients has been followed closely during the entire period of the treatment and the eventual toxicity has been documented. These cohort studies are meant to provide information about specific subgroups of patients, i.e., with specific bladder shape and anatomical characteristics, who are more prone to side effects. Also, they aim to provide clinical researchers with new knowledge about strategies that are more robust, or more advisable for specific subgroups of patients. Several studies have been conducted in the past [CMMH*17, MLK*07, ZLH*08]. However, they have not been conclusive, since the current exploratory means for the involved data are limited and provide restricted insight. Once the bladder

shape variation and its correlation to dose distribution deviations and bladder toxicity has been well-established in clinical research, then it may influence its incorporation in clinical practice.

For the purpose of this work, we obtained a dataset of 29 patients from a cohort study conducted by our collaborators at UC San Diego, USA. Before the beginning of the treatment, an initial 3D dose plan was designed and the same RT strategy, i.e., fractionation scheme, was selected for all patients. During the treatment period, 13 significant time points were determined. The first five correspond to the first five days of treatment, and the other eight to the following eight weeks. At these 13 time points, the patients were scanned again to acquire the actual bladder position and shape. Additional adaptations of the initial 3D dose plan were performed – if necessary – to realign the tumor. A schematic depiction of the cohort study data is presented in Figure 2. The data were already registered across patients and time points by our collaborators, using the rigid approach described by Casares et al. [CMMH*17]. Also, the bladders of all patients and at all time points were manually delineated by the same radiologist. The registration and segmentation steps might entail some limitations and introduce uncertainty into our procedure. Although uncertainty is an interesting topic to consider, addressing the issues of registration and segmentation is not within our scope.

Clinical researchers, and especially medical physicists, working on the investigation of RT-induced bladder toxicity showed interest in visualization strategies that would enable them to visually explore and analyze the impact of the bladder shape variation on the dose distribution, during the course of RT treatment in their cohort studies. After an intensive discussion with our clinical collaborators, we defined together the most relevant open tasks and research questions, schematically depicted also in Figure 2. These reflect exploratory needs in the clinical research domain. As already mentioned, in clinical practice this entire topic is not often regarded. Our tasks and research questions can be summarized as follows:

(T1) Individual patient exploration and analysis through the course of RT treatment (Figure 2, blue) :

- **(T1-a) Bladder Variation Quantification and Exploration** – *How much does the bladder of an individual patient vary through RT treatment, i.e., through time?*
- **(T1-b) Dose Distribution Exploration and Analysis** – *How much does the administered, adapted RT dose vary through RT treatment with respect to the bladder shape?*
- **(T1-c) Toxicity Risk Exploration and Analysis** – *How much does the administered, adapted RT dose deviate from the initial plan through RT treatment?*

(T2) Cohort exploration and analysis through the course of RT treatment (Figure 2, orange) :

- **(T2-a) Bladder Variation Quantification and Exploration** – *How much does the bladder vary through RT treatment, i.e., through time, within the cohort?*
- **(T2-b) Dose Exploration and Analysis** – *How much does the administered, adapted RT dose vary through RT treatment with respect to the bladder shape within the cohort?*

3. Related Work

To the best of our knowledge, there is no previous work specifically designed for exploring and analyzing the impact of bladder shape variation on the dose distribution during the course of RT treatment in a cohort of patients. However, there are several works that cover topics related to the research questions presented in Section 2, which we review in this section.

Traditional visualization frameworks employed during RT treatment planning include volume rendering of the involved organs, along with the 3D calculated dose distributions and several interaction techniques [PLW*12, SFA*17]. However, these approaches do not incorporate any information about the impact of the organ shape variation on the dose distribution and the outcome of RT treatment. Frameworks for the *assessment of the outcome* of RT treatment [RCMM*16] and for the *optimization* of tumor treatment [LDM*17] have also been proposed. These are limited to targeting the tumors, without taking into consideration the healthy surrounding tissues. Additionally, in the work of Raidou et al. [RCMM*16], the entire analysis is performed based on Tumor Control Probability (TCP) curves, and not on the patient anatomy, which comprises the same disadvantages as DVHs. Other frameworks address modeling bladder variation, but they do not offer exploratory capabilities of the impact on the dose distribution and the RT outcome [CvHvdK*11, LvHB*05, RDCO*17, CvHHB12].

With respect to the *quantification and exploration of the organ shape variation*, Hermann et al. [HSSK16] covered the topic of visualization of anatomic covariances in ensembles of data. They also published a state of the art report with future prospects on the visual analysis of shapes [HK15], which we took into consideration during the design phase of the *Bladder Runner*. Other techniques for visualizing shape space and comparing shapes based on their similarity were proposed by Busking et al. [BBP10], where a shape space scatterplot is employed as a central exploratory view of the system. This concept was later adapted by von Landesberger et al. [VLBK*13]. Also, Klemm et al. [KLR*13] designed a system for the visual analysis of spine canal variability in a cohort.

Previous work on the *visual analysis of cohorts of patients*, such as the work of Steenwijk et al. [SMB*10], Zhang et al. [ZGP14] and Klemm et al. [KOJL*14] proposed frameworks for the interactive visual analysis of cohorts that go also beyond shape analysis. However, these methods assume that the structure of interest has a spatial correspondence between patients, which is not valid in our case. More recent work on cohort analysis was conducted by Preim et al. [PKH*16], Bernard et al. [BSM*15] and Alemzadeh et al. [AHN*17], but the applications and the involved data differ from our intended scope.

In the field of *comparative visualization* [KCK17], Busking et al. [BBF*11] proposed a set of visualization designs for the comparison of two surfaces, using visual or graphical variables. Simple overlays and extensions of checkerboard visualizations on 2D imaging slices [MHG10, SGB13], and comparisons of 3D volumes have also been proposed [AWH*12]. Visual variables, deformations, glyphs [ZSL*16], and combinations of these [PF96] have also been employed. MeshLab [CCR08] and PolyMeCo [SMS09] are examples of comparative visualizations, designed particularly for mesh comparison. Most of these papers address the comparison of only

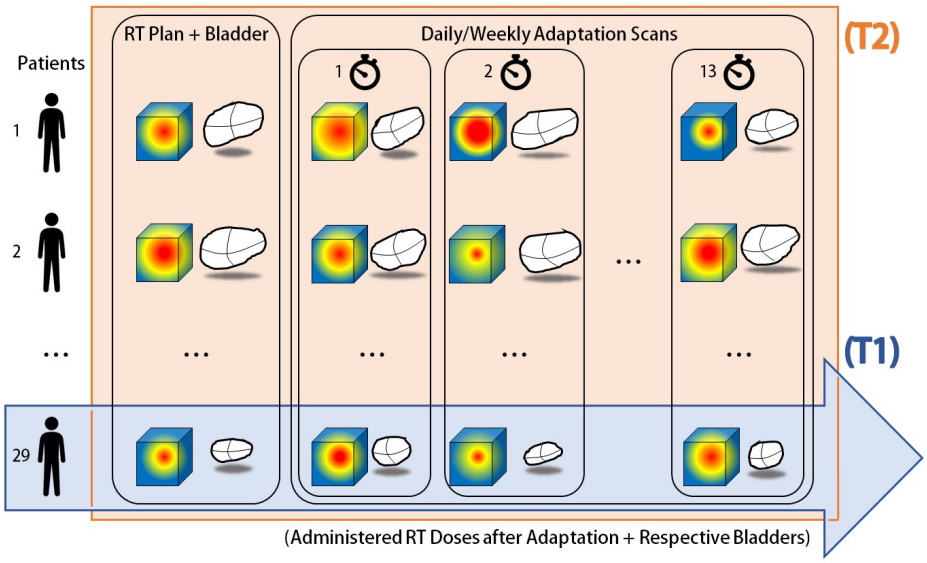

Figure 2: *Schematical representation of our available cohort data and links to exploratory tasks **(T1)** and **(T2)** of the Bladder Runner.*

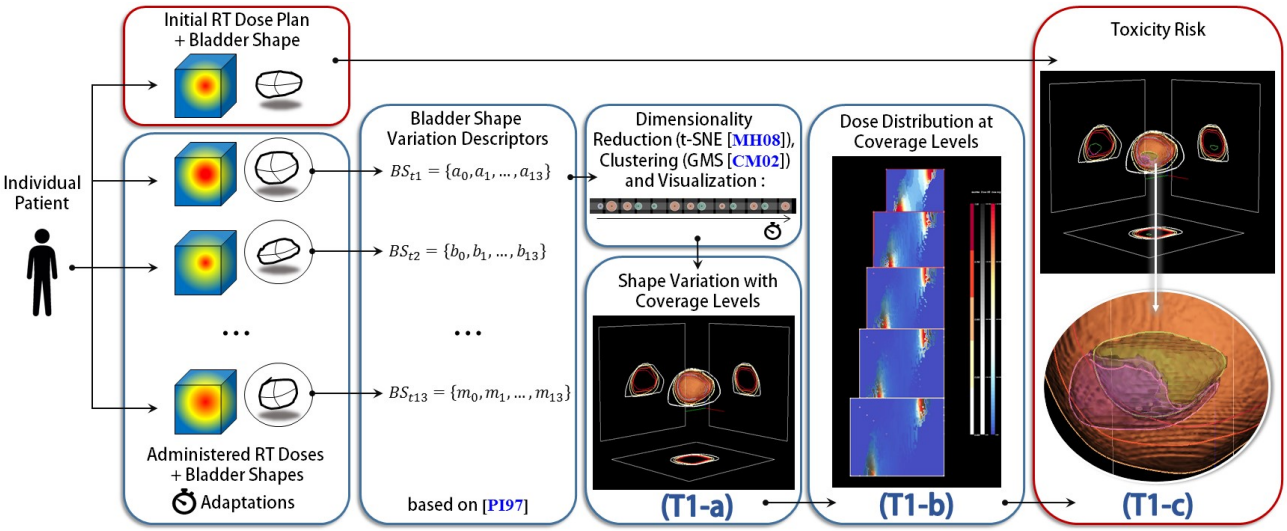

Figure 3: *Steps conducted in our approach for the exploration and analysis of an individual patient through the treatment period **(T1)**.*

two subjects, which was extended to multiple subjects by Schmidt et al. [SPA*14] and, later, by Raidou et al. [RMB*16]. However, the work of Schmidt et al. is meant for evaluating meshes generated by different algorithms with respect to a reference mesh, and the work of Raidou et al. assumes point-to-point correspondences of the involved data, which cannot be applied to our case.

4. Design of the *Bladder Runner*

We designed and implemented a visual analytics tool, the *Bladder Runner*, which enables clinical researchers to perform their analysis for individual patients and in the entire cohort, facilitating the tasks described in Section 2. In this section, we describe in detail the design choices and employed encodings within our visual tool.

(T1) Individual patient exploration and analysis – For the exploration and analysis of an individual patient through RT treatment, we propose the following steps, also depicted in Figure 3:

(T1-a) Bladder Variation Quantification and Exploration – For examining the variation in the shape characteristics of the bladder of an individual patient through the available time points of treatment, we employ a computationally simple approach, proposed by Peura et al. [PI97]. In this approach, a number of *primitive shape descriptors* is used to build a feature vector, which is adequate for grouping objects with similar shapes. In our case, we compute a 14-D feature vector for each bladder at each time point, using the following descriptors: the three eigenvalues of the inertia tensor to denote the three principal axes, the convexity, variance and elliptic variance for each principal axis, the compactness [PI97], and the bladder volume. These shape descriptors are simple and flexibly used, they

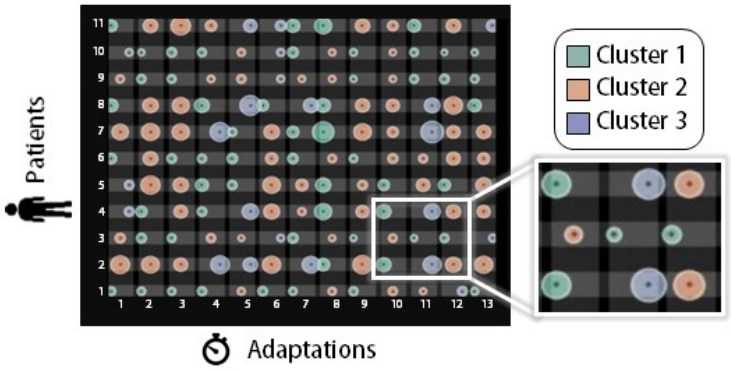

Figure 4: *Contingency matrix for the exploration of the bladder variation of patients (rows) through adaptation time (columns)* ***(T1-a)****. The area of the bubbles indicates the bladder volume. The color assignment results from the Gaussian Mean Shift (GMS) clustering.*

can be computed in real-time and perform well even in complex shape recognition tasks.

To incorporate the resulting 14-D space in one simple view that enables visual exploration of all shape features, dimensionality reduction is employed – in particular, *t-Distributed Stochastic Neighbor Embedding* (t-SNE) [MH08]. t-SNE is a non-linear dimensionality reduction technique, which maps data from a high dimensional feature space into a 2D embedding space, preserving the local structure of the initial space. Each high dimensional data point – i.e., each 14-D bladder shape feature vector – is embedded in the abstract 2D space in such a way that similar points from the 14-D space are plotted close together in the 2D space, as well. In this way, the embedding can be represented in a simple 2D scatterplot, the axes of which do not have a direct association to the features. Normalization had to be performed on the shape descriptor vector prior to the embedding, and also the parametrization of t-SNE (perplexity value choice) was done based on established guidelines [WVJ16].

For the visualization of the bladder variation of the individual patients through treatment, we initially design a *contingency matrix* representation, depicted in Figure 4. It enables the user to have an overview on the changes of the bladder volume and shape of each patient through time. It also serves as a selection mechanism to obtain more details about the bladder anatomy in a linked window. Each row represents a patient and each column corresponds to a discrete adaptation time point. The bladder of one patient at one time point has been abstracted to a bubble glyph. The area of a bubble encodes the volume of the organ [BKC*13], facilitating an overview on which patients have a highly variable bladder volume through time. To facilitate the identification and exploration of bladder shapes, we have to take a look at the projected 2D shape descriptor. To do so, it is easier to partition the cohort into subgroups of bladders that share the same shape characteristics with a *Gaussian Mean Shift (GMS) clustering* [CM02] and show these subgroups and their characteristics, instead of each bladder separately. GMS is a reliable clustering algorithm, building upon the concept of kernel density estimation (KDE), which does not assume any predefined distribution of the data clusters. The clustering is performed in real-time and it requires a kernel bandwidth parameter, which we leave user-defined based on knowledge from an initial inspection of the bladders.

After GMS clustering of the shape characteristics, each bladder is assigned to a cluster. This is encoded in the colors of the bubbles. The categorical color scheme has been selected using ColorBrewer. The displacement of the bubbles is additionally employed for better discriminating the clusters, as shown in Figure 4. In this figure, we depict only a subset of the entire cohort. Three clusters are present (orange, green and purple) and it is noticeable how the bladder of specific patients, e.g. patient 7, changes volume and shape, moving between clusters. Cohort partitioning is expected to facilitate understanding of similarities and changes in shape characteristics and to guide the selection of individual cases. It can be useful with an increasing number of subjects and provides a link to the upcoming cohort exploration and analysis task **(T2)**.

To investigate the evolution of anatomical shapes of the bladder in an individual patient through time, we design an approach similar to the previous work of Klemm et al. [KLR*13]. For this, we take advantage of the concept of *coverage probability*, which is well-known in the RT field [HSMA11]. The coverage probability denotes a probability density function, calculated in the individual patient data volumes through the time points. The value at any given voxel within the data volume of the patient indicates the relative likelihood to encounter part of a bladder volume at this spatial location. For visualization purposes we abstract it to five descriptive statistical levels that correspond to the minimum, 25%, 50%, 75% and maximum probability. Although these levels do not reflect anatomically significant regions, nor actual bladder shapes, they represent – in simple terms – the likelihood that a voxel belongs to bladder volume. They also serve the purpose of providing a benchmark for comparison between entirely different patients. Other approaches, such as shape morphing, could also have been employed, but we prefer the coverage levels due to previous familiarity of our intended users.

The *five coverage probability levels* are encoded into silhouettes with distinct colors for discrimination purposes (Figure 5). The color choice was based on guidelines from ColorBrewer. This encoding provides an easy interpretation of the limits within which the bladder shape of a patient evolves, as well as an indication of the bladder parts where a higher anatomical variation is encountered. An additional surface rendering is employed to denote the 50% probability, as this value also denotes the mean bladder shape through its temporal evolution. In Figure 5, we show a patient with a large bladder variation and a patient with a small variation.

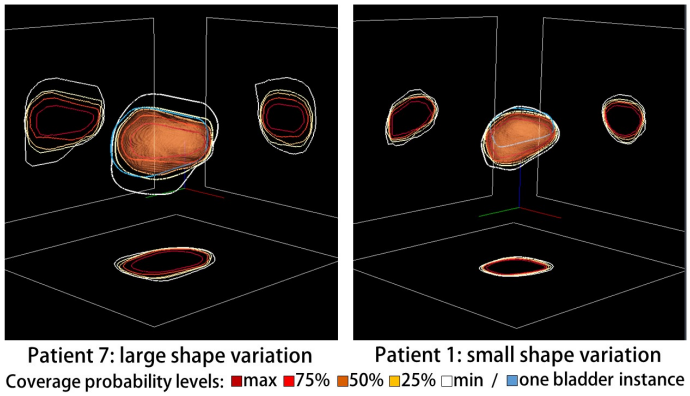

Figure 5: *Coverage probability levels showing one patient with large bladder shape variation **(T1-a)**, and one with small variation.*

The silhouettes change *interactively*, as the user changes the viewpoint. Additional *projections on the three main planes* – axial, coronal and sagittal – are provided for better interpretability of the bladder shape variation along the main anatomical axes. *Interaction* between the contingency matrix and the coverage levels view enables the exploration and analysis of each individual patient, separately. Clicking on the contingency matrix allows the expert to select one patient and closely examine the temporal evolution of the bladder shape. Additional probing enables the user to see each bladder instance through time (Figure 5 - blue silhouette).

(T1-b) Dose Distribution Exploration and Analysis – After the exploration of the bladder shape variation across the treatment period, the user is interested to inspect the *respective dose distribution through the treatment process*. For this, we compute the distribution of the administered dose across all time points for an individual patient, i.e., the average μ and standard deviation σ, assuming a normal distribution. This dose distribution needs to be visualized within the bladder variation limits. We take advantage of the previous view comprising the five levels of coverage probability and we encode the dose distribution with respect to these levels.

Visualizing the dose distribution directly on the five nested coverage levels would result in a cluttered and overwhelming view for the user. Mapping the coverage levels onto planes and visualizing on them the dose distribution facilitates the exploration of the latter, with respect to bladder shape variation. The *mapping* is performed either by unfolding the coverage probability surfaces or by flattening them, as shown in Figure 6. For the *unfolding*, we use a simple equirectangular mapping [Sny94], after we have inflated each coverage probability level to a sphere equivalent. For the *flattening*, we map half of each bladder coverage level onto a plane.

We leave the choice among these two mapping methods to the user, as they both have strengths and limitations. The unfolding uses the equirectangular projection, which comes with distortions at the north and south pole of the bladder. Yet, it is able to visualize in a natural and intuitive way the entire anatomical structure in a simple and familiar view. The flattening does not entail the limitation of distortions at the poles and it gives more information about the relative shape of each coverage level, but visualizes only – the most significant – half of each bladder coverage level.

After the mapping of each coverage probability level, the average dose through the RT treatment process is *color-encoded* on each flattened or unfolded surface, using a divergent red-to-blue colormap from ColorBrewer. The same color encoding is commonly employed in clinical practice for encoding the dose and we preserve it for familiarity reasons. The standard deviation is shown with an additional transparency encoding. Parts of the bladder with higher standard deviation, i.e., lower confidence about the true value of the administered average dose through time, will be assigned a black semi-opaque layer to denote uncertainty. Parts of the bladder with lower standard deviation, i.e., higher confidence about the true value of the administered average dose through time, will be assigned a layer with increasing transparency, as shown in Figure 6. Having done this for all coverage levels, we stack the five mappings in a pyramid-like configuration to save screen space (Figure 6). This stacked representation has been inspired by the Russian nested dolls, and we will refer to it as *matryoshka pyramid*.

(T1-c) Toxicity Risk Exploration and Analysis – For the exploration and analysis of the toxicity risk in each individual patient, we need to visualize the amount of deviation of the administered, adapted doses from the initial plan. To this end, we visualize the parts of the bladder that are *under-radiated*, and – more importantly – *over-radiated* with respect to the initial plan. Initially, we compare per-voxel the distribution of the administered, adapted doses with respect to the scalar dose values of the initial plan, as shown in Figure 7. We set as boundary the interval (-1.5σ, +1.5σ), where σ is the standard deviation around the average dose μ. This boundary was provided by our clinical collaborators as a rule of thumb, but can be fine-tuned. If the initial plan dose is above the limit, then the specific voxel has been under-radiated, which is obviously not a problem. If the initial plan dose is below this limit, then it has been over-radiated. Otherwise, it is within acceptable limits.

For the visualization of the over- and under-radiated areas, we depict – within the coverage level view – the *silhouettes* of the two areas with two distinct colors: green for under-radiated and magenta for over-radiated areas, as shown in Figure 7. The areas within acceptable radiation limits are implied through the coverage level distribution and are not shown, to reduce clutter. These silhouettes *interactively* change when the user rotates the view. Also their

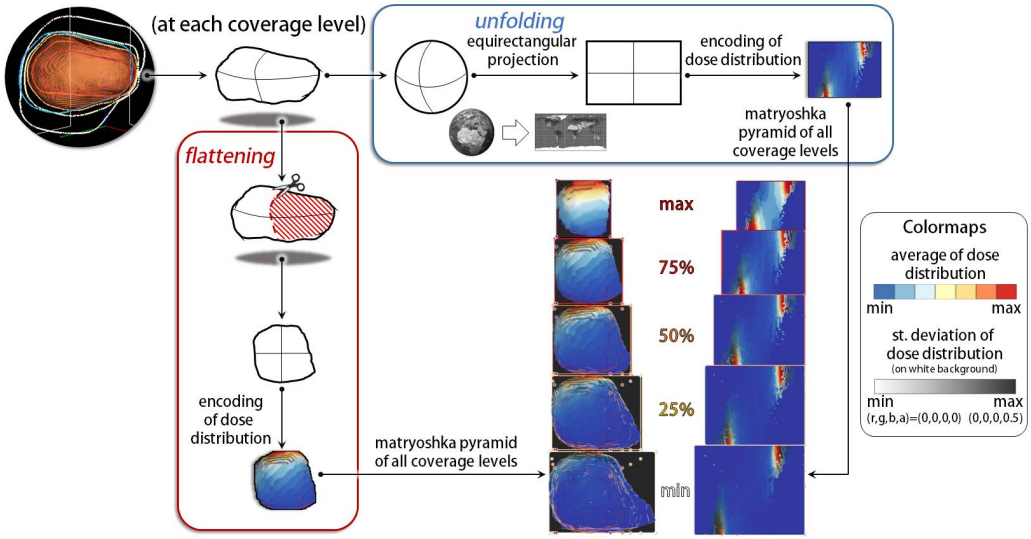

Figure 6: *Matryoshka pyramids of the five coverage probability levels, depicted in two ways (unfolded and flattened)* ***(T1-b)****.*

projections are shown on the three main planes, as previously for the coverage levels.

(T2) Cohort exploration and analysis – For the inter-patient exploration and analysis through the course of the treatment, we propose the following steps, also depicted in Figure 8:

(T2-a) Bladder Variation Quantification and Exploration – For the exploration of the shape variation in a cohort of bladders through the available time points of treatment, we employ the representation depicted in Figure 8-②. In this representation, we show the results of the t-SNE and GMS steps, performed in the same way as for task **(T1-a)**. At each time point, we denote in a *stacked bar chart* representation the number of patients that belong to each cluster. We mark with lines across time points whether some bladders are changing clusters, i.e., shape characteristics. The amount of patients changing clusters between time points is encoded in the transparency of the lines. The *contingency matrix* used in **(T1-a)** can be employed, as well (Figure 4). Probing the time points of the stacked bar chart will reveal an additional view on the underlying t-SNE (scatterplot points) and GMS (kernel density estimation contours) (Figure 8-①). This is meant for quality assurance of the t-SNE and GMS results,

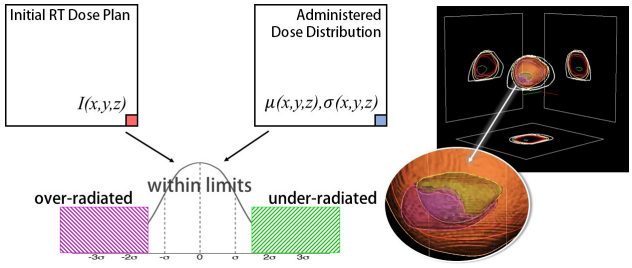

Figure 7: *Computation, and visualization for the exploration and analysis of the toxicity risk in an individual patient* ***(T1-c)****.*

and for probing each bladder instance. Probing the stacked bar chart reveals also a view on the *distribution of the bladder shapes of the entire cohort* (Figure 8-③), using the same approach of coverage levels as for task **(T1-a)**. Also, a view on the *distribution of the bladder shapes of each of the clusters* is provided (Figure 8-④). This is based on the coverage levels of task **(T1-a)**, as well. For easier discrimination, the 50% coverage level is assigned the same color as in the cluster evolution view (orange, green or purple), as shown in Figure 8-④.

(T2-b) Dose Exploration and Analysis – For the inspection and analysis of the respective dose distribution through the treatment process with respect to the bladder shape within the cohort, different encoding designs could have been used. For the comparative visualization of the dose distributions along the coverage levels across clusters, approaches such as juxtaposition, superimposition, animation, or explicit encoding of differences would be possible [KCK17]. Since our cohort is expected to be partitioned into a small number of clusters, we follow a similar approach to task **(T1-b)**. This is done separately for each of the clusters, after which we *juxtapose the resulting matryoshka pyramids*. Juxtaposition was considered to be the only possible option in this case, as each matryoshka pyramid is an already dense representation that carries a lot of information. Therefore, superimposition and animation would be overwhelming for the user. Also, it would be difficult to obtain a meaningful and understandable representation with the explicit encoding, given that we are comparing differently sized and shaped structures. The juxtaposed view provides adequate information to compare the shape and dose distribution at each coverage level of the clusters in a standardized manner, as shown in Figure 8-⑤. Also in this case, both flattening and unfolding can be employed.

Implementation Details – *Bladder Runner* was implemented as a stand-alone application in Python, using the Visualization Toolkit and PyQt, together with other libraries, such as sklearn, scipy, numpy and matplotlib.

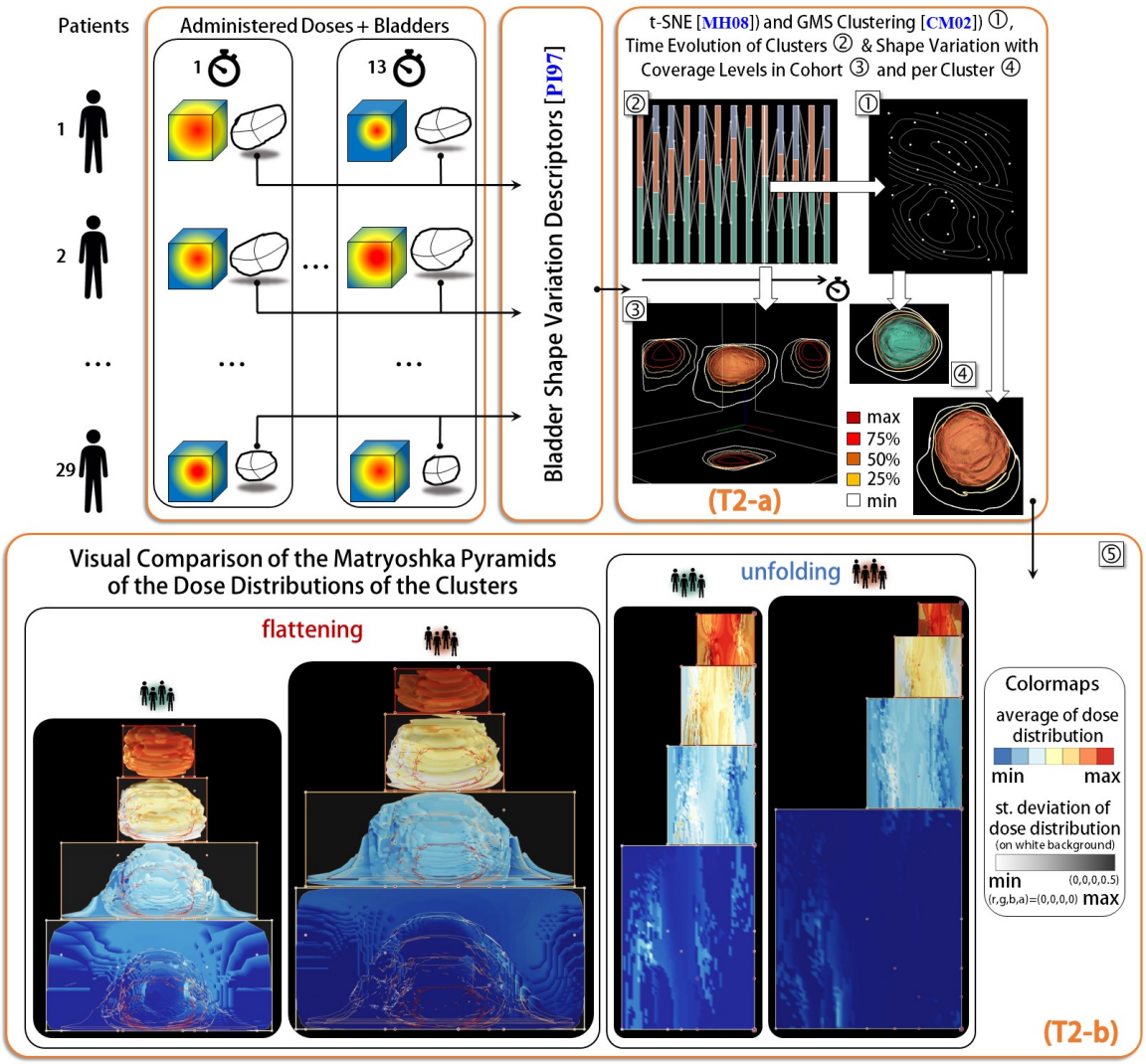

Figure 8: *Steps in our approach for the inter-patient exploration and analysis of a bladder cohort through the course of RT treatment* ***(T2)****.*

5. Evaluation

To assess the value of *Bladder Runner*, we performed an evaluation following the guidelines proposed in the paper of Lam et al. [LBI*12]. The evaluation was performed with three domain experts – medical physicists – from two clinical institutions. Their field experience varies from medium high (<10 years) to very high (>10 years). All participants have normal vision, two with and one without glasses, and nobody is colorblind. One of the participants (I) was actively involved in the design of the *Bladder Runner*. Table 1 summarizes information about the evaluation participants.

Before the evaluation, we gave an introduction to the *Bladder Runner*, where we explained basic notions and main components of the tool. We simulated the visual environment for the exploration and analysis of the tasks described in Section 2, as the clinical researchers would do in a real-life study and we conducted a usage scenario. The visual tool was operated by the first author, while the clinical researchers observed the demonstration of the individual components and the functionality for accomplishing the tasks. At the same time, an active discussion about the employed visualizations and the achieved insights took place. Each of the tasks of Section 2 was performed with the thinking-out-loud method, as the clinical researchers explained and reasoned on their opinion about the visualizations and on findings in the data. This particular setup for the evaluation was chosen for two main reasons: first, the same collaborative process and discussion would occur in a real-life study, where multiple researchers would be involved and would discuss their findings; second, the lack of other means of exploration and

Table 1: *Evaluation participants.*

Participants	Clinical Institution	Field	Field Experience	Involved in Design	Vision
I	A	Medical Physicist	<7 years	Yes	Normal, with glasses
II	A	Medical Physicist	>10 years	No	Normal
III	B	Medical Physicist	>10 years	No	Normal, with glasses

analysis of the available cohort data – apart from rudimentary DVHs – does not allow for comparison to an existing baseline. After the interviews, the participants completed a questionnaire. The usage scenario and the evaluation results are documented in this section.

5.1. Usage scenario

In this subsection, we present an initial usage scenario, which illustrates the functionality and some basic results that can be achieved with the *Bladder Runner*. This usage scenario has been guided by our clinical collaborators, based on their previous knowledge and expectations. A more thorough clinical study would demand more time from our clinical collaborators, which is out of the purpose of this work. The goal of this scenario is to demonstrate that the *Bladder Runner* can be used to explore the available cohort data, to discover new knowledge within the cohort and to confirm or generate hypotheses about the impact of bladder variations on the dose delivery and the resulting toxicity. For this scenario, the data described in Section 2 – a cohort of 29 patients across 13 adaptation time points – was used.

(T1) Individual patient exploration and analysis – Although all patients were treated under a presumably full-bladder daily image-guided protocol to minimize bladder variations, the *contingency matrix* shows substantial bladder volume changes during the treatment course. For example, in Figure 4, several patients (e.g., 2, 4, 5, 7, 8 and 11) exhibit drastic changes in their bladder volumes, and also in their shape characteristics, as reflected by the largely varying bubble glyph sizes and color encodings. Of particular interest is patient 7. During the first three adaptations, his bladder volume and shape characteristics remain constant, but afterwards both volume and shape change to a great extent. This indicates at a glance that inaccuracies between planned and administered dose might have occurred and bladder toxicity might have been induced. This case contrasts with patient 0, who exhibits slight variations in bladder volume and significantly smaller variation in bladder shape characteristics. The bladder shape variation of patients 1 and 7 can be also examined in the *coverage levels view* **(T1-a)** (Figure 5), which depicts clearly the bladder shape variation of patient 7, especially in the anterior-posterior and superior-inferior directions. Even in patient 1, where a slight variation can be noticed, these same two directions are more prone to shape and volume variation than the lateral-medial. Exploration of the other patients also confirms this. In general, we noticed that smaller bladders exhibit also smaller shape variations. With respect to the impact on the dose **(T1-b)**, patients with smaller bladder volumes or smaller bladder variations exhibit a trend of higher average dose at the interface with the prostate and low standard deviation, i.e., less uncertainty about the true administered dose, as shown in the *matryoshka pyramids* of Figure 9. Opposed to this, larger bladder volumes or patients with large bladder variations reflect high uncertainty about the true administered dose. In any case, this uncertainty seems to increase towards the outer coverage level of the bladder shape. Finally, bladders with higher shape/volume variation tend to have larger under- or over-radiated areas, which gives a first indication of *potential toxicity risks* within this subgroup of patients **(T1-c)** (Figure 7).

(T2) Cohort exploration and analysis – The exploration and analysis of the cohort throughout the entire RT treatment time using the *cluster evolution stacked bar charts* reveals three main clusters of bladder shapes **(T2-a)**. The green cluster corresponds to small bladders, or to narrow bladder shape distributions, where the five coverage levels are close together. The opposite holds for the orange cluster, while the outliers are grouped together in the purple cluster (Figure 10). Also, the green cluster, which corresponds to narrow shape distributions, reflects slightly higher average dose values, but lower standard deviations, i.e., lower uncertainty about the true administered dose **(T2-b)**. High uncertainty occurs more often in the orange cluster, or the purple cluster (outliers), as well as close to the interface with the prostate of the bladder coverage levels. These findings are depicted in the *juxtaposed matryoshka pyramids* of Figure 8-⑤.

Our initial usage scenario showed that all linked views were deemed to be useful and to provide complementary insight for the users, which could not be achieved with the previous means of exploration, i.e., DVHs. An additional session for a more thorough

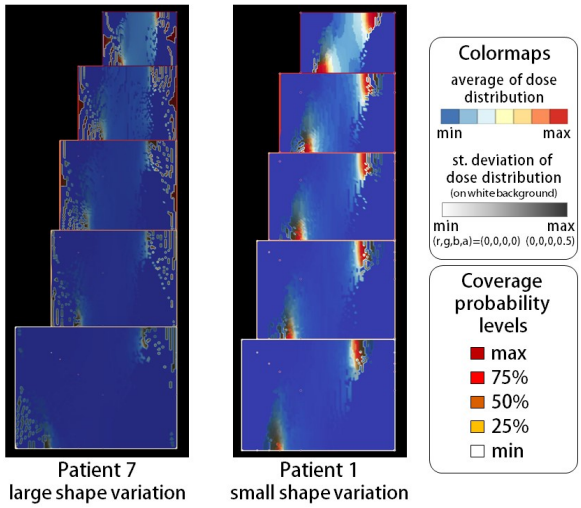

Figure 9: *Individual patient exploration and analysis of the impact of bladder shape variation on the dose distribution **(T1-b)**, as performed in the usage scenario.*

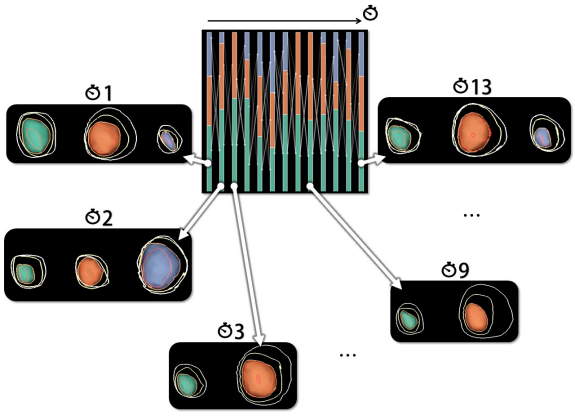

Figure 10: *Inter-patient exploration and analysis of the bladder shape variation across the treatment period **(T2-a)**, as performed in the usage scenario.*

exploration has been planned for the future, with the purpose of conducting a proper clinical study of the available cohort data, with significant clinical inferences.

5.2. Evaluation Results

For the evaluation, a questionnaire was designed. The first questions were related to the main tasks of Section 2. Each question required an open answer, but also grading using Likert scales (1-5) for the perceived effectiveness, efficiency and satisfaction. To avoid compromising the results, we separate in our analysis participant I, who was involved in the design. In Figure 11, we summarize the results of the evaluation. The designed visualizations have received high grades in perceived effectiveness, efficiency and satisfaction (≥ 4) for all tasks. Tasks **(T1-b)** and **(T2-b)**, which concern the dose distribution exploration and analysis in an individual patient and in the cohort, received lower grades (but not lower than 4), maybe due to the complexity of the approach. When completing the questionnaire, the evaluation participants commented about the functionality of task **(T1-a)** that *"individual exploration is straightforward with this tool. The user can have an accurate overview of variations in bladder volume* [..and..] *the directions more prone to expand/contract"*. Also, *"this set of tools is very useful for physicists, physicians and therapists. It provides an insight to day-to-day variations of the bladder shape/size for a patient on treatment"*. With respect to task **(T1-b)**, they commented that *"individual patient information of the spatial dose delivered will allow a more accurate analysis of the administered dose"*. For task **(T1-c)**, they remarked that *"this is a very strong component of the framework. Consequences of day-to-day variations in terms of delivered dose are presented in an easy to interpret manner, even though the underlying methodology is complex"*, and that *"this information will potentially widen our knowledge about patients more prone to develop treatment-induced toxicity"*. In the cohort exploration and analysis, for task **(T2-a)**, they stated that *"to date,* [..there is..] *no study accounting for bladder variation in a cohort. This tool has the potential to classify the patients into different groups beyond purely dosimetric parameters, providing new knowledge of inter-patient variation"*. Finally, for task **(T2-b)**, they noted that *"the proposed framework will allow us to retrospectively reanalyze outcome data and design prospective trials"* [..] *"more efficiently and less time-consuming"*.

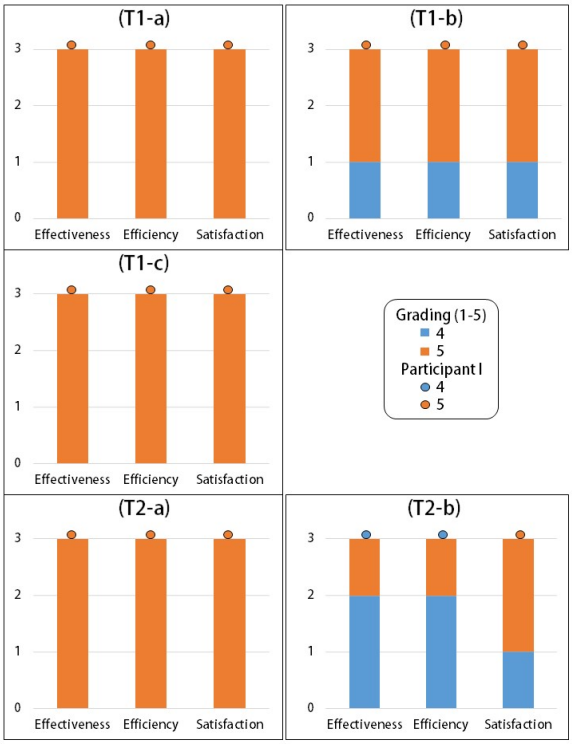

Figure 11: *Evaluation results. We denote particularly the grading of Participant I, as he was involved in the design of our tool.*

The three participants were also asked to compare the *Bladder Runner* to what they are currently using (DVHs) and to evaluate the overall usefulness of our tool. They commented that *"currently, software tools capable of handling all the tasks performed by the Bladder Runner do not exist"*. According to one participant, *"there are two main clear uses of this tool. Firstly, this tool presents an excellent way to understand how bladder changes happen during the radiotherapy course. It can be easily used for teaching purposes for new physicians and physicists. Secondly, this tool can be extremely useful for the analysis of delivered dose vs. planned dose. This will potentially provide new knowledge about mechanisms behind developing toxicity after radiotherapy"*. Another participant stated that *"it provides a completely new way to analyze patterns in variation of the shape within and across patients"*. All participants agreed that the visual tool is overall understandable and useful. The major potential of the *Bladder Runner* is, according to the evaluation participants, *"in the cohort analysis,* [..providing..] *the possibility to carry out hypothesis-driven analysis"*. Another strong point is *"the flexibility and holistic approach to evaluate changes in bladder volume and dose distributions during the treatment course"*. Improvement suggestions were related to including views *"to perform some statistical analysis, to know if differences between groups or between days are significant or not"*, and also to *"a way to summarize observations and translate them into an actionable quantity"* (link to clinical practice).

6. Conclusions and Future Work

We presented a novel Visual Analytics tool, the *Bladder Runner*. It facilitates the exploration and analysis of the impact of bladder shape variations on the accuracy of dose delivery, during the course of prostate cancer RT, in individual patients and in entire cohorts. Correlating bladder shape variations to dose deviations and toxicity risk in cohort studies was not possible with the existing means of exploration. We demonstrated the applicability of the *Bladder Runner* with a usage scenario guided by our clinical collaborators, who also positively evaluated the functionality and designed visualizations. With adequate adaptations, the visualizations designed for this specific application could be possibly extended to others.

Although the evaluation showed that *Bladder Runner* provides better understanding and new opportunities for the exploration and analysis of the involved cohort data, it also opened new directions for future work. Generalizing the use of the tool to enable the investigation of additional scenarios, such as RT-induced toxicity to other

organs with entirely different anatomy – e.g., the rectum –, would be an interesting continuation of our work. This extension would require modifications in our proposed encodings, especially for the simultaneous exploration of more than one anatomical structure. Another direction would be to link the *Bladder Runner* to actual bladder toxicity data. In this case, a retrospective analysis of side effects in the cohort data would be possible, but information about bladder complications are not yet available in our collaborating clinical settings. Moreover, the clinical evaluators pointed out that they would be interested in adding visualization functionalities to build and analyze a risk prediction model for their incoming patients, such as the patient-centered models by Hakone et al. [HHO*17].

As already stressed, our proposed approach is meant to address the exploratory needs of clinical researchers, such as medical physicists – not radiologists. To address the latter group, a simplified and/or automatized version would be required. The goal, in this case, would not be the exploration of the data and hypothesis generation, but actual decision making. Therefore, a summarization and quantification of the results achieved with the *Bladder Runner* would be necessary, as well as appropriate linking between the two approaches. Still, we are far from this, as current clinical practice prioritizes tumor position, and only very evident variations of the bladder are taken into consideration. All in all, the *Bladder Runner* offers new potential to clinical researchers working in the field of RT-induced toxicity, and new perspectives for their future studies.

7. Acknowledgements

This paper was partly written in collaboration with the VRVis Competence Center. VRVis is funded by BMVIT, BMWFW, Styria, SFG and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (854174), managed by FFG.

References

[AHN*17] ALEMZADEH S., HIELSCHER T., NIEMANN U., CIBULSKI L., ITTERMANN T., VÖLZKE H., SPILIOPOULOU M., PREIM B.: Subpopulation Discovery and Validation in Epidemiological Data. In *EuroVis Workshop on Visual Analytics (EuroVA)* (2017), The Eurographics Association. 3

[AWH*12] ALABI O. S., WU X., HARTER J. M., PHADKE M., PINTO L., PETERSEN H., BASS S., KEIFER M., ZHONG S., HEALEY C.: Comparative visualization of ensembles using ensemble surface slicing. In *IS&T/SPIE Electronic Imaging* (2012), International Society for Optics and Photonics, pp. 82940U–82940U. 3

[BBF*11] BUSKING S., BOTHA C. P., FERRARINI L., MILLES J., POST F. H.: Image based rendering of intersecting surfaces for dynamic comparative visualization. *The visual computer 27*, 5 (2011), 347–363. 3

[BBP10] BUSKING S., BOTHA C. P., POST F. H.: Dynamic Multi-View Exploration of Shape Spaces. In *Computer Graphics Forum* (2010), vol. 29, Wiley Online Library, pp. 973–982. 3

[BKC*13] BORGO R., KEHRER J., CHUNG D. H., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN M.: Glyph-based visualization: Foundations, design guidelines, techniques and applications. *Eurographics State of the Art Reports* (May 2013), 39–63. 5

[BSM*15] BERNARD J., SESSLER D., MAY T., SCHLOMM T., PEHRKE D., KOHLHAMMER J.: A visual-interactive system for prostate cancer cohort analysis. *Computer Graphics and Applications (CG&A), IEEE 35*, 3 (2015), 44–55. 3

[CCR08] CIGNONI P., CORSINI M., RANZUGLIA G.: MeshLab: An open-source 3D mesh processing system. *Ercim news 73*, 45-46 (2008), 6. 3

[CM02] COMANICIU D., MEER P.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on pattern analysis and machine intelligence 24*, 5 (2002), 603–619. 5

[CMMH*17] CASARES-MAGAZ O., MOISEENKO V., HOPPER A., PETTERSSON N., THOR M., KNOPP R., DEASY J., MUREN L., EINCK J.: Oc-0489: Variation in bladder volume and associated spatial dose metrics in prostate and pelvic radiotherapy. *Radiotherapy and Oncology 123* (05 2017), S260. 1, 2, 3

[CvHHB12] CHAI X., VAN HERK M., HULSHOF M. C., BEL A.: A voxel-based finite element model for the prediction of bladder deformation. *Medical physics 39*, 1 (2012), 55–65. 3

[CvHvdK*11] CHAI X., VAN HERK M., VAN DE KAMER J. B., HULSHOF M. C., REMEIJER P., LOTZ H. T., BEL A.: Finite element based bladder modeling for image-guided radiotherapy of bladder cancer. *Medical physics 38*, 1 (2011), 142–150. 3

[DJFB05] DELANEY G., JACOB S., FEATHERSTONE C., BARTON M.: The role of radiotherapy in cancer treatment. *Cancer 104*, 6 (2005), 1129–1137. 1, 2

[HHO*17] HAKONE A., HARRISON L., OTTLEY A., WINTERS N., GUTHEIL C., HAN P. K., CHANG R.: Proact: Iterative design of a patient-centered visualization for effective prostate cancer health risk communication. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 601–610. 11

[HK15] HERMANN M., KLEIN R.: A visual analytics perspective on shape analysis: state of the art and future prospects. *Computers & Graphics 53* (2015), 63–71. 3

[HSMA11] HYSING L. B., SÖHN M., MUREN L. P., ALBER M.: A coverage probability based method to estimate patient-specific small bowel planning volumes for use in radiotherapy. *Radiotherapy and Oncology 100*, 3 (2011), 407–411. 5

[HSSK16] HERMANN M., SCHUNKE A. C., SCHULTZ T., KLEIN R.: Accurate interactive visualization of large deformations and variability in biomedical image ensembles. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 708–717. 3

[KCK17] KIM K., CARLIS J. V., KEEFE D. F.: Comparison techniques utilized in spatial 3d and 4d data visualizations: A survey and future directions. *Computers & Graphics 67* (2017), 138–147. 3, 7

[KLR*13] KLEMM P., LAWONN K., RAK M., PREIM B., TÖNNIES K. D., HEGENSCHEID K., VOLZKE H., OELTZE S.: Visualization and analysis of lumbar spine canal variability in cohort study data. In *VMV* (2013), pp. 121–128. 3, 5

[KOJL*14] KLEMM P., OELTZE-JAFRA S., LAWONN K., HEGENSCHEID K., VOLZKE H., PREIM B.: Interactive visual analysis of image-centric cohort study data. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1673–1682. 3

[LBI*12] LAM H., BERTINI E., ISENBERG P., PLAISANT C., CARPENDALE S.: Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics 18*, 9 (2012), 1520–1536. 8

[LDM*17] LIU J., DWYER T., MARRIOTT K., MILLAR J., HAWORTH A.: Understanding the relationship between interactive optimisation and visual analytics in the context of prostate brachytherapy. *IEEE Transactions on Visualization and Computer Graphics* (2017). 3

[LvHB*05] LOTZ H. T., VAN HERK M., BETGEN A., POS F., LEBESQUE J. V., REMEIJER P.: Reproducibility of the bladder shape and bladder shape changes during filling. *Medical physics 32*, 8 (2005), 2590–2597. 3

[MH08] MAATEN L. V. D., HINTON G.: Visualizing data using t-SNE. *Journal of Machine Learning Research 9*, Nov (2008), 2579–2605. 5

[MHG10] MALIK M. M., HEINZL C., GROELLER M. E.: Comparative visualization for parameter studies of dataset series. *IEEE Transactions on Visualization and Computer Graphics 16*, 5 (2010), 829–840. 3

[MLK*07] MOISEENKO V., LIU M., KRISTENSEN S., GELOWITZ G., BERTHELET E.: Effect of bladder filling on doses to prostate and organs at risk: a treatment planning study. *Journal of Applied Clinical Medical Physics 8*, 1 (2007), 55–68. 1, 2

[PF96] PANG A., FREEMAN A.: Methods for comparing 3D surface attributes. In *Electronic Imaging: Science & Technology* (1996), International Society for Optics and Photonics, pp. 58–64. 3

[PI97] PEURA M., IIVARINEN J.: Efficiency of simple shape descriptors. In *Proceedings of the third international workshop on visual form* (1997), vol. 443, p. 451. 4

[PKH*16] PREIM B., KLEMM P., HAUSER H., HEGENSCHEID K., OELTZE S., TOENNIES K., VÖLZKE H.: Visual analytics of image-centric cohort studies in epidemiology. In *Visualization in Medicine and Life Sciences III*. Springer, 2016, pp. 221–248. 3

[PLW*12] PINTER C., LASSO A., WANG A., JAFFRAY D., FICHTINGER G.: SlicerRT: Radiation therapy research toolkit for 3D Slicer. *Medical physics 39*, 10 (2012), 6332–6338. 3

[RCMM*16] RAIDOU R. G., CASARES-MAGAZ O., MUREN L., VAN DER HEIDE U. A., RØRVIK J., BREEUWER M., VILANOVA A.: Visual analysis of tumor control models for prediction of radiotherapy response. In *Computer Graphics Forum* (2016), vol. 35, Wiley Online Library, pp. 231–240. 3

[RDCO*17] RIOS R., DE CREVOISIER R., OSPINA J. D., COMMANDEUR F., LAFOND C., SIMON A., HAIGRON P., ESPINOSA J., ACOSTA O.: Population model of bladder motion and deformation based on dominant eigenmodes and mixed-effects models in prostate cancer radiotherapy. *Medical image analysis 38* (2017), 133–149. 3

[RMB*16] RAIDOU R., MARCELIS F., BREEUWER M., GRÖLLER M. E., VILANOVA A., VAN DE WETERING H.: Visual analytics for the exploration and assessment of segmentation errors. *Eurographics Workshop on Visual Computing for Biology and Medicine* (Sept. 2016), 193–202. 4

[SFA*17] SCHLACHTER M., FECHTER T., ADEBAHR S., SCHIMEK-JASCH T., NESTLE U., BÜHLER K.: Visualization of 4D multimodal imaging data and its applications in radiotherapy planning. *Journal of Applied Clinical Medical Physics* (2017). 3

[SGB13] SCHMIDT J., GRÖLLER M. E., BRUCKNER S.: VAICo: Visual analysis for image comparison. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2090–2099. 3

[SMB*10] STEENWIJK M. D., MILLES J., BUCHEM M., REIBER J., BOTHA C. P.: Integrated visual analysis for heterogeneous datasets in cohort studies. In *IEEE VisWeek Workshop on Visual Analytics in Health Care* (2010), vol. 3. 3

[SMS09] SILVA S., MADEIRA J., SANTOS B. S.: PolyMeCo - An integrated environment for polygonal mesh analysis and comparison. *Computers & Graphics 33*, 2 (2009), 181–191. 3

[Sny94] SNYDER J. P.: Flattening the earth: Two thousand years of map projections. *Bibliovault OAI Repository, the University of Chicago Press* (10 1994). 6

[SPA*14] SCHMIDT J., PREINER R., AUZINGER T., WIMMER M., GRÖLLER M. E., BRUCKNER S.: YMCA-Your mesh comparison application. In *Visual Analytics Science and Technology (VAST), 2014 IEEE Conference on* (2014), IEEE, pp. 153–162. 4

[VLBK*13] VON LANDESBERGER T., BREMM S., KIRSCHNER M., WESARG S., KUIJPER A.: Visual analytics for model-based medical image segmentation: opportunities and challenges. *Expert Systems with Applications 40*, 12 (2013), 4934–4943. 3

[WVJ16] WATTENBERG M., VIÉGAS F., JOHNSON I.: How to Use t-SNE Effectively. *Distill 1*, 10 (2016), e2. 5

[ZGP14] ZHANG Z., GOTZ D., PERER A.: Iterative cohort analysis and exploration. *Information Visualization* (2014), 1473871614526077. 3

[ZLH*08] ZELEFSKY M. J., LEVIN E. J., HUNT M., YAMADA Y., SHIPPY A. M., JACKSON A., AMOLS H. I.: Incidence of late rectal and urinary toxicities after three-dimensional conformal radiotherapy and intensity-modulated radiotherapy for localized prostate cancer. *International Journal of Radiation Oncology* Biology* Physics 70*, 4 (2008), 1124–1129. 1, 2

[ZSL*16] ZHANG C., SCHULTZ T., LAWONN K., EISEMANN E., VILANOVA A.: Glyph-based comparative visualization for diffusion tensor fields. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 797–806. 3