BMC
Medical Imaging

## RESEARCH ARTICLE

**Open Access**

# A semiautomatic tool for prostate segmentation in radiotherapy treatment planning

Jörn Schulz[1*], Stein Olav Skrøvseth[2], Veronika Kristine Tømmerås[3], Kirsten Marienhagen[3] and Fred Godtliebsen[1]

### Abstract

**Background:** Delineation of the target volume is a time-consuming task in radiotherapy treatment planning, yet essential for a successful treatment of cancers such as prostate cancer. To facilitate the delineation procedure, the paper proposes an intuitive approach for 3D modeling of the prostate by slice-wise best fitting ellipses.

**Methods:** The proposed estimate is initialized by the definition of a few control points in a new patient. The method is not restricted to particular image modalities but assumes a smooth shape with elliptic cross sections of the object. A training data set of 23 patients was used to calculate a prior shape model. The mean shape model was evaluated based on the manual contour of 10 test patients. The patient records of training and test data are based on axial T1-weighted 3D fast-field echo (FFE) sequences. The manual contours were considered as the reference model. Volume overlap (Vo), accuracy (Ac) (both ratio, range 0-1, optimal value 1) and Hausdorff distance (HD) (mm, optimal value 0) were calculated as evaluation parameters.

**Results:** The median and median absolute deviation (MAD) between manual delineation and deformed mean best fitting ellipses (MBFE) was Vo ($0.9 \pm 0.02$), Ac ($0.81 \pm 0.03$) and HD ($4.05 \pm 1.3$)mm and between manual delineation and best fitting ellipses (BFE) was Vo ($0.96 \pm 0.01$), Ac ($0.92 \pm 0.01$) and HD ($1.6 \pm 0.27$)mm. Additional results show a moderate improvement of the MBFE results after Monte Carlo Markov Chain (MCMC) method.

**Conclusions:** The results emphasize the potential of the proposed method of modeling the prostate by best fitting ellipses. It shows the robustness and reproducibility of the model. A small sample test on 8 patients suggest possible time saving using the model.

**Keywords:** Delineation, Ellipse model, Empirical Bayes, Prostate, Radiotherapy treatment planning, Statistical shape analysis

## Background

Prostate cancer is the second most diagnosed cancer accounting for 14 percent of all cancers diagnosed worldwide [1]. It is most common in males over the age of 50, and has the highest incidence rate in the developed countries. Aggressive tumors are usually treated with extern radiotherapy or brachytherapy which requires a precise treatment plan for the target volume. In any type of radiotherapy treatment, radiation of healthy tissue should be minimized while maintaining the desired dose to the target volume. Therefore, a successful treatment of prostate

cancer relies on an accurate segmentation of the prostate from the surrounding tissue, by image-based description of the shape and location of the target volume. The volume of interest is characterized by a smooth shape, and for this reason an algorithmic description of the volume is feasible.

Transrectal ultrasound (TRUS), magnetic resonance (MR) and computed tomography (CT) images are the three main imaging techniques used in diagnosis, treatment planning and follow-up examination of prostate cancer. Smith et al. [2] investigated the effects of these imaging techniques on the properties of the prostate volume. A collection of methods available for prostate segmentation is reviewed by Ghose et al. [3]. In addition to the methods presented by Ghose et al., alternative

*Correspondence: jorn.schulz@uit.no
[1] Department of Mathematics and Statistics, University of Tromsø, 9037 Tromsø, Norway
Full list of author information is available at the end of the article

approaches are available in the literature, such as the medial or skeleton representation of the prostate [4-8]. The present work proposes a segmentation method which falls into the category of deformable meshes in Ghose et al. [3], but refers to the term geometrical parametrization as described in Dryden and Mardia [9]. The main focus of this paper is the development of a statistical shape model for the prostate. An overview about this type of models in 3D medical image segmentation is presented for example by Davies et al. [10] and Heimann and Meinzer [11].

The works of Saroul et al. [12] and Mahdavi et al. [13] are related to the stacked ellipses parametrization method used in this paper. Mahdavi et al. [13] proposes a 3D ellipsoid shape of the prostate in warped transrectal 3D ultrasound images based on control points. This method extends the warping idea proposed in Badiei et al. [14] from 2D to 3D ultrasound images. On the contrary, we focus on slice-wise best fitting ellipses which will introduce more flexibility into the model, e.g., between the positions and lengths of the first and second axes of the ellipses between neighbor slices. The approach of slice-wise best fitting ellipses has similarities to a tubular medial representation [15].

Beside the single segmentation of the prostate, several attempts have been tried out for a joint segmentation of neighbor organ and structure to gain improved segmentation results [16-18].

To our knowledge, despite the substantial effort in this area, no widely implemented algorithm exists. In oncology departments this means that the physician has to delineate the prostate slice by slice. This is time-consuming and inefficient. We propose a less ambitious approach compared to more sophisticated models, such as skeletal models as discussed above, in that we use a method that gives a useful starting point for the physician after the definition of few control points. Given the initial estimate of the volume of interest, the physician can adjust the estimate according to their evaluation of the image rather than starting from scratch. By this approach, we obtain the same accuracy with less effort. The main points in our approach are as follows: First, we accept that the algorithm cannot give a fully precise description of the volume. Our main aim is therefore to give a good estimate which can be used as a starting point for the physician. Second, we use a simple ellipse model that is easy to interpret and understand. Our hypothesis is that a more efficient use of physicians in Radiotherapy Treatment Planning (RTP) of patients with prostate cancer can be obtained by an easy-to-interpret semiautomatic tool.

Figure 1 shows an example of the initial estimate we typically obtain for a single image slice. The dashed line in (a) to (e) describes the manual contour while the solid line shows the best fitting ellipse including the two principal axes for the observed data of this slice. Note that the fitted model is very much in agreement with the manual line, indicating that the stacked ellipses model gives a good description of the object of interest. The solid lines in (f)-(j) shows the outcome from our model in this situation together with few defined control points.

This result shows a typical performance of the method, and that the estimate is close to the best fit we can obtain with the ellipse model. The full processing demands little computational resources, such that the suggested delineation can be presented immediately. The example is discussed further in the Methods and Results and discussion section.

The rest of the paper is organized as follows. In the Methods section, we introduce the data sources and the proposed stacked ellipses model, and discuss the shape space and statistics along with constraints and parameters. Results are presented in the Results and discussion section using a test data set to show the potential of the mean shape model, followed by a Conclusions section. Additional file 1 with further detailed discussion is available online.

## Methods

### Preliminaries

Each prostate must be described by a shape model in order to calculate statistics, e.g., by stacked ellipses as a parametric shape model. The parameters of a parametric shape model can be estimated from a training set. The training set models also the geometric variability of anatomical structures by a shape probability distribution. The training set contains volume and contour information of segmented prostates from $N$ patients. The volume information describes the image modalities (e.g., CT or MR) and the contour information the volume of interest as defined in the following.

The volume information of each training set $n = 1, \ldots, N$ is defined by a 3-dimensional matrix $V_n$ where $V_n(i, h)$ contains the observed gray level in voxel $(i, h)$, $i = (i_1, i_2) \in \{1, \ldots, I_1\} \times \{1, \ldots, I_2\}$ are the pixel indices in a slice, where typically $I_1 = I_2$, and $h \in \{1, \ldots, H\}$ is the number of slices per data set. The number of slices $H$ is not necessarily the same for all patients in the training data sets. Therefore, we indicate $H$ by $H_n$ and in the same manner $I_1$ by $I_{n1}$ and $I_2$ by $I_{n2}$, but for simplicity we use $H$, $I_1$ and $I_2$ if the meaning is clear.

In addition to the volume information, each training set $n = 1, \ldots, N$ consists of contour information of the prostate, manually drawn by a physician. The contour information can be modeled by a $(M \times K_n)$ configuration matrix $X_n := (X_{n1}, \ldots, X_{nK_n})$ with $X_{nk} = (x_{1k}^n, x_{2k}^n, x_{3k}^n)^T \in \mathbb{R}^3$, $k = 1, \ldots, K_n$, where $K_n$ defines the total number of available contour information points in a data set and $M = 3$ defines the dimension. We assume the contour information for an object is defined
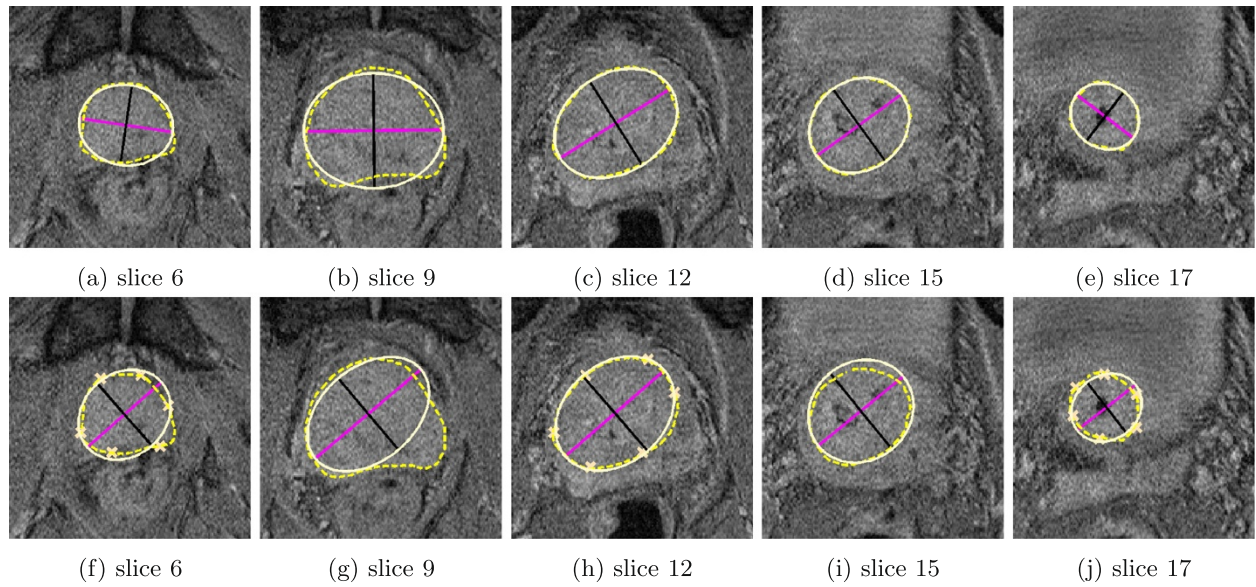
**Figure 1 Selected slices of MR data set 3 from the test data set. (a)-(e)** Manual delineation of the prostate (dashed line) and best fitting ellipse (solid line). **(f)-(j)** Manual delineation of the prostate (dashed line), deformed mean shape (solid line) and defined control points in the first, center and last slice.

in a sequentially sorted number $L_n$ of equidistant slices whereas each contour slice contains $\tilde{K}_{nl}$ contour points, $l = 1, \ldots, L_n$. Hence it follows $K_n = \sum_l \tilde{K}_{nl}$ and $X_n = (\tilde{X}_{n1}, \ldots, \tilde{X}_{nL_n})$. The image information in slice $l$ is denoted by $S_{nl}$ and $S_n = \{S_{n1}, \ldots, S_{nL_n}\} \subseteq V_n$ and $\tilde{X}_{nl}$ defines the configuration matrix in slice $S_{nl}$.

In summary, the training population is given by the set $\{\mathbb{V}, \mathbb{X}\}$, with a set of volume information $\mathbb{V} = \{V_1, \ldots, V_N\}$ and configuration matrices $\mathbb{X} = \{X_1, \ldots, X_N\}$. We assume $X_n$ defines the configuration matrix for the corresponding data set $V_n$ and matches the volume information $V_n$ exactly.

The contour information is often defined in a Patient based Coordinate System (PCS) whereas the volume information is defined in an Image based Coordinate System (ICS). The ICS can be transformed to PCS by a transformation matrix $\Lambda_{DCM}$, which transform an image coordinate $p^{im} = (i_1, i_2, h)^T$ to patient coordinate $p^p = (x, y, z)^T$. The definition of $\Lambda_{DCM}$ and the relation between PCS and ICS (see Figure 2) is discussed in detail in the Additional file 1. In addition, we introduce a de-rotated PCS where volume and contour information are aligned to each other.

**Modeling**

The prior information inferred from the training set is incorporated into a shape model. We assume a stacked ellipse model as a shape prior for the prostate. Specifically,
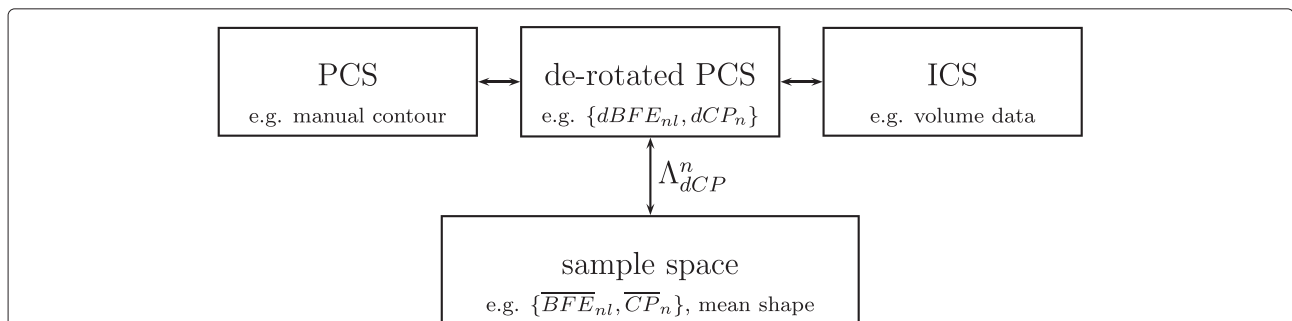


**Figure 2 Visualization of different coordinate systems with example data.** *PCS:* Patient based coordinate system (manual delineation line). *ICS:* Image coordinate system (volume data). *de-rotated PCS:* de-rotated patient based coordinate system with same scale and origin as PCS but same orientation as ICS (de-rotated best fitting ellipse $dBFE_{nl}$, de-rotated control points $dCP_n$). *sample space:* The transformation matrix $\Lambda_{dCP}^n$ maps the de-rotated data $\{dBFE_n, dCP_n\}$ to $\{\overline{BFE}_n, \overline{CP}_n\}$ in the sample space.

the prostate outline in slice $S_{nl}$, $l = 1, \ldots, L_n$, $n = 1, \ldots, N$ is modeled by a slicewise best-fitting ellipse, as visualized in Figure 3. An ellipse in slice $S_{nl}$ can be uniquely described by $\rho_{nl} = (\theta^{nl}, \alpha^{nl}, \phi^{nl})^T \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times (-\frac{\pi}{2}, \frac{\pi}{2}]$ with

- **position** $\theta^{nl} = (\theta_1^{nl}, \theta_2^{nl})^T \in \mathbb{R}^2$ defines the center in slice $S_{nl}$,
- **length of principle axes** $\alpha^{nl} = (\alpha_1^{nl}, \alpha_2^{nl})^T \in \mathbb{R}^2_+$ and
- **rotation angle** $\phi^{nl} \in (-\frac{\pi}{2}, \frac{\pi}{2}]$ in slice $S_{nl}$.

The rotation parameter $\phi^{nl}$ is defined corresponding to the ICS with origin $\theta^{nl}$ in slice $S_{nl}$. The boundary of an ellipse $\rho_{nl}$ centered at $\theta^{nl} \in \mathbb{R}^2$ in slice $S_{nl}$ is defined by

$$C(\rho_{nl}) = \left\{ Rx + \theta^{nl} : \frac{x_1^2}{(\alpha_1^{nl})^2} + \frac{x_2^2}{(\alpha_2^{nl})^2} = 1, x \in \mathbb{R}^2 \right\} \tag{1}$$

$$\text{and } R = \begin{pmatrix} \cos\phi^{nl} & -\sin\phi^{nl} \\ \sin\phi^{nl} & \cos\phi^{nl} \end{pmatrix}$$

is a rotation matrix in $\mathbb{R}^2$ with rotation angle $\phi^{nl}$ and $x = (x_1, x_2)^T$.

The shape model described in this section requires the best fit of an ellipse $C(\rho_{nl})$ to the contour information $\tilde{X}_{nl}$ in each slice, i.e., we model $\tilde{X}_{nl} = C(\rho_{nl}) + \epsilon$ where $\epsilon$ is an error with mean zero. The best-fitting ellipses provide us with a slice-by-slice parametrization of the prostate for all slices in each training shape.

The problem of fitting an ellipse to geometric features like the contour is discussed widely in the literature (e.g., [19,20]). This work follows Ahn et al. [19], who proposed a least-square minimizer for $\tilde{X}_{nl}$. The nonlinear estimate of parameters $\rho_{nl} = (\theta_1^{nl}, \theta_2^{nl}, \alpha_1^{nl}, \alpha_2^{nl}, \phi^{nl})^T$ given $\tilde{X}_{nl}$ must minimize the error

$$g(\hat{\rho}_{nl}) = \left(\tilde{X}_{nl} - \tilde{C}(\hat{\rho}_{nl})\right)^T \left(\tilde{X}_{nl} - \tilde{C}(\hat{\rho}_{nl})\right)$$

where $\tilde{C}(\hat{\rho}_{nl})$ is a set of nearest orthogonal points of $\tilde{X}_{nl}$ to $C(\hat{\rho}_{nl})$.
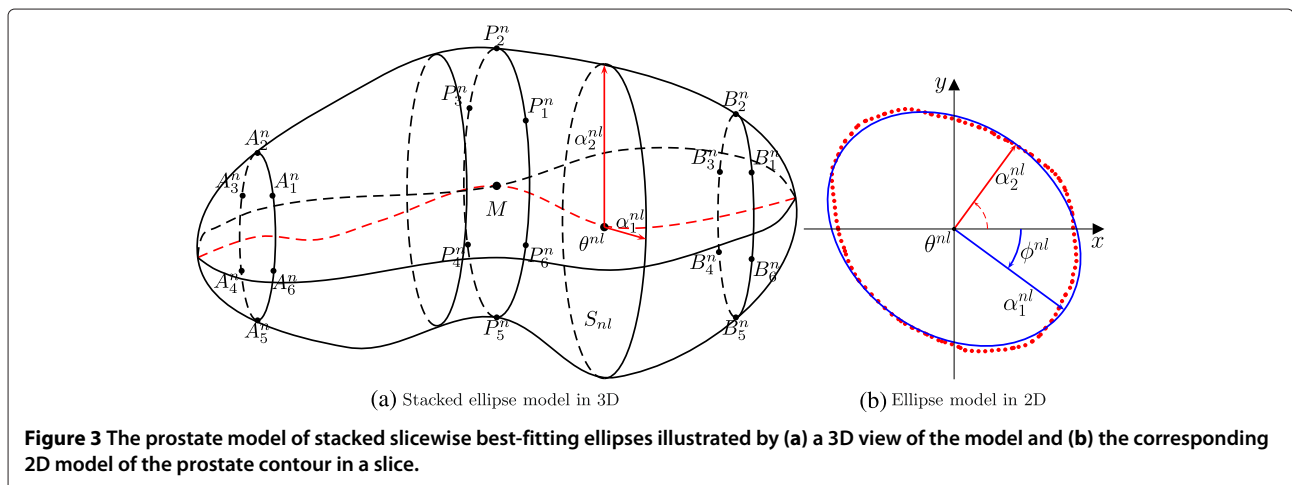
**Definition 1** (Best fitting ellipse (BFE)). *A best fitting ellipse for slice $S_{nl}$ is defined by the set $BFE_{nl} := (\theta^{nl}, \alpha^{nl}, \phi^{nl})^T \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times (-\frac{\pi}{2}, \frac{\pi}{2}]$, $l = 1, \ldots, L_n$, $n = 1, \ldots, N$ and minimizes the error function g, i.e., $BFE_{nl} = \hat{\rho}_{nl}$ with*

$$g(\hat{\rho}_{nl}) = \min_{\rho_{nl} \in \mathbb{R}^2 \times \mathbb{R}^2_+ \times (-\frac{\pi}{2}, \frac{\pi}{2}]} g(\rho_{nl}). \tag{2}$$

The first and second principal axes must be reordered after calculation of $BFE_n = \{BFE_{n1}, \ldots, BFE_{nL_n}\}$ in order to establish correspondence between parameters of adjacent slices and across the population. Improved correspondence will support accurate statistics. The basic idea in our reordering procedure is to carry out the reordering corresponding to the lowest rotation angle of both principal axes to the first principal axis of the neighbor slice where the center slice is chosen as the basis. The rotation between the center slice $M$ and an arbitrary slice is constrained by $\max(|\phi^i - \phi^M|) = \pi, i \in \{1, \ldots, L\}$ after reordering. Therefore, the set $BFE_n$ of reordered best-fitting ellipses is an element of $(\mathbb{R}^2 \times \mathbb{R}^2_+ \times (-\pi, \pi])^{L_n}$.

A further improvement of correspondence is achieved by the introduction of two additional constraints in the parameter model.

First, we relax the rotation parameter $\phi^{nl}$ in case of circularity. If both principal axes have the same length, the orientation of an ellipse is undefined. Therefore we penalize $\phi^{nl}$ in the case of high circularity by taking $\phi^{nl^l}$ from the neighboring slices into account. Second, smoothing is performed between neighboring slices to avoid large forward and backwards rotations between $\phi^{n(l-1)}, \phi^{nl}$ and $\phi^{n(l+1)}$. The reordering algorithm and implementation of constraints are described in detail in Additional file 1.



**Figure 3 The prostate model of stacked slicewise best-fitting ellipses illustrated by (a) a 3D view of the model and (b) the corresponding 2D model of the prostate contour in a slice.**

The current implementation assumes the definition of control points $CP_n$ in the training data set $\{V_n, X_n, BFE_n\}$, where $BFE_n \in (\mathbb{R}^2 \times \mathbb{R}_+^2 \times (-\pi, \pi])^{L_n}$ is a reordered set of best fitting ellipses, $n = 1, ..., N$. Furthermore, the control points have to be defined manually by a physician in a new patient data set. The control points are used to make the best fitting ellipses $BFE_n$ comparable and to transform the parametrized ellipses model to a common position, scale and orientation by a transformation matrix $\Lambda_{dCP}^n$. The transformation matrix $\Lambda_{dCP}^n$ maps the de-rotated prior data $\{dBFE_n, dCP_n\}$ to $\{\overline{BFE_n}, \overline{CP_n}\}$ in the sample space, as depicted in Figure 2. In this article, we assume 6 control points in the first, center and last slice at the boundary of the prostate, i.e.,

$$CP_n = \{A_1^n, \ldots, A_6^n, P_1^n, \ldots, P_6^n, B_1^n, \ldots, B_6^n\}$$

as visualized in Figure 3. In addition, we have tested alternative control point configurations. They are described together with the construction of $\Lambda_{dCP}^n$ in Additional file 1.

After transformation we have obtained a reordered and comparable set of best fitting ellipses

$$\overline{BFE}_n = \{\overline{BFE}_{n1}, \ldots, \overline{BFE}_{nL_n}\}$$

with $\overline{BFE}_{nl} = (\overline{\theta}^{nl}, \overline{\alpha}^{nl}, \overline{\phi}^{nl})^T$, $n = 1, \ldots, N$, $l = 1, \ldots, L_n$. The statistical analysis of the training data requires an equal number $L_1 = \ldots = L_N$ to establish correspondence between the parameters of the best fitting ellipses. Therefore, we interpolate the set $\overline{BFE}_{nl}$ to a common number $L$.

When $L$ is chosen, interpolation is done by independent cubic interpolation in each dimension, i.e., we find points of a one-dimensional function that underlies the data $\overline{\theta}_1^{nl}, \overline{\theta}_2^{nl}, \overline{\theta}_3^{nl}, \overline{\alpha}_1^{nl}, \overline{\alpha}_1^{nl}$ and $\overline{\phi}^{nl}$. The final interpolated best fitting ellipses are denoted by

$$iBFE_n = \{iBFE_{n1}, \ldots, iBFE_{nL}\}. \quad (3)$$

These ellipses are used for the statistical analysis and computation of a mean shape model. To keep things simple, we denote such a reordered, transformed and interpolated set of best-fitting ellipses by $BFE_{nl} = (\theta^{nl}, \alpha^{nl}, \phi^{nl})^T$ for the number $L$ of contour slices with $l = 1, \ldots, L$ and $n = 1, \ldots, N$. The comparable set of best fitting ellipses $BFE_n$ is an element of the shape space $(\mathbb{R}^2 \times \mathbb{R}_+^2 \times (-\pi, \pi])^L$.

### Statistical analyses

After reconstruction of our shape space we estimate the expectation and variance of the parameters of a mean shape model $\mu_{BFE} = \{\mu_{BFE}^1, \ldots, \mu_{BFE}^L\}$ with $\mu_{BFE}^l = (\mu_\theta^l, \mu_\alpha^l, \mu_\phi^l)$ from the training set $BFE_{nl}$, $l = 1, \ldots, L$. We

denote the mean shape mean best fitting ellipses (MBFE). In addition to the described ellipse parameters we define the position $\theta^{nl} = (\theta_1^{nl}, \theta_2^{nl}, \theta_3^{nl})^T$ in terms of a distance vector $\eta^{nl}$ of $\theta^{nl}$ to a center curve defined by the control points. We model $\theta^l = \xi^l + \eta^l$, where $\xi^l$ is analytically defined by $L$ intersection points of the curve within each slice. Thereby, we are describing the mean shape which is closest to the control points. This approach is reasonable under the assumption that the control points are well defined. In Additional file 1 we explore various ways of describing the position parameter for different control point methods.

The mean curve of the expected location is given by

$$\mu_{\theta_j}^l = \frac{1}{N} \sum_{i=1}^N \theta_j^{il}, \quad j \in \{1, 2, 3\}, \quad (4)$$

where $\mu_\theta^l = (\mu_{\theta_1}^l, \mu_{\theta_2}^l, \mu_{\theta_3}^l)^T$, $l = 1, \ldots, L$. The variance and covariance are estimated by

$$(\sigma_{\theta_j}^l)^2 = \frac{1}{N-1} \sum_{i=1}^N (\theta_j^{il} - \mu_{\theta_j}^l)^2, \quad j \in \{1, 2, 3\}, \text{ and } \quad (5)$$

$$\Sigma_\theta^l = \frac{1}{N-1} \sum_{i=1}^N (\theta^{il} - \mu_\theta^l)(\theta^{il} - \mu_\theta^l)^T. \quad (6)$$

The length parameter is modeled by a log-normal distribution because $\alpha \in \mathbb{R}_+^2$. Thus we estimate the mean and variance of $a = \log(\alpha) \in \mathbb{R}^2$. The estimation of means and variances of the remaining parameters $a, \phi, \eta$ is according to (4-5).

Following Dryden and Mardia [9] we suggest a prior distribution for a new data set as

$$\theta_1^l \sim N\left(\mu_{\theta_1}^l, (\sigma_{\theta_1}^l)^2\right), \quad \theta_2^l \sim N\left(\mu_{\theta_2}^l, (\sigma_{\theta_2}^l)^2\right),$$

$$a_1^l \sim N\left(\mu_{a_1}^l, (\sigma_{a_1}^l)^2\right) \Longleftrightarrow$$

$$\alpha_1^l \sim \log\text{-}N\left(\mu_{a_1}^l, (\sigma_{a_1}^l)^2\right) \text{ with } a_1^l = \log(\alpha_1^l),$$

$$a_2^l \sim N\left(\mu_{a_2}^l, (\sigma_{a_2}^l)^2\right) \Longleftrightarrow$$

$$\alpha_2^l \sim \log\text{-}N\left(\mu_{a_2}^l, (\sigma_{a_2}^l)^2\right) \text{ with } a_2^l = \log(\alpha_2^l),$$

$$\phi^l \sim N\left(\mu_\phi^l, (\sigma_\phi^l)^2\right),$$

$l = 1, \ldots, L$. If $\theta^l$ is defined according to the center curve given by the control points as described above, we model $\eta_i^l \sim N\left(\mu_{\eta_i}^l, (\sigma_{\eta_i}^l)^2\right)$, $i = 1, 2$. Since the rotational parameter is expected to have small variance it is not necessary to apply a circular distribution, and we assume normality.

After constructing the shape model we estimate the best fitting ellipse $BFE_l$ parametrized by $\rho_l = (\theta^l, \alpha^l, \phi^l)^T$, $l = l, \ldots, L$ in a new data set given the control points $CP$. This

is obtained through the posterior $\pi(\rho \mid S)$ where $s_{il} \in S \subseteq V$ is the volume information and $i = (i_1, i_2) \in I(\rho)$ is a set of indices within the ellipses $\rho$. The control points $CP$ are used to deform the prior model $\pi(\rho)$. Therefore we model the posterior by an empirical Bayes approach [21]. The posterior

$$\pi(\rho \mid S, CP) \propto L(S \mid \rho) * \pi(\rho \mid CP) \qquad (7)$$

defines the posterior density of the deformed template $\pi(\rho \mid CP)$ given the the observed image. The Likelihood or image model $L(S \mid \rho)$ is the joint probability density function of the gray levels given the parametrized object $\rho|_{CP}$, while $\rho|_{CP}$ defines the ellipses $\rho$ deformed by the control points $CP$. The prior $\pi(\rho \mid CP)$ models realistic variations from our mean shape $\mu_{BFE} \in (\mathbb{R}^2 \times \mathbb{R}_+^2 \times (-\pi, \pi])^L$ given the control points. We are estimating the posterior distribution using a Markov chain Monte Carlo (MCMC) approach. The method and results are discussed in detail in Additional file 1.

**Evaluation**
We have evaluated the proposed method using 33 patient case studies. The training data set consists of $N = 23$ T1-weighted Fast Field Echo (FFE) 3D Magnetic Resonance (MR) data. The mean shape model and variance is calculated from the training data set and applied to a test data set of 10 MR FFE case studies. The splitting in test and training data is done according to the sequence of data acquisition. Each data set consists of $H_n$ Digital Imaging and Communications in Medicine (DICOM) image files and one DICOM region structure file, while the contour information of the prostate is stored in the header of a DICOM file without any image information. The voxel size $(l_x, l_y, l_z)$ is $(0.559\text{mm}, 0.559\text{mm}, 3\text{mm})$ with a slice distance of 3.3mm of the data sets $V_n$. Each slice consists of $288 \times 288$ voxels. The average number of slices with manual prostate contour information is $10.478 \pm 2.626$ (mean±standard deviation) in the training set and $10.5 \pm 2.799$ in the test set where the mean is given by $\mu_L = \frac{1}{N} \sum_{n=1}^{N} L_n$ and the standard deviation by $(\frac{1}{N-1} \sum_{n=1}^{N} (L_n - \mu_L)^2)^{\frac{1}{2}}$. Figure 1 illustrates test patient 3, whose image set consists of 24 MR FFE slices whereas 12 slices contain contour information. Slice 6 is the first slice where contour information of the prostate is available and the last slice is 17.

An ethics approval was not required for this study under Norwegian law because the aim of this study was to develop a tool and not to obtain new knowledge on medicine or diseases. The study uses solely data that were collected during routine medical treatment independent of this study at the University hospital Northern Norway (UNN). Data was provided after a full anonymization following the required guidelines at UNN.

Three metrics are used to compare the manual and the semi-automatic contours. In the axial slices, where the expert manual delineations are present, we calculate the Hausdorff distance (HD) by

$$d_{hd}(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y) \right\}.$$

$$(8)$$

The Hausdorff distance measures the maximum distance of a point in a set $X$ to the nearest point in $Y$ or vice versa. Generalization to 3D uses mean or median over all slices. The measure indicates how much manual corrections are required. An ideal value of HD equal to zero reflects complete agreement of the contours. In addition, the number of slices with a HD greater than 3 mm is reported and compared to the total number of slices with prostate information. A threshold of 3 mm is often seen as clinically acceptable [22]. A second criteria is the volume overlap (or Dice similarity coefficient) defined by

$$d_{vol}(X, Y) = 2 \frac{|X \cap Y|}{|X| + |Y|}, \qquad (9)$$

where $| \cdot |$ is the number of voxels contained in a region. Finally, accuracy is defined as

$$d_{acc}(X, Y) = 1 - \frac{|FP| + |FN|}{|TP| + |FN|} \qquad (10)$$

with $TP = X \cap Y$ volume included in both $X$ and $Y$ (true positive), $FN = X \cap (\neg Y)$ volume of $X$ not included in $Y$ (false negative) and $FP = (\neg X) \cap Y$ volume of $Y$ not included by $X$ (false positive). Both values range from 0 to 1, with optimal value 1. Volume overlap indicates how much of the prostate has been detected by the approach while accuracy shows how incapable the method is to select the true prostate pixels.

In addition to the quantitative metrics, we have performed a small pilot test on 8 new patients comparing time expenditure using the proposed method and manual delineation. The time expenditure for the proposed method includes the definition of control points and the correction of the contour obtained by the method for each patient. The used mean shape model and variance was calculated from the training data as described above. Time measurements were obtained by two independent physicians for each case. Manual delineations, definition of control points and corrections were performed using the treatment planning system Eclipse$^{\text{TM}}$.

**Results and discussion**
The evaluation is performed on the deformed mean best fitting ellipses, i.e., on $\pi(\rho \mid CP)$ in formula (7). Additional evaluations are done in Additional file 1 for $\pi(\rho \mid S, CP)$.

**Table 1 Evaluation metrics comparing BFE to manual delineations, and comparing deformed MBFE to manual delineations**

| Test set | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | **BFE** | | | | | |
| Dice 3D | 0.96 | 0.97 | 0.95 | 0.96 | 0.94 | 0.94 | 0.96 | 0.96 | 0.97 | 0.93 |
| Accuracy | 0.93 | 0.94 | 0.91 | 0.91 | 0.88 | 0.88 | 0.92 | 0.93 | 0.94 | 0.85 |
| HD mean | 1.32 | 1.09 | 2.24 | 1.34 | 1.94 | 1.61 | 1.41 | 1.65 | 1.59 | 2.49 |
| #HD > 3mm | 0 (6) | 0 (8) | 3 (12) | 0 (8) | 2 (11) | 0 (13) | 0 (9) | 0 (15) | 1 (13) | 2 (10) |
| | | | | | **Deformed MBFE** | | | | | |
| Dice 3D | 0.92 | 0.93 | 0.90 | 0.91 | 0.88 | 0.84 | 0.92 | 0.88 | 0.89 | 0.88 |
| Accuracy | 0.84 | 0.84 | 0.81 | 0.82 | 0.74 | 0.70 | 0.84 | 0.73 | 0.80 | 0.74 |
| HD mean | 2.38 | 2.58 | 6.12 | 2.76 | 3.79 | 4.32 | 2.73 | 5.59 | 4.88 | 4.63 |
| #HD > 3mm | 1 (6) | 1 (8) | 9 (12) | 4 (8) | 7 (11) | 8 (13) | 2 (9) | 13 (15) | 8 (13) | 9 (10) |

(unit: Dice 3D and accuracy in percentage, HD mean in mm, number of slices with HD $\geq$ 3mm compared to the total number of slices in brackets).
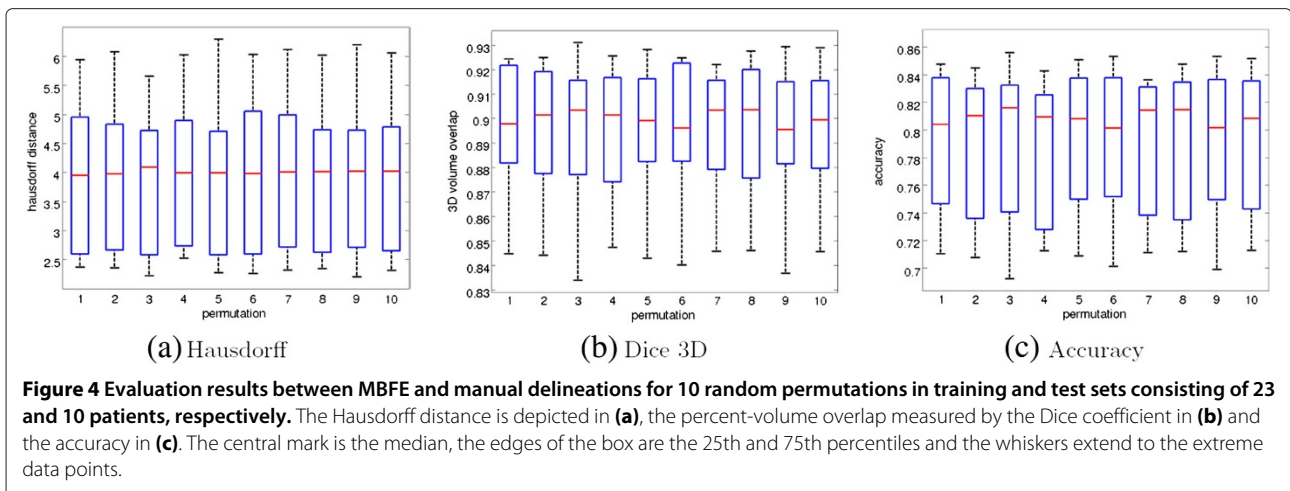
Table 1 contains the distance metrics defined in (8) - (10) comparing the manual delineation and BFE for each test data set, and comparing the manual delineation and the deformed MBFE described by $\pi(\rho \mid CP)$. The high Dice similarity coefficient and accuracy values and small Hausdorff distances between manual delineations and BFE confirm the stacked ellipses model. This is also reflected by the count of slices with a HD greater than 3 mm in each test data set. Only 7.6% of the in total 105 slices have a HD greater than 3 mm. The values show the best possible description of the test cases by the proposed model. Furthermore, Table 1 presents the metrics comparing the manual delineation and the deformed MBFE for each test data set. The distance metrics reveal the fairly accurate results. The values indicate that the deformed MBFEs used as initial contours for final delineations will lower the time expenditure of the delineation procedure. However, more slices with a HD greater than 3 mm can be observed, 59% of the 105 slices. Particularly, the test sets 3, 7 and 10 have a ratio greater than 75%. Figures 1a to 1e illustrate 5 slices of the BFE evaluation of test patient 3 from Table 1. Figures 1f to 1j illustrate 5 slices of the deformed MBFE evaluation of test patient 3 from Table 1 with a volume overlap of the manual delineation line and the deformed mean shape of 0.90 and accuracy 0.81.

Table 2 shows the median and median absolute deviation (MAD) for the data groups "test data", "training data" and "all data". A BFE volume overlap for all data of $0.954 \pm 0.010$ (median$\pm$MAD) and accuracy of $0.908 \pm 0.020$ confirm the model further (Table 2). Similar values in the subset of test data and training data are indicating model robustness. In addition to the BFE results, Table 2 summarizes the results by median and MAD of the distances between manual delineation and the deformed MBFE. A median volume overlap of $0.899 \pm 0.021$ and accuracy of $0.807 \pm 0.035$ of the test data show further the power of the prior. The deformation of the prior is done by the control points and can be computed directly since there is no sampling or estimation involved at this point.

To evaluate the robustness of the model, we randomly split 10-times the set of 33 patients into a training set with 23 cases and a test set with 10 cases. Figure 4 shows the evaluation distances between the manual delineation and the deformed MBFE. The central mark is the median, the edges of the box are the 25th and 75th percentiles and the whiskers extend to the extreme data points. The figure shows only small variation between the different permutations, thereby demonstrating robustness of the stacked ellipses model.

**Table 2 Evaluation metrics comparing BFE/MBFE and manual delineations for different data groups**

| Data group | | Test data | | Training data | | All data | |
|---|---|---|---|---|---|---|---|
| | | BFE | MBFE | BFE | MBFE | BFE | MBFE |
| HD mean [mm] | median | 1.604 | 4.052 | 1.840 | 3.806 | 1.810 | 3.806 |
| | MAD | 0.274 | 1.305 | 0.286 | 0.427 | 0.289 | 0.538 |
| Dice 3D [pct] | median | 0.959 | 0.899 | 0.954 | 0.903 | 0.954 | 0.903 |
| | MAD | 0.008 | 0.021 | 0.007 | 0.013 | 0.010 | 0.019 |
| Accuracy [pct] | median | 0.918 | 0.807 | 0.908 | 0.800 | 0.908 | 0.800 |
| | MAD | 0.015 | 0.035 | 0.014 | 0.031 | 0.020 | 0.036 |

**Figure 4 Evaluation results between MBFE and manual delineations for 10 random permutations in training and test sets consisting of 23 and 10 patients, respectively.** The Hausdorff distance is depicted in **(a)**, the percent-volume overlap measured by the Dice coefficient in **(b)** and the accuracy in **(c)**. The central mark is the median, the edges of the box are the 25th and 75th percentiles and the whiskers extend to the extreme data points.

Results from using MCMC to further optimize the delineation, as described in the Methods section, are only presented in Additional file 1 since a slight improvement comes at the cost of large computation time.

The time comparisons indicated an average of 30% time saving using the proposed method compared to manual delineation. The time measurement of the proposed method includes the definition of controls points as well as the correction of the estimated contour by the physician.

## Conclusions

The presented results demonstrate the potential of the proposed method in modeling the prostate by slicewise best fitting ellipses. Deformation of the mean shape using control points gives very good results with little computational cost. Hence we believe that providing physicians with a good initial contour is beneficial in the clinical praxis of radiotherapy treatment.

The corrections of generated delineations based on few control points were not streamlined in the workflow of the physicians, and the task of correcting contours is not part of their everyday activity. Furthermore, corrections were not done directly after the definition of the control points and sometimes by different physicians, and physicians had to deal with a different orientation of the data set in the treatment planning system than in the diagnostic MRI. These issues must and can be solved for a well designed system. Therefore, a time saving of 30% likely represent a lower limit, and has to be validated in a well designed and properly powered study. Furthermore, we expect larger time savings in data sets where the prostate is imaged in a higher number of slices. In the extreme case, if the prostate is visible in only three slices, the BFE approach would not give any benefit using the current control point method. The study of inter/intra-observer variability using the proposed method compared to manual delineation was considered to be beyond the

scope of this study and is left open for interesting future work.

In addition, the results show a precise description of the prostate by the BFE model with an average volume overlap of 95%. The high performance of the deformed mean shape model using the control points explains the small improvement by applying MCMC. Nevertheless, an improvement of the likelihood in the posterior distribution or by an active appearance model [23] is a field of further research as elaborated in Additional file1. A clear disadvantage of an additional method like MCMC is the extra computation time.

Further improvements can be achieved in the constraint and regularization terms, e.g., by considering the surface curvature versus changes of the ellipses parameters. We do not expect abrupt changes between neighboring slices around the central slice, but larger changes between slices towards the ends can be permitted, particularly in the length of the first and second principal axis. Also, the reduction of manual interaction in the proposed method is left for future work.

## Additional file

**Additional file 1: Supplementary materials.** Article containing *i.)* a detailed description of the relative coordinates systems (e.g. ICS, PCS) on the basis of the DICOM file structure, *ii.)* post-processing procedures as for example reordering and introduction of constraints, *iii.)* a discussion of different control point method with construction of the transformation matrix $\Lambda_{dCP}^n$ and the parameter $\eta^{nl}$, *iv.)* elaboration of the posterior distribution, and *v.)* a section with additional data analysis.

**Authors' contributions**
JS conducted the design, development of the methodology and the implementation of the study resulting in a stacked ellipsoid model generation from a training data set and the registration of the model in a new patient

data set. Furthermore, JS drafted and revised the manuscript. SOS and FG contributed to the conception and design of the study, and helped to draft the manuscript. VKT and KM coordinated and carried out the acquisition of patient data. All authors read and approved the final manuscript.

### Author details
[1]Department of Mathematics and Statistics, University of Tromsø, 9037 Tromsø, Norway. [2]Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, 9038 Tromsø, Norway. [3]Department of Oncology, University Hospital of North Norway, 9038 Tromsø, Norway.

### References
1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D: **Global cancer statistics.** *CA- Cancer J Clin* 2011, **61**(2):69–90.
2. Smith WL, Lewis C, Bauman G, Rodrigues G, D'Souza D, Ash R, Ho D, Venkatesan V, Downey D, Fenster A: **Prostate volume contouring: a 3D analysis of segmentation using 3DTRUS, CT, and MR.** *Int J Radiat Oncol* 2007, **67**(4):1238–1247.
3. Ghose S, Oliver A, Martá R, Lladó X, Vilanova JC, Freixenet J, Mitra J, Sidibé D, Meriaudeau F: **A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images.** *Comp Meth Prog Bio* 2012, **108**:262–287.
4. Crouch J, Pizer SM, Chaney EL, Hu Y, Mageras GS, Zaider M: **Automated finite element analysis for deformable registration of prostate images.** *IEEE T Med Imaging* 2007, **10**:1379–1390.
5. Dam EB, Fletcher P, Pizer SM: **Automatic shape model building based on principal geodesic analysis bootstrapping.** *Med Image Anal* 2008, **12**(2):136–151.
6. Merck D, Tracton G, Saboo R, Levy J, Chaney E, Pizer S, Joshi S: **Training models of anatomic shape variability.** *Med Phys* 2008, **35**(8):3584–3596.
7. Siddiqi K, Pizer SM: *Medial Representations*. Springer: Dordrecht; 2008.
8. Pizer SM, Jung S, Goswami D, Zhao X, Chaudhuri R, Damon JN, Huckemann S, Marron JS: **Nested sphere statistics of skeletal models.** In *Innovations for Shape Analysis: Models and Algorithms*. Maragos, P. Edited by Breuß M, Bruckstein A. Berlin: Springer; 2013:93–115.
9. Dryden IL, Mardia KV: *Statistical Shape Analysis*. Chichester: John Wiley & Sons; 1998.
10. Davies R, Twining C, Taylor C: *Statistical Models of Shape - Optimisation and Evaluation*. London: Springer; 2008.
11. Heimann T, Meinzer HP: **Statistical shape models for 3D medical image segmentation: A review.** *Med Image Anal* 2009, **13**:543–563.
12. Saroul L, Bernard O, Vray D, Friboulet D: **Prostate segmentation in echographic images: A variational approach using deformable super-ellipse and Rayleigh distribution.** In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro: May 14-17 2008*. Paris: IEEE; 2008:129–132.
13. Mahdavi SS, Chng N, Spadinger I, Morris WJ, Salcudean SE: **Semi-automatic segmentation for prostate interventions.** *Med Image Anal* 2011, **15**(2):226–237.
14. Badiei S, Salcudean SE, Varah J, Morris WJ: **Prostate segmentation in 2D ultrasound images using image warping and ellipse fitting.** In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006, Volume 4191 of Lecture Notes in Computer Science*. Springer Berlin: Heidelberg; 2006:17–24.
15. Saboo R, Levy J, Chaney E, Pizer SM: **Medial models of populations of nearly tubular objects.** In *Proceedings of the MICCAI Workshop on Probabilistic Models for Medical Image Analysis*. London: MICCAI; 2009:232–243.
16. Betrouni N, Dewalle AS, Puech P, Vermandel M, Rousseau J: **Retraction notice to " 3D delineation of prostate, rectum and bladder on MR images.** *Comput Med Imag Grap* 2009, **32**(7):622–630.
17. Lu C, Chelikani S, Papademetris X, Knisely JP, Milosevic MF, Chen Z, Jaffray DA, Staib LH, Duncan JS: **An integrated approach to segmentation and nonrigid registration for application in image-guided pelvic radiotherapy.** *Med Image Anal* 2011, **15**(5):772–785.
18. Pasquier D, Lacornerie T, Vermandel M, Rousseau J, Lartigau E, Betrouni N: **Automatic segmentation of pelvic structures from magnetic resonance images for prostate cancer radiotherapy.** *Int J Radiat Oncol* 2007, **68**(2):592–600.
19. Ahn SJ, Rauh W, Warnecke HJ: **Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola, and parabola.** *Pattern Recogn* 2001, **34**(12):2283–2303.
20. Mulchrone KF, Choudhury KR: **Fitting an ellipse to an arbitrary shape: implications for strain analysis.** *J Struct Geol* 2004, **26**:143–153.
21. Casella G: **An Introduction to empirical bayes data analsysis.** *Am Stat* 1985, **39**(2):83–87.
22. Oguro S, Tokuda J, Elhaway H, Haker S, Kikinis R, Tempany CM, Hata N: **MRI Signal intensity based B-spline nonrigid registration for pre- and intraoperative imaging during prostate brachytherapy.** *J Magn Reson Imaging* 2009, **30**(5):1052–1058.
23. Cootes T, Taylor C: **Statistical models  of appearance for computer vision.** Tech. Rep., Imaging Science and Biomedical Engineering 2004. [http://www.isbe.man.ac.uk/~bim/refs.html].