HHS Public Access

Author manuscript

IEEE Trans Biomed Eng. Author manuscript; available in PMC 2021 February 01.

Published in final edited form as:

IEEE Trans Biomed Eng. 2020 February ; 67(2): 565–576. doi:10.1109/TBME.2019.2918731.

Open Source Platform for Transperineal In-bore MRI-guided Targeted Prostate Biopsy

Christian Herz,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Kyle MacNeil,
Queen's University, Kingston, ON, Canada.

Peter A. Behringer,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Junichi Tokuda,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Alireza Mehrtash,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Parvin Mousavi,
Queen's University, Kingston, ON, Canada.

Ron Kikinis,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Fiona M. Fennessy,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Clare M. Tempany,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Kemal Tuncali,
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Andriy Fedorov
Brigham and Women's, Hospital, Harvard Medical School, Boston, MA, USA

Abstract

Objective: Accurate biopsy sampling of the suspected lesions is critical for the diagnosis and clinical management of prostate cancer (PCa). Transperineal in-bore MRI-guided prostate biopsy (tpMRgBx) is a targeted biopsy technique that was shown to be safe, efficient and accurate. Our goal was to develop an open source software platform to support evaluation, refinement and translation of this biopsy approach.

Methods: We developed *SliceTracker*, a *3D Slicer* extension to support tpMRgBx. We followed modular design of the implementation to enable customization of the interface, and interchange of

correspondence fedorov@bwh.harvard.edu.

image segmentation and registration components to assess their effect on the processing time, precision and accuracy of the biopsy needle placement. The platform and supporting documentation were developed to enable the use of software by an operator with minimal technical training to facilitate translation. Retrospective evaluation studied registration accuracy, effect of the prostate segmentation approach, and re-identification time of biopsy targets. Prospective evaluation focused on the total procedure time and biopsy targeting error (BTE).

Results: Evaluation utilized data from 73 retrospective and 10 prospective tpMRgBx cases. Mean Landmark Registration Error (LRE) for retrospective evaluation was 1.88±2.63 mm and was not sensitive to the approach used for prostate gland segmentation. Prospectively, we observed target re-identification time of 4.60±2.40 min, and BTE of 2.40±0.98 mm.

Conclusion: *SliceTracker* is modular and extensible open source platform for supporting image processing aspects of the tpMRgBx procedure. It has been successfully utilized to support clinical research procedures at our site.

Keywords

image-guided interventions; image registration; magnetic resonance imaging; prostate cancer; software evaluation; targeted biopsy

I. Introduction

Prostate Cancer (PCa) remains one of the most common causes of cancer deaths among men in the USA and worldwide [1]. While the clinical standard for diagnosis of suspected cancer in the prostate continues to be systematic sextant biopsy guided by transrectal ultrasound (TRUS), the inherent limitations of this approach, such as low sensitivity, overdiagnosis of clinically insignificant PCa, and underdiagnosis of clinically significant PCa, motivated development of targeted biopsy techniques that are increasingly used in the clinic [2]. Such targeted techniques utilize multiparametric Magnetic Resonance Imaging (mpMRI) for localizing the suspected cancer sites [3] and rely on various approaches to re-identify those targets in either ultrasound or MRI used for needle guidance [2]. In-bore targeted biopsies are conducted entirely within the MRI. Variations of this technique utilize transrectal (most common), transperineal, or transgluteal biopsy needle insertion. The in-bore technique is not as commonly used as MRI-ultrasound fusion targeted biopsy [4]; however, it has a major advantage of allowing for direct visualization of both the needle and the suspected lesion [5]. The transperineal MRI-guided targeted prostate biopsy (tpMRgBx) approach [6]–[8] offers the additional advantage of reduced infection risk and improved access to the anterior areas of the prostate gland [8]. Although the advantages of targeted biopsy over sextant sampling are well-documented [9], [10], there is limited evidence of which specific approach to targeted biopsy is preferred [2], [11]. The information on how variations of a given targeted biopsy approach (e.g., the choice of the image registration approach) affect its overall performance [12]–[14] is conflicting, prompting further research. Furthermore, commercial platform used for targeted biopsy often operate as “black box” solutions complicating access to target annotations and processing results, and not allowing to fine-tune processing components. These factors provided the primary motivation for the development of a free open source platform for tpMRgBx to streamline implementation and

enable clinical research on refining the prostate biopsy related analysis tools. We believe that such a platform can facilitate the evaluation of the overall efficacy of the technique and its comparison with the alternative approaches.

Clinical research on tpMRgBx was initiated at our institution in 1998 [6], [7]. Since 2010, the procedure has been performed over 700 times in a 3.0T, wide bore (70 cm) Siemens Magnetom Verio [8]. At other institutions, the tpMRgBx approach has also been performed utilizing the commercially available DynaTRIM system (Invivo, Gainesville, FL) [5], [15]. To the best of our knowledge, no publicly available open source platform supporting tpMRgBx is currently available.

In this work we present *SliceTracker* - an open source extension for *3D Slicer* (https://slicer.org) [16]. *SliceTracker* integrates the individual processing components developed earlier (e.g., deformable registration mapping biopsy plan to the intra-procedural data [17], registration of the guidance template [18] and automatic segmentation of the prostate gland [19]) with a user workflow interface and versatile image visualization, providing a complete end-to-end user-oriented open research platform for transperineal MR-guided in-bore prostate biopsy.

II. Materials and Methods

SliceTracker was developed as a versatile platform capable of meeting the requirements of different user groups: procedure support personnel, developers of image processing tools, and clinical researchers utilizing the data for knowledge discovery and gaining new insights into the disease. From the perspective of the procedure support personnel, we provide an easy to learn and use interface that enables robust and powerful visualization and decision support capabilities, with a powerful back-end. The platform was designed to allow streamlined integration by the developers and testing of new tools within the workflow. Finally, to enable clinical research and reuse of the data collected during tpMRgBx procedures, *SliceTracker* captures relevant procedure data and makes it accessible using open and documented representation.

The implementation closely follows the two-phase clinical workflow of the tpMRgBx procedure. In this section, an introduction to the clinical workflow is provided, followed by the analysis of the requirements and our approach to the evaluation of the software.

A. Clinical Procedure

The tpMRgBx procedure is performed in our institution by a team consisting of four to five staff members:

- an interventional radiologist utilizing the information provided by the imaging and post-processing equipment to perform the biopsy;
- an MR technologist operating the MRI scanner;
- a research assistant operating a workstation that provides image processing and visualization functionality to the interventionalist;

- nurses dedicated to supporting the clinical side of the procedure;
- an anesthesiologist joining the team when needed (the procedure is typically conducted with the patient sedated, but not requiring full anesthesia).

The biopsy workflow is summarized in Fig.1. It includes two phases, both of which take place on the day of the biopsy and within the interventional MR suite: the planning (preprocedural) phase, and the intervention (intra-procedural) phase.

The planning phase consists of the review of mpMRI DICOM study acquired before the procedure for localization and grading of suspicious targets. Target(s) are placed on the pre-procedural axial T2-weighted (T2w) image series. Patients can be imaged and referred for biopsy from practices outside of our institution. Therefore, pre-procedural datasets can be heterogeneous in acquisition protocols (as an example, in the dataset used in the evaluation, in-plane resolution and slice thickness ranges were 0.3-0.7 mm and 3.5-4.0 mm, respectively). This phase is not time-critical, since it can be performed before the patient is positioned in the scanner bore.

Interventional phase starts with the positioning of the patient within the scanner using a custom-made MR compatible table top and leg support, and setup of the needle insertion grid template device [18]. Image acquisition starts with a T2w image of the z-frame calibration device [20] rigidly attached to the needle insertion template. Image of the z-frame is then used to register the needle insertion template to the patient space. Next, a T2w image covering the prostate gland (~4 min scan time) is acquired for re-identification of biopsy target(s) localized earlier in the planning dataset. The latter is done by means of deformable registration between the pre- and intra-procedural T2w images [17], which in turn relies on the approximate segmentation of the prostate gland to define the region of interest. Once the images are registered and the accuracy of the registration is visually confirmed by the interventionalist, a registration transformation is applied to propagate the locations of the biopsy targets from the planning (pre-procedural) to the intra-procedural T2w image. Given target locations, template hole and needle insertion depth are automatically calculated, and the interventionalist proceeds with the needle placement and tissue sampling for the individual target locations. Upon needle insertion, another quick T2w scan (~1 min scan time) is performed with the needle in place to evaluate needle position with respect to the target. This verification step is necessary due to the possibility of needle deflection and motion of the prostate gland. If necessary, needle adjustments are made, and the process is repeated until the interventionalist is satisfied with the targeting accuracy and quality of the biopsy sample [8].

B. Software Requirements

1) Functional Requirements—The key requirement for the software is to provide support for the clinical research procedure while accounting for the possible variations in the clinical workflow or imaging data, and accommodating preferences of the interventional radiologist. The main capabilities that should be supported by the software include automatic receipt and volumetric reconstruction of the DICOM data, support of target localization in the pre- and intra-procedural images, automatic assignment of the image type,

registration of the z-frame device, re-identification of the biopsy targets, and flexible visualization modes to support target localization, evaluation of needle placement, and assessment of automated processing steps.

The interface to the processing tools should expose just the minimum required parameters necessary, aiming to simplify the interface for the user. The software should guide the user over the steps of the workflow, prompting for input where automatic updates are not available, or where manual verification is needed. Advanced functionality of the software should be accessible and configurable using customizable workflow components and user settings.

The software must be able to seamlessly integrate existing tools into the workflow and user interface. The platform should be extensible to accommodate integration of new processing tools in the future, as improved and more automated approaches are developed, and new features are added. In practice, failures of automatic processing tools are inevitable. The software should handle failures gracefully, implementing alternative processing approaches. In the situation when all automatic processing tools fail, the software should implement a fallback step allowing for manual remediation with minimum burden to the operator.

Reuse of the data collected during the clinical procedures is critical in research. In medical imaging research, such data can be used to establish performance and its limits for the automatic processing tools (as applied to the biopsy procedure; those can include localization of the targets, segmentation of the prostate gland, deformable registration), identify bottlenecks, and quickly locate problematic datasets. In clinical research, data collected during the biopsy procedures can be used for investigating correlations between pathology samples and the imaging phenotype. It is therefore critical that the software maintains the provenance of the collected image data, as well as the data produced by the computational tools or by the software operator. The resulting data should be captured using structured, machine-readable form to allow for its subsequent analysis and reuse.

2) Non-functional Requirements—Non-functional requirements refer to the specific criteria that the resulting system should satisfy. Here we describe those criteria that need to be followed in the process of the software development, and how they map to the procedure the software intends to support.

Usability: The system operator should not be expected to possess detailed technical knowledge about the workings of the system. It is also important to make the system usable by the contributors of the automated analysis components, who are expected to have the domain expertise, but may not be familiar with the intricacies and requirements of the clinical procedure.

Maintainability: The software is intended for supporting clinical research in tpMRgBx at our site and others. It is therefore important to develop the code with the goal of supporting future contributors of new functionality and bug fixes. To achieve this goal, it is critical to use a code versioning system optimized for collaborative development, follow consistent style and provide documentation. To encourage reuse and contributions, the software needs

to be freely available and distributed under the *3D Slicer* license. Furthermore, dependent software libraries need to be compatible with the *3D Slicer* license.

Robustness: The proposed software is based on the *3D Slicer* platform. *3D Slicer* is neither an FDA-approved nor FDA-cleared medical device and does not meet all the requirements for robustness that would be expected from a commercial product. Under certain conditions, it is possible (if not inevitable) to experience failures of the application. It is essential to develop the application while accounting for the possibility of such failures and implementing features to communicate those failures to the operator, and support recovery of the application state during clinical research procedures (i.e., a failing registration where no output volumes or transforms were produced, or complete crash of the application).

Documentation: Dedicated documentation is necessary to provide introduction to main functionalities of the software, and to train new operators of the software. Documentation should be developed to accommodate prospective users that may or may not have the technical background or understanding of the procedure workflow. Documentation should be accompanied by sample data to support training of the new users.

Portability: Constraints on the operating system required to run the software can limit its adoption. *3D Slicer* can be used on either Windows, Linux, or macOS operating systems. It is important the proposed extension does not introduce any new components that would restrict portability.

Reusability: The proposed software solution is tailored for the specific clinical procedure, and as such it is not expected that the software would be reusable directly for an unrelated procedure. Despite that, the components that emerge during development should be designed to support reuse for similar image-guided interventions (e.g., needle guided cryoablation).

C. Software evaluation

The imaging data used in this evaluation were collected as part of a HIPAA-compliant prospective study that was approved by the IRB at BWH. Signed consent form permitting the use of images for research purposes was obtained from every participant of the study.

The software was extensively evaluated to establish conformance of the implementation to the key functional requirements related to its accuracy and reliability, and to evaluate the performance of the automatic processing components. We performed both retrospective and prospective evaluation. Retrospective evaluation of a system, using data collected during past cases, allows for manually annotating such data and performing quantitative assessment. We also used retrospective evaluation to compare interchangeable components of the platform and inform the choice of the default configuration for its prospective use. The software was further refined and developed based on this evaluation. The prospective part utilized the final version of *SliceTracker* with no modifications introduced in the course of the evaluation, with the data being collected during "live" clinical procedures.

1. Retrospective evaluation

Registration accuracy: We used in-plane Landmark Registration Error (LRE) for quantifying registration accuracy. A biopsy core has a length of about 1 cm and is collected while the needle is being inserted longitudinally. The out-of-plane error has less impact on the sampling accuracy because of this orientation [17]. Between one and three localized anatomical features of the prostate (e.g., entry of the urethra, centroids of calcifications and cysts) were selected as landmarks. LRE was computed for all sets of landmarks by calculating the in-plane Euclidean distance between the corresponding intra-procedural landmarks and the pre-procedural landmarks after applying the registration transformation produced by *SliceTracker*. A specific case was considered suitable for annotation with landmarks if both the pre-procedural and intra-procedural images had anatomical features that could be localized visually by the operator performing image annotation.

Effect of the prostate segmentation approach: The image registration approach that proved effective for biopsy target reidentification relies on the selection of the region of interest corresponding to the prostate. This region of interest can be segmented by the user, or it can be segmented automatically. To inform the choice of the default approach for prostate segmentation, we compared LRE values for the registration results obtained using the two different segmentation approaches.

In addition to measuring LRE for the individual segmentation approaches, we also evaluated variability in the registered target location as a function of the segmentation method. We quantified that measure using the Target Registration Sensitivity (TRS) measure, which was defined as the absolute in-plane Euclidean distance between the locations of targets obtained using the two segmentation approaches. TRS was evaluated in the axial plane.

Biopsy target re-identification time: We define target reidentification time as the duration between receipt of the intra-procedural T2w image of the prostate and the availability of the biopsy targets propagated to the intra-procedural image by means of deformable registration. This time includes receipt and volumetric reconstruction of the DICOM data, prostate segmentation and image registration, visual confirmation of registration accuracy, troubleshooting of the registration if needed, and application of the resulting transformation to the pre-procedural targets.

Overall reliability: In addition to the quantitative measures of performance, we collected data regarding the reliability of the individual components of the system, measured as the number of failures of the automatic processing components (namely, automatic segmentation and deformable registration).

Statistical analysis: Normality distribution testing of a sample was done using Shapiro-Wilk test. Non-parametric comparison of two samples was performed using the Mann-Whitney U-test. Statistical significance level was fixed at 0.05. All statistical testing was done using the scipy.stats (v.1.2.1)Python package. Plots were prepared using matplotlib (v.3.0.2), seaborn (v.0.9.0) and/or bokeh (v.1.0.4) Python packages.

2. Prospective evaluation—Our prospective evaluation utilized the data collected during consecutive clinical cases supported with the finalized implementation of the software.

Procedure time: The total procedure time was measured as the time between receiving the T2w image of the z-frame calibration device and the time when the operator explicitly indicated the completion of the procedure.

Biopsy target re-identification time: This measure was defined in the same fashion as for the retrospective evaluation.

Targeting accuracy: During the procedure, needle confirmation images are obtained after each needle insertion to assess accuracy of placement. Locations of biopsy targets are transformed automatically by means of deformable registration. We used the orthogonal distance from the needle trajectory to the planned biopsy target location as a measure of Biopsy Targeting Error (BTE). This measure provides complementary information to LRE. While LRE allows one to retrospectively quantify the error in the locations of the image that contain recognizable image landmarks, BTE quantifies how close is the needle to the target. The needle trajectory was computed by using the *3D Slicer* extension *DeepInfer* [19] and the methodology presented earlier for automatic segmentations of the needle [21] in every needle confirmation image. For each target, we chose the shortest distance to the target and needle trajectory obtained over the course of the procedure. It is expected that over the course of the procedure the interventionalist is refining the needle position to adjust for needle deflection and prostate motion, iteratively approaching the target closer.

Overall reliability: Data characterizing the reliability of the individual components of the workflow utilized in the prospective evaluation were captured, as was done for the retrospective evaluation.

III. Results

A. Implementation

The *SliceTracker* platform is available to the user as an extension for the *3D Slicer* application. As summarized in Fig. 2, *SliceTracker* leverages the infrastructure provided by *3D Slicer* to access and utilize the capabilities of the existing external libraries, such as DCMTK [22], ITK [23], and VTK. It also leverages relevant extensions and modules of *3D Slicer*, such as *SegmentEditor* (for semi-automatic image segmentation) and BRAINSFit [24] (for deformable image registration). The individual steps of the workflow integrate previously developed computational tools available as *3D Slicer* plugins and extensions, streamlined for use in clinical research applications. Source code of *SliceTracker*, accompanied by the documentation, is available at https://github.com/SlicerProstate/SliceTracker.

Functional requirements—*SliceTracker* was designed to follow the two-phase workflow of the clinical procedure as summarized in Fig. 1, and to meet the functional requirements outlined earlier. To initiate the workflow for a new clinical case, the operator creates a new

"case" within the software. All of the information pertinent to the processing is automatically stored to the file system for the active case, until the operator indicates that the case is completed. In the planning phase, individual mpMRI series are automatically parsed from the files corresponding to the DICOM multiparametric MRI study, volume-reconstructed, and opened for review. Following the PI-RADS guidelines [19], T2w, DCE and DWI series are automatically displayed and linked to allow for correlation of these modalities and identification of suspicious areas. Biopsy targets are marked by the interventionalist on the axial T2w image as point fiducials.

Automatic segmentation of the prostate gland is performed using the deep learning approach integrated within *DeepInfer*. Segmentation is triggered automatically upon completion of target placement, with the result being displayed as an overlay in the axial view. The user has the option to perform segmentation manually, should the automatic result prove unsatisfactory. All the aforementioned steps can be completed prior to the start of the procedure. If the pre-procedural data is not available, the software can proceed directly to the biopsy phase, allowing for target placement on the intra-procedural T2w images.

During the intra-procedural phase *SliceTracker* is configured to monitor the imaging data that is sent from the MRI console to the research workstation. The type of the incoming image (i.e., z-frame, T2w planning scan of the prostate, or needle confirmation image) is automatically recognized from the content of the DICOM SeriesDescription tag using configurable rules. Receipt of an image of a known type automatically triggers the corresponding step in the processing workflow. Steps that involve automatic processing are always followed by a verification step, allowing the operator to remedy an unsatisfactory result, if necessary.

Upon receipt of the z-frame image, the *SliceTracker* operator is prompted to identify the bounding box region of interest (ROI) containing the z-frame artifacts. Once the ROI is identified, segmentation of the z-frame is done automatically, followed by automatic registration of the template to the MRI coordinate system. Segmentation of the prostate in the initial intra-procedural T2w image is performed automatically using the same segmentation approach as in the pre-procedural step, with the segmentation mask dilated by 5 mm to include the area outside the prostate capsule (this proved effective to make registration more robust). Prostate segmentation regions are used to automatically calculate the initial transformation for the intensity-based deformable registration. The deformable registration approach is based on the mutual information similarity metric and hierarchical transformation model implemented using ITK filters within the BRAINSFit extension of *3D Slicer*, as described earlier [17].

Upon completion of the deformable registration, *SliceTracker* switches to the registration evaluation mode that provides interactive tools for inspection of the result at any level of the prostate gland (e.g., checkerboard and split views of the registered images, cross-fade between the intra-procedural and registered planning images). The interventional radiologist visually inspects the quality of the registration and the locations of the transformed biopsy targets, and once they confirm the result to be satisfactory, the transformation produced by the registration is applied to the targets in the planning image.

Subsequently, re-registration of the targets to the images collected for confirmation of needle positioning (needle confirmation images) is applied automatically. This is done to account for the possible motion of the patient and displacement and deformation of the prostate gland, as discussed in [25], [26]. Needle confirmation image registration is done fully automatically first, without the need to do any segmentation. User interaction is required only if the automatic attempt fails.

In the situations where deformable registration fails, the operator has various options to rectify the possible reasons for automatic registration failure. The operator has the option to resegment the gland using semi-automatic tools by placing points on the boundary of the gland, with the overall segmentation interpolated from those points. Improved accuracy of the segmentation can lead to improved registration initialization. The operator also has a choice to examine and accept the affine or rigid registration result, if only the deformable registration stage failed. If all attempts fail, the user can manually modify the target location directly on the intra-procedural images. Fig. 3 displays an illustration of the user interface (UI) in the registration evaluation step.

Usability and Reusability—*SliceTracker* integrates the various components of the workflow under a unified interface. Mundane operations (e.g., parsing and loading of the DICOM image data or re-configuration of the image visualization layouts) that can take significant time from the operator and can lead to mistakes are performed automatically. The parameters for the individual processing steps are available in the advanced configuration settings and in the source code but are not exposed in the user interface to simplify the interaction.

To support reuse of the data collected during the biopsy procedures, *SliceTracker* automatically stores the DICOM images, pre-processed data (targets, segmentations, volumes), and registration results in dedicated directories created with initialization of a new biopsy case. All data entities are stored using open formats. JavaScript Object Notation (JSON) file format in conjunction with JSON-Schema [27] for keeping track of persistent data location and including qualitative details pertaining to:

- procedure event times: start, stop, resume, and completion;
- series information: receipt times and types of series (cover template/cover prostate/guidance);
- segmentation information: start and completion time, approach (manual or automatic), modification start and completion times (if applicable);
- registration results: start and completion, approval status, person responsible for the approval of the result (clinician or operator).

To make the code more usable by software developers, we implemented *SliceTracker* following a plugin architecture. Individual tools such as image segmentation, registration, and management of the case data are implemented as plugins that can be replaced or modified. Processing, visualization and annotation of the biopsy targets are implemented in

a separate extension, *mpReview*[1], which is integrated within *SliceTracker*, but can also be used independently, maximizing its applicability.

SlicerDevelopmentToolbox (SDT)[2] is another extension developed as part of *SliceTracker* implementation, providing a toolbox to extension developers that aim to maintain reusable components of the code. SDT provides a Python API offering mixins, widgets, helpers, constants, decorators, and other useful classes. The main goal of SDT is to assist developers in creating new extensions more quickly by providing a collection of frequently used components. Besides this, it provides assistive decorators helpful during debugging.

Maintainability and Portability—*SliceTracker* is implemented primarily in Python, with some of the components, such as z-frame registration, implemented in C++. *SliceTracker* and all of the components it utilizes are available under an open source license that does not restrict commercial or academic use. We use GitHub for maintaining the development process (revision history, integration of external contributions, and issue tracking). Being a *3D Slicer* extension, we are leveraging the standard capabilities provided by the *3D Slicer* platform to generate packaged extension for Windows, Linux and macOS operating systems.

Robustness—The software implements several levels of fallbacks to improve robustness in case individual components fail. To support overall fault-tolerance, the state of the workflow is automatically checkpointed to the file system at the completion of the individual processing steps. If the software fails completely during the procedure, the state of the case can be restored from the checkpoint. Evaluation of *SliceTracker* using the data collected in clinical procedures, as presented in the following section, establishes specific evidence about the robustness of the software.

Documentation and training—*SliceTracker* is accompanied by a user guide detailing each of the workflow steps with accompanying figures[3], and a de-identified sample dataset. Implementation includes “Training mode”, which allows to mimic receipt of the data collected in any of the past cases, allowing the new operators to be trained on real data.

Development timeline—Prototyping and requirements analysis started in February 2015. The initial feature-complete version was successfully integrated into the clinical research workflow in July 2016. Following the initial evaluation period, the source code underwent a major refactoring process to modularize its architecture and simplify implementation of new features. Software workflow steps outlined earlier were organized in separate Python classes implemented as event-based plugins. *SliceTracker* v2 was the result of the refactoring process and led to new features being integrated, including target displacement charts for tracking intra-procedural prostate gland motion and automatic segmentation of the prostate gland in pre- and intra-procedural images via the *3D Slicer* extension *DeepInfer* [16]. *SliceTracker* successfully met the requirements described in the Requirements section with the introduction of version 2.

[1] https://github.com/SlicerProstate/mpReview
[2] https://github.com/QIICR/SlicerDevelopmentToolbox
[3] https://slicerprostate.gitbooks.io/slicetracker/

B. Retrospective evaluation

Registration accuracy—To evaluate accuracy of the deformable registration used for target re-identification, 73 retrospective cases were selected and pre- and intra-procedural images from retrospective cases were manually annotated with landmarks (1-3 per case, for a total of 168 landmarks). Mean LRE was 1.88±2.63 (range 0.05-12.68) mm. In 60 cases an endorectal coil (ERC) was used during pre-procedural image acquisition. Mean LRE for cases with ERC was higher ($p<0.01$) with 2.17±2.83 (0.05-12.68) mm, when compared to 0.59±0.39 (0.09-2.12) mm for cases with no ERC present. The number of biopsy targets in this cohort was 1.7±0.8 (1-4) per case, and the number of needle confirmation images per case was 7.3±4.0 (1-21).

Effect of segmentation approach—Automatic segmentation was retrospectively applied to 44 of the 73 cases selected initially. Cases were only selected if all of the following conditions were satisfied: the image was from the same domain as the dataset used for training the deep learning model (ERC was used in the pre-procedural scan, and pixel spacing was in the range of 0.27-0.275 mm), and the case was not used for training and/or validation of the model. One case was excluded from this subset due to failure of the automatic segmentation, resulting in a total of 43 cases used for evaluation. LRE was computed for all landmarks (n=96) using both the manual and automatic segmentations (see per-case averaged LRE in Fig. 4.). Distributions of the LRE were strongly skewed towards 0 and non-normal based on visual checks, and as confirmed by the Shapiro-Wilk test ($p<0.0005$). Therefore, statistical significance testing for comparing LRE samples was done using the Mann-Whitney U-test. No significant difference ($p>0.05$) was observed between the mean LRE for the two approaches compared (1.62±2.12 (range 0.10-13.28) mm for automatic segmentation vs 1.76±2.15 (0.05-8.83) mm for manual segmentation). Deformable B-spline registration failed to produce a result in 5 cases where the manual approach was used, and in 3 cases (different from the aforementioned 5) where the automatic approach was used. Affine registration was applied for these cases in place of the B-spline approach.

TRS was calculated for all targets (n=70). Mean TRS was 2.51±2.19 (range 0.17-8.94) mm. TRS averaged over the targets in each of the cases is summarized in Fig. 4 alongside average LRE for each respective case. Deformable B-spline registration failed in one case. As discussed above, affine registration was applied for this case in place of the B-spline registration.

Target re-identification time—A total of 56 retrospective cases supported with *SliceTracker* were analyzed to calculate the target re-identification time. The subset of 56 cases that had procedure timestamps were selected from the 73 used in the registration accuracy evaluation. The mean time to provide target locations to the operator was 4.98±5.00 (1.17-25.17) minutes.

C. Prospective evaluation

A total of 10 consecutive cases were used to collect the data for prospective evaluation. In 3 cases ERC was used during the pre-procedural image acquisition. The number of biopsy

targets per case was 1.2±0.4 (1-2), and the number of needle confirmation images was 7.3±1.9 (4-10).

Intra-procedurally, a total of 84 images were taken (11 cover prostate, since on one occasion the initial intra-procedural scan had to be re-taken, and 73 needle confirmation images). A total of 91 registration results were produced. In 7 of 91 of registrations, the result was deemed unsatisfactory and registration was retried. In only 1 of the 81 image pairs that were registered did the situation arise where none of the fallback approaches for automated registration produce a satisfactory result for the operator, and the registration was completed manually. The comparison of times for different types of images, with and without retries, is shown in Fig.5.

Automatic segmentation in the biopsy planning images was modified by the operator in 7 out of 10 cases. Automatic segmentation of the intra-procedural images was more consistent, with only one of the 16 intra-procedural automatic segmentations being modified by the operator.

Target re-identification time was 4.60±2.40 (2.21-8.31) min. Mean procedure length from receiving the z-frame template image to completion of the case was 53.06±9.08 (22.24-82.59) min. BTE was 2.4±0.98 (1.4-4.4) mm.

Importantly, computation contributed only a small fraction of the total procedure time. This can be seen from Fig.6, which provides a timeline view for each of the 10 prospective cases, summarizing various relevant events and their duration from the receipt of the template image to the completion of the case by the operator.

IV. Discussion

It was our goal to investigate and develop an extensible open source platform for tpMRgBx research. The result of this work is *SliceTracker*, an open source extension to the *3D Slicer* application, and can be used on Windows, macOS, and Linux platforms. Based on our evaluation and experience, the software provides satisfactory registration in most of the evaluated cases, includes support of both automatic and manual prostate segmentation approaches, and helps streamline uniform collection of the research data, which is critical for supporting secondary analysis studies.

Since its initial release in September 2016, *SliceTracker* has been used routinely to support biopsy procedures at our institution on a weekly basis. Close to 300 tpMRgBx procedures have been supported by *SliceTracker* as of March 2019, with over 5 operators trained to use the software. The software has also been reused to customize a similar interventional workflow to support prostate cryoablation research procedures[4], providing extra evidence of reusability.

While *SliceTracker* proved useful in practice, it was important to evaluate its individual components and quantify its performance, where possible. The average LRE of 1.88 mm we

[4]https://github.com/SlicerProstate/SlicerProstateAblation

observed is smaller than the image slice thickness of 3 mm used for our intra-procedural MRI data. Higher mean LRE was reported for cases that used ERC for pre-procedural image acquisition. This trend may be explained by stronger deformation of the prostate gland, when compared to imaging without ERC. We note that we are unable to quantify error across the whole gland. As such, we recognize that a low LRE does not necessarily measure registration accuracy at the biopsy targets, as registration could be poor in the regions of the prostate where landmarks were not selected. It will be interesting and appropriate to study more thoroughly the uncertainty of target registration, and its relation to the uncertainty of landmark registration. Such investigation was deemed out of scope of this manuscript. At the same time, we note that each of the registration results accepted during the course of the procedure was deemed satisfactory by the interventional radiologist with extensive expertise in the procedure (Dr. Kemal Tuncali, 15 years of experience in targeted in-bore MRI-guided prostate biopsy). The observed LRE is also comparable with the technical registration accuracy of 1.9 mm that is required for correctly assessing high-grade focal spots of PCa tumors, and is below 3.1 mm TRE required for detecting 95% of the 0.5 cc tumors, based on simulations by van de Ven et al. [28].

Our comparison of the segmentation approaches did not reveal statistically significant difference in LRE as a result of the use of automatic segmentation. We also observed that automatic segmentation does not perform well on the data that has different resolution as compared to that used in training of the model. The majority of data for training of the model came from within our institution, which may contribute to the sensitivity of the model to outside data. Considering the heterogeneity of pre-procedural data collected at different sites, this justifies support of the manual segmentation as a fallback in case automatic segmentation fails.

While we are unable to quantify registration errors at the biopsy target locations, TRS values can serve as an indirect measure for the sensitivity of the registration in the clinically important areas of the image. In several of the targets, sensitivity values exceed 2 mm. We do not know which of the segmentation approaches leads to more accurate registration at the target locations, since error could be evaluated at the locations of anatomical landmarks (cysts or calcifications) which typically do not coincide with the locations of biopsy targets. We also do not know the effect of the TRS on the accuracy of the biopsy sample collection. However, this result may indicate that the effect of the segmentation approach used may be important and should motivate further studies investigating registration sensitivity both in the tpMRgBx approach, and in the commercial tools implementing fusion biopsy support.

To the best of our knowledge, no similar investigation evaluating registration sensitivity and landmark registration accuracy in the commercial tools has been published. Furthermore, many of the commercially available systems restrict access to the data necessary for the kind of analysis we performed. The open source nature of *SliceTracker* enables investigations of such aspects of the processing pipeline and can enable incremental development to support more accurate biopsy targeting. As an example, refinement of the workflow with the new processing tools, such as automatic segmentation of the biopsy needle [21], becomes possible and is planned to be integrated in *SliceTracker* in the future. Another direction for the future work will investigate the possibility of comparing the performance of the

commercially available MRI/ultrasound fusion guided biopsy systems with the tpMRgBx approach. We hope that availability of *SliceTracker* will make it more feasible to conduct such comparisons at other research sites.

Mean procedure time we observed in the prospective evaluation was 53 min, compared to 67 min reported by Penzkofer et al. [8], where individual processing steps had lower degree of automation and were available as in-house standalone tools not integrated into a workflow. Although the patient populations used in the analysis by Penzkofer et al. and in our study are not the same, our experience shows that availability of all of the processing steps in a single workflow significantly reduces processing time and possibility of making mistake by the operator. One should be careful, however, drawing any conclusions about the procedure time being reduced by improved computational tools. Detailed *SliceTracker* logs that were summarized in Fig.6 highlight that the contribution of computational steps to the procedure time is very small. The exact sources of delays in acceptance of the registration result were not tracked, since those are related to the human component of the procedure. We can speculate, however, for the possible reasons of those delays. First, we required explicit verbal confirmation of registration result being satisfactory by the clinical personnel. In some instances, clinical personnel could make the decision about the next step in the procedure without consulting with the research software. Due to the logistics of the procedure, verbal confirmation could not be possible to obtain immediately in all cases. Second, BWH is a teaching hospital, and targeted biopsy procedures are often supported by a clinical resident assisting the interventionalist and learning about the procedure. The procedure time includes discussions between the trainee and the senior interventionalist. Finally, since the in-bore targeted biopsy procedure has now been performed for over 5 years (over 700 procedures conducted with the 3T magnet), it become routine at our site, and is often combined with additional research studies (e.g., evaluating new imaging sequences or new devices), which also require extra time.

Overall, we believe reduction of the procedure time is more likely to be gained from the optimization or revision of the overall workflow, or development of tools that enable accurate target sampling by reducing the number of needle insertion attempts. At this time, further optimization aimed at reducing processing time for the computational steps is not warranted.

Evaluation of the specific sources of biopsy targeting error was out of scope of our study. Some of those sources include needle deflection, as investigated by Moreira et al. [29], and motion of the prostate gland (see preliminary analysis in [25]). We also did not attempt to compare either the workflow or duration of the individual steps implemented in *SliceTracker* with the commercial counterparts. Such comparison could be valuable to consider in the future.

We emphasize that *SliceTracker*, similar to *3D Slicer*, is a research software and is not an FDA-approved medical device. It is designed for the use by trained individuals as part of clinical research studies, in compliance with the appropriate regulations and safeguards. It is not intended to replace commercial biopsy systems.

Commercial solutions that are approved medical devices can be used for targeted biopsy procedures as part of routine clinical care, and without the need to establish research study protocols or acquire patient consent. Research studies utilizing those systems can be conducted with relatively minimal involvement of technical personnel. A disadvantage of using a commercial system is the limited ability to learn about the implementation of the internal components, modify the processing workflows to evaluate or compare the utility of alternative processing tools, or to adapt or reuse individual components for a different procedure. Those limitations are expected from commercial systems. However, another set of limitations stems from the lack of adoption of interoperable formats for communicating detailed information about target locations and segmentation. Although it is possible to utilize the DICOM standard to communicate, for example, segmentations and measurements, as was demonstrated in particular for prostate imaging [30], adoption of the standard in commercial tools is currently lacking. Such information may be absolutely critical in comparison of alternative solutions, or for assessing the targeting accuracy of the system or performance of individual components. In our experience, access to such information may require establishing special research agreements with the manufacturer, or may not be possible at all.

The situation is reversed for the open source solutions, which provide ultimate configuration flexibility, unconstrained access to the results of image annotation and processing, and intermediate analysis results. However, due to the research nature of the solution, their use can only be possible as part of an institutionally-approved research study. Furthermore, in the present manuscript we describe only the software component of the solution, which also includes various assistive hardware devices, such as the z-frame, needle insertion template and leg holder, as presented earlier by Tokuda et al. [18], which would need to be procured by the adopting site. Naturally, the use of research components necessitates specialized training and experience. Adoption of *3D Slicer* for clinical research will require involvement of research software engineers, and different expectations for stability and level of support as compared to the commercial products.

The choice of the optimal platform for targeted prostate biopsy research will depend on the specific research interests and resources availability at the given site. In our view, our work fills an important gap, since there are numerous commercial tools available for targeted prostate biopsy (mostly utilizing MR-US fusion approach), but none available as open source.

V. Conclusion

In conclusion, we presented an open source software application for supporting transperineal in-bore MRI-guided prostate biopsies, which can be readily extended and modified based on future needs, and may facilitate translation of the tpMRgBx procedure to other sites. No engineering expertise is required for the operator of the system. We have evaluated the application of the system in a clinical setting with a retrospective and prospective study of cases that were supported with the software. To the best of our knowledge, *SliceTracker* is the only end-to-end solution available as free open source software for supporting transperineal in-bore MRI-guided prostate biopsy.

Acknowledgment

We would like to thank the clinical staff supporting the clinical research biopsy procedure, and especially Janice Fairhurst, Angela Kanan, Marina Taziyakova, Carol Callahan, Sean Jackson and Nancy Trane for their support, patience and encouragements through the course of this study.

This work was supported in part by the U.S. National Institutes of Health grants P41 EB015898, U24 CA180918 and R01 CA111288.

References

[1]. Siegel RL, Miller KD, and Jemal A, "Cancer statistics," CA Cancer J Clin, vol. 66, no. 1, pp. 7–30, 2016. [PubMed: 26742998]

[2]. Verma S et al., "The Current State of MR Imaging-targeted Biopsy Techniques for Detection of Prostate Cancer.," Radiology, vol. 285, no. 2, pp. 343–356, 2017. [PubMed: 29045233]

[3]. Hegde JV et al., "Multiparametric MRI of prostate cancer: An update on state-of-the-art techniques and their performance in detecting and localizing prostate cancer," J Magn Reson Imaging, vol. 37, no. 5, pp. 1035–1054, 2013. [PubMed: 23606141]

[4]. Marks L, Young S, and Natarajan S, "MRI-ultrasound fusion for guidance of targeted prostate biopsy.," Curr. Opin. Urol, vol. 23, no. 1, pp. 43–50, 1 2013. [PubMed: 23138468]

[5]. Trilisky I, Oto A, and Eggener S, "In-Bore Magnetic Resonance Imaging-Targeted Prostate Biopsy," in Imaging and Focal Therapy of Early Prostate Cancer, 2017, pp. 205–218.

[6]. V D'Amico A et al., "Transperineal magnetic resonance image guided prostate biopsy.," J. Urol, vol. 164, no. 2, pp. 385–387, 8 2000. [PubMed: 10893591]

[7]. Hata N et al., "MR imaging-guided prostate biopsy with surgical navigation software: device validation and feasibility.," Radiology, vol. 220, no. 1, pp. 263–8, 7 2001. [PubMed: 11426008]

[8]. Penzkofer T et al., "Transperineal in-bore 3-T MR imaging-guided prostate biopsy: a prospective clinical observational study.," Radiology, vol. 274, no. 1, pp. 170–80, 2015. [PubMed: 25222067]

[9]. Kasivisvanathan V et al., "MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis," N. Engl. J. Med, p. NEJMoa1801993, 2018.

[10]. Vanden Berg RNW, McClure TD, and Margolis DJA, "A Review of Prostate Biopsy Techniques," Semin. Roentgenol, 4 2018.

[11]. Wegelin O et al., "Comparing Three Different Techniques for Magnetic Resonance Imaging-targeted Prostate Biopsies: A Systematic Review of In-bore versus Magnetic Resonance Imaging-transrectal Ultrasound fusion versus Cognitive Registration. Is There a Preferred Technique?," Eur. Urol, vol. 71, no. 4, pp. 517–531, 2017. [PubMed: 27568655]

[12]. Delongchamps NB et al., "Prebiopsy magnetic resonance imaging and prostate cancer detection: comparison of random and targeted biopsies.," J. Urol, vol. 189, no. 2, pp. 493–9, 2013. [PubMed: 22982424]

[13]. Puech P et al., "Prostate cancer diagnosis: multiparametric MR-targeted biopsy with cognitive and transrectal US-MR fusion guidance versus systematic biopsy--prospective multicenter study.," Radiology, vol. 268, no. 2, pp. 461–9, 8 2013. [PubMed: 23579051]

[14]. Monda SM et al., "Cognitive Versus Software Fusion for MRI Targeted Biopsy: Experience Before and After Implementation of Fusion," Urology, 6 2018.

[15]. Wolter K, Decker G, and Willinek W, "Transperineal MR-Guided Stereotactic Prostate Biopsy Utilizing a Commercially Available Anorectal Biopsy Device," RöFo - Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgeb. Verfahren, vol. 185, no. 02, pp. 116–120, 1 2013.

[16]. Fedorov A et al., "3D Slicer as an image computing platform for the Quantitative Imaging Network.," Magn. Reson. Imaging, vol. 30, no. 9, pp. 1323–1341, 7 2012. [PubMed: 22770690]

[17]. Fedorov A et al., "Image registration for targeted MRI-guided transperineal prostate biopsy," J. Magn. Reson. Imaging, vol. 36, no. 4, pp. 987–992, 2012. [PubMed: 22645031]

[18]. Tokuda J et al., "In-bore setup and software for 3T MRI-guided transperineal prostate biopsy.," Phys. Med. Biol, vol. 57, no. 18, pp. 5823–40, 9 2012. [PubMed: 22951350]

[19]. Mehrtash A et al., “DeepInfer: Open-source deep learning deployment toolkit for image-guided therapy,” in Progress in Biomedical Optics and Imaging - Proceedings of SPIE, 2017, vol. 10135.

[20]. DiMaio SP et al., “Dynamic MRI scan plane control for passive tracking of instruments and devices.,” Med. Image Comput. Comput. Assist. Interv, vol. 10, no. Pt 2, pp. 50–8, 1 2007. [PubMed: 18044552]

[21]. Mehrtash A et al., “Automatic Needle Segmentation and Localization in MRI with 3D Convolutional Neural Networks: Application to MRI-targeted Prostate Biopsy,” IEEE Trans. Med. Imaging, vol. PP, no. X, pp. 1–1, 2018.

[22]. Eichelberg M, Riesmeier J, Wilkens T, Hewett AJ, Barth A, and Jensch P, “Ten years of medical imaging standardization and prototypical implementation: the DICOM standard and the OFFIS DICOM toolkit (DCMTK),” in Medical Imaging 2004, 2004, no. 4 2004, p. 57.

[23]. Ibanez L and Schroeder WJ, The ITK Software Guide 2.4. Kitware, Inc., 2005.

[24]. Johnson HJ, Harris G, and Williams K, “BRAINSFit: Mutual Information Registrations of Whole-Brain 3D Images, Using the Insight Toolkit,” Insight J, no. Jul-Dec, 2007.

[25]. Fedorov A et al., “Quantification of intra-procedural gland motion during transperineal MRI-guided prostate biopsy,” in Proc. of ISMRM’13, 2013, vol. 36, no. 4.

[26]. Behringer PA, Herz C, Penzkofer T, Tuncali K, Tempany CM, and Fedorov A, Open-source platform for prostate motion tracking during in-bore targeted MRI-guided biopsy, vol. 9401 2016.

[27]. McIntyre J, Using JSON Schema. 2004.

[28]. van de Ven WJM, Hulsbergen-van de Kaa C. a, Hambrock T, Barentsz JO, and Huisman HJ, “Simulated required accuracy of image registration tools for targeting high-grade cancer components with prostate biopsies.,” Eur. Radiol, 11 2012.

[29]. Moreira P, Patel N, Wartenberg M, Li G, Tuncali K, and Heffter T, “Evaluation of robot-assisted MRI-guided prostate biopsy : needle path analysis during clinical trials Evaluation of robot-assisted MRI-guided prostate biopsy : needle path analysis during clinical trials,” 2018.

[30]. Fedorov A et al., “An annotated test-retest collection of prostate multiparametric MRI,” Sci. Data, vol. 5, p. 180281, 12 2018. [PubMed: 30512014]

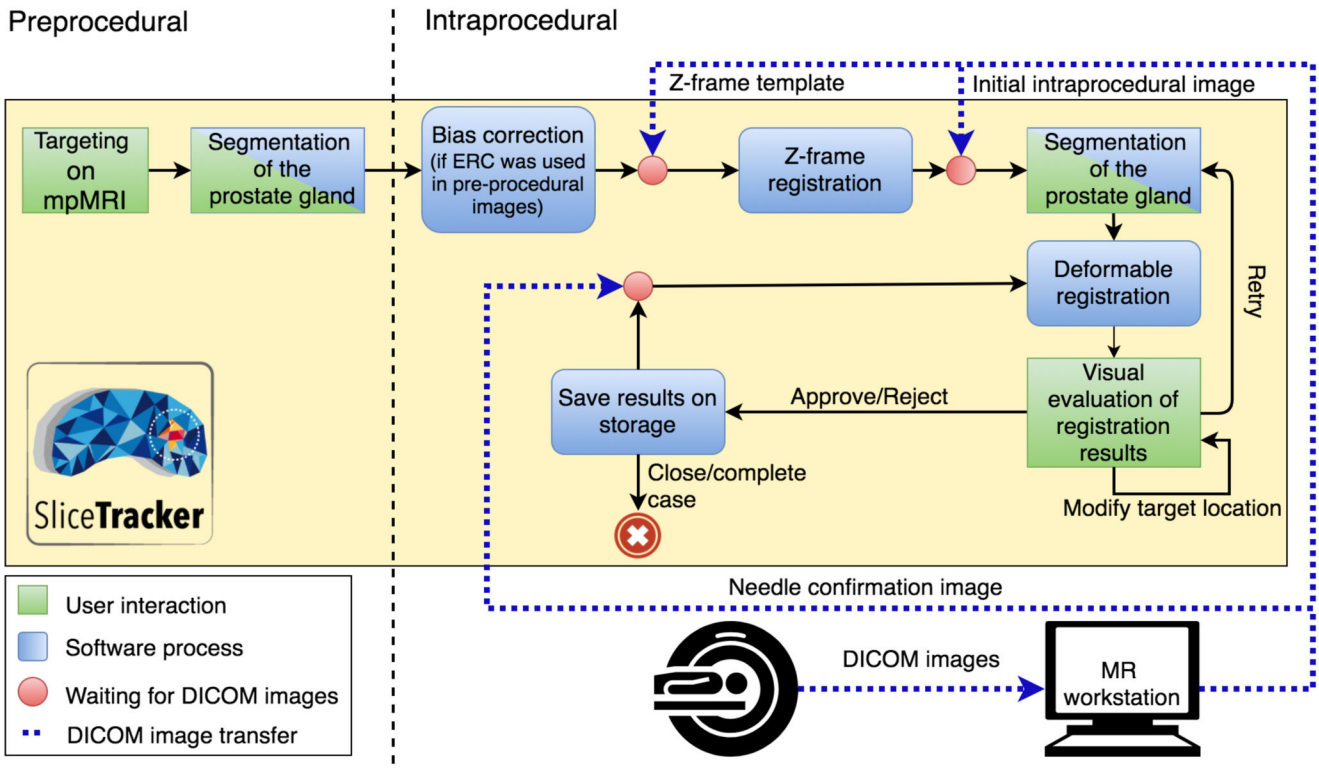

Figure 1.

Procedure workflow implemented in SliceTracker. Green boxes represent steps that require user interaction. Blue boxes correspond to the software processes that are performed automatically. Segmentation of the prostate gland can be done either automatically or under user guidance.

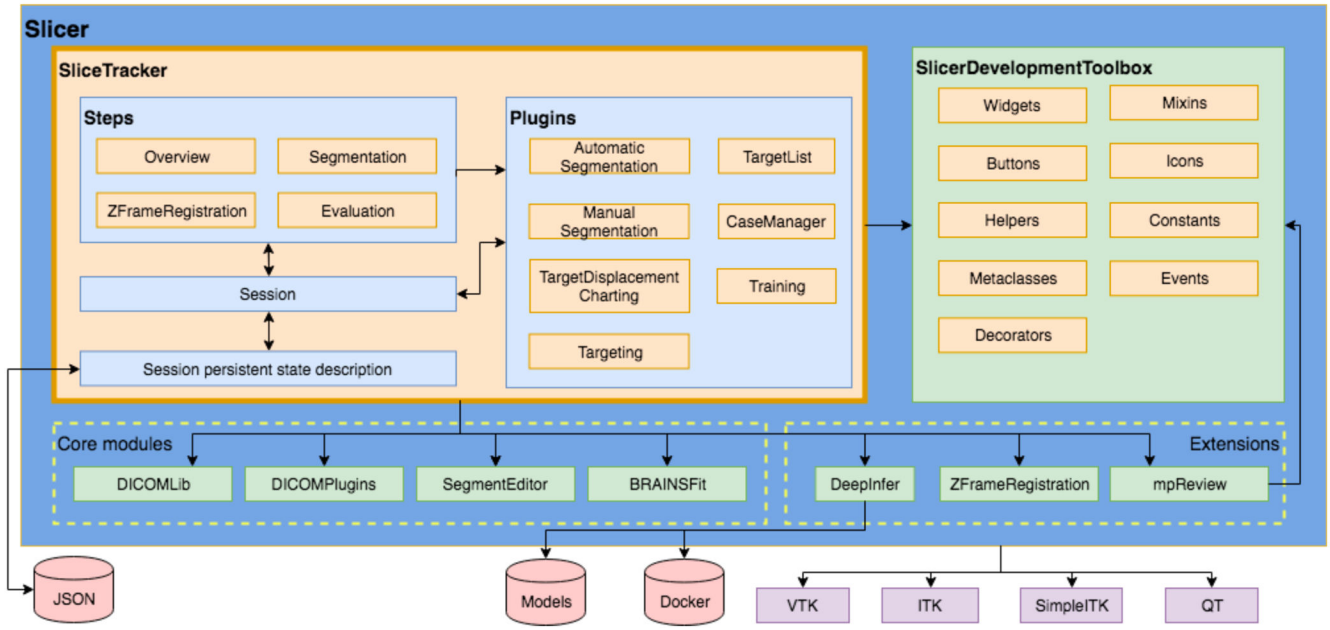

Figure 2.

Overall architecture of SliceTracker and its relationship with the supporting components of the 3D Slicer platform. SliceTracker core dependency is SlicerDevelopmentToolbox, introduced as part of SliceTracker development, implementing reusable components of general utility. A number of 3D Slicer modules and extensions were developed independently from SliceTracker (e.g., DeepInfer and SegmentEditor) but provide key functionality to the processing workflow.

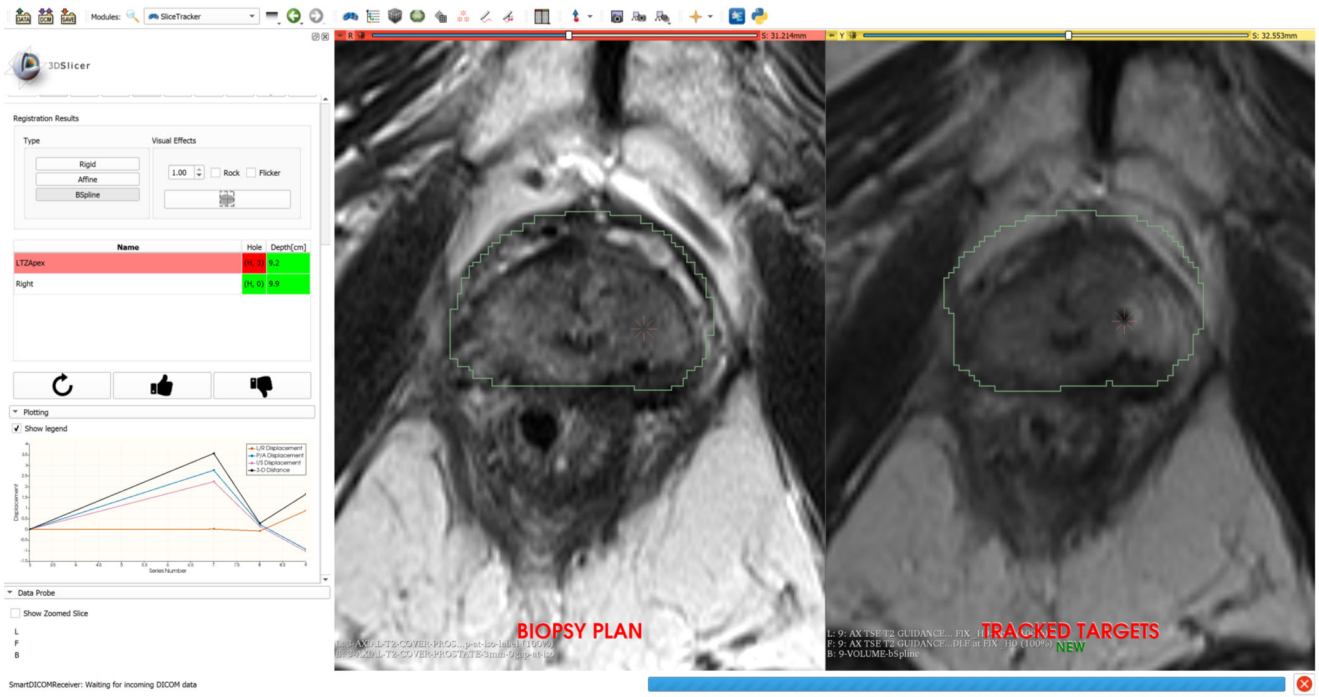

Figure 3.

SliceTracker interface during visualization of the tracked target location. On the left, the user control panel shows the list of targets, template insertion hole and depth, and registration type. The plots summarize the motion of the target tissue through the course of the procedure. On the right, planned position of the target is shown side by side with the result of propagating the target location by means of registration to the needle confirmation image. Correlation of the needle artifact and target location assist in evaluating targeting accuracy by the operator.

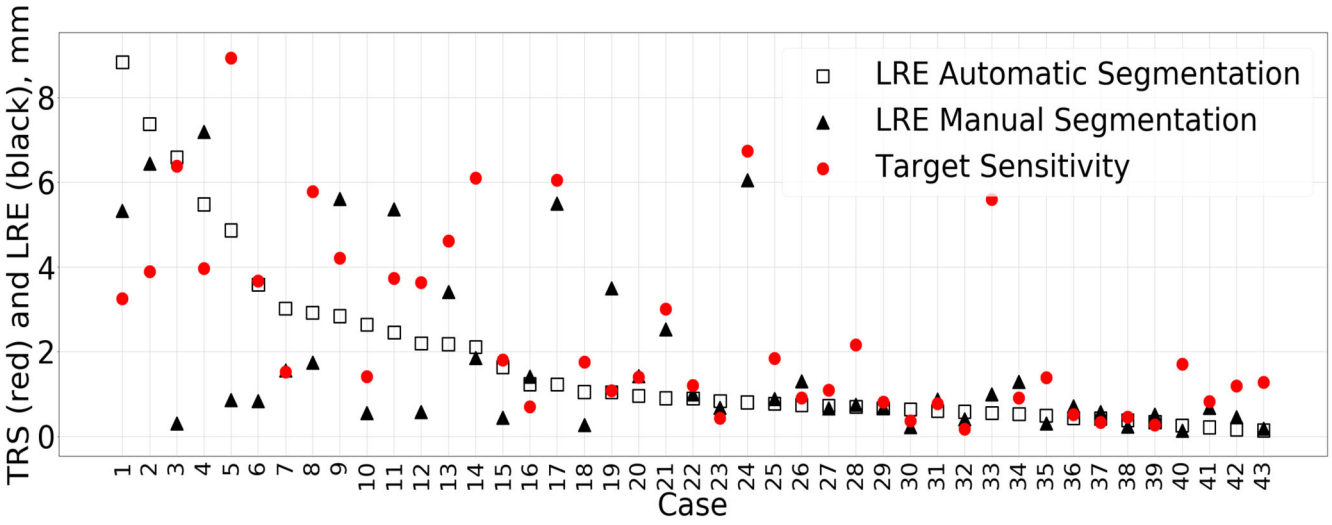

Figure 4.

Summary of the Landmark Registration Error (LRE) and Target Registration Sensitivity (TRS) for the data used in the evaluation. Cases are ordered by the LRE Automatic Segmentation values to facilitate visual assessment of the error distribution.

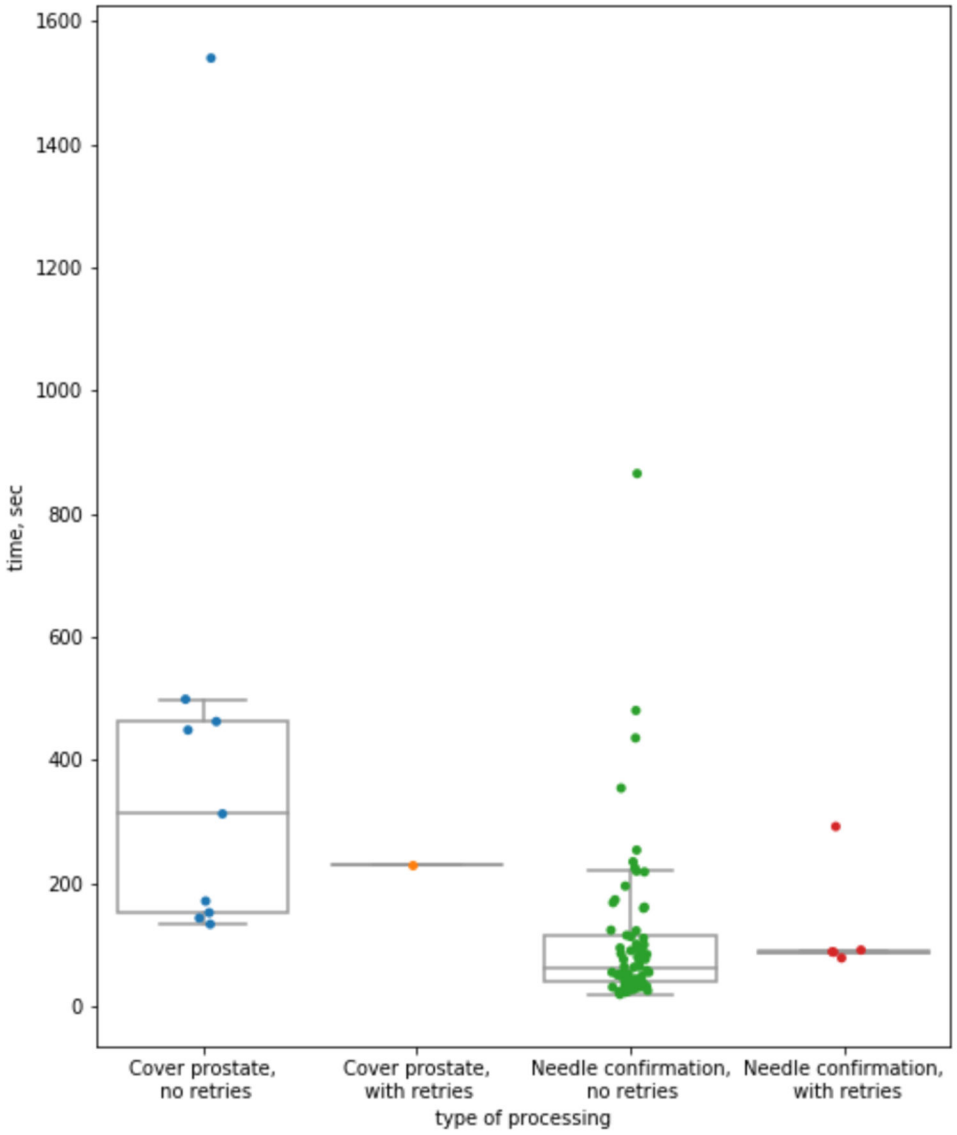

Figure 5.

Boxplot summary of the processing time for different image types for the prospective evaluation component. Time was measured as the difference in automatically recorded timestamps between the acceptance of the registration result by the operator and the receipt of the image by SliceTracker.

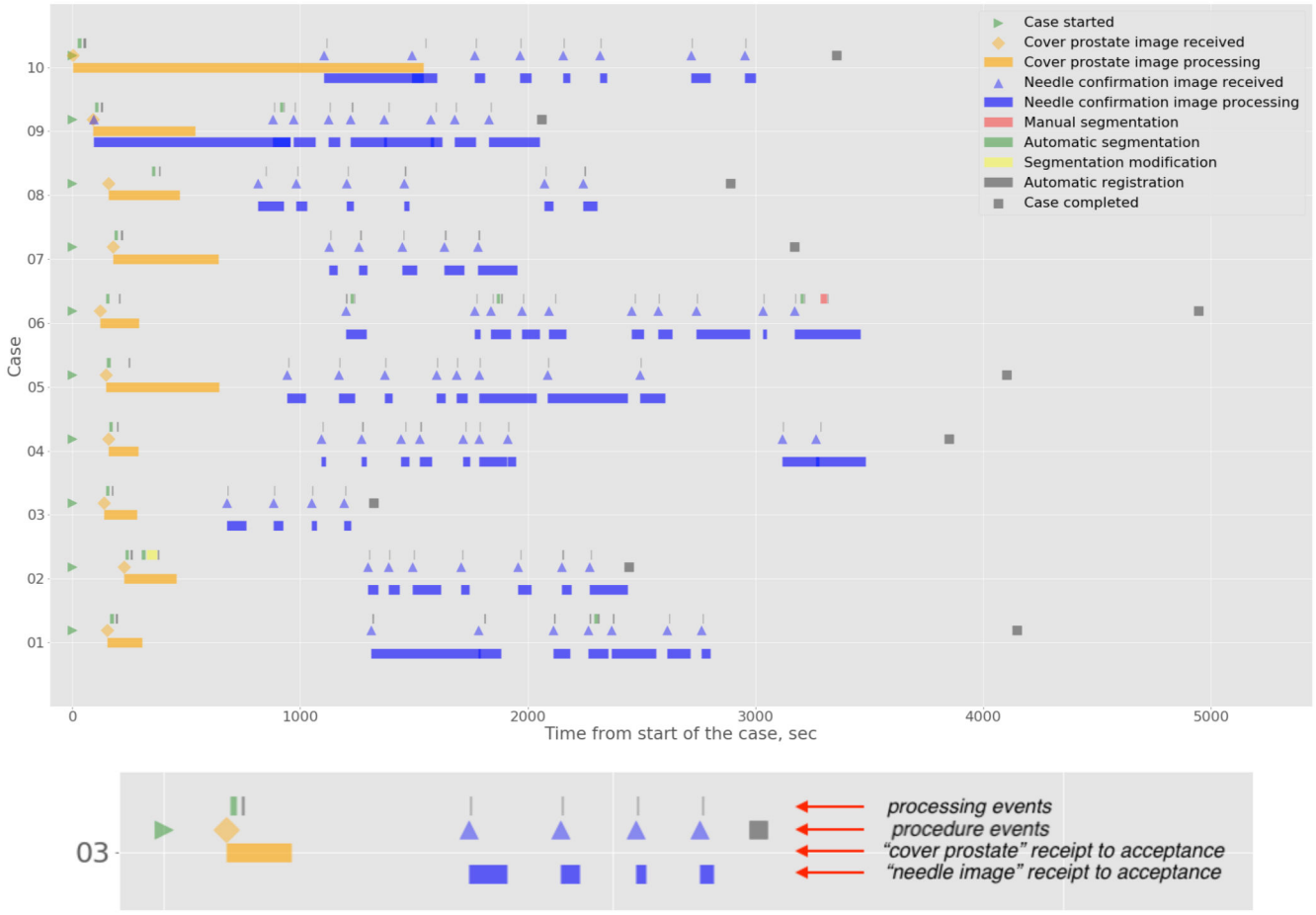

Figure 6.

Procedure event timeline for the cases used in the prospective evaluation. Each of the 10 cases is accompanied by four lines (as shown in the annotated inset for Case 3), from top to bottom: processing events (registration and segmentation, both automatic and manual), procedure events (start and stop, and receipts of the images), and the total time from image receipt by SliceTracker to the eventual approval of the registration result by the operator (separately for the "cover prostate " and "needle confirmation" images). In Case 10, approval was most likely delayed due to reasons unrelated to the software. In Case 9, workflow was most likely restarted, resulting in coinciding receipt of the cover prostate and the first needle confirmation images. Case 2 was the only one where retry of the cover prostate image registration was triggered. As evident from these timelines, computational processing steps comprise only a very small fraction of the overall procedure time.