Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

VAPOR: Visual Analytics for the Exploration of Pelvic Organ Variability in Radiotherapy

Katarína Furmanová[a,b], Nicolas Grossmann[a], Ludvig P. Muren[b], Oscar Casares-Magaz[b], Vitali Moiseenko[c], John P. Einck[c], M. Eduard Gröller[a,d], Renata G. Raidou[a,*]

[a] *TU Wien, Vienna, Austria*
[b] *Department of Medical Physics, Aarhus University Hospital, Denmark*
[c] *Department of Radiation Medicine and Applied Sciences, UC San Diego, United States*
[d] *VRVis Research Center, Vienna, Austria*

ARTICLE INFO

Article history:
Received July 2, 2020

Keywords: Medical Visualization, Visual Analytics, Comparative Visualization, Ensemble Visualization, Radiotherapy Planning, Cohort Study

ABSTRACT

In radiation therapy (RT) for prostate cancer, changes in patient anatomy during treatment might lead to inadequate tumor coverage and higher irradiation of healthy tissues in the nearby pelvic organs. Exploring and analyzing anatomical variability throughout the course of RT can support the design of more robust treatment strategies, while identifying patients that are prone to radiation-induced toxicity. We present *VAPOR*, a novel application for the exploration of pelvic organ variability in a cohort of patients, across the entire treatment process. Our application addresses: (i) the global exploration and analysis of anatomical variability in an abstracted tabular view, (ii) the local exploration and analysis thereof in anatomical 2D/3D views, where comparative and ensemble visualizations are integrated, and (iii) the correlation of anatomical variability with radiation doses and potential toxicity. The workflow is based on available retrospective cohort data, which include segmentations of the bladder, the prostate, and the rectum through the entire treatment period. *VAPOR* is applied to four usage scenarios, which were conducted with two medical physicists. Our application provides clinical researchers with promising support in demonstrating the significance of treatment adaptation to anatomical changes.

1. Introduction

Prostate cancer is one of the most frequent malignancies in the male population [1]. Radiation therapy (RT) is a common therapeutic approach for prostate cancer patients, requiring detailed treatment planning to identify where the tumor is located and how to treat the disease effectively [2, 3]. In RT, high radiation doses are administered to treat the tumor. Although current dose delivery techniques allow for precise treatment, the surrounding healthy tissues may still be affected by radiation [4, 5, 6]. This can potentially lead to severe side effects—commonly known as *toxicity*.

Recent clinical research suggests that the healthy bladder or rectum tissues of certain patients might be receiving increased radiation doses, due to high anatomical variability [4, 5, 6]. The RT dose is not delivered all at once, but it is split into multiple sessions over a period of weeks [3]. During this time, anatomical variations of the organs occur naturally. As it is not practically feasible to recalculate the entire treatment plan before each session, only alignment corrections are made before dose administration [2]. During these corrections, the main goal is to prioritize the irradiation of the tumor location. Thus, discrepancies between planned and administered doses occur. In *adaptive RT*, adapting the workflow to encompass changes in organ shape is anticipated to enable higher precision with less damage to healthy tissues [7], but this is not widely incorporated into clinical practice.

*Corresponding author: Tel.: +43-1-58801-18684
e-mail: rraidou@cg.tuwien.ac.at (Renata G. Raidou)

The overall robustness of specific treatment options is currently evaluated by means of retrospective cohort studies, while individual patient exploration accounts for particular cases. Clinical researchers and medical physicists working on the design of robust treatment strategies require a better understanding of the anatomical, i.e., shape and positional, variability of all pelvic organs in a cohort of patients, and an indication of the correlations between anatomical variability and toxicity manifestation [8, 9, 10, 11]. In the past, visual analytics approaches for treatment strategy evaluation have been proposed for the bladder [4, 12, 13], without considering other pelvic organs. Other previous work [14] does not support the correlation of anatomical variability to RT doses and toxicity. By incorporating the relation between anatomical variability, dose variability, and toxicity effects in the pelvic region, we aim to support clinical researchers in demonstrating the significance of treatment plan adaptation to anatomical changes.

Our contribution is the design and development of *VAPOR*. This is a novel visual analytics application for the exploration of pelvic organ variability during RT treatment. We focus on:

- the *global* exploration and analysis of the *positional and shape variability* of all pelvic organs in a cohort of patients (`T1`)
- the *local* exploration and analysis of all pelvic organs in individual patients or cohort partitions (`T2`), and
- the *correlation* of anatomical variability to RT dose variability and potential toxicity effects (`T3`).

For *VAPOR*, we retrospectively employ pelvic organ data from a cohort of 24 prostate cancer patients, for whom detailed cone-beam computed tomography (CBCT) and dose plan data are available for 13 treatment sessions. The application allows exploration of the entire pelvic anatomy of a cohort of patients in a quick and easy way, and also enables in-depth exploration of particular patients or cohort partitions, with regard to the administered dose and potentially induced toxicity.

2. Clinical Background

For patients diagnosed with prostate cancer, a common treatment method is external beam radiotherapy (EBRT) [3]. EBRT follows a complex workflow, which involves an interdisciplinary team and incorporates several processes from imaging to pre-processing, and from treatment plan simulation to evaluation [2]. Radiation doses are delivered using multiple beams, aimed at the tumor location. When superimposed, these beams sum up to a high dose applied to the targeted tumor area and a lower dose to the surrounding tissue. The planned dose is not administered at once, but it is instead distributed over several weeks, to allow the recovery of healthy tissue, while minimizing tumor growth [3]. This process is called *fractionation*, and its distinct sessions are called *fractions*. Recent techniques effectively spare healthy tissue while delivering the desired high dose to the tumor volume [15]. However, parts of healthy organs of the pelvis are still unavoidably irradiated and this can lead to side-effects affecting the patient's quality of life.

The anatomy of the male pelvis is depicted in Figure 1. In every human, unique variations occur naturally across individuals, or are cased by pathological factors, or day-to-day changes in the same person. The latter occurs because the pelvic organs are soft deformable tissues, which are flexible and their shapes are affected by filling changes [8, 9, 10, 11, 16]. Organs, such as the bladder and the rectum are especially prone to this effect and their positions and shape vary significantly on a daily basis [6]. Recent studies suggest a link between pelvic organ motion/deformation and increased toxicity risks [4]. The inherent complexity of the RT workflow makes it impossible to adapt the treatment plan before every fraction. Usually, tumor irradiation is prioritized.

The standard treatment procedure is to generate one initial treatment plan and to use it as a basis for all subsequent sessions. To facilitate this, the setting of the initial planning is reproduced during the treatment. For example, prostate treatment commonly requires a full bladder regimen [3], while positioning inaccuracies are addressed with simple translational adaptations. There are many different factors that lead to shape deformations and position variations over the course of the treatment. These cannot be entirely covered by small adaptations to the initial plan [4]. Actual adjustment of the target volume in prostate cancer therapy on a per-treatment basis needs to be considered in the future [6, 8, 9, 10, 11]. Prostate cancer research has started looking into adaptive treatment approaches—similarly to lung cancer treatment, where breathing motion is considered [17]. These adaptive approaches take into account the shape variability and movement of all pelvic organs through treatment [7].

3. User Task Analysis

3.1. Intended Users

In the course of RT treatment, several clinical experts are involved [2, 18]. The present work is targeting clinical researchers and medical physicists, i.e., scientists who evaluate the robustness of different treatment regimes. The aim is to advise on the best treatment strategy to follow, and research new, more effective ways of treatment.

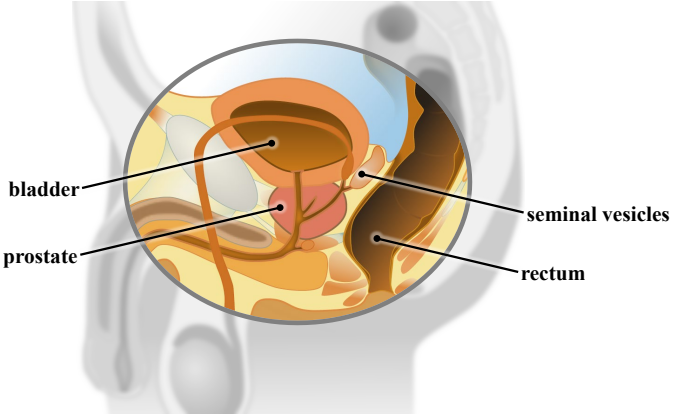

Fig. 1. Pelvis anatomy of the male body. We depict the main organs targeted in this work.

3.2. Current Workflow

In clinical practice, the evaluation of a treatment plan is currently done in two ways [2]. Both are shown in Figure 2. First, *spatial 2D/3D views* (Figure 2 (a)) allow the experts to see how the dose affects the tumor and its surrounding organs for a given point in the treatment period [19]. This approach does not support an easy exploration of multiple patients or multiple fractions at the same time—an important aspect for judging the robustness of treatment strategies. Second, *dose volume histograms (DVHs)* (Figure 2 (b)) show how much radiation is received by the volume of each organ and allow the experts to quickly identify organs at risk of toxicity [3]. Although DVHs scale well for a large number of patients, they do not allow for an easy link to individual patient anatomy.

Adequate tools for the inspection and analysis of pelvic organ variability within the context of RT do not exist—with the exception of the *Bladder Runner* [12] and the *Pelvis Runner* [14]. The former application has demonstrated its clinical usefulness in a retrospective clinical study with a single focus on bladder toxicity in cohorts of patients [13]. However, the *Bladder Runner* does not support the exploration of anatomical variability of *all pelvic organs* during the entire RT treatment period. It also does not support the exploration of *motion* of the pelvic organs. The *Pelvis Runner* supports the exploration of the anatomical variability of all pelvic organs, but it does not provide functionality for the correlation of the anatomical variability to *dose administration and potential RT-induced toxicity*. As we will demonstrate in the upcoming sections, *VAPOR* builds upon our previous work on the *Bladder Runner* [12] and the *Pelvis Runner* [14], to explore the entire pelvis anatomy of a large patient cohort in a quick and easy way, with regard to the administered dose and potentially induced toxicity.

3.3. Available Dataset

For this work, we had access to data from a cohort of 24 patients undergoing RT for prostate cancer. The provided data includes 13 treatment sessions for each patient. The first five are from the five daily sessions of the first week, while the subsequent datasets were evenly sampled from the following treatment weeks [4]. The initial treatment plan was calculated for patients with an empty rectum and full bladder. At each treatment session, the patients were instructed to have roughly the same organ fillings. Before each treatment, a CBCT acquisition was done for patient alignment using rigid translations. For each of these sessions, pelvic organ delineations in the form of contour lines are available. For all patients, the bladder and rectum delineations are included. Additionally, delineations of either the prostate, or the prostate and seminal vesicles, or the prostate, seminal vesicles, and lymph nodes might also be included. In the context of this work, we use for simplicity the term "prostate" for the first category (prostate only) and "clinical target volume" (CTV) for the other two. The dataset is depicted schematically in Figure 3.

3.4. Requirements and Tasks

Clinical researchers and medical physicists working on the design of robust treatment strategies require functionality that can provide them with a better understanding of the general shape and positional variability of all pelvic organs within the cohort, as well as the anatomical variability of subgroups of patients. Correlating anatomical variability with administered vs. planned RT doses and the resulting toxicity is also a required functionality. These functionalities, combined in one comprehensive tool, are not available in other applications, as we will discuss in Section 4. Another requirement is to aim for a general setup and interface that is easily understandable for a user from the medical community, where representations are not unnecessarily complex [2]. Although the clinical experts, for whom the application is designed, are visualization-literate, they still prefer representations that are common practice in the domain. Finally, interaction schemes, such as selection and filtering, as well as zooming, panning, rotation, and *F+C* are welcome. To ensure that all these requirements are met, one of our domain experts has been involved in the early design phases of *VAPOR*.

With regard to the tasks, the clinical co-authors of this paper have been initially interested in extracting the *amount of variability* of the available pelvic organs among all patients and across time (`T1`). Therefore, for each organ class, we need to quantify organ similarity and estimate the variability of each organ. Subsequently, we need to visualize the variability of the organ classes within the *whole cohort*. This provides a quick overview of the entire cohort, as well as capabilities to identify patients or organs with high variability, i.e., outliers. At

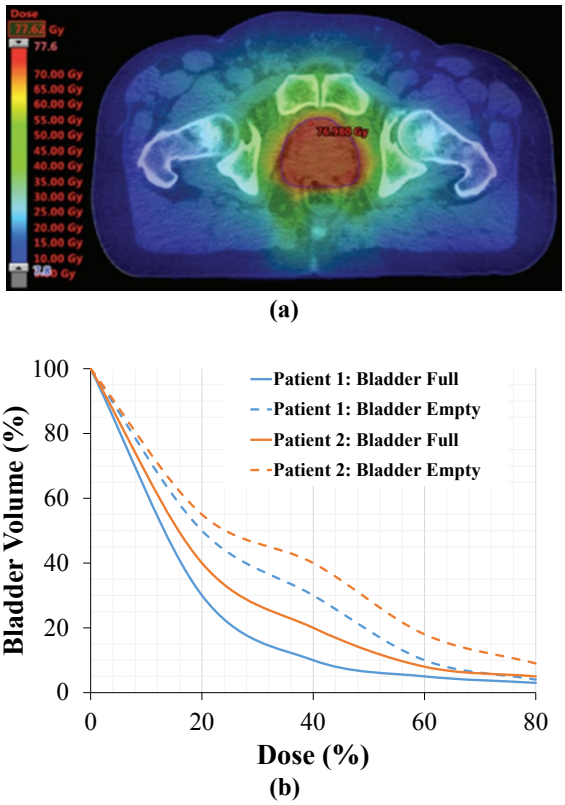

Fig. 2. (a) Spatial 2D view on the RT plan of one patient. The employed rainbow colormap represents the dose distribution, and it is used commonly in the clinical practice of RT. (b) Dose Volume Histogram (DVH) of two patients for two treatment regimes (empty and full bladder).

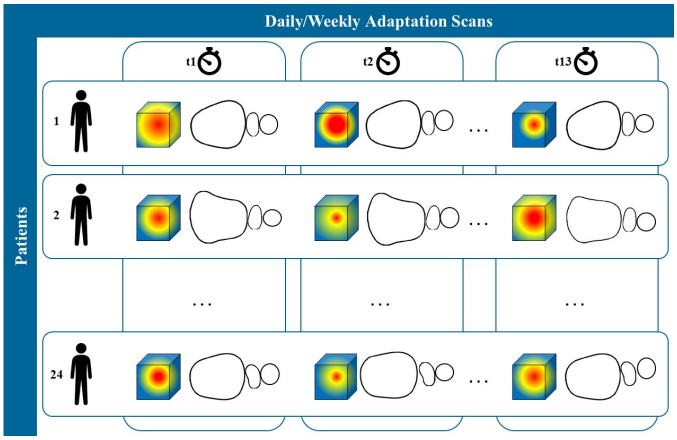

Fig. 3. Schematic depiction of the cohort data used in this work. The delineations of pelvic organs (bladder, prostate, and rectum) of 24 patients are available. Each of them had 13 sessions throughout treatment.

this point, patient and time correspondences should not be lost. When interesting parts of the cohort are identified, a more *detailed exploration* needs to be conducted (T2). Drilling down to individual objects should be possible, i.e., exploring individual patients and/or organs, to understand which regions of certain organs are prone to variations and how large these differences are. Changes in position and shape should be displayed. Finally, the anatomical variability needs to be explored in relation to the *administered RT dose*, and its *variability throughout the treatment period* (T3). This exploration, *steered by the domain experts*, is anticipated to provide useful insights about why and when potential toxicity may occur.

4. Related Work

Some studies [4, 19] facilitate the understanding of the daily occurring shape variations in pelvic organs and especially their correlation to toxicity. These are, however, limited to the exploration of spatial 2D/3D views or DVH analysis, as discussed in the previous section. These studies give insight into what kind of visualizations are commonly used in the domain of RT. They also show that looking at more than one patient or more than one time point of treatment simultaneously is a tedious process that does not scale well. Wentzel et al. [20] presented a visual computing approach for the estimation of RT plans in head and neck cancer patients, where anatomical similarity based on topology and measures of image fidelity were considered. With this approach, it is still not possible to derive any information with regard to potential RT-induced toxicity. Solutions for the visualization of many pelvic organs in a cohort of patients through the entire treatment period can be provided by *shape space* and *cohort analysis*, and with *comparative and ensemble visualization*.

VAPOR is building upon the previous work of the *Bladder Runner* [12] and the *Pelvis Runner* [14]. The *Bladder Runner* provides information about the amount of radiation delivered to the bladder across the treatment for a cohort of patients. The entire approach is based on a 14-D shape descriptor vector for the bladder cohort [21]. The 14-D shape descriptors undergo a t-Distributed Stochastic Neighbor Embedding (t-SNE) [22] followed by clustering [23] to detect cohort partitions with similar bladder shapes and evolutions through the treatment period. Using multiple coordinated views, the users analyze the bladder cohort through the RT treatment sessions, while the dose distributions and toxicity information are also incorporated in the views.

Extending the *Bladder Runner* to include multiple organs resulted into the *Pelvis Runner*. Different subsets of organs are supported in the data (e.g., for one patient we have the delineations of the bladder, rectum, and prostate and for another one we have additionally the seminal vesicles). Changes in the shape descriptor were made, as the 14-D vector of the *Bladder Runner* is not adequate for describing other than spherical shapes, e.g., it is not suitable for the rectum. However, the *Pelvis Runner* still does not support the correlation to dose administration, the analysis of its variability and the investigation of potential RT-induced toxicity. This functionality is the main addition, which resulted into *VAPOR*.

Other previously proposed frameworks include the work of Reiter et al. [24] to explore and analyze the variability in multiple pelvic organs. Their approach is based on spherical harmonics [25]. To distinguish clusters across organ classes, they employ t-SNE [22]. To distinguish clusters within organ classes (and more importantly, outliers) they use Principal Component Analysis (PCA) [26]. Their data is derived from automatic segmentation algorithms where a triangle-to-triangle correspondence can be ensured across the individual structures. Yet, the approach does not support multi-timestep analysis. Also, the 8-D descriptor from the spherical harmonics frequencies that was employed in this work is not sufficient to describe non-spherical organs, such as the rectum. Generally, the use of descriptors, as presented in the former works, supports the efficient differentiation between diverse shapes, but it lacks the ability to synthesize arbitrary elements in their shapes.

In *shape space analysis*, Hermann et al. [27, 28, 29] investigate anatomic covariances in ensembles of data, providing also a state of the art report with prospects on the visual analysis of shapes [30]. Busking et al. [31] propose a 2D scatter plot to represent the distribution of elements inside a cohort and to synthesize additional arbitrary objects in the shape space. For comparing objects, they later deal with visualizing intersecting 3D surface meshes [32]. Landesberger et al. [33] extend the scatter-plot concept to parameter sensitivity analysis in segmentation and the link to the segmentation outcomes. Considering the high learning curve for many complex visualizations of high dimensional data, such as cohort data, Blumenschein et al. [34] propose concepts aimed at people who are not from the visualization domain.

More specifically for *cohort analysis*, Klemm et al. [35] focus on the extraction of spine-canal variability and the exploration of clusters of similarly shaped spines. This work has been extended to incorporate additional patient information [36], demonstrating how to effectively reduce and visualize image cohort data and to facilitate their understanding on a broader basis. Steenwijk et al. [37] also go beyond shape analysis by proposing a framework for the interactive and structured

Table 1. Schematic comparison of *VAPOR* and the most relevant previous work, with regard to the task analysis of Section 3.

	Multiple Organs	Possibly Different Organs	Multiple Patients	Multiple Time Points	Relation to Dose & Toxicity
VAPOR	✓	✓	✓	✓	✓
[12]	✗	✗	✓	✓	✓
[14]	✓	✓	✓	✓	✗
[20]	✓	✓	✓	✗	✗
[24]	✓	✗	✓	✗	✗
[27, 28, 30]	✗	✗	✓	✗	✗
[31]	✗	✗	✓	✗	✗
[33]	✗	✗	✓	✗	✗
[34]	✗	✓	✓	✗	✗
[35, 36]	✗	✗	✓	✗	✗
[37]	✗	✗	✓	✗	✗
[43]	✗	✗	✓	✗	✗
[44, 45, 46]	✗	✗	✓	✓(in [46])	✗
[48]	✗	✗	✓	✓	✗
[49]	✗	✗	✗	✓	✗
[50]	✗	✗	✓	✗	✗

visual analysis of cohort data. Cohort analysis has also been tackled by Preim et al. [38], Bernard et al. [39], and Alemzadeh et al. [40], for various purposes.

Given the available data, which are contour delineations of the pelvic organs, we consider the previous work in *ensemble visualization* [41]. Our work relates to contour boxplots by Whitaker et al. [42], their extension for streamline ensemble data by Mirzargar et al. [43], and the recent techniques of Ferstl et al. [44, 45, 46]. The latter are applied on weather simulation ensemble data, covering 2D lines, 3D volumes, and also the time evolution thereof. In *comparative visualization* [47], for the investigation of jaw movement, Keefe et al. [48] introduce small juxtaposed representations, where the movement is explicitly encoded giving a good overview of all the data, while parallel coordinates allow for an in-depth search. Tory et al. [49] investigate a superposition approach for tracking brain lesions extracted at different time points from MRI images. Explicit encoding to highlight structural differences is used by Schmidt et al. [50], where they compare a large number of similar meshes and can quickly identify regions of differences in multiple linked views.

Previous literature includes approaches that process a multitude of individual objects (in our case, either multiple patients or multiple organs). In some cases, different object sets, i.e., sets missing some instances (in our case, organs), are also handled. Also, previous work visualizes the development of structures through time (in our case, multiple timesteps). The most relevant works and their characteristics are summarized in Table 1. However, there is no approach with comprehensive functionality that covers all aspects of our problem. As described in Section 3, these span from the quantification and visualization of multiple organs in a patient cohort throughout the treatment time, to the correlation of anatomical variability and toxicity manifestation. We cover this literature gap with *VAPOR*.

5. Methods in *VAPOR*

VAPOR focuses on three main objectives: the *global* exploration and analysis of pelvic anatomy variability across the treatment period and across a cohort of patients (`T1`), the *local* exploration and analysis of pelvic anatomy variability across the treatment period for individual patients or cohort partitions (`T2`), and the *correlation* of anatomical variability to delivered radiation and toxicity (`T3`).

The general workflow of *VAPOR* is presented in Figure 4. Our approach starts with data processing, and with quantifying the similarity of the organ shapes in order to estimate their anatomical variability. For visualizing the variability in the organ shapes, an aggregation approach based on Ferstl et al. [44] is employed. For (`T1`), a low dimensional embedding of each organ is used to calculate the variability on a per-patient basis and to visualize the whole cohort. After grouping, a tabular plot is employed to explore the cohort partitioning in a flexible and intuitive manner. For (`T2`), information on the anatomical space is shown on demand. We enable the user to drill down to selected patient groups from the cohort and to perform a detailed inspection of the organ variations. This is achieved by reconstructing the initial 3D objects from their low dimensional embeddings. By sampling the embedding space for the median and the standard deviation of the organs, we reconstruct the shape variations and we show them in a representation similar to contour boxplots [42]. For (`T3`), we compute and visualize the distribution of the administered RT dose, i.e., the average and standard deviation, for selected groups of patients. The clinical co-authors of this work are interested mainly in pelvic organ regions with high anatomical variability and high radiation dose. *VAPOR* provides the option to guide and restrict the anatomical variability computation to regions with doses that exceed a user-selected threshold. More details on each step of our workflow are provided in the upcoming subsections.

5.1. Data Processing, Linearization, and Reduction

The first step in the organ shape analysis *transforms* the organ data into a format that is easier to handle and to visualize. The organs in the cohort are manually delineated by medical experts, through contours at individual slices of CBCT scans of each patient. We initially convert the contours to volumetric coverage masks, i.e., volumes. The resolution of our volumes is given by the resolution of the CBCT scans. In our data, this is $2.5 \times 2.5 \times 2.5$ mm per voxel. Each organ for each patient and timestep is stored in a separate volume, which initially covers the entire pelvic region, i.e., the entire volume captured in the CT scans. This is done to preserve the original position with respect to other organs. We store each organ in a separate volume for convenience, as the shape analysis is later performed separately for each organ class. Additionally, by storing all organs in separate volumes, we avoid the risk of overlaps at neighboring voxels of different organs.

In the second step, we *register* the volumes. For each patient, the individual timesteps are already pre-aligned manually by medical experts, using the prostate as the reference organ—still, some per-patient positional variations of the prostate can be observed. This is a common approach in prostate cancer treatment, as the radiation dose is also centered around the prostate, but it also has limitations. It only allows us to analyze the average between-timestep (inter-fraction) organ motion of the groups of patients with respect to the prostate, which is a mobile organ itself. While for some treatment methods, such as

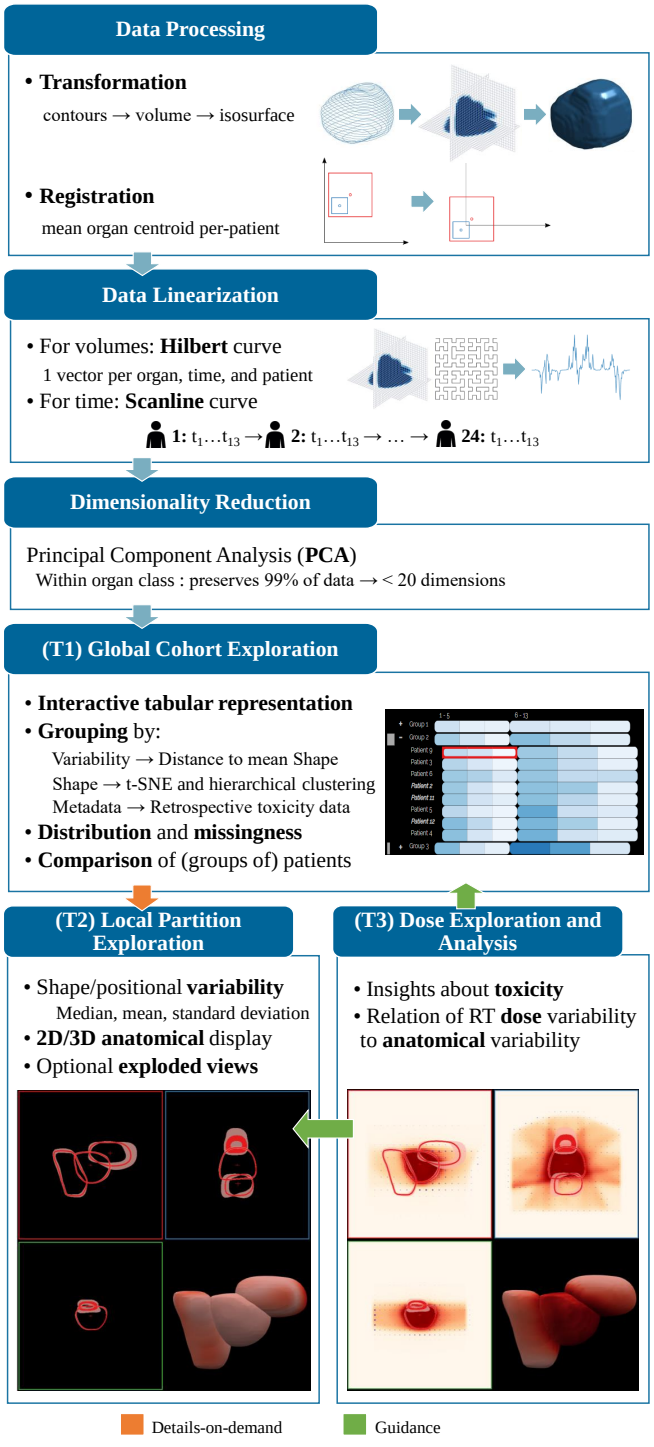

Fig. 4. Schematic depiction of the workflow, the main components of *VA-POR* and their in-between links.

photon-based RT, this is not an issue, for other, such as proton-based RT, the motion of prostate can become also an important factor in treatment planning. For a more robust analysis of positional changes, registration based on the position of pelvic bones or femoral heads would be necessary, as bones are the most rigid structures in the human body. This approach would preserve the positional variations of all pelvic organs. Unfortunately, this approach was not feasible for us. Segmentation of the bones would require additional contouring from medical experts (or, at the very least, corrections if automatized segmentation was used) which is a very time-consuming process.

Also, we want to preserve the persisting positional variations between individual timesteps of a single patient, as they indicate how the organs move during the treatment. However, we still need to align different patients to each other. To do this, we compute the mean centroids across all timesteps separately for each organ and patient, i.e., for 24 patients and three organs, we compute 72 mean centroids. We then align the organs so that the mean centroid for a given organ and patient is translated to the center of the coordinate system. Although this approach adds small translational variations, it preserves the volume changes and the main growth directions. After registration, the volumes are cropped to a uniform size based on the bounding box containing all of the volumes. We store the translation vectors for all organs, in order to be able to retrieve their original positions and to compute new mean positions for subgroups of the cohort. For the computation of shape and positional changes, the organs are aligned individually. For rendering, we align the groups based on the mean centroid of all organs.

In the third step, the 3D volumetric patient data are *linearized*, before we can employ the dimensionality reduction step. At the same time, we map the two dimensions of our cohort, i.e., patients and timesteps, into a single one without losing correspondences within the data. For this, we employ linearization strategies along two types of curves: Scanline Curve and Hilbert Curve [51]. The volumes, which initially correspond to binary coverage masks, are converted to signed distance maps representing the distance to the organ's surface. The distance volumes are then linearized into 1D vectors using the 3D space-filling Hilbert Curve that allows us to analyze how the shape differentiation capabilities of our method changes if the sampling density is reduced. This has also been employed by Weissenböck et al. [52] and by Demir et al. [53] for volume data comparison. After volume linearization, there is a unique vector for each organ, patient, and timestep. The vectors representing organs from the same class are then organized following the Scanline principle, as we are interested in preserving the temporal order within the data. We create a data structure where all timesteps of the first patient are followed by the timesteps of the second patient, and so forth. This allows us to easily select patients and their timesteps, while we can also efficiently add new patients in the analysis. Each organ class is stored and processed separately.

In the fourth step, the vectors containing the volumetric data (without losing patient and timestep correspondence within the cohort) are *reduced* into a low dimensional vector representation that allows us to create a computationally efficient way to store and process large cohorts of patient data. The dimensionality reduction step creates a low dimensional embedding of the structure of the high dimensional space. Each cohort data point, i.e., an individual patient's organ at a specific timestep, is represented by one position in space, where similar shapes are placed nearby. As discussed in Section 4, the approaches used in our previous works (e.g., 14-D space based on shape descriptors from Bladder Runner [12]) are not easily generalizable to

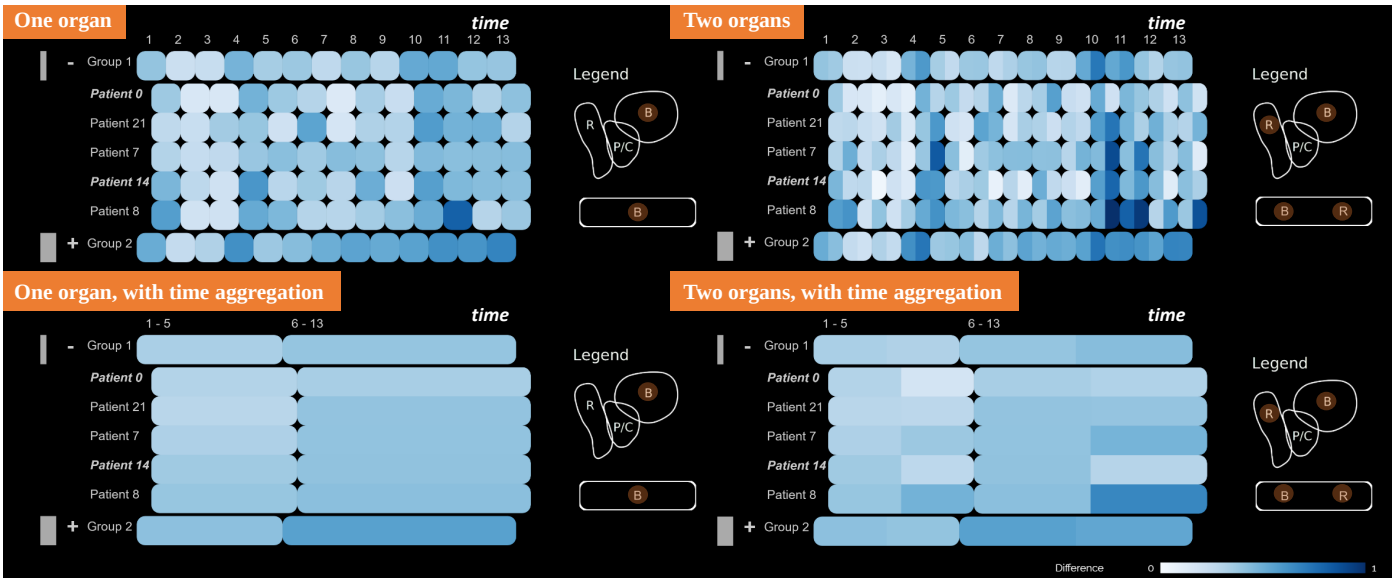

Fig. 5. Some of the possible configurations of the tabular view—with one or multiple organs, and with or without time aggregation.

other pelvic organs, e.g., rectum, seminal vesicles, or bowel loops, which can have vastly varying shapes. This led us to a different approach. We employ Principal Component Analysis (PCA) [26] to create a low dimensional embedding of the data and use only as many components as are needed to ensure the preservation of 99% of the original data. In our case, we need up to 20 dimensions, depending on the organ class. The low dimensional embedding allows us to efficiently store the data and to perform further calculations and analyses. The accurate representation of the patient anatomy is also a vital part of any medical visualization software. We can always reconstruct the volumetric data from the low dimensional space, but the visualization thereof is computationally very expensive. Thus, for the visualization components, we employ the triangular meshes that are generated on-demand from the reconstructed volumes, as iso-surfaces.

5.2. (T1) Global Exploration of Anatomy within a Cohort

For task (T1), we enable clinical researchers to compare the *different pelvic organs* from *multiple patients* throughout *several timesteps*. In some cases, the patient data also incorporate *different sets of organs*, as the delineations include either the prostate, or the prostate and seminal vesicles, or the prostate, vesicles, and lymph nodes.

We first provide users with an *overview of the whole cohort data*. The main idea behind this is to generate a high-level representation that conveys the general patterns present in the data. Afterwards, the user starts a detailed investigation of individual interesting cases. This is based on the low dimensional outcome of the dimensionality reduction step and we offer two possibilities here. The first option is based on the distance of each organ to the mean per-patient organ shape in low dimensional space. The distance calculation between data points enables the explicit estimation of outliers on a per-patient basis. It also indicates the shape variation across the treatment time points for each patient. For this, we calculate the *Euclidean distance*, similar to Klemm et al. [35]. In the second option, clustering can be used for the extraction of the main shape groups within patients. The drawback of clustering is that subtle differences between shapes are obscured. Clustering only offers a binary variability option—either the shape belongs to a cluster, or not. The analysis and comparison of the clusters can offer an understanding of what shape types are to be expected in patients and how prominent they are. To get a better separation between the shapes, we first perform a t-Distributed Stochastic Neighborhood Embedding (t-SNE) [22] on the low dimensional data from the PCA (Section 5.1). We, then, employ a *hierachical clustering with complete linkage* [54]. This is done similarly to the work of Klemm et al. [35], with which the clustering tasks are very similar. We chose this method, as hierarchical clustering is more flexible, gives more intuitive results, and has fewer assumptions about the distribution of the underlying data than other clustering techniques, e.g., *k*-means, which are essential requirements for a generally applicable system. Regarding the cluster proximity measure, we selected complete linkage. Klemm et al. [35] showed that complete linkage performs best for this type of task. In their work, single and average linkage approaches led to big clusters containing dissimilar shapes, due to the chaining effect. Another advantage of hierarchical clustering is that the generated number of clusters is easily adjustable. Therefore, we give the users the option to set and adjust the number of clusters, interactively. Alternatively, we offer the option of automatic selection for the optimal number of clusters, which can be different for each organ. For this, we employ the cluster analysis method by Calińsky and Harabasz [55].

From the previous calculations, we receive a single distance metric and/or cluster value per combination of patient, timestep and organ. To visualize this, we employ a *tabular representation* similar to the contingency matrix of the *Bladder Runner* [12] or the representation in the work of Blumenschein et al. [34]. This representation (Figure 5) shows the shape

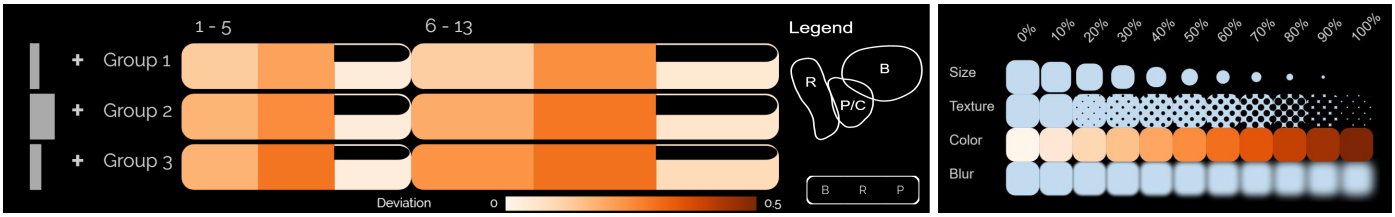

Fig. 6. Left: Encodings for the standard deviation from the mean shape (orange colormap) and for missing data (partially filled cells). Bladder (B), Rectum (R), Prostate (P). Right: Alternative encodings considered for the standard deviation of each organ from the mean value (size, texture, color, and blur).

change information, while at the same time preserving information about time and patient correspondences. We also aim at visualization readily understandable by users who do not employ visual analytics tools on a regular basis. In the tabular view, patients are depicted on the vertical axis and timesteps on the horizontal one, to enable comparisons across both timesteps and patients. The encoded values represent the similarity distance shown with a sequential white(low)-to-blue(high) colormap (Figure 5), or the cluster membership denoted with a qualitative colormap (Figure 10 (a)). Both of these maps have been taken from Colorbrewer [56]. To extend the approach to multiple organs, we split each cell of the tabular view into equally sized parts—one for each organ (Figure 5, right). With this encoding, the users can directly compare the values of multiple organs and detect patterns and correlations. This is similar to a glyph-based representation, as also demonstrated by Blumenschein et al. [34]. The users manually decide which organs are shown every time, as well as whether they want to depict the Euclidean distance or the clustering. Labels and legends complete the representation.

The tabular representation can accommodate additional information with regard to the *underlying data distribution* and to the amount of *missing data*, i.e., missing organ delineations, as both of these indicate trustworthiness. The former is visualized with additional distribution histograms accompanying the groups and positioned to the left-hand side of the tabular plots, as shown with the gray bars in Figure 5. The latter is represented with a "partially filled glass" metaphor at each cell in the tabular plot. As shown in Figure 6 (left), the less filled a cell, the less data it contains and the partition is less trustworthy. For example, in Figure 6 (left), Groups 1 and 2 have less available data for the prostate than Group 3. The prostate data is visualized in the third part of the glyph, which is also indicated in the legend. Going one step further, the user might also be interested in finding out how different *shape group types compare* to each other. For this, several encodings, i.e., size, texture, color, and blur, have been investigated. An example is given in Figure 6 (right) for encoding the standard deviation of each observation from the mean value.

The initial layout of the overview visualization provides the option to see the whole cohort, at once. The analysis process in this case requires the user to scan row-by-row the representation to detect similarities or outliers. This can be time-consuming even for a small cohort of patients. For improvement, we enable Focus+Context (F+C) [57], sorting and grouping [58], and visual aggregations of patients and timesteps as shown in the bottom row of Figure 5. Patients can be split into groups based on organ shape clustering, organ variability, or categorical patient metadata (e.g., available retrospective toxicity data). With the clustering option, the patients are aggregated into groups based on their prevalent organ shape type identified by the clustering algorithm. For organ variability-based grouping, we estimate the variability as the average Euclidean distance of organ shapes over time to the patient's mean organ shape (in the low dimensional PCA embedding). The patients are then grouped based on their average shape distance. Four different groups are automatically generated, based on low $< 25\%$, medium $25\% - 75\%$, and high $> 75\%$ average distance in interquartile range, as well as one group for patients with missing values in case no data for the given organ are present.

5.3. (T2) Local Exploration of Anatomy in Cohort Partitions

During the exploration and analysis of the entire cohort, the users identify specific interesting cases, i.e., individual patients or partitions of the cohort, which require further investigation. We enable the users to drill down to individual patients or partitions, for local exploration. Up to this point, only abstract information with regard to the cohort and its shape properties have been displayed in the tabular view. We provide an additional view of anatomical shapes for selected patients or partitions. Multiple patients or subgroups within the cohort are selected respectively by clicking on a cell or a row label in the cohort visualization. Each selection is assigned a unique color from a qualitative scheme from Colorbrewer [56].

For the *summarization of shape variations*, we first extract the geometric median in the low dimensional embedding of the shape space as a general representative of the group. In this way, we retrieve a representative shape that exists in our cohort—as opposed to the mean shape. We then employ the approach proposed by Ferstl et al. [45] for the analytical transformation of confidence intervals from the low dimensional PCA embedding to the spatial domain. This way we retrieve representatives of the shape distribution. We are using this method with the interval $(\mu-\sigma, \mu+\sigma)$, where μ is the mean shape and σ is the standard deviation. However, these confidence intervals can be adjusted by the user, as we show in Figure 7, to show instead the 90% confidence interval.

The analysis of the center-point variations is indicative of the organ movement. For this, we also use the mean and standard deviation of the center point of each organ to calculate the main variation directions for groups of organs. This is also in accordance with our registration method, where we also took the

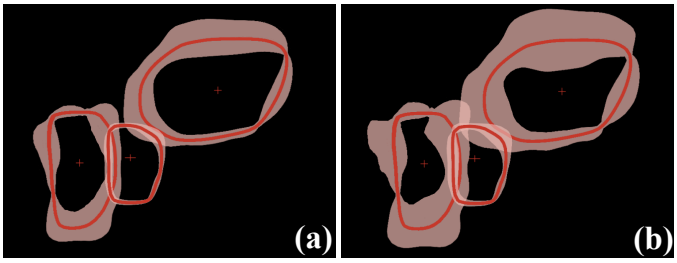

Fig. 7. Example of different settings for the confidence intervals (denoted with the bands) around the organ medians (denoted with the red lines) in the anatomical view. (a) Standard deviation. (b) 90% confidence interval.

average center point for each patient to align his organs before the analysis. In advance, we have already performed a Kolmogorov-Smirnoff test to confirm that the distribution of the shapes within the cohort is indeed close to a normal distribution. This combined approach has also been employed by Ferstl et al. [44, 45].

To *display the shape and positional variability*, we employ the common combination of three anatomical 2D planes (sagittal, coronal and axial) with a 3D view, as given at the bottom of Figure 4. Standard interaction, e.g., zooming, panning, and slicing through the volume, is possible. For the comparative visualization of the pelvic organs of multiple patients within a 2D view, two alternatives are possible [47]: (i) superposition of stacked contours, where each patient instance is denoted with a distinct color, (ii) superposition of contour boxplots [42], where each patient or cohort partition is denoted with a distinct color. The latter is shown in Figure 8 (a). A combination of the two is also possible, e.g., when comparing one patient instance to a specific partition. We additionally display the center-point variation for each organ. This is explicitly encoded by drawing a cross, the bars of which extend to indicate the main directions of organ motion, as shown in Figure 8 (a).

In the 3D views, we superimpose the median shapes of all selected groups (Figure 8 (b)). The lighting in the scene and the surface material aim at highlight the organ structure, while transparency is not employed. Instead, if a specific group is selected, it is brought forward with a F+C strategy in the 2D (Figure 8 (c)) and the 3D views. On demand, the 3D view can show the explicit encoding of the surface variations (Figure 8 (f)). In this case, the surface color encodes the amount of surface variation, using a sequential colormap based on the organs' group color. With this view, we support users in finding regions with interesting shape changes. As the adjacency of the organs may cause overplotting and difficulties in judging the shape variations, we provide also an optional *exploded view* [59], where the user can extrude the organs in the display (Figure 8 (d,e)). In the exploded view, the same organs of all groups are taken and placed in such a way that they do not overlap with any other shape, while at the same time being centered at a common point. To preserve parts of the initial context, a line connects the center of the extruded organ to its original position (Figure 8 (e)).

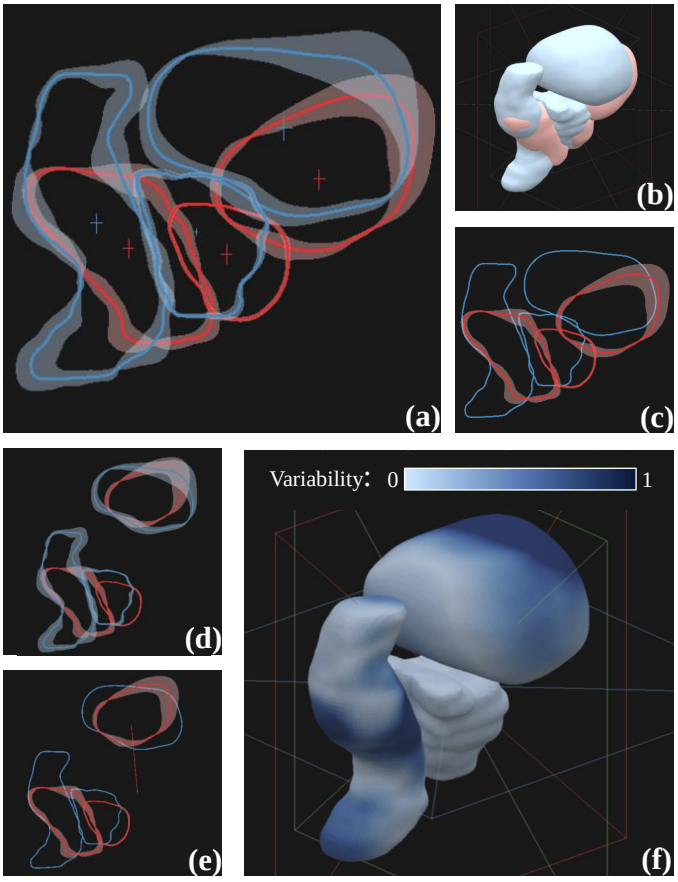

Fig. 8. Comparison of two cohort partitions (red and blue) in the anatomical view. (a) Shape (contour boxplots) and positional (cross glyphs) variability are shown in 2D. (b) Superposed 3D view. (c) F+C for shape variability with focus on the red partition. Positional variability has been hidden. (d) Exploded view for the extrusion of bladders in 2D. (e) F+C for the exploded bladder view with an indication of the extent of the extrusion to see the red partition in focus. (f) Explicit encoding of variability in the 3D view for the blue group.

5.4. (T3) Dose Exploration and Analysis

In RT, it is important to administer a high enough dose to the target volume, i.e., the volume that covers the tumor area. At the same time the dose to the healthy tissues should be minimized. Healthy tissue close to (or within) the target volume are particularly affected by anatomical variations, which may lead to higher dose delivered compared to the planned. Clinical researchers need a functionality that supports dose exploration and analysis. They need functionality for relating dose administration, anatomical variability, and toxicity effects, in a global and a local way—complementing tasks (T1) and (T2).

Not all regions of the pelvic organs are equally important. The most critical regions are those, where anatomical variability and radiation dose are both high. For a constrained navigation, the domain experts can guide the global anatomical variability exploration and analysis of (T1) by restricting the RT dose. A *user-selected threshold* can be employed, e.g., by determining that the "maximum acceptable dose is 67 Gy". The constrains are linked to the methods used for (T1). The data, as they result from the low dimensional embedding described in Section 5.1, are reconstructed in the 3D space. A mask containing the thresholded RT dose, e.g., all voxels receiving a dose above 67 Gy, removes the organ regions where the dose is below the threshold. This is performed for each patient and each

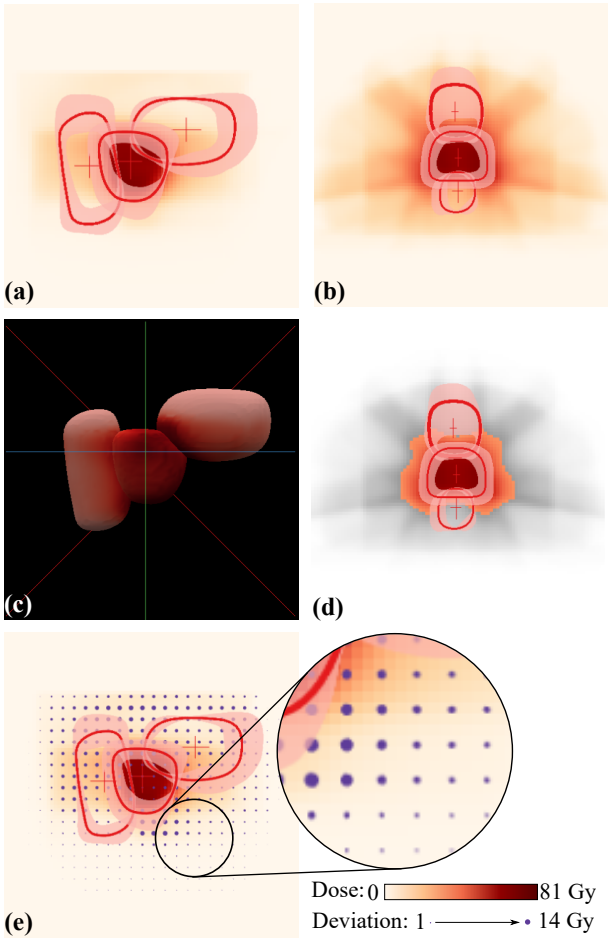

Fig. 9. Anatomical views incorporating the RT dose mapping (a) in the sagittal plane, (b) in the coronal plane, and (c) in 3D. (d) F+C employed to gray out the RT dose below a user-defined threshold. (e) Dose deviation mapped on the area of the superimposed circular glyphs.

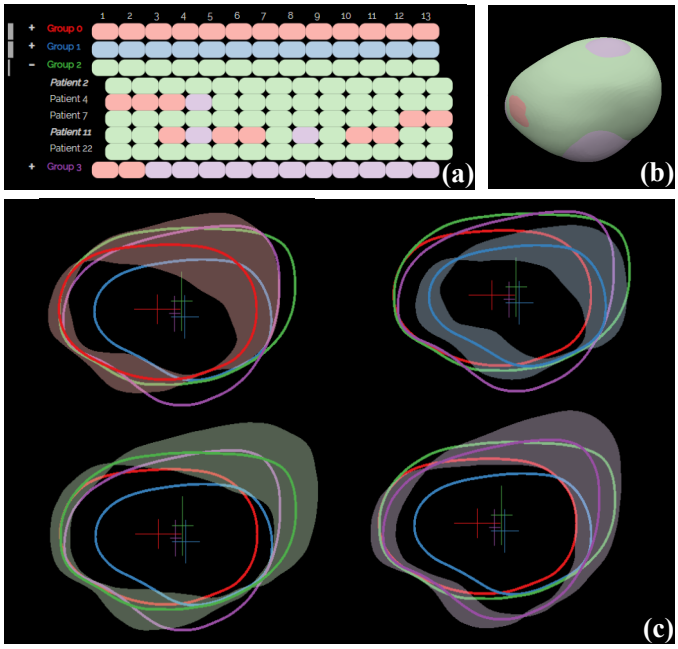

Fig. 10. Scenario for shape type identification, applied to a bladder analysis for the completion of `(T1)`. Four clusters are identified and denoted with four distinct colors, representing bladder groups with different shape characteristics and different kinds of anatomical variability. (a) Tabular view showing the patients grouped by their prevalent bladder shape type. (b) Superimposed median shapes from each cluster in 3D view. (c) Shape variations of each cluster shown in sagittal plane.

treatment session. The data are subsequently linearized using the Hilbert Curve and then processed in the same way as the low dimensional embedding described in Sections 5.2 and 5.3. The *updated tabular representation* depicts now the anatomical variability information, but only in regions where the RT dose exceeds the user-determined threshold. As the tabular representation also incorporates retrospective toxicity information, it is possible to *relate toxicity with the anatomical variability and the locations of high dose administration.*

In addition to knowing the locations of high radiation dose and high anatomical variability, it is necessary to have a more localized view on these regions of interest. In `(T2)`, when a group of patients is selected, the anatomical views show the local organ variability thereof. To link this to the RT dose and its variability, we compute the distribution of the administered RT dose, i.e., the average and the standard deviation. We subsequently show the average dose as a *background colormap* in the 2D anatomical planes, as given in Figure 9 (a-b). This follows a sequential white (low dose)-to-red (high dose) color scale [56], but can be changed by the user to match domain conventions [2]. In the 3D view, we encode the average dose on the mean organ shape using the same color scheme (Figure 9 (c)). The standard deviation is mapped to the area of *superimposed circular glyphs* [60], similarly to Raidou et al. [61] (Figure 9 (e)). As an alternative encoding, we considered the approach of Ristovski et al. [62], but for two reasons we decided not to adopt it. First, our clinical experts were already familiar with the superimposed circular glyphs, and they are already working with this encoding [63]. Second, the approach of Ristovski et al. would require from the user to zoom into the treatment plan to obtain details on the variability, which involves more interaction than our approach. In the future, it would be interesting to investigate alternative encodings for the dose deviation. To preserve anatomical context, *F+C* is employed [57]. Regions that have been discarded by the dose thresholding are kept in the view, but are grayed out, as shown in Figure 9 (d).

5.5. Implementation

VAPOR is designed as a server-client application. A web server in conjunction with MATLAB performs the computationally expensive operations, including data processing, linearization, and dimensionality reduction. A client-side browser application written in JavaScript receives the shape information and creates the visualizations using three.js [64] and D3.js [65].

6. Results

In this section, we present four scenarios of increasing complexity, as conducted together with two medical physicists to assess how well tasks `(T1)`, `(T2)`, and `(T3)` are supported with *VAPOR*. We further document the feedback from the domain

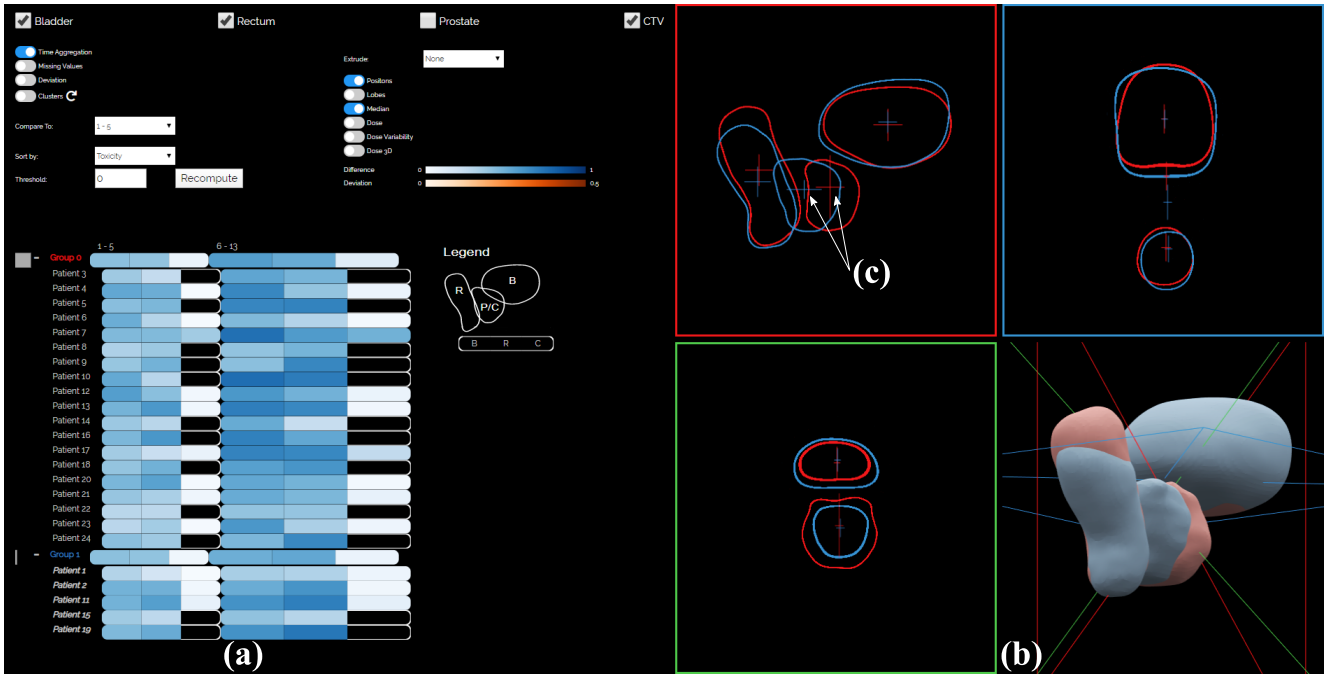

Fig. 11. Scenario for retrospective toxicity analysis, to compare patients with toxicity (blue) against patients without (red). This scenario addresses all three tasks. (a) A preliminary analysis indicates that the shape variability does not differ significantly between the two groups. (b) There are also no significant anatomical differences. (c) However, the positional variability of the CTV looks vastly different between the two groups.

experts giving an initial indication of the strengths and weaknesses of *VAPOR*, and directions for future improvements.

6.1. Shape Type Identification in a Cohort

Shape type identification in a cohort is depicted in Figure 10. It investigates possible organ shape types resulting from the clustering. Therefore, it focuses only on the first task (`T1`) for exploring the anatomical variability of organs within a cohort. In the case of the bladder, four groups (Figure 10 (a): red, green, blue, and purple) are obtained. All groups are selected to inspect their median shapes, confidence bands, and positions, as shown in Figure 10 (c). The green and purple groups contain bladders with bigger sizes. Bladders from the green group are rather convex, while purple bladders protrude further in the direction of the prostate (bottom left side of the shapes in Figure 10 (c)). This is visible in the 2D views and also in the superimposed 3D view (Figure 10 (b)). The red and blue groups contain smaller bladders, which are again split into convex bladders (red) with a rather flat interface towards the prostate (bottom left side of the shapes in Figure 10 (c)) and bladders with concave shapes (blue). In general, all bladders indicate the largest variation at their upper side. There the bladder is the least constrained by other internal organs and can freely extend. Most of the bladders move predominantly along the vertical axis. The red group also exhibits large positional variability along the sagittal axis, i.e., left-to-right in Figure 10 (c). This verifies findings of previous clinical work [4, 66].

6.2. Retrospective Toxicity Analysis

Retrospective toxicity analysis is depicted in Figure 11. It investigates possible correlations of organ shapes to toxicity manifestation, i.e., addresses tasks (`T1`) and (`T2`) of Section 3. Figure 11 also showcases the comprehensive interface of *VAPOR*. For the toxicity, retrospective data of all patients are available. The patients are sorted based on toxicity, as seen in Figure 11 (a). The red group presents no toxicity and the blue group presents toxicity (`T1`). In the toxicity group, there are patients with high (2, 11, and 19) and low (1, 15) shape changes (`T2`). Also, there are patients whose average shape in the first five days is similar to the rest of the treatment (1, 2, and 15), and those whose average shape is not (11 and 19), leading to higher variations. Both of these findings do not indicate a connection between shape variability and induced toxicity, but the number of patients is too small for a conclusive statement. When looking at the anatomical views, there are no large differences in the shapes themselves, although the group with toxicity (blue) seems to have slightly bigger bladder shapes (Figure 11 (b)) (`T1`). However, the positional changes of the CTV look vastly different for the two groups of patients. The sagittal view (Figure 11 (c)) indicates that the group with toxicity (blue) seems to move more in the sagittal direction than the one without (red), as shown by the cross glyphs. Increasing the number of patients might provide in the future more information about these preliminary findings.

6.3. Single Organ Exploration in a Cohort

Single organ exploration in a cohort is depicted in Figures 12 and 13. It addresses all three tasks of Section 3. The exploration starts with grouping patients based on their average bladder shape changes (`T1`). When comparing each shape to the first treatment day (Figure 12 (a)), all bladders change significantly through the treatment period. This is indicated by the different shades of blue for all groups in the tabular representation.

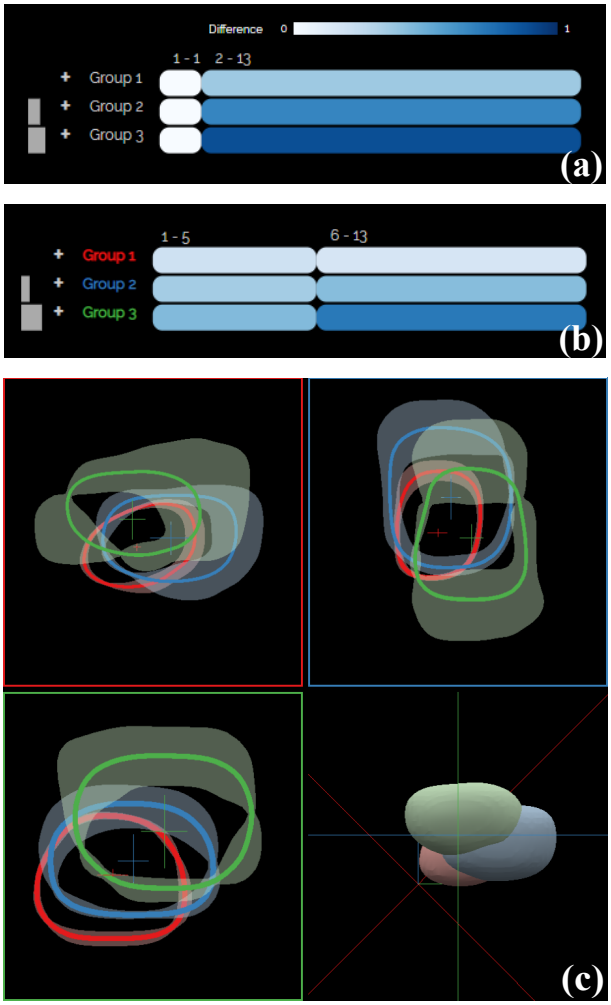

Fig. 12. Scenario for single organ cohort exploration, showing the shape and positional variability of bladders. This scenario addresses the first two tasks. It indicates that performing the planning based on (b) the first five timesteps instead of (a) only the first one may more precisely model the bladder shape over time. (c) *VAPOR* may allow to early identify patients with high organ shape variability in critical regions (Group 3, green), and account for this information in treatment planning.

It is an important argument in favor of adaptive RT. The current clinical practice uses only the first timestep for treatment planning, and our finding confirms that simple translational adaptations of the initial treatment plan will not suffice. When comparing each shape to the mean of the first five treatment days (Figure 12 (b)), the variability is lower. This is an indication that performing the planning based on the first five timesteps instead of only the first one may more precisely model the bladder shape over time. The anatomy concerning the respective shape variations can also be seen in the contour boxplots of Figure 12 (c). All groups have similar shapes, which can be due to the fact that patients with high average variability are found all over the shape space and have no individually distinctive shape. The group with low shape variability (Group 1, red) has also small local shape variations, i.e., smaller bands. The group with high shape variability (Group 3, green) has also large local shape variations, i.e., larger bands. With regard to positional variations, higher shape variability correlates with larger positional variations, as denoted by the cross glyphs in Figure 12 (c). The positions largely vary along the sagittal and vertical axes (red square in the figure, horizontal and vertical direction respectively), which corresponds to previous findings [4].

The contour boxplots in the sagittal view of Figure 12 (in (c), red square) indicate that Group 1 and 2 present the lowest shape variability in the area of the prostate (lower left corner of the sagittal view). In Group 3, this is not the case. Expanding the tabular representation helps inspecting individual patients (Figure 13 (a)) (`T2`). Patients from Group 3 are particularly interesting, as high shape variability can potentially lead to complications. When looking at the individual patients from this group, some patients, e.g., Patient 7 (Figure 13 (b)), exhibit a similar local shape variability pattern to patients from Group 1 and 2, i.e., the shape changes mostly outside of the high dose region. However, some patients, e.g., Patient 13 (Figure 13 (c)), exhibit high shape variability also in the area of high dose. For such cases, the dose-masking feature of our tool can be used to recompute the shape variability only based on the regions, where the RT dose exceeds the user-determined threshold (`T3`). Figures 13 (d) and (e) show Patient 7 and 13, respectively, after dose masking. After the recalculation, the tabular representation shows that the order and grouping of patients has changed (Figure 13 (f)). Patient 7 has moved from Group 3 to Group 1, as he exhibits low organ shape variability in the masked area. Patient 13 stayed in Group 3. This indicates that our tool can be used to separate patients with high organ variability in high dose regions from patients with low overall shape variability or low variability in high dose regions. Also, there is a clearer separation between Group 1 and 2. This is visible already in the first five timesteps of the treatment and is even more apparent in the remaining timesteps. This supports the hypothesis that a few initial plans obtained over the first few days of treatment (e.g., five) may allow to early identify patients with high organ shape variability in critical regions. This information can be taken into the account in treatment planning.

6.4. Multi-Organ Exploration in a Cohort

Multi-organ exploration in a cohort is depicted in Figure 14, and targets all three tasks of Section 3. The explorative tasks of the scenario presented in Section 6.1 can be repeated for all the available organs (`T1`). In Figure 14 (a), the tabular representation encodes the average variability values of the three organs side-by-side. In Figure 14 (b), it presents their deviations. The prostate volumes (in the rightmost cells) do not undergo large shape variations. These low values are encoded with almost white color for the respective cells of all groups. The anatomical view of Group 3 (Figure 14 (c)), which is the one with the highest shape variability, shows all shape and positional changes of the organs (`T2`). While the prostate and the bladder undergo positional changes mostly along the vertical axis, as indicated by the cross glyphs, the motion of the rectum is predominantly along the sagittal axis, i.e., the back-to-front axis of a patient. Overlaps between the prostate shape and other organs may happen as the CTV includes an additional safety margin [2]. Regarding the shape changes, the bladder extends mostly away from the prostate, similar to the results of

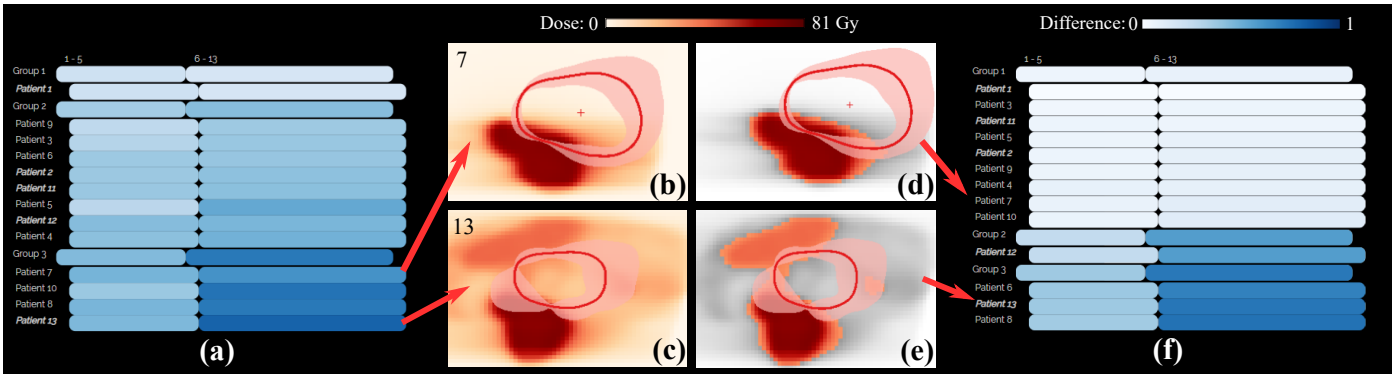

Fig. 13. Scenario for single organ cohort exploration along with the radiation, showing the variability before (left) and after dose masking (right). (a) Tabular view for the cohort partitioning before dose masking. (b),(c) Anatomical view with dose overlay for Patients 7 and 13 before dose masking, and (d),(e) after dose masking. (f) Tabular view for the cohort partitioning after dose masking. Patients from Group 3 are particularly interesting, as high shape variability in combination with high RT dose administration can potentially lead to complications.

Section 6.1. For the rectum, there is no predominant direction of change, which might be due to its inherently high overall anatomical variability. The dose distribution in the same group (Figure 14 (d)) indicates that both bladder and rectum are exposed to high RT dose, as seen in the 3D view (`T3`). The circular glyphs superimposed on the anatomical planes denote a high RT dose variability and higher doses outside of the prostate. A possible explanation is that some patients in this group received also lymph node irradiation to reduce recurrence, therefore the irradiation field was much larger.

6.5. Initial Feedback

We address here the strengths, weaknesses, limitations, and future improvements of our work. The domain experts commented that the application provides a flexible and systematic way to explore the data. It allows them to aggregate information in different ways and inspect the most interesting aspects of the data. The approach is "*a promising and useful decision-making tool for radiation oncologists*". As they stated, "*there are many possibilities, and many features*" and this allows them to approach their data in many different ways—depending on their specific hypotheses or exploratory tasks. It allows them to see individual organs, multiple organs, multiple patients, and also subgroups of the cohort, at the same time. Although this was not an intended functionality, they commented that "*the tool offers a way of identifying the setup uncertainty of the entire treatment*". This follows from providing an overview of the motion, i.e., uncertainty, of the prostate. The exploded views have been created to allow the users to "drag apart" the different organs so that the overlaps would not interfere with their understanding of the variability at organ interfaces The reaction of experts to this functionality was rather neutral. It was seen as an additional (neutral) feature—neither absolutely necessary nor useless. The 2D views seemed to be more useful than the 3D views, which is a common observation in radiation therapy treatment [2]. 3D views are, in general, not very common in clinical practice, and all representations are mainly 2D-based. We included the 3D view for completeness and context. The domain experts would like to explore further the data in the frame of their future clinical research. They expect that working more with the application will bring forward improvement suggestions, particularly for treatment planning. For example, the application could give "*indications of patients that will fail or that may develop toxicity at the beginning of the treatment*", allowing the experts to adapt the employed strategy. Potentially, it could help "*creating thresholds* [i.e., guidelines] *for patient treatment*". For future work, the domain experts proposed the addition of functionality to easily add annotations and perform measurements concerning, e.g., the confidence bands of the contour boxplots. This would quantify the up-to-now qualitative inspection of the variability and could be done, for example, by probing along the median contour. The initial feedback is informal in nature. In the future, we will conduct an extensive evaluation, also in the scope of a retrospective clinical study with a larger cohort.

7. Conclusions and Future Work

We present *VAPOR*, a visual analysis application for the exploration of pelvic organs in multiple patients, across the whole RT treatment procedure. *VAPOR* focuses on the global exploration and analysis of pelvic organ variability in an abstracted tabular view and on the local exploration and analysis of shape and positional variability in a combined 2D/3D anatomical view. The application integrates functionality for the analysis of the irradiated dose with regard to the anatomical variability. It includes the possibility to relate the analysis to retrospective toxicity information within cohort studies. We showcased the functionality of *VAPOR* with four usage scenarios conducted with two domain experts.

Future work includes a thorough evaluation with the intended users, as well as a quantitative evaluation to assess the robustness of the current partitioning approach. For this, a larger cohort would also be needed. The registration part of the workflow could also be evaluated and improved to yield more robust results. Also, for the exploration of dose deviations other encodings, such as those proposed by Ristovski et al. [62], could be investigated. In its current state, *VAPOR* has been designed for domain experts—namely, medical physicists. They are familiar with the implemented analysis and are also (up to a certain extent) visualization and machine learning literate. For

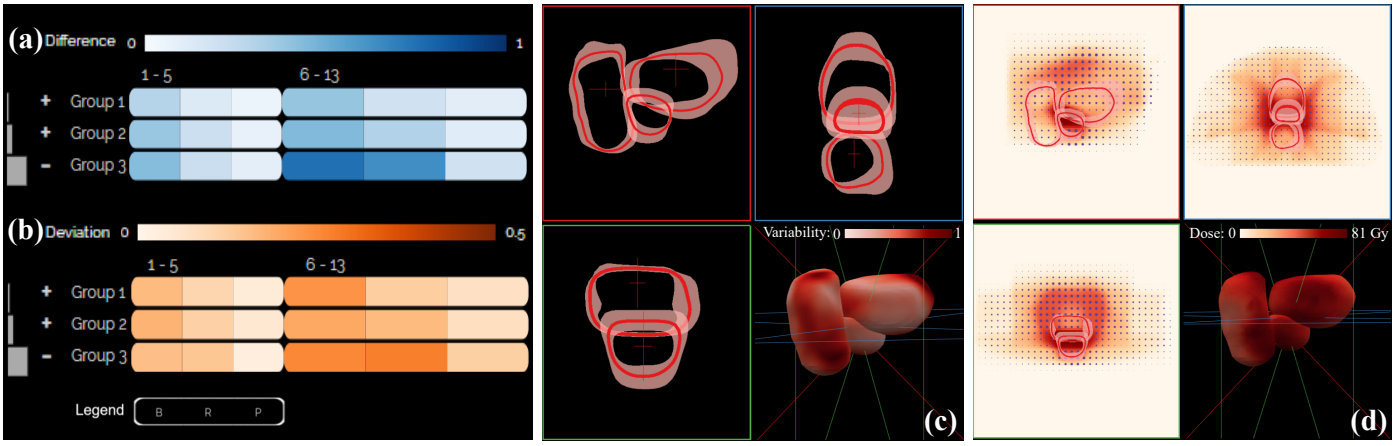

Fig. 14. Scenario for multi-organ cohort exploration along with the radiation. (a) The average anatomical variability of the three involved organs and (b) their deviation. Bladder (B), Rectum (R), Prostate (P). (c) The shape and positional variability of all pelvic organs. (d) The dose variability in the most varying group. Group 3 manifests the highest shape and positional variability. Within this group, both bladders and rectums are exposed to high RT dose.

clinicians, who are more involved in the design and administration of treatment plans, the application is not yet suitable. This group might significantly benefit from a version that focuses more on describing the organ shape variations of individual patients. While *VAPOR* supports different possibilities of grouping patients, organs or timesteps, each option is suitable for different types of tasks. For each task, the exploration is straightforward—if the user has a specific hypothesis or exploratory task in mind. Without a clear task in mind, the number of options could be overwhelming. In this case, guidance [67] and a higher degree of automatization should be considered.

VAPOR is a first step towards the analysis of variability in multi-organ patient cohorts, the investigation of the effects of anatomical variability on dose administration and potential RT-induced toxicity, and inclusion of these effects in adaptive RT.

Acknowledgments

This work was supported by Varian Medical Systems of Palo Alto, California, USA in the frame of a research project entitled "*A machine learning centered visualization system for model-based decision making in image-guided and adaptive radiotherapy of cancer*" (principal investigator L. P. Muren). This paper was partly written in collaboration with the VRVis Competence Center. VRVis is funded by BMVIT, BMWFW, Styria, SFG and Vienna Bussiness Agency in the scope of COMET—Competence Centers for Excellent Technologies (854174) which is managed by FFG.

References

[1] Delaney, G, Jacob, S, Featherstone, C, Barton, M. The role of radiotherapy in cancer treatment. Cancer 2005;104(6):1129–1137.

[2] Schlachter, M, Raidou, R, Muren, L, Preim, B, Putora, P, Bühler, K. State-of-the-art report: Visual computing in radiation therapy planning. In: Computer Graphics Forum; vol. 38. 2019, p. 753–779.

[3] Washington, CM, Leaver, DT. Principles and practice of radiation therapy. Book. Elsevier Health Sciences; 2015.

[4] Casares-Magaz, O, Moiseenko, V, Hopper, A, Pettersson, NJ, Thor, M, Knopp, R, et al. Associations between volume changes and spatial dose metrics for the urinary bladder during local versus pelvic irradiation for prostate cancer. Acta Oncologica 2017;56(6):884–890.

[5] 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 2007;8(1):55–68.

[6] Viswanathan, AN, Yorke, ED, Marks, LB, Eifel, PJ, Shipley, WU. Radiation dose–volume effects of the urinary bladder. International Journal of Radiation Oncology* Biology* Physics 2010;76(3):S116–S122.

[7] Thariat, J, Hannoun-Levi, JM, Myint, AS, Vuong, T, Gérard, JP. Past, present, and future of radiotherapy for the benefit of patients. Nature reviews Clinical oncology 2013;10(1):52.

[8] Chai, X, van Herk, M, van de Kamer, JB, Hulshof, MC, Remeijer, P, Lotz, HT, et al. Finite element based bladder modeling for image-guided radiotherapy of bladder cancer. Medical physics 2011;38(1):142–150.

[9] Lotz, HT, van Herk, M, Betgen, A, Pos, F, Lebesque, JV, Remeijer, P. Reproducibility of the bladder shape and bladder shape changes during filling. Medical physics 2005;32(8):2590–2597.

[10] Chai, X, van Herk, M, Hulshof, MC, Bel, A. A voxel-based finite element model for the prediction of bladder deformation. Medical physics 2012;39(1):55–65.

[11] Rios, R, De Crevoisier, R, Ospina, JD, Commandeur, F, Lafond, C, Simon, A, et al. Population model of bladder motion and deformation based on dominant eigenmodes and mixed-effects models in prostate cancer radiotherapy. Medical image analysis 2017;38:133–149.

[12] Raidou, RG, Casares-Magaz, O, Amirkhanov, A, Moiseenko, V, Muren, LP, Einck, JP, et al. Bladder runner: Visual analytics for the exploration of rt-induced bladder toxicity in a cohort study. In: Computer Graphics Forum; vol. 37. 2018, p. 205–216.

[13] Casares-Magaz, O, Raidou, R, Pettersson, N, Moiseenko, V, Einck, J, Hopper, A, et al. Bladder changes during first week of RT for prostate cancer determine the risk of urinary toxicity. European SocieTy for Radiation & Oncology (ESTRO) 38 2019;.

[14] Grossmann, N, Casares-Magaz, O, Muren, LP, Moiseenko, V, Einck, JP, Gröller, E, et al. Pelvis Runner: Visualizing Pelvic Organ Variability in a Cohort of Radiotherapy Patients. In: Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM 2019). The Eurographics Association; 2019, p. 69–78.

[15] Quan, EM, Li, X, Li, Y, Wang, X, Kudchadker, RJ, Johnson, JL, et al. A comprehensive comparison of imrt and vmat plan quality for prostate cancer treatment. International Journal of Radiation Oncology* Biology* Physics 2012;83(4):1169–1178.

[16] Muren, LP, Smaaland, R, Dahl, O. Organ motion, set-up variation and treatment margins in radical radiotherapy of urinary bladder cancer. Radiotherapy and Oncology 2003;69(3):291–304.

[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;18(6):183–193.

[18] Aselmaa, A, Goossens, R, Laprie, A, Ken, S, Fechter, T, Ramkumar,

A, et al. Workflow analysis report. Delft University of Technology 2013;.
[19] Nejad-Davarani, SP, Sevak, P, Moncion, M, Garbarino, K, Weiss, S, Kim, J, et al. Geometric and dosimetric impact of anatomical changes for MR-only radiation therapy for the prostate. Journal of applied clinical medical physics 2019;20(4):10–17.
[20] Wentzel, AP, Hanula, P, Luciani, T, Elgohari, B, Elhalawani, H, Canahuate, G, et al. Cohort-based t-ssim visual computing for radiation therapy prediction and exploration. IEEE Transactions on Visualization and Computer Graphics 2019;26:949–959.
[21] Peura, M, Iivarinen, J. Efficiency of simple shape descriptors. Aspects of visual form 1997;:443–451.
[22] Maaten, Lvd, Hinton, G. Visualizing data using t-sne. Journal of machine learning research 2008;9(Nov):2579–2605.
[23] Comaniciu, D, Meer, P. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis & Machine Intelligence 2002;(5):603–619.
[24] Reiter, O, Breeuwer, M, Gröller, ME, Raidou, RG. Comparative visual analysis of pelvic organ segmentations. In: Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers. Eurographics Association; 2018, p. 37–41.
[25] Kazhdan, M, Funkhouser, T, Rusinkiewicz, S. Rotation invariant spherical harmonic representation of 3 d shape descriptors. In: Symposium on geometry processing; vol. 6. 2003, p. 156–164.
[26] Shlens, J. A tutorial on principal component analysis. arXiv preprint arXiv:14041100 2014;.
[27] Hermann, M, Schunke, AC, Klein, R. Semantically steered visual analysis of highly detailed morphometric shape spaces. In: 2011 IEEE Symposium on Biological Data Visualization (BioVis). IEEE; 2011, p. 151–158.
[28] Hermann, M, Schunke, AC, Schultz, T, Klein, R. A visual analytics approach to study anatomic covariation. In: 2014 IEEE Pacific Visualization Symposium. IEEE; 2014, p. 161–168.
[29] Hermann, M, Schunke, AC, Schultz, T, Klein, R. Accurate interactive visualization of large deformations and variability in biomedical image ensembles. IEEE Transactions on Visualization and Computer Graphics 2016;22(1):708–717.
[30] Hermann, M, Klein, R. A visual analytics perspective on shape analysis: state of the art and future prospects. Computers & Graphics 2015;53:63–71.
[31] Busking, S, Botha, CP, Post, FH. Dynamic multi-view exploration of shape spaces. In: Computer Graphics Forum; vol. 29. 2010, p. 973–982.
[32] Busking, S, Botha, CP, Ferrarini, L, Milles, J, Post, FH. Image-based rendering of intersecting surfaces for dynamic comparative visualization. The visual computer 2011;27(5):347–363.
[33] 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 2013;40(12):4934–4943.
[34] Blumenschein, M, Behrisch, M, Schmid, S, Butscher, S, Wahl, DR, Villinger, K, et al. Smartexplore: Simplifying high-dimensional data analysis through a table-based visual analytics approach. In: IEEE Conference on Visual Analytics Science and Technology (VAST) 2018. 2018,.
[35] Klemm, P, Lawonn, K, Rak, M, Preim, B, Tönnies, KD, Hegenscheid, K, et al. Visualization and analysis of lumbar spine canal variability in cohort study data. In: Vision, Modeling and Visualization. 2013, p. 121–128.
[36] Klemm, P, Oeltze-Jafra, S, Lawonn, K, Hegenscheid, K, Völzke, H, Preim, B. Interactive visual analysis of image-centric cohort study data. IEEE Transactions on Visualization and Computer Graphics 2014;20(12):1673–1682.
[37] Steenwijk, MD, Milles, J, Buchem, M, Reiber, J, Botha, CP. Integrated visual analysis for heterogeneous datasets in cohort studies. In: IEEE VisWeek Workshop on Visual Analytics in Health Care; vol. 3. 2010, p. 3.
[38] Preim, B, Klemm, P, Hauser, H, Hegenscheid, K, Oeltze, S, Toennies, K, et al. Visual analytics of image-centric cohort studies in epidemiology. In: Visualization in Medicine and Life Sciences III. Springer; 2016, p. 221–248.
[39] 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 2015;35(3):44–55.
[40] Alemzadeh, S, Hielscher, T, Niemann, U, Cibulski, L, Ittermann, T, Völzke, H, et al. Subpopulation Discovery and Validation in Epidemiological Data. In: EuroVis Workshop on Visual Analytics (EuroVA). The Eurographics Association; 2017,.
[41] Wang, J, Hazarika, S, Li, C, Shen, HW. Visualization and visual analysis of ensemble data: A survey. IEEE Transactions on Visualization and Computer Graphics 2018;25(9):2853–2872.
[42] Whitaker, RT, Mirzargar, M, Kirby, RM. Contour boxplots: A method for characterizing uncertainty in feature sets from simulation ensembles. IEEE Transactions on Visualization and Computer Graphics 2013;19(12):2713–2722.
[43] Mirzargar, M, Whitaker, RT, Kirby, RM. Curve boxplot: Generalization of boxplot for ensembles of curves. IEEE Transactions on Visualization and Computer Graphics 2014;20(12):2654–2663.
[44] Ferstl, F, Kanzler, M, Rautenhaus, M, Westermann, R. Visual analysis of spatial variability and global correlations in ensembles of iso-contours. In: Computer Graphics Forum; vol. 35. 2016, p. 221–230.
[45] Ferstl, F, Bürger, K, Westermann, R. Streamline variability plots for characterizing the uncertainty in vector field ensembles. IEEE Transactions on Visualization and Computer Graphics 2016;22(1):767–776.
[46] Ferstl, F, Kanzler, M, Rautenhaus, M, Westermann, R. Time-hierarchical clustering and visualization of weather forecast ensembles. IEEE Transactions on Visualization and Computer Graphics 2017;23(1):831–840.
[47] Kim, K, Carlis, JV, Keefe, DF. Comparison techniques utilized in spatial 3d and 4d data visualizations: A survey and future directions. Computers & Graphics 2017;67:138–147.
[48] Keefe, D, Ewert, M, Ribarsky, W, Chang, R. Interactive coordinated multiple-view visualization of biomechanical motion data. IEEE Transactions on Visualization and Computer Graphics 2009;15(6):1383–1390.
[49] Tory, MK, Möller, T, Atkins, MS. Visualization of time-varying mri data for ms lesion analysis. In: Medical Imaging 2001: Visualization, Display, and Image-Guided Procedures; vol. 4319. International Society for Optics and Photonics; 2001, p. 590–599.
[50] Schmidt, J, Preiner, R, Auzinger, T, Wimmer, M, Gröller, ME, Bruckner, S. YMCA—Your mesh comparison application. In: 2014 IEEE Conference on Visual Analytics Science and Technology (VAST). IEEE; 2014, p. 153–162.
[51] Hilbert, D. Über die stetige Abbildung einer Linie auf ein Flächenstück. In: Dritter Band: Analysis· Grundlagen der Mathematik· Physik Verschiedenes. Springer; 1935, p. 1–2.
[52] Weissenböck, J, Fröhler, B, Gröller, E, Kastner, J, Heinzl, C. Dynamic volume lines: Visual comparison of 3d volumes through space-filling curves. IEEE Transactions on Visualization and Computer Graphics 2019;25(1):1040–1049.
[53] Demir, I, Dick, C, Westermann, R. Multi-Charts for Comparative 3D Ensemble Visualization. IEEE Transactions on Visualization and Computer Graphics 2014;20(12):2694 – 2703.
[54] Everitt, B, Landau, S, Leese, M. Cluster analysis. A member of the Hodder Headline Group, London 2001;:429–438.
[55] Caliński, T, Harabasz, J. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods 1974;3(1):1–27.
[56] Harrower, M, Brewer, CA. Colorbrewer. org: an online tool for selecting colour schemes for maps. The Cartographic Journal 2003;40(1):27–37.
[57] Buja, A, Cook, D, Swayne, DF. Interactive high-dimensional data visualization. Journal of computational and graphical statistics 1996;5(1):78–99.
[58] Furmanova, K, Gratzl, S, Stitz, H, Zichner, T, Jaresova, M, Ennemoser, M, et al. Taggle: Scalable visualization of tabular data through aggregation. arXiv preprint arXiv:171205944 2017;.
[59] Balabanian, JP, Viola, I, Gröller, ME. Interactive illustrative visualization of hierarchical volume data. Proceedings of Graphics Interface 2010 2010;:137–144.
[60] Borgo, R, Kehrer, J, Chung, DH, Maguire, E, Laramee, RS, Hauser, H, et al. Glyph-based visualization: Foundations, design guidelines, techniques and applications. In: Eurographics (STARs). 2013, p. 39–63.
[61] Raidou, RG, Casares-Magaz, O, Muren, LP, Heide, Uvd, Roervik, J, Breeuwer, M, et al. Visual analysis of tumor control models for prediction of radiotherapy response. Computer Graphics Forum 2016;35(3):231–240.
[62] Ristovski, G, Garbers, N, Hahn, HK, Preusser, T, Linsen, L. Uncertainty-aware visual analysis of radiofrequency ablation simulations. Computers & Graphics 2019;79:24–35.
[63] Casares-Magaz, O, Raidou, RG, Rørvik, J, Vilanova, A, Muren, LP.

Uncertainty evaluation of image-based tumour control probability models in radiotherapy of prostate cancer using a visual analytic tool. Physics and Imaging in Radiation Oncology 2018;5:5–8.

[64] JavaScript 3D Library. `https://threejs.org/`; 2020.

[65] Bostock, M. Data-Driven Documents. `https://d3js.org/`; 2019.

[66] Pinkawa, M, Asadpour, B, Gagel, B, Piroth, MD, Holy, R, Eble, MJ. Prostate position variability and dose–volume histograms in radiotherapy for prostate cancer with full and empty bladder. International Journal of Radiation Oncology* Biology* Physics 2006;64(3):856–861.

[67] Ceneda, D, Gschwandtner, T, May, T, Miksch, S, Schulz, HJ, Streit, M, et al. Characterizing guidance in visual analytics. IEEE Transactions on Visualization and Computer Graphics 2016;23(1):111–120.