Technical Note: Error metrics for estimating the accuracy of needle/instrument placement during transperineal magnetic resonance/ultrasound-guided prostate interventions

Technical Note: Error metrics for estimating the accuracy of needle/instrument placement during transperineal MR/US-guided prostate interventions

Ester Bonmati[a], Yipeng Hu[a], Barbara Villarini[a,c], Rachael Rodell[a], Paul Martin[a], Lianghao Han[a,d], Ian Donaldson[b], Hashim U. Ahmed[b,e,f], Caroline M. Moore[b], Mark Emberton[b], Dean C. Barratt[a].

[a] *UCL Centre for Medical Image Computing, Department of Medical Physics & Biomedical Engineering, University College London, UK*

[b] *Division of Surgery and Interventional Science, University College London, UK*

[c] *Department of Computer Science, University of Westminster, UK*

[d] *Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, PR China*

[e] *Division of Surgery, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, UK*

[f] *Imperial Urology, Charing Cross Hospital, Imperial College Healthcare NHS Trust, London, UK*

e-mail: e.bonmati@ucl.ac.uk

Abstract

Purpose: Image-guided systems that fuse magnetic resonance imaging (MRI) with three-dimensional (3D) ultrasound (US) images for performing targeted prostate needle biopsy and minimally-invasive treatments for prostate cancer are of increasing clinical interest. To date, a wide range of different accuracy estimation procedures and error metrics have been reported, which makes comparing the performance of different systems difficult.

Methods: A set of 9 measures are presented to assess the accuracy of MRI-US image registration, needle positioning, needle guidance, and overall system error, with the aim of providing a methodology for estimating the accuracy of instrument placement using a MR/US-guided transperineal approach.

Results: Using the SmartTarget fusion system, an MRI-US image alignment error was determined to be 2.0±1.0 mm (mean ± SD), and an overall system instrument targeting error of 3.0±1.2 mm. Three needle deployments for each target phantom lesion was found to result in a 100% lesion hit rate and a median predicted cancer core length of 5.2 mm.

Conclusions: The application of a comprehensive, unbiased validation assessment for MR/TRUS guided systems can provide useful information on system performance for quality assurance and system comparison. Furthermore, such an analysis can be helpful in identifying relationships between these errors, providing insight into the technical behaviour of these systems.

Keywords: Accuracy Validation, Needle Placement, Prostate Cancer, Targeted Biopsy, Focal Therapy, Image-guided Interventions

Introduction

The widespread introduction of magnetic resonance imaging (MRI) for detecting, staging, and localising prostate cancer has led to an increasing clinical interest in MRI-ultrasound (US) image fusion systems to guide tumour-targeted needle biopsy and minimally-invasive treatments[1,2]. The accuracy of such systems has been the subject of a number of previous phantom studies, with mean errors between 2 and 3mm commonly reported[3–8]. However, the error metrics adopted vary considerably. Furthermore, few studies have attempted to estimate needle/instrument placement accuracy for procedures where the needle/instrument is inserted via the perineum. Among such transperineal procedures are a wide range of minimally-invasive surgical treatments, such as cryotherapy and injectable drug therapies. A further problem is that accuracy measures are typically difficult to relate to clinically meaningful measures, such as tumour/lesion hit-rate. Although phantom experiments are generally performed under idealised conditions, where, for example, the phantom motion does not represent tissue motion encountered *in vivo*, it is still important to estimate the accuracy of a fusion system to provide an indication of its performance under "perfect" conditions (i.e. where sources of error are well controlled) to enable comparison between different systems and for the purposes of quality assurance. In this paper, we present a series of error metrics to characterise the accuracy of fusion systems for guiding transperineal procedures.

Materials and Methods

Accuracy Validation Method

Figure 1 illustrates a typical workflow for an MRI-US fusion system used to place one or more needles/instruments into US-visible, target lesions in the phantom. To eliminate bias introduced when an operator targets an US-visible lesion, the overall gain on the B-mode of the US scanner may be reduced until lesions are no longer visible but the needles or similarly strongly-reflective instruments remain identifiable in the US images. This procedure is also effective at reducing the visibility of needle tracks from previous instrument insertions, which can be an additional source of bias. Once the desired number of instruments have been inserted, 3D US imaging of the phantom is performed to enable the 3D position of each instrument to be determined with respect to the target. Here, we refer to this image as a "*validation volume*". Two validation volumes are obtained after instrument placement – one with the US scanner gain set low so that the instrument artefacts are minimised, and one after the gain has been increased so that lesions are clearly visible – this helps to ensure that both the instrument tip, lesions, and lesion centres are determined as accurately as possible. To calculate the error metrics introduced in the next section, the validation volumes are analysed to define the 3D co-ordinates of the surface and the "*true target centre*" for each target lesion; the term "true target centre" is used here to refer to the ground truth point target location, which for convenience is assumed to be at the centroid of the target phantom lesion.

Error Metrics

In this section, a set of 3D error metrics for quantifying the system accuracy of an MR/TRUS fusion system related to different error sources are defined. These metrics are illustrated graphically in **Figure 2**.

Target Registration Error (Measure 1)

The MRI-TRUS *target registration error* (TRE) is defined as the Euclidean distance between the centre of the TRUS true target lesion (*defined in the validation image volume*) and the centre of the registered target lesion. We assume here that the target lesion centre point is not used in the registration of the MRI and TRUS images. Most commonly, the prostate surfaces, defined in both MRI and TRUS images, are registered. The TRE calculated for one or more target lesions is then independent of the points used to calculate the registration transformation. Since this measure depends on the image resolution, all imaging parameters must be reported.

Surface Overlap Error (Measure 2)

The MRI-TRUS lesion overlap assesses the accuracy of the registration in predicting the location of regions of interest. Here we adopt the commonly used Dice Similarity Coefficient (DSC), which is calculated in the validation image volume. Additionally, the agreement between the MRI and US surfaces of the prostate gland, in terms of distance between points on the gland (identified in TRUS images) and the registered gland surfaces, is also calculated to provide a reference point on the registration algorithmic performance. Unlike the lesion overlap (Measure 2) or TRE (Measure 1), it is important to note that this measure does not provide an independent measure to assess registration accuracy for the purposes of instrument targeting when MRI and TRUS surfaces are registered. The surface distance is defined as the root-mean-square (RMS) distance between the points identified in both surfaces (TRUS and registered MRI).

Overall Targeting Error (Measure 3)

The *overall targeting error*, based on physical needle-tip placement, is defined as the Euclidean distance between the needle tip position and the physical centre of the true target lesion identified in the same TRUS image volume. This measure represents the overall error of the system as it describes the physical needle placement accuracy.

Procedural Errors (Measures 4 - 6)

The position of the biopsy needle is subject to bending, deflection and uncertainty in insertion depth, which can be measured collectively by the distance between the “true”- and the “planned” needle tips. This is termed the *needle deflection error* (Measure 5). In case of using a fixed brachytherapy template grid, the needle placement is further restricted by the position of the grid relative to the phantom. This *template grid error* can be measured by the distance between the “planned” needle tip and the registered lesion centre, i.e. the target upon which the calculation of the needle position is based (Measure 6). The overall *procedural error* – i.e. the error due to the uncertainty in needle placement during the procedure – is defined as the distance between the registered target centre and the “true” needle tip (Measure 4), combining Measures 5 and 6.

Decoupled Targeting Errors (Measures 7 - 8)

The overall targeting error (Measure 3) includes the registration error (TRE; Measure 1) and the overall procedural error (Measure 4). In some applications, such as biopsy or brachytherapy, needle deflection may be significant[9]. For such applications, de-coupling the overall procedural error into template and needle deflection (Measures 6 and 5 respectively) enables these two component errors to be excluded from the overall targeting error. A decoupled guidance error, including only the template grid error (Measure 6) without the contribution of the needle deflection error

(Measure 5), may be defined as the distance between the true target lesion and the planned needle- (or instrument-) tip. This error measure is referred to here as the *targeting error excluding needle/instrument deflection error* and is shown as Measure 7 in **Figure 2**. Similarly, another decoupled targeting error including only deflection error (Measure 5), which is equivalent to the targeting error, *excluding the template grid error*, may be measured as the distance between the true target lesion and a virtual needle tip location, assuming an idealised needle/instrument trajectory without deflection (following registration, if the template grid was adaptable and moved to the same position of the registered target in the transverse plane). The position of this virtual needle/instrument tip position is estimated by adding the 3D position of the registered target to the independently measured deflection error in 3D. This error is denoted as Measure 8 and provides an estimate of the optimal overall system guidance error only with needle/instrument deflection.

Biopsy-specific Measures

For needle biopsy, the *cancer core length* (CCL), defined as the physical length of cancerous tissue in a tissue sample, and the *lesion hit rate* are two important clinical measures that have a direct bearing on the performance of targeted prostate biopsy as a diagnostic test. Such measures are straightforward to calculate from the phantom data. Therefore, if possible, we propose to estimate additional measures including the predicted CCL (for each needle insertion), lesion hit rate, maximum CCL, and total CCL (for multiple needle insertions), based on the point-to-point distance errors calculated for the measures defined above. In this work, the CCL is calculated by finding the length of the intersection between the needle trajectory and the true lesion segmented from the validation volume(s).

Error Analysis

A comprehensive error analysis in which distributions and summary statistics were calculated for each error and each needle insertion. Statistical tests comparing each decoupled targeting errors (Measure 7 and 8) and the overall targeting error (Measure 3) were performed to reveal the effect of needle deflection or fine-adjustment of the template grid position on the overall system accuracy. We assessed the distribution of the results using normality tests (χ^2 goodness-of-fit (GoF) and Kolmogorov-Smirnov (K-S) GoF). When the normality tests failed, in addition to the Student's t-test, results from a non-parametric test (Wilcoxon rank sum test) were also reported.

Example Application

We tested a research version of the commercially available SmartTarget guidance system, developed by our research group (SmartTarget Ltd., London, UK). The equipment was set up as per a template-guided transperineal biopsy using a disposable brachytherapy template grid (CIVCO Medical Solutions, Iowa, USA) attached to an US probe stepper. The template contained a 13 by 13 grid of holes, spaced 5 mm apart in both directions. We used the CIRS 053-MM prostate training phantom (Computerized Imaging Reference Systems, Inc. Norfolk, Virginia, USA), compatible with US, CT and MR imaging. This phantom contains three hypoechoic spherical objects, with a volume of 0.5 cc, randomly-placed within the prostate to represent three lesions. The phantom was scanned with a Philips Archieva 3.0 Tesla MR machine to get the T2-weighted MRI images for the pre-operative planning with a voxel size of 0.38 × 0.38 × 1.00 mm and a total of 55 transversal slices. Three Bard® Max-Core® disposable core biopsy needle guns were used, each with a 16 cm, 18-gauge needle (Bard Biopsy Systems, Arizona, USA). Briefly, a patient specific model was generated during pre-operative planning[10] and registered to a reconstructed 3D TRUS volume[11]. Then, three biop-

sy needles were inserted into the three optimal 3 template grid co-ordinates calculated and displayed, following the validation procedure outlined earlier.

Results

In this section, accuracy results for the system tested based on the measures introduced earlier are summarised.

MR-US image Registration Accuracy

The mean (± SD) TRE and lesion surface overlap (Measures 1 and 2) were 2.03 ± 0.98 mm and 68.77 ± 14.25%, respectively. Histograms of these errors are shown in **Figure 3** (a) and (b). The mean distance between the manually-defined TRUS prostate boundary points and the registered MR model was 0.67 ± 0.04 mm (see **Figure 3** (c)).

A χ^2 GoF test confirms that the cross-tabulated TRE and point-to-surface distance values are statistically independent ($p<0.0001$). This suggests that the measure of gland surface agreement may not be appropriate to replace the role of the independent TRE in assessing registration accuracy. However, the point-to-surface distance provides useful real-time feedback to indicate the algorithmic performance of the MR-TRUS registration (i.e. how well the algorithm performs at the task of fitting the MRI-derived model surface to the US points).

Procedural Errors

The overall procedural error (Measure 4) for the first targeted needle-tip was 2.89 ± 1.24 mm (mean ± SD), and 4.01 ± 1.45 mm over all three inserted needles. As expected, the second and the third targeted points resulted in larger errors due to the physical constraints of the template grid, which means that these tend to be further from the centre of the lesion. Procedural errors, including overall error, template and needle deflection errors are summarised in **Table 1**.

Both χ^2 GoF and K-S GoF tests failed to reject the null hypothesis that the *difference* between the procedural error and the overall targeting error follow a normal distribution (χ^2 $p=0.1229$, K-S $p=0.9226$). Paired t-tests between procedural errors and overall targeting error (Measure 3) show that: a) there is statistically significant difference between the overall procedural error and the overall targeting error ($p=0.0282$), which means that the procedural error contributed towards the overall targeting error, and b) there is statistically significant difference between the template error and the overall targeting error ($p<0.0001$) and also between the needle deflection error and the overall targeting error ($p=0.0273$), indicating that neither the template nor the needle deflection errors are negligible.

A further χ^2 GoF test failed to reject that there is statistical significant dependence between the procedural error and TRE ($p=0.3391$). This suggests that the needle positioning may not be necessarily independent of registration error. This is probably because both errors may be affected by some properties of the target, such as the phantom material properties and the relative spatial locations of the targets of interest.

Targeting Errors

Targeting errors (Measures 3, 7 and 8) are summarised in **Table 2**. A histogram of the corresponding errors for each needle placement is shown in **Figure 4**. In 93% of the cases, the lesion was hit with the first needle; a 100% lesion hit rate was achieved when all 3 deployed needles were taken into account.

χ^2 GoF tests were performed between the overall targeting error and procedural error ($p=0.2381$) and between overall targeting error and TRE ($p=0.3391$). This confirms that both the TRE and procedural error contributed significantly towards the overall targeting error. The χ^2 GoF test rejected the null hypothesis that the difference between the targeting error excluding needle deflection error (Measure 7) and the overall targeting error follows a normal distribution ($p=0.0188$), while the single sample Kolmogorov-Smirnov GoF test failed to reject the hypothesis

($p=0.1042$). While, both the χ^2 GoF test and the sample Kolmogorov-Smirnov GoF test, failed to reject the null hypothesis that the difference between the targeting error excluding template error and the overall targeting error follows a normal distribution (χ^2 $p=0.1736$, K-S $p=0.6960$). Paired t-tests and non-parametric Wilcoxon rank sum tests were performed to test if there is significant difference between the overall targeting error (Measure 3) and Measures 7 and 8. Results show that the targeting error excluding needle deflection error is statistically significant different than the overall targeting error (both tests $p<0.0001$). Additionally, although results show a significant difference between the targeting error excluding template error (Measure 8) and the overall targeting error (both tests $p<0.0001$) considering all needles, there is no significant increase in the overall targeting error considering only the first needle (t-test $p=0.9095$, Wilcoxon test $p=1.0$). The test also accepts the hypothesis that the difference is significantly smaller than 2 mm between these two measures for all the needles.

Biopsy-specific Measures

A summary of the mean predicted CCL, the maximum CCL, and total CCL (for multiple needle insertions) are shown in **Table 3**.

Discussion

In this paper we propose a practical validation procedure that enables a comprehensive set of quantitative measures to be estimated to assess the accuracy of a prostate MR/TRUS fusion system.

We chose point-to-point distance errors instead of target-point to needle-track-segment distances so that the tip of the instrument was guided to a point target (generally, the centre of the lesion or other location). Distances in 2D/3D between a line (or line segment) that represents an instrument trajectory, and a point target, or between two lines, provide less informative measures of targeting error than the distance between two points.

The phantom used in this study provides a basis for examining system accuracy under idealised "laboratory" conditions. However, its mechanical properties may differ significantly from human tissue and motion encountered *in vivo*. This poses a challenge for estimating the "real-world" registration accuracy, as well as the overall *in vivo* targeting accuracy, but arguably the validation under controlled conditions using a phantom is still important for quality assurance purposes, system comparison, and understanding the contributions of different error sources. Note that, the set of measures can be used with different phantoms. However, if the intention is to compare the accuracy of different fusion systems, the same model of phantom should be used in the evaluation of each system to ensure a fair comparison.

We have reported results using non-parametric Wilcoxon sum rank tests when the normality tests rejected the null hypothesis that the data followed a normal distribution. However, we would like to note that both the normality tests and the non-parametric tests may lack of statistical power with small data size, which has to be addressed in the experiment design prior to obtaining the results. Nevertheless, such validation experiments complement further clinical validation studies, which focus on clinical outcomes and typically require a reasonable level of system accuracy to be established to ensure safety and maximise the achievable efficacy of the procedure[12–15].

Another limitation relates to the way in which 3D TRUS images are used for validation purposes, which is known to sometimes present problems in accurately localising needle/instrument tips due to image artefacts. On balance, US provides a simple and immediate means of acquiring a 3D image of instruments in situ, which avoids errors introduced after transferring to, and imaging with, other modalities, such as CT.

An interesting finding from the application of the validation procedure was the dependency between registration error and procedural error using the investigated guidance system. This may be due to the fact that the lesion has some material or mechanical properties that may bias both registration and needle positioning errors. This may lead to a propagation model being formulated to predict the TRE, which is not easily available in real patient data validation or during the procedure as an accuracy feedback.

To date, neither the relationship between different types of errors, nor the contribution of each error to the overall targeting error of a system, has been reported systematically. Therefore, it is difficult to draw conclusions that can inform decisions on the adoption or improvement of a particular guidance system for specific clinical applications.

Conclusions

In this paper, we have proposed and demonstrated the application of a comprehensive, unbiased validation assessment for MR/TRUS guided targeting systems for prostate transperineal biopsy and focal therapy. The error analysis indicated that the proposed procedure can provide useful information on system performance for quality assurance, system comparison, evaluating the magnitudes of different sources of errors, by comparing these errors (for instance to identify workflow and algorithmic improvements), and identifying relationships between these errors that provide insight into the technical behaviour of these systems.

Acknowledgements

This publication presents independent research supported by the Health Innovation Challenge (HIC) Fund (Grant Ref. HICF-T4-310), a parallel funding partnership between the Department of Health and the Wellcome Trust. The views expressed in this publication are those of the author(s) and not necessarily those of the Department of Health or the Wellcome Trust. The research was undertaken at UCL/ULCH who received a proportion of funding from the Department of Health's NIHR Biomedical Research Centres funding scheme.

The authors would like to thank Dr David Atkinson from UCL Centre or Medical Imaging for his assistance in acquiring the MRI images of the phantom.

Conflicts of interest

SmartTarget Ltd., is a UCL spin-out company that has commercialised the fusion system used in the application example showed in this paper. Dean Barratt is a Director and shareholder in SmartTarget Ltd. Yipeng Hu is also a shareholder in SmartTarget Ltd. Rachael Rodell is employed by SmartTarget Ltd. Hashim Ahmed's research is supported by core funding from the United Kingdom's National Institute of Health Research (NIHR) Imperial Biomedical Research Centre. Hashim Ahmed currently receives funding from the Wellcome Trust, Prostate Cancer UK, Sonacare Inc., Trod Medical and Sophiris Biocorp for trials in prostate cancer. Hashim Ahmed is a paid medical consultant for Sophiris Biocorp and Sonacare Inc.

References

1. Marks L, Young S, Natarajan S. MRI-ultrasound fusion for guidance of targeted prostate biopsy. *Current*

opinion in urology. 2013;23(1):43-50. doi:doi:10.1097/MOU.0b013e32835ad3ee.

2. Moore CM, Robertson NL, Arsanious N, et al. Image-Guided Prostate Biopsy Using Magnetic Resonance Imaging–Derived Targets: A Systematic Review. *European Urology*. 2013;63(1):125-140. doi:10.1016/j.eururo.2012.06.004.
3. Xu S, Kruecker J, Turkbey B, et al. Real-time MRI-TRUS fusion for guidance of targeted prostate biopsies. *Computer Aided Surgery*. 2008;13(5):255-264. doi:10.3109/10929080802364645.
4. Bax J, Cool D, Gardi L, et al. Mechanically assisted 3D ultrasound guided prostate biopsy system. *Medical Physics*. 2008;35(12):5397-5410. doi:10.1118/1.3002415.
5. Ukimura O, Desai MM, Palmer S, et al. 3-Dimensional Elastic Registration System of Prostate Biopsy Location by Real-Time 3-Dimensional Transrectal Ultrasound Guidance With Magnetic Resonance/Transrectal Ultrasound Image Fusion. *The Journal of Urology*. 2012;187(3):1080-1086. doi:10.1016/j.juro.2011.10.124.
6. Kuru TH, Roethke M, Popeneciu V, et al. Phantom Study of a Novel Stereotactic Prostate Biopsy System Integrating Preinterventional Magnetic Resonance Imaging and Live Ultrasonography Fusion. *Journal of Endourology*. 2012;26(7):807-813. doi:10.1089/end.2011.0609.
7. Tokuda J, Song SE, Fischer GS, et al. Preclinical evaluation of an MRI-compatible pneumatic robot for angulated needle placement in transperineal prostate interventions. *International Journal of Computer Assisted Radiology and Surgery*. 2012;7(6):949-957. doi:10.1007/s11548-012-0750-1.
8. Blumenfeld P, Hata N, DiMaio S, et al. Transperineal prostate biopsy under magnetic resonance image guidance: A needle placement accuracy study. *Journal of Magnetic Resonance Imaging*. 2007;26(3):688-694. doi:10.1002/jmri.21067.
9. Sadjadi H, Hashtrudi-Zaad K, Fichtinger G. Needle deflection estimation: prostate brachytherapy phantom experiments. *International Journal of Computer Assisted Radiology and Surgery*. 2014;9(6):921-929. doi:10.1007/s11548-014-0985-0.
10. Hu Y, Gibson E, Ahmed HU, Moore CM, Emberton M, Barratt DC. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. *Medical Image Analysis*. 2015;26(1):332-344. doi:10.1016/j.media.2015.10.006.
11. Hu Y, Ahmed HU, Taylor Z, et al. MR to ultrasound registration for image-guided prostate interventions. *Medical Image Analysis*. 2012;16(3):687-703. doi:10.1016/j.media.2010.11.003.
12. Pinto PA, Chung PH, Rastinehad AR, et al. Magnetic Resonance Imaging/Ultrasound Fusion Guided Prostate Biopsy Improves Cancer Detection Following Transrectal Ultrasound Biopsy and Correlates With Multiparametric Magnetic Resonance Imaging. *The Journal of Urology*. 2011;186(4):1281-1285. doi:10.1016/j.juro.2011.05.078.
13. Natarajan S, Marks LS, Margolis DJA, et al. Clinical application of a 3D ultrasound-guided prostate biopsy system. *Urologic Oncology*. 2011;29(3):334-342. doi:10.1016/j.urolonc.2011.02.014.
14. Schouten MG, Bomers JGR, Yakar D, et al. Evaluation of a robotic technique for transrectal MRI-guided prostate biopsies. *European Radiology*. 2012;22(2):476-483. doi:10.1007/s00330-011-2259-3.
15. Shoji S, Hiraiwa S, Endo J, et al. Manually controlled targeted prostate biopsy with real-time fusion imaging of multiparametric magnetic resonance imaging and transrectal ultrasound: An early experience. *International Journal of Urology*. 2015;22(2):173-178. doi:10.1111/iju.12643.

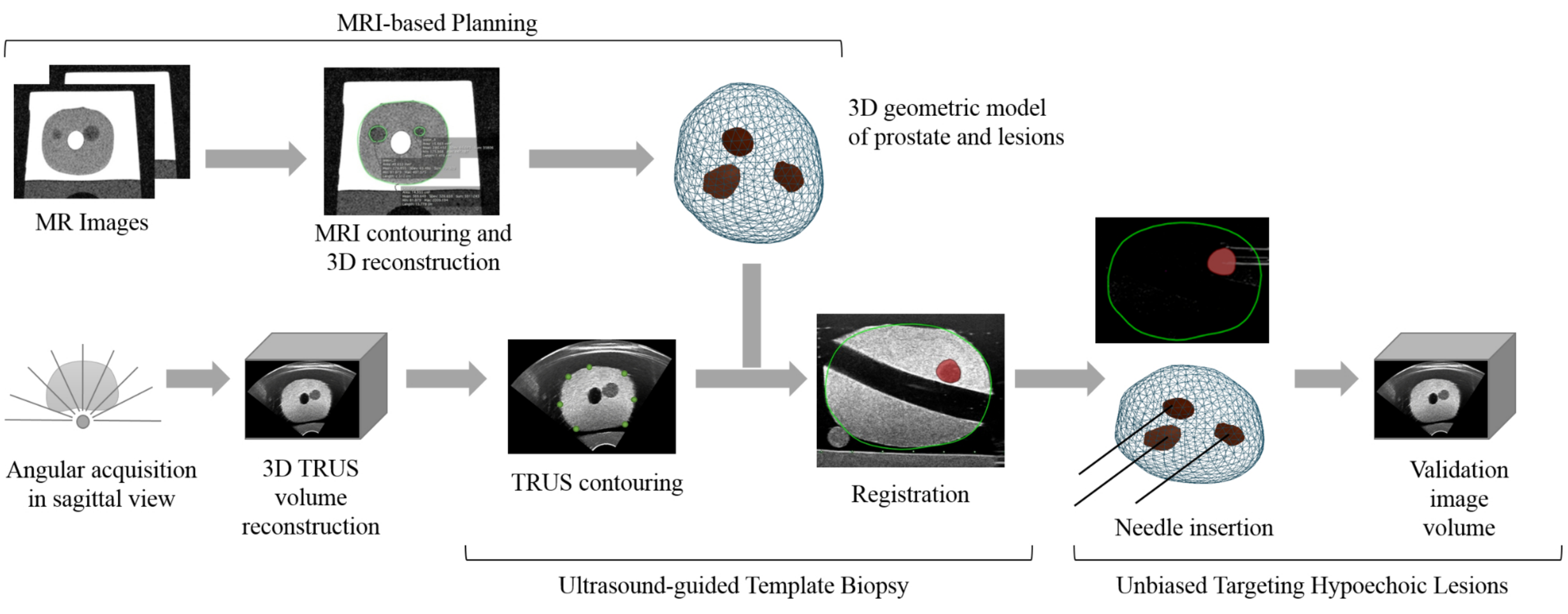
Figure legends

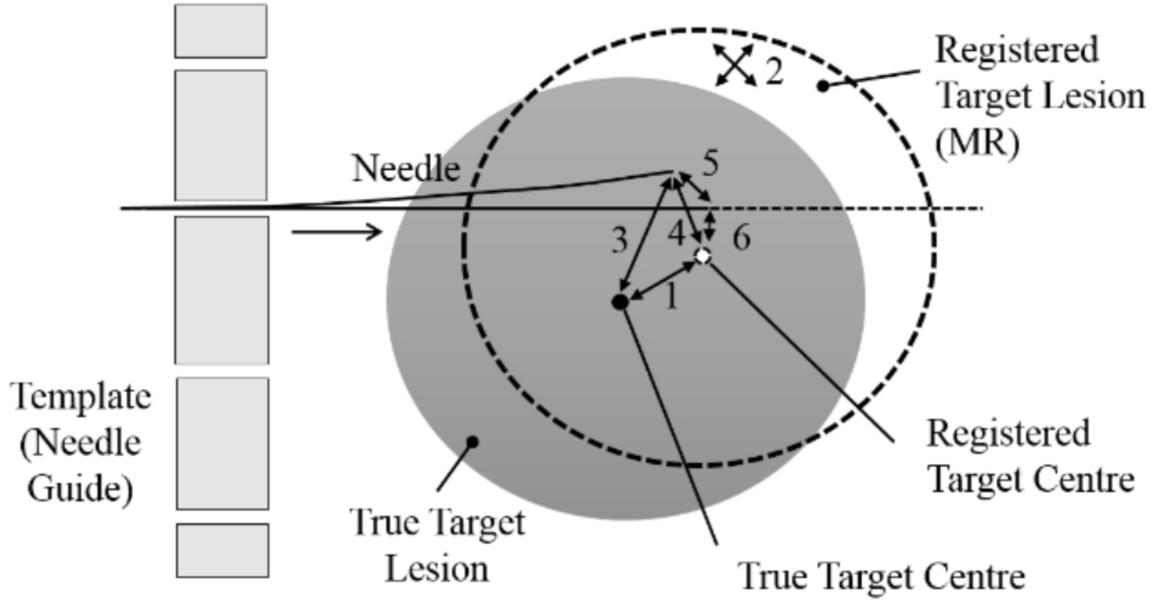
Figure 1. Schematic overview of the workflow of a MR-US guidance software.

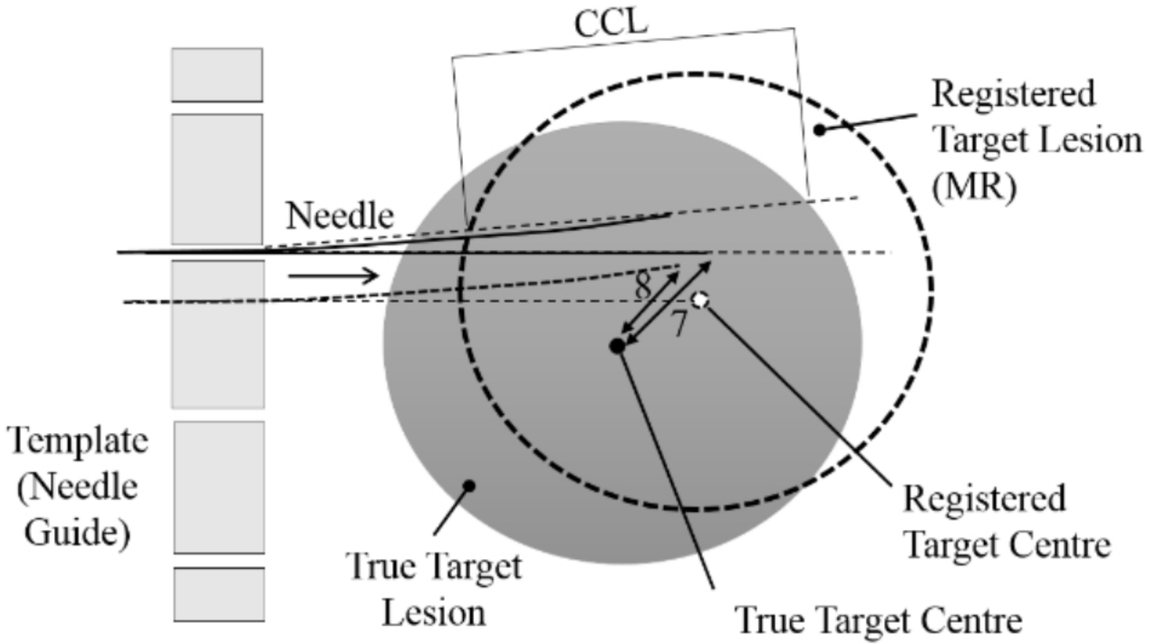
Figure 2. Measures used for the validation: 1 - Target registration error (TRE). 2 - Surface overlap error. 3 - Overall targeting error. 4 – Overall procedural error. 5 – Needle/instrument deflection error. 6 – Template grid error. 7 - Targeting error excluding needle deflection error. 8 - Targeting error excluding template grid error.

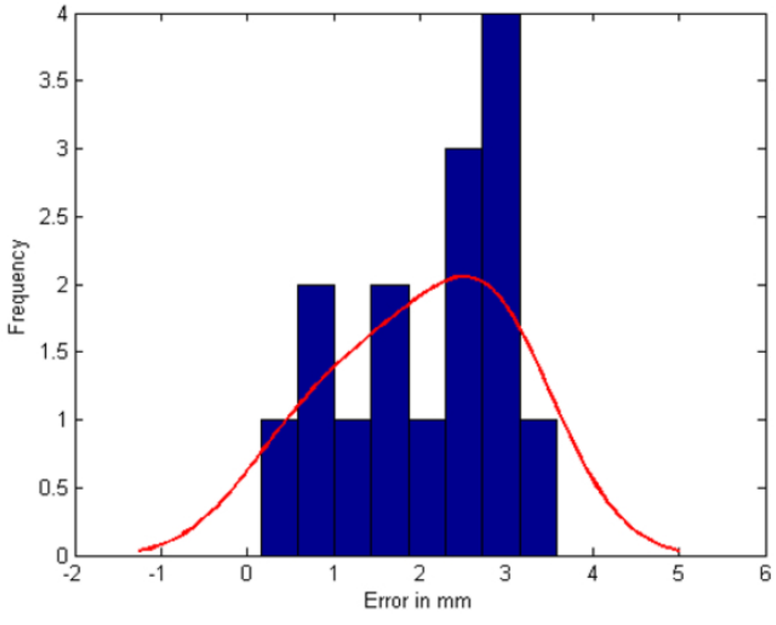
Figure 3. Histograms and estimated error distribution of: a) the TRE (Measure 1); b) the surface overlap error between registered and true lesion (Measure 2); and c) the point-to-surface distance between the registered model and the user-defined US points.

Figure 4. Histogram and estimated error distribution of overall targeting errors for: a) all data; b) the 1st needle, c) the 2nd needle, and d) the 3rd needle (in mm).

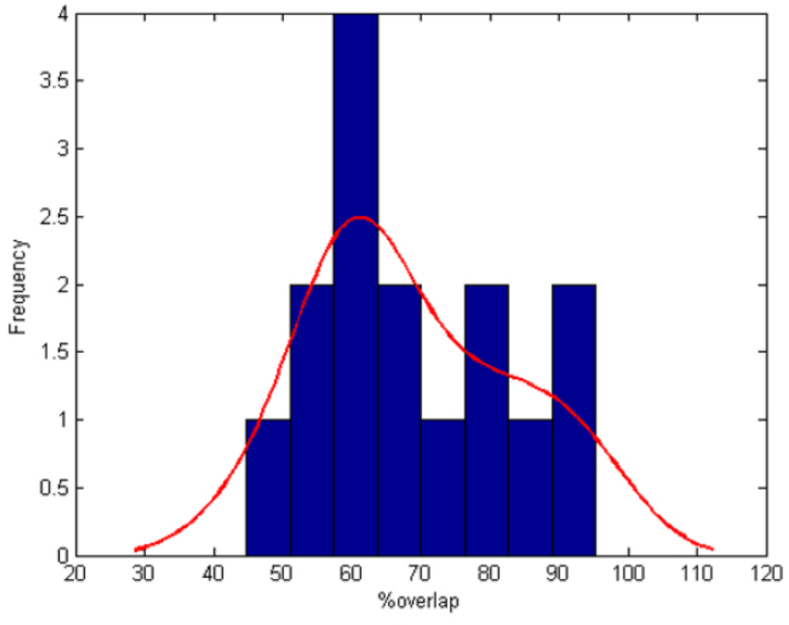

MRI-based Planning
MR Images
MRI contouring and
3D reconstruction
3D geometric model
of prostate and lesions
Angular acquisition
in sagittal view
3D TRUS
volume
reconstruction
TRUS contouring
Registration
Needle insertion
Validation
image
volume
Ultrasound-guided Template Biopsy
Unbiased Targeting Hypoechoic Lesions


2
Registered
Target Lesion
(MR)
Needle
5
3
4
6
1
Template
(Needle
Guide)
Registered
Target Centre
True Target
Lesion
True Target Centre

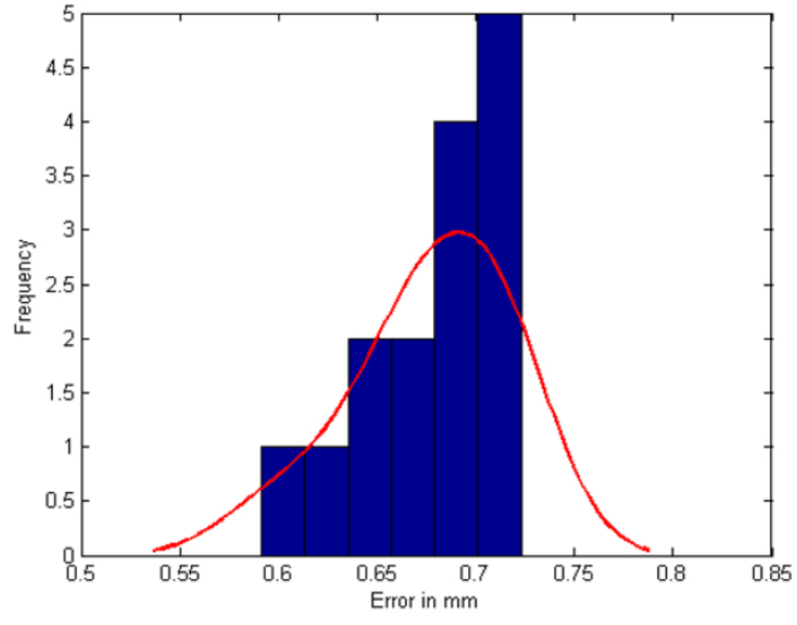

CCL
Registered
Target Lesion
(MR)
Needle
8
7
Template
(Needle
Guide)
Registered
Target Centre
True Target
Lesion
True Target Centre


Frequency
Error in mm

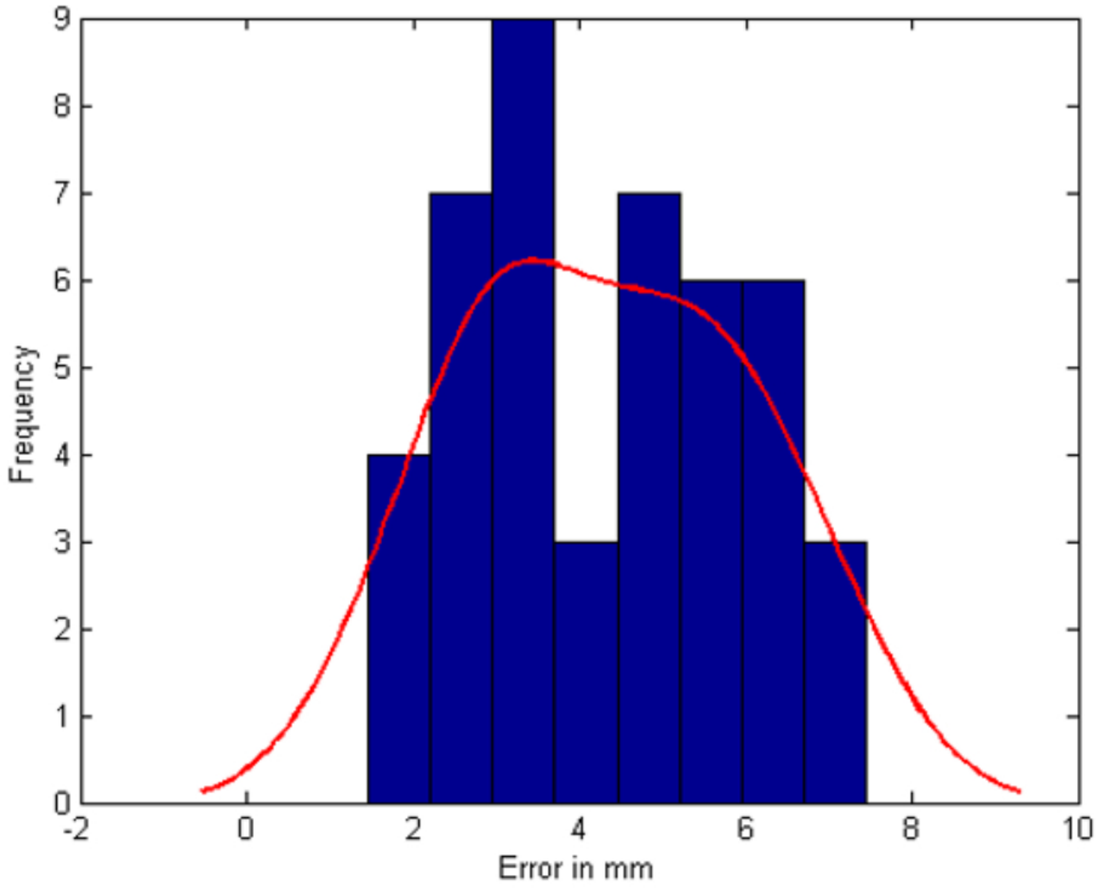
(a)


Frequency
%overlap

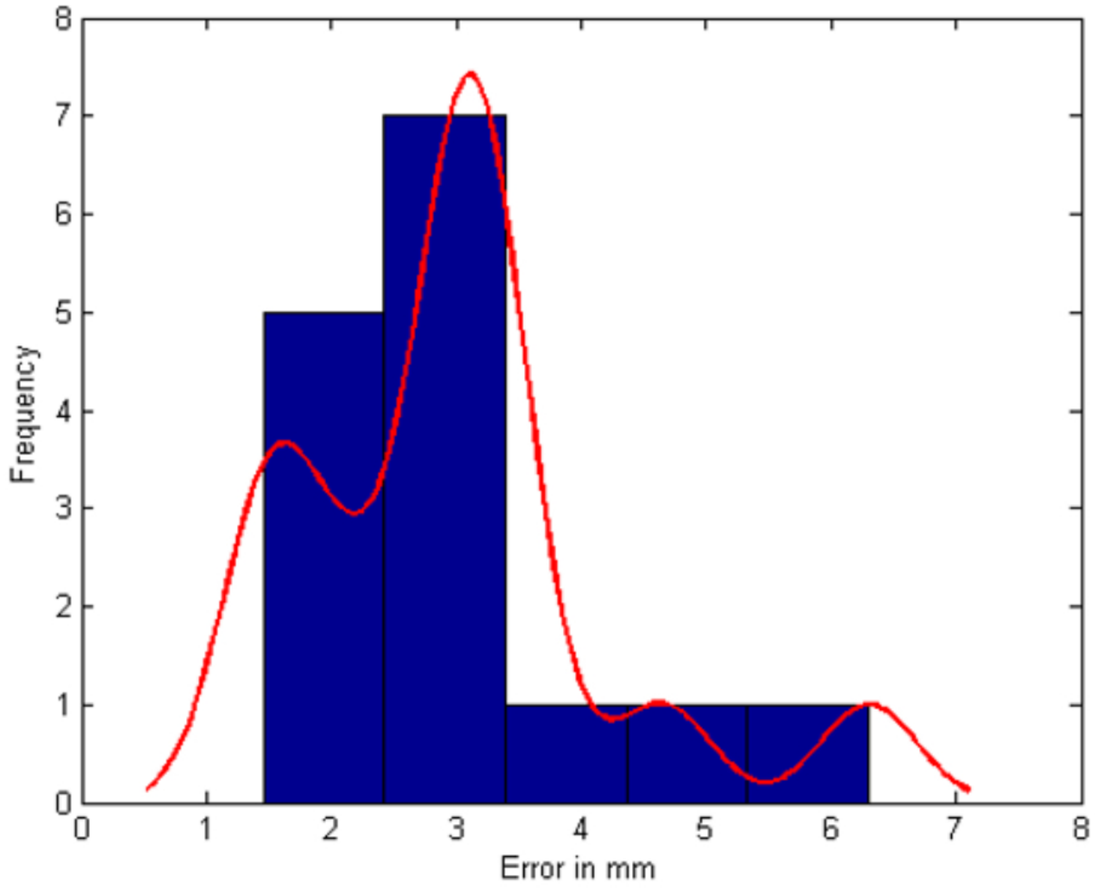
(b)


Frequency
Error in mm

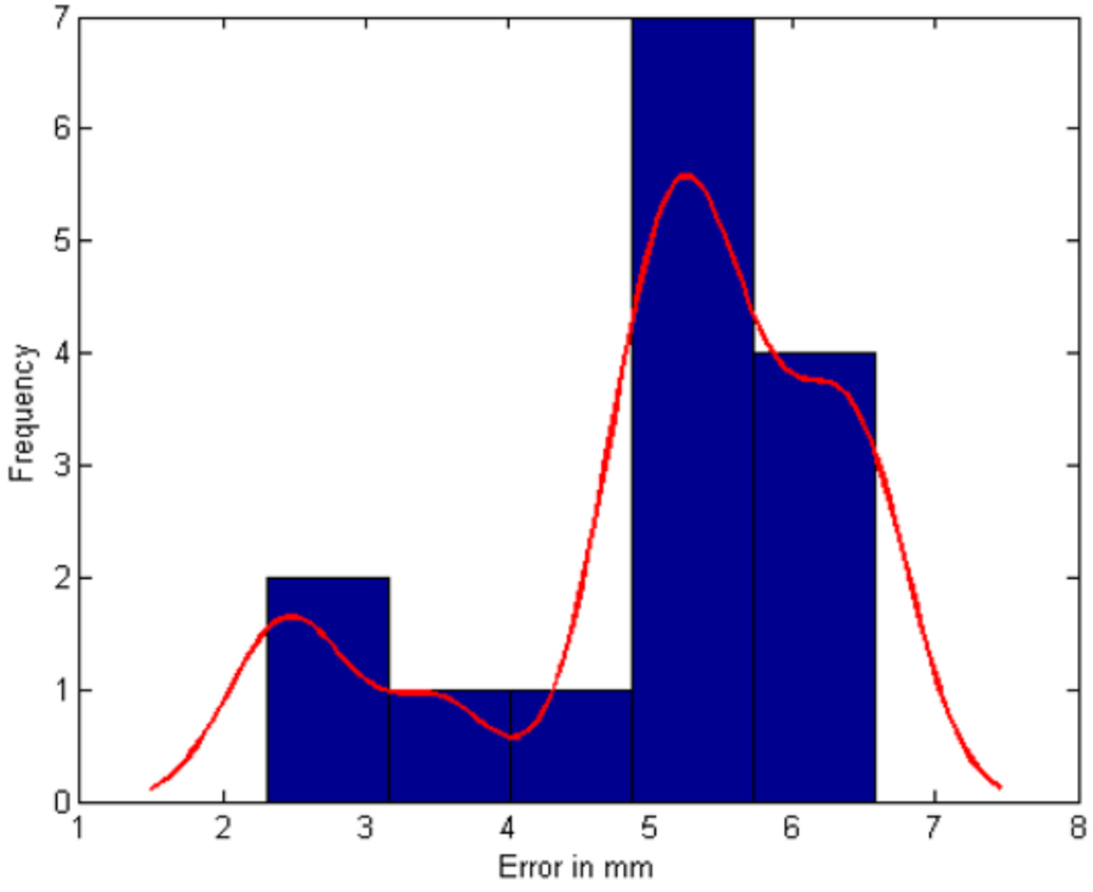
(c)


Frequency
Error in mm

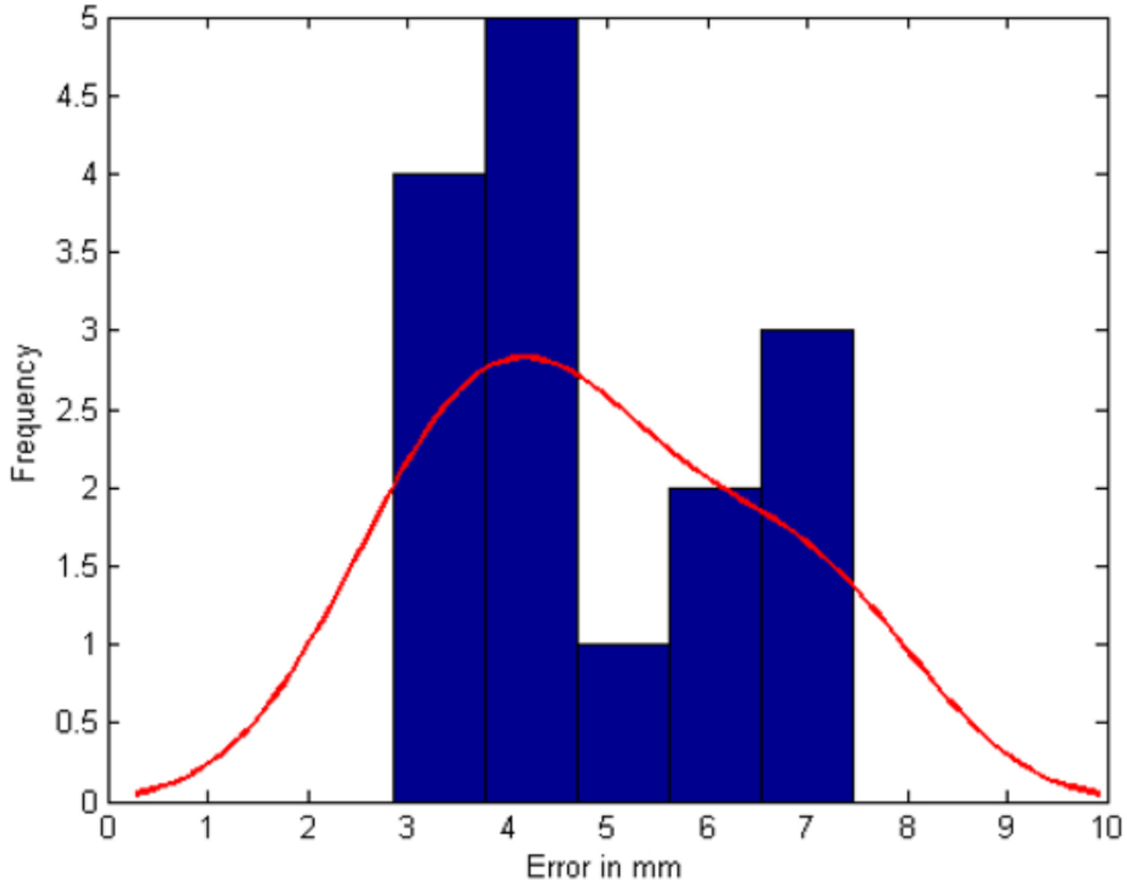
(a)


Frequency
Error in mm

(b)


Frequency
Error in mm

(c)


Frequency
Error in mm

(d)

Table 1. A summary of the mean (± SD) and median [25th 75th percentiles] numerical procedural errors (in mm).

Order of needle placement	Overall procedural error		Needle deflection error		Template error	
	Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**
1st needle	2.9±1.2	3.3 [1.8 4.0]	3.7±0.7	3.7 [3.2 4.0]	4.4±1.4	4.6 [3.0 5.3]
2nd needle	4.7±1.1	4.5 [4.1 5.3]	3.6±1.0	3.6 [3.3 4.2]	6.2±2.0	7.0 [5.2 7.4]
3rd needle	4.5±1.3	4.3 [3.5 4.9]	3.8±0.9	3.7 [3.2 4.3]	5.5±3.0	6.7 [2.6 7.7]
All	4.0±1.5	4.2 [3.1 4.8]	3.7±0.9	3.7 [3.2 4.2]	5.4±2.3	5.5 [3.0 7.2]

Table 2. A summary of the mean (± SD) and median [25th 75th percentiles] numerical targeting errors (in mm).

Order of needle placement	Overall targeting error		Targeting error excluding needle deflection error		Targeting error excluding template error	
	Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**
1st needle	3.0±1.2	3.1 [2.2 3.2]	5.5±1.0	5.2 [5.0 5.5]	3.0±1.1	2.9 [2.2 3.6]
2nd needle	5.1±1.3	5.3 [4.9 5.9]	7.3±2.1	8.4 [6.6 8.6]	2.6±0.9	2.4 [2.1 3.0]
3rd needle	4.9±1.5	4.6 [3.7 6.0]	6.6±3.1	8.9 [2.9 9.2]	3.1±1.3	3.1 [2.5 3.6]
All	4.3±1.6	4.4 [3.0 5.6]	6.4±2.4	6.2 [5.0 8.9]	2.9±1.1	2.7 [2.3 3.4]

Table 3. A summary of the median CCL [25th 75th percentiles], maximum CCL (MCCL), and total CCL (in mm)

Lesion	CCL	MCCL	Total CCL
1st lesion	2.7 [0.0 6.7]	8.14	8.92
2nd lesion	5.2 [2.3 7.4]	7.61	13.85
3rd lesion	6.9 [0.0 7.8]	8.03	15.27
All	5.2 [0.0 7.6]	8.14	39.69