ORIGINAL ARTICLE

Assessing the accuracy of a new 3D2D registration algorithm based on a non-invasive skin marker model for navigated spine surgery

Bas J. J. Bindels[1] · Rozemarijn A. M. Weijers[3] · Martijn S. van Mourik[3] · Robert Homan[3] · Jan J. Rongen[3] · Maarten L. J. Smits[2] · Jorrit-Jan Verlaan[1]

Abstract

Purpose   We assessed the accuracy of a new 3D2D registration algorithm to be used for navigated spine surgery and explored anatomical and radiologic parameters affecting the registration accuracy. Compared to existing 3D2D registration algorithms, the algorithm does not need bone-mounted or table-mounted instruments for registration. Neither does the intraoperative imaging device have to be tracked or calibrated.

Methods   The rigid registration algorithm required imaging data (a pre-existing CT scan (3D) and two angulated fluoroscopic images (2D)) to register positions of vertebrae in 3D and is based on non-invasive skin markers. The algorithm registered five adjacent vertebrae and was tested in the thoracic and lumbar spine from three human cadaveric specimens. The registration accuracy was calculated for each registered vertebra and measured with the target registration error (TRE) in millimeters. We used multivariable analysis to identify parameters independently affecting the algorithm's accuracy such as the angulation between the two fluoroscopic images (between 40° and 90°), the detector-skin distance, the number of skin markers applied, and waist circumference.

Results   The algorithm registered 780 vertebrae with a median TRE of 0.51 mm [interquartile range 0.32–0.73 mm] and a maximum TRE of 2.06 mm. The TRE was most affected by the angulation between the two fluoroscopic images obtained ($p < 0.001$): larger angulations resulted in higher accuracy. The algorithm was more accurate in thoracic vertebrae ($p = 0.004$) and in the specimen with the smallest waist circumference ($p = 0.003$). The algorithm registered all five adjacent vertebrae with similar accuracy.

Conclusion   We studied the accuracy of a new 3D2D registration algorithm based on non-invasive skin markers. The algorithm registered five adjacent vertebrae with similar accuracy in the thoracic and lumbar spine and showed a maximum target registration error of approximately 2 mm. To further evaluate its potential for navigated spine surgery, the algorithm may now be integrated into a complete navigation system.

Keywords Surgical navigation · Image-guidance · Computer-assisted surgery · Fluoroscopy · Spine · Vertebra

Introduction

Minimally invasive spine surgery is associated with better patient outcomes and lower overall costs than open spine surgery [1–5]. During minimally invasive spine surgery, surgeons strongly depend on intraoperative imaging to visualize relevant anatomical structures, and surgical hardware like screws and rods.

Compared to intraoperative two-dimensional (2D) fluoroscopic imaging, intraoperative three-dimensional (3D) navigation has large potential as spine surgeons can place pedicle screws more accurately while maintaining a short operation time, also (or rather, especially) in anatomically

✉ Bas J. J. Bindels
b.j.j.bindels@umcutrecht.nl

1   Department of Orthopedic Surgery, University Medical Center Utrecht – Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

2   Department of Radiology, University Medical Center Utrecht – Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands

3   Philips Healthcare, Veenpluis 6, 5684 PC Best, The Netherlands

challenging cases [6–8]. However, the required 3D imaging and navigational equipment is often heavy, cumbersome, and expensive [9–11].

We developed a new 3D2D registration algorithm for spine surgery that registers vertebrae from a preoperatively acquired CT to an intraoperative situation using 2D fluoroscopic imaging. In future, the new 3D2D registration algorithm may facilitate low-cost and easy-to-use 3D navigation without disrupting the routine fluoroscopic-guided workflow. The algorithm is based on non-invasive hybrid skin markers (radiopaque and optical), which are used to register the navigated optical space to the fluoroscopic space. Compared to existing 3D2D registration algorithms for navigated spine surgery, the algorithm does not need any bone-mounted or table-mounted instruments for registration. Neither does the algorithm require additional equipment attached to the intraoperative fluoroscopic imaging device to calibrate or track the imaging device [8, 12, 13].

In this study, we assessed the accuracy of the new 3D2D registration algorithm based on a non-invasive skin marker model, and explored anatomical and radiologic parameters affecting the registration accuracy.

Materials and methods

The study was conducted at the Department of Radiology and a surgical suite for experimental surgery of a university affiliated hospital in the Netherlands. The experiments were performed in compliance with the ethical guidelines for human cadaveric studies. All donors had provided written permission that their remains were to be used for research purposes. The study subjects were three fresh-frozen human torsos with no history of spinal surgery (Table 1).

Imaging data, marker model, and 3D2D registration algorithm

A baseline CT scan (Philips Brilliance 64 CT scanner, Philips, Best, Netherlands) was used to obtain 3D data (slice thickness 0.67 mm, contiguous slices, reconstruction matrix 512×512).

The (2D) fluoroscopic images for registration were acquired with a mobile C-arm system (Philips Zenition 70, Philips, Best, Netherlands). The imaging settings were set to the spine protocol (variable kV, typical dose-level 0.408 mGy 20 cm PMMA) to achieve optimal image quality of the vertebrae, which was part of the regular software (version 5.1.7: IQ NA HC R5.1.7).

All imaging data files were transferred to a secured portable computer in Digital Imaging and Communications in Medicine (DICOM) format.

Table 1 Characteristics of study subjects

	Subject A	Subject B	Subject C
Gender	Male	Male	Female
Age (years)	88	78	51
Waist circumference (centimeter)	121	96	73
Radiologic findings on baseline CT	Multiple sclerotic lesions throughout the vertebral column, possibly (prostate cancer) metastases Compression fracture 12th thoracic vertebra	None	Multiple lytic lesions 9th and 10th thoracic vertebrae Compression fractures 10th and 11th thoracic vertebrae

The non-invasive marker model consisted of a randomly applied pattern of prototype hybrid skin markers (radiopaque and optical), which were an update of previously used optical markers [14]. The update consisted of a radiopaque sphere added to the marker's center to make them visible on fluoroscopy (Fig. 1).

The 3D2D registration was performed offline by running image data through a prototype algorithm (Philips Healthcare, Best, the Netherlands) on a regular computer (Intel® Core™ i7-9750H processor, NVIDIA Quadro® T1000 graphics card).

3D2D registration process

The prototype 3D2D registration algorithm contains three different functionalities: a segmentation algorithm [15], a registration algorithm, and a pose-estimation algorithm.

After the anatomical level of one vertebra was manually indicated on the baseline CT, the model-based segmentation algorithm automatically segmented all vertebrae present in the CT. Subsequently, the registration algorithm processed the segmented vertebrae into digitally reconstructed radiograph (DRR) images using a forward projection algorithm [16].

Then, the anatomical level of one vertebra was manually indicated on one fluoroscopic image and the registration algorithm matched the segmented vertebrae to their position in the fluoroscopic images (2D) with a rigid registration per vertebrae to correct for inter-vertebrae deformation. The vertebrae were matched based on their gradients, and the gradient dif-
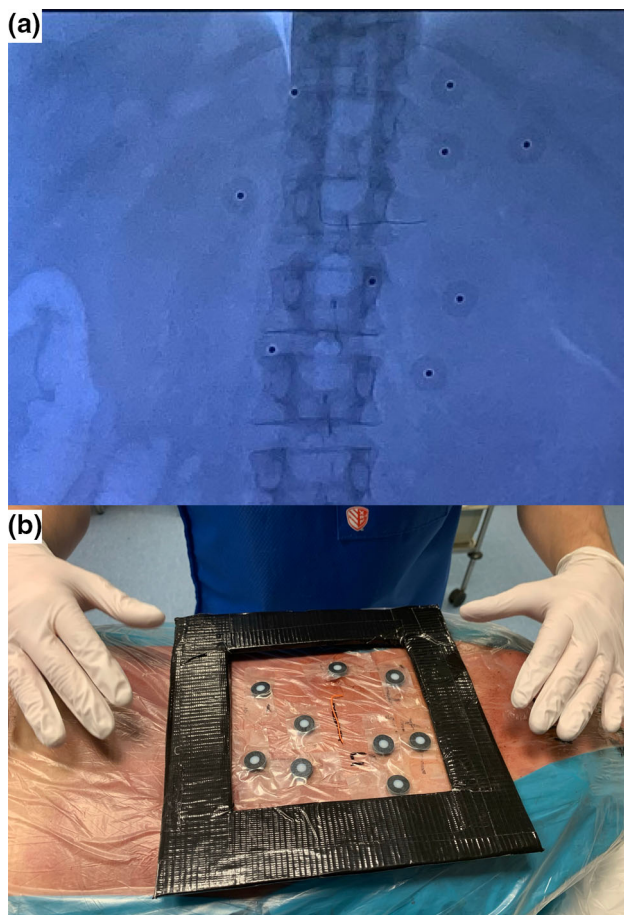
Fig. 1 Examples of the hybrid skin markers. **a** Fluoroscopic image capturing nine markers containing a radiopaque sphere, **b** nine markers applied to the skin

ferences were used as similarity measure (S) to compare the DRRs and fluoroscopic images (Eq. 1):

$$S = \sum_{i,j} \frac{A_v}{A_v + (I_{\text{diffV}}(i,j))^2} + \sum_{i,j} \frac{A_h}{A_h + (I_{\text{diffH}}(i,j))^2} \quad (1)$$

$$I_{\text{diffV}}(i,j) = \frac{dI_{\text{fl}}}{di} - s\frac{dI_{\text{DRR}}}{di} \quad (2)$$

$$I_{\text{diffH}}(i,j) = \frac{dI_{\text{fl}}}{dj} - s\frac{dI_{\text{DRR}}}{dj} \quad (3)$$

In Eq. (1), A_v represents a constant value for the maximum vertical similarity, and A_h a constant value for the maximum horizontal similarity. Equation (1) is structured according to the form $\frac{1}{1+x^2}$ to normalize the similarity between 0 (minimum similarity) and 1 (maximum similarity) [12, 17]. In Eqs. 2 and 3, the gradient-based measure first differentiates I_{fl} and I_{DRR} and then takes the vertical difference image (I_{diffV}) (Eq. 2) or the horizontal difference image (I_{diffH}) (Eq. 3). The gradient-based registration optimizes the geometrical position, $q = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T$, which is the input of the DRR calculation, using the Covariance Matric Adaption Evolution Strategy (CMA-ES) to avoid local minima's [18].

The registration algorithm executed the 3D2D registration twice. First, using one fluoroscopic image, a rough rigid registration was performed to reduce the search space for the second, more complicated per vertebrae registration based on two fluoroscopic images. During the first registration, two adjacent vertebrae were simultaneously matched.

Before the second registration, the pose-estimation algorithm calculated the pose difference—or relative angulation—between the two fluoroscopic images. The pose-estimation algorithm is an Umeyama-based algorithm which determines the pose based on the 3D model of the radiopaque spheres from the skin markers [19]. The automatic pose-estimation avoided that the absolute angulation and rotation of the mobile C-arm had to be calibrated or that its relative position to the study subject had to be tracked [20]. During the second registration, each vertebra was separately matched to both fluoroscopic images that have a fixed position with respect to each other. Per vertebrae registration corrects for possible spinal curvature changes and shifting of adjacent vertebrae relative to each other that could have occurred between the baseline CT scan and fluoroscopic image acquisition. We assessed the accuracy of the prototype 3D2D registration algorithm based on the second registration.

Reference standard

The 3D2D registration algorithm's accuracy was assessed by comparing it to a 3D3D registration based on a reference imaging device using cone-beam CT (CBCT) (Fig. 2). The CBCT was performed by a calibrated, motorized, ceiling-mounted, C-arm system installed in the surgical suite (Philips AlluraClarity FD20, Philips, Best, Netherlands).

The 3D3D registration was a rigid intensity-based registration that registered the 3D volume of each vertebra from the baseline CT scan to the position of the same vertebra as determined by the CBCT to correct for any local vertebral deformation regarding the 3D volume that had occurred between the baseline CT scan and the CBCT [12, 15]. After 3D3D registration and 3D2D registration, the 3D volumes of corresponding vertebrae were identical and because the marker model for both registrations was also identical, the registration accuracy of the 3D2D registration algorithm could be assessed (Fig. 2).

Accuracy measurement

The accuracy measure was the Target Registration Error (TRE) in millimeters. The TRE is a recommend measure to evaluate the accuracy of 2D to 3D registration [12] and is calculated by measuring the distance between similar points
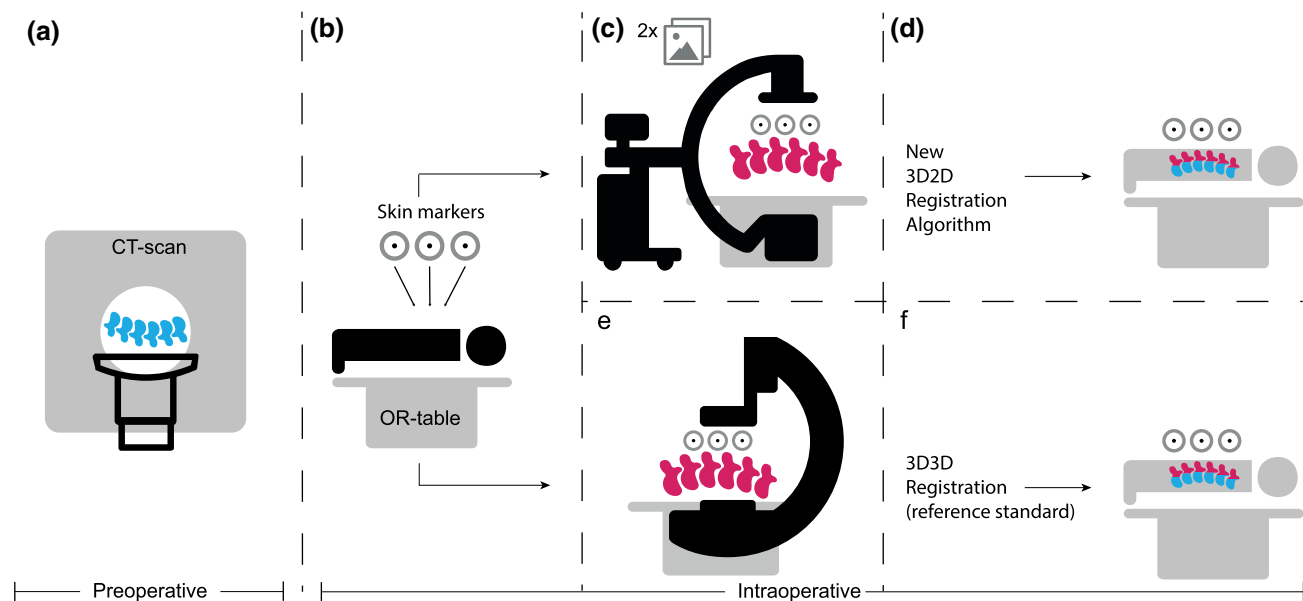
Fig. 2 Schematic overview of how we could assess the Target Registration Error (TRE) between the two registration algorithms. **a** Baseline Computed Tomography (CT) scan of the subject in supine position. **b** The study subject is placed prone on the surgical table and the radiopaque adhesive skin markers are randomly placed on the back. **c** Two fluoroscopic images are acquired capturing the vertebrae and the marker model. **d** The 3D2D registration algorithm performs the registration offline. **e** Cone-beam CT scan is acquired capturing the same vertebrae and identical marker model. **f** The 3D3D registration algorithm (reference standard) performs the registration offline
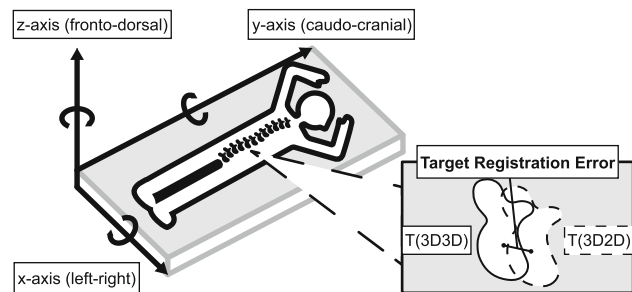

Fig. 3 Schematic overview of the three-dimensional coordinate system that was used for calculating the Target Registration Error (TRE). The TRE was the three-dimensional distance in millimeters between the centers of corresponding vertebrae as registered by the two registration methods

within corresponding vertebrae as registered by two registration methods in a 3D coordinate system [21]. The TRE was calculated between the centers of vertebral bodies (Fig. 3). The registration error was also visually inspected for every registration and was on sub-voxel level [12].

Additionally, although the TRE is a translational error measure, an indication of the rotational error was calculated. The rotational error was estimated by the TRE difference between two different points within one vertebra for which the TRE was separately calculated: if the 3D2D registration algorithm registered two different points in one vertebra with the same TRE (no TRE difference), it is unlikely that a rotational error occurred, but if the TRE differed between the two

points, some rotational error had occurred. The location of the two points, the centers of the vertebral body and the left pedicle, were automatically determined in the model-based segmentation of the spine. We chose the centers of the vertebral body and the left pedicle because they are usually within the trajectory of a pedicle screw.

Anatomical and radiologic parameters

Radiologic parameters included: the relative angulation between the fluoroscopic images, the detector-skin distance, the number of skin markers applied, and the captured level. Anatomical parameters included: the study subject and the anatomical region.

Fluoroscopic images were taken at different angles from the anterior–posterior (AP) position of the cadaveric specimen. The mobile C-arm always remained in neutral position toward the cranio-caudal direction (transversal plane). The mobile C-arm was manually rotated toward the latero-lateral direction (sagittal plane) to acquire fluoroscopic images in angles of $-45°$, $-32°$, $-30°$, $-28°$, $-20°$, $+20°$, $+28°$, $+30°$, $+32°$, or $+45°$. Pairs of fluoroscopic images with opposed angulations (e.g., $-20°$ and $+20°$) were used for 3D2D registration resulting in relative rotation angle differences (RAD) of 40°, 56°, 60°, 64°, and 90° between the 2D images. The detector-skin distance was either 20 cm or 30 cm measured from AP position. The number of skin
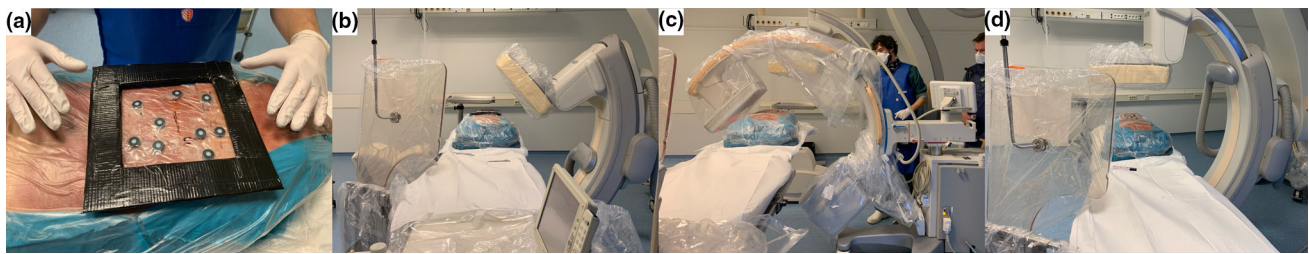
Fig. 4 Workflow of a test case. **a** Study subject is placed prone on the surgical table and the adhesive skin markers are randomly placed within a frame of 15 × 15 cm. **b** A Cone-Beam Computed Tomography Scan (CBCT) is performed. **c** Fluoroscopic images are acquired. **d** Another CBCT is performed from the exact same position as the previous CBCT

markers varied between seven to nine markers. Anatomical regions included three centered vertebrae in the thoracic (T4/T7/T10) and two in the lumbar spine (L1/L4). Each fluoroscopic image fully captured five vertebrae: a centered vertebra, and two adjacent vertebrae above and below. From the fluoroscopic images with L4 centered, only four vertebrae were registered because the algorithm did not register sacral vertebrae.

Workflow of the experiment

All frozen cadaveric specimens underwent CT in a supine position. Before transfer to the operating room, each torso was thawed at room temperature for 72 h to allow for spine curvature changes and the shifting of vertebrae. In the operating room, the cadaveric specimens were placed in a prone position on the surgical table (representing a realistic surgical position). All subjects underwent identical test cases consisting of a combination of anatomical and radiologic parameters during which the 2D images were acquired.

At the beginning of each test case, the centered vertebra was identified through fluoroscopy. Then, skin markers were placed randomly on the back of each torso within a frame of 15 × 15 cm around the centered vertebra. First, a CBCT was performed with the fixed C-arm system. Subsequently, fluoroscopic images were obtained with the mobile C-arm from various angles depending on the test case. At last, an additional CBCT was performed from the exact same position as the previous. This cycle was repeated for every test case (Fig. 4).

The second CBCT was performed to confirm that neither the study subject nor the markers had moved between the first CBCT and fluoroscopic image acquisition to ensure that both the 3D2D and the 3D3D registration had registered vertebrae with an identical marker model. The two CBCTs were compared offline and only the first CBCT was used for 3D3D registration.
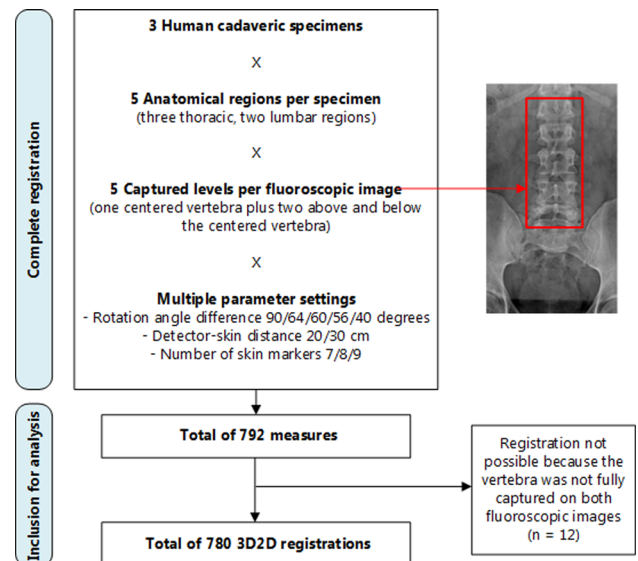

Fig. 5 Flowchart of 3D2D registrations

Statistical analysis

The primary outcome was the TRE of which the median, interquartile range (IQR), outliers, and the maximum value were assessed.

Descriptive statistics were summarized for all parameters divided into groups. Bivariate analysis was performed with Spearman's rank correlation test for the continuous parameter RAD. Inter-group differences for the other parameters were assessed using the unpaired Wilcoxon rank-sum test for two groups and the Kruskal–Wallis test for more than two groups. Post-hoc analysis was conducted with a Bonferroni correction.

The parameters study subject, RAD, detector-skin distance, anatomical region, and captured level were included in a multiple regression analysis with backward elimination to determine parameters independently affecting the regis-

tration accuracy. Model assumptions were checked using histograms and quantile–quantile plots of residuals. The number of skin markers was not part of the multivariable analysis because this parameter was not assessed for every test case. Test cases with seven or nine markers were only assessed with a RAD of 60° and a detector-skin distance of 20 cm.

Additionally, an indication of the rotational error was calculated and the TRE was assessed for all three 3D axes to analyze the direction of the error in more depth [21].

p values of <0.05 were used to denote statistical significance. TRE distributions were displayed with histograms and boxplots. Outliers were defined as the values above or below the upper or lower fences. The upper and lower fences represent values more and less than the 3rd and 1st quartiles, respectively, by 1.5 times the difference between the 3rd and 1st quartiles [22]. All statistical analyses were performed with R statistical software (R, version 4.0.3).

Results

In total, the 3D2D registration algorithm successfully registered 780 vertebrae. Twelve vertebrae (two thoracic and ten lumbar) could not be registered as they were not fully captured on both fluoroscopic images. The unsuccessful registrations could not be corrected because the 3D2D registration was performed offline at a different moment (Fig. 5).

The algorithm had a median TRE of 0.51 mm [IQR 0.32–0.73 mm], and a maximum TRE of 2.06 mm. The algorithm had eighteen outliers, which all occurred in registrations performed with a RAD of 40° and in the two subjects with a larger waist circumference: subject A had a circumference of 121 cm and subject B had a circumference of 96 cm (Fig. 6) (Table 1 shows the waist circumference of the subjects).

In bivariate analysis, no inter-group difference was observed regarding the TRE for the parameter captured level ($p = 0.473$). The parameters study subject ($p < 0.001$), anatomical region ($p = 0.003$), detector-skin distance ($p = 0.007$) and number of markers ($p = 0.021$) all showed statistically significant inter-group differences (Table 2). Post-hoc analysis revealed that the median TRE of subject C was lower than that of subjects A and B. The registrations with eight skin markers had a higher median TRE than the registrations with seven and nine markers.

The TRE was best predicted by the multiple regression model with RAD, anatomical region, and study subject as predictors (F-statistic (4, 491) = 16.22, $p < 0.001$, adj $R^2 = 0.11$). The parameter RAD was the most important predictive factor with a standardized coefficient β of -0.41 ($p < 0.001$), i.e., the larger the RAD (with a current maximum of 90°), the lower the TRE. The 3D2D registration algorithm was

more accurate in thoracic vertebrae ($p = 0.004$) and subject C ($p = 0.003$). The standardized coefficients β were small for both thoracic vertebrae (-0.10), and subject C (-0.12) (Table 3).

Because the registration accuracy of the algorithm increased (thus, a lower TRE) with the increase in the RAD (larger angulations), the registration accuracy was explored by cumulatively excluding registrations with a RAD from small to large. Excluding registrations performed with a RAD of 40° ($n = 143$) resulted in a median TRE of 0.47 mm [IQR 0.30–0.67 mm] for the remaining 637 registrations. Additionally excluding registrations performed with a RAD of 56° ($n = 141$) resulted in a median TRE of 0.45 mm [IQR 0.28–0.63 mm] for 496 registrations, and subsequently excluding registrations performed with a RAD of 60° ($n = 265$) resulted in a median TRE of 0.42 mm [IQR 0.25–0.59 mm] for 213 registrations. The median TRE was 0.33 mm [IQR 0.22–0.45 mm] for the 72 registrations performed with a RAD of 90°. The registrations in thoracic vertebrae ($n = 493$) had a median TRE of 0.50 mm [IQR 0.29–0.71 mm] and the registrations in lumbar vertebrae ($n = 287$) a median of 0.53 mm [IQR 0.38–0.80 mm].

Translational errors for the x-axis (left–right direction) and the y-axis (caudo-cranial direction) were small, and normally distributed. The algorithm was the least accurate on the z-axis (fronto-dorsal direction): the TRE was skewed toward the dorsal side of the study subjects with a median of 0.39 mm [IQR 0.17–0.65 mm] (Fig. 7). The TRE difference between two points within one vertebra (rotational error) was normally distributed with a median of 0.01 mm [-0.04 to 0.05 mm] (Fig. 8).

Discussion

Key results

We assessed the accuracy of a new 3D2D registration algorithm based on a non-invasive skin marker model to be used for navigated spine surgery. The algorithm registered five adjacent vertebrae in the thoracic and lumbar spine from three human cadaveric specimens. When all 780 registrations were included, the algorithm had a median TRE of 0.51 mm [IQR 0.32–0.73 mm] and a maximum TRE of 2.06 mm. The algorithm registered all five adjacent vertebrae with similar accuracy.

Interpretation

Few navigation systems for spine surgery using an integrated 3D2D registration algorithm are commercially available, for example, the ExcelsiusGPS by GlobusMedical and the Mazor X Stealth Edition by Medtronic [13]. The registra-
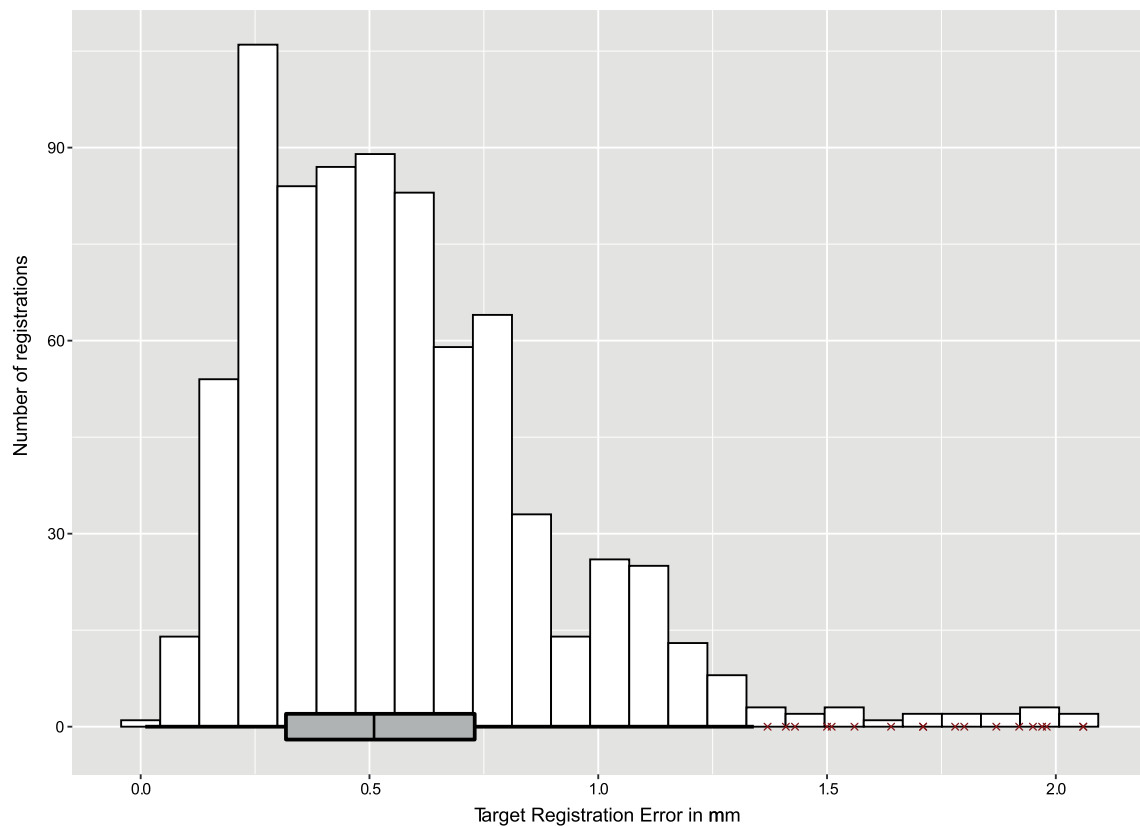
Fig. 6 Distribution of the TRE in millimeter ($n = 780$). The median TRE was 0.51 mm [IQR 0.32–0.73 mm]. All 18 outliers were performed with a rotation angle difference of 40°, 12 were in thoracic vertebrae, 11 in subject B and 7 in subject A
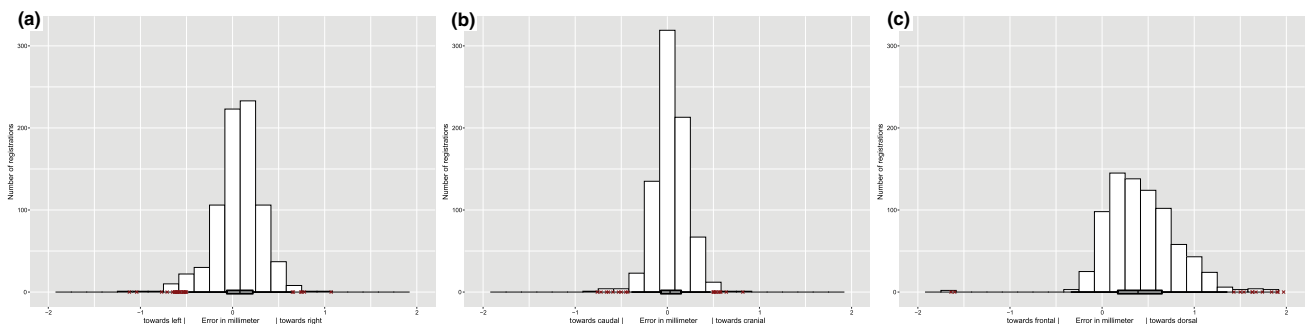

Fig. 7 Distributions of the TRE in millimeter for all three dimensions ($n = 780$). The largest TREs were in the z-axis and were skewed toward the dorsal side of the patient. **a** The TRE in millimeter regarding the x-axis. The median TRE was 0.08 mm [IQR $-$ 0.07 to 0.22 mm]. In total, 31 outliers occurred, which were in the registrations performed with rotation angle differences of 40° ($n = 11$), 56° ($n = 2$), 60° ($n = 7$), 64° ($n = 8$), and 90° ($n = 3$), thoracic vertebrae ($n = 11$), subject A ($n = 14$) and subject B ($n = 17$). **b** The TRE in millimeter regarding the y-axis. The median TRE was 0.03 mm [IQR $-$ 0.07 to 0.15 mm].

In total, 19 outliers occurred, which were in the registrations performed with rotation angle differences of 40° ($n = 9$), 56° ($n = 4$), 60° ($n = 2$), 64° ($n = 2$), and 90° ($n = 2$), thoracic vertebrae ($n = 1$), subject A ($n = 15$), subject B ($n = 2$) and subject C ($n = 2$). **c** The TRE in millimeter regarding the z-axis. The median TRE was 0.39 mm [IQR 0.17–0.65 mm]. In total, 14 outliers occurred, which were in the registrations performed with a rotation angle difference of 40° ($n = 14$), thoracic vertebrae ($n = 12$), subject A ($n = 7$) and subject B ($n = 7$)

tion setup of the algorithm under study may be more simple and intuitive than the existing algorithms because no bone-mounted or table-mounted instruments are needed, nor is any equipment required for calibrating and tracking the C-arm. The non-invasive skin markers do not require surgeons to

make additional incisions for registration instruments and, because the C-arm does not have to be tracked or calibrated, allow for quick registration and re-registration with minimum disruption of the routine surgical workflow. In addition, surgeons can easily replace the markers to register other ver-

Table 2 Descriptive statistics of all parameters on TRE

	Number of registrations	Median TRE in mm ($\pm$IQR)	Statistical test	p value
Study subject				
Subject A	259	0.56 [0.33–0.75]	Kruskal–Wallis	<u><0.001</u>
Subject B	259	0.53 [0.38–0.81]		
Subject C	262	0.45 [0.27–0.64]		
Anatomical region				
Thoracic	493	0.50 [0.29–0.71]	Mann–Whitney U	<u>0.003</u>
Lumbar	287	0.53 [0.38–0.80]		
*Rotation angle difference**				
40°	143	0.79 [0.52–1.12]	Spearman's rank correlation	<u><0.001</u> ($\rho = -0.34$)
56°	141	0.51 [0.38–0.74]		
60°	283	0.47 [0.32–0.69]		
64°	141	0.49 [0.27–0.64]		
90°	72	0.33 [0.22–0.45]		
Detector-skin distance from AP position				
20 cm	499	0.49 [0.31–0.69]	Mann–Whitney U	<u>0.007</u>
30 cm	281	0.56 [0.33–0.79]		
Captured level				
Centered vertebra	165	0.48 [0.28–0.75]	Mann–Whitney U	0.473
One level from centered vertebra	328	0.49 [0.33–0.71]		
Two levels from centered vertebra	287	0.53 [0.32–0.77]		
*Number of skin markers***				
7 markers	70	0.48 [0.35–0.63]	Kruskal–Wallis	<u>0.021</u>
8 markers	72	0.6 [0.32–0.92]		
9 markers	71	0.42 [0.29–0.57]		

*To calculate the median and IQR, rotation angle differences were divided into groups according to the amount of degrees; however, to assess the correlation with the TRE, the rotation angle difference was considered a continuous variable

**For a direct comparison, only the registrations ($n = 213$) with a rotation angle difference of 60° and a detector distance-skin of 20 cm were included for this analysis. All other registrations ($n = 639$) were performed with 8 markers and were performed with a different relative angulation and/or detector-skin distance

Underlined p-values indicate a stastically significant difference ($p < 0.05$)

tebrae and expand the vertebral levels they can treat during one procedure.

Other non-invasive 3D2D registration algorithms for spine surgery serve a different intended clinical application. For instance, some algorithms aim to automatically label vertebrae at the start of surgery to prevent wrong-level interventions [23, 24], while other algorithms aim to intraoperatively verify the 3D position of pedicle screws [25–27]. Our non-invasive 3D2D registration algorithm aims to provide surgeons navigation for interventions such as pedi-

cle screw insertion. However, in future, applications from existing algorithms may be added to the algorithm under study, such as automatically labeling vertebrae to prevent wrong-level surgery and 3D2D registration for intraoperatively verifying screw positions.

When the 3D2D registration algorithm is integrated into a complete navigation system, the system will contain new elements such as a camera unit and specific surgical tools that could potentially affect the overall precision. Based on the accuracy of available navigation systems, 3D2D naviga-

Table 3 Multiple regression analysis of parameters influencing the accuracy of 3D2D registration after backward selection

Independent variable	Unstandardized coefficients		Standardized coefficients β	p value
	B	Standard error		
Intercept	1.24	0.06		
Rotation angle difference	− 0.01	0.00	− 0.41	<u><0.001</u>
Anatomical region				
Thoracic	− 0.07	0.03	− 0.10	<u>0.004</u>
Lumbar	Reference value	Reference value	Reference value	Reference value
Study subject				
Subject A	Reference value	Reference value	Reference value	Reference value
Subject B	0.06	0.3	0.08	0.06
Subject C	− 0.09	0.03	− 0.12	<u>0.003</u>

Adjusted $R^2 = 0.11$, F-statistic $(4, 491) = 16.22$, p value of the model ≤ 0.001
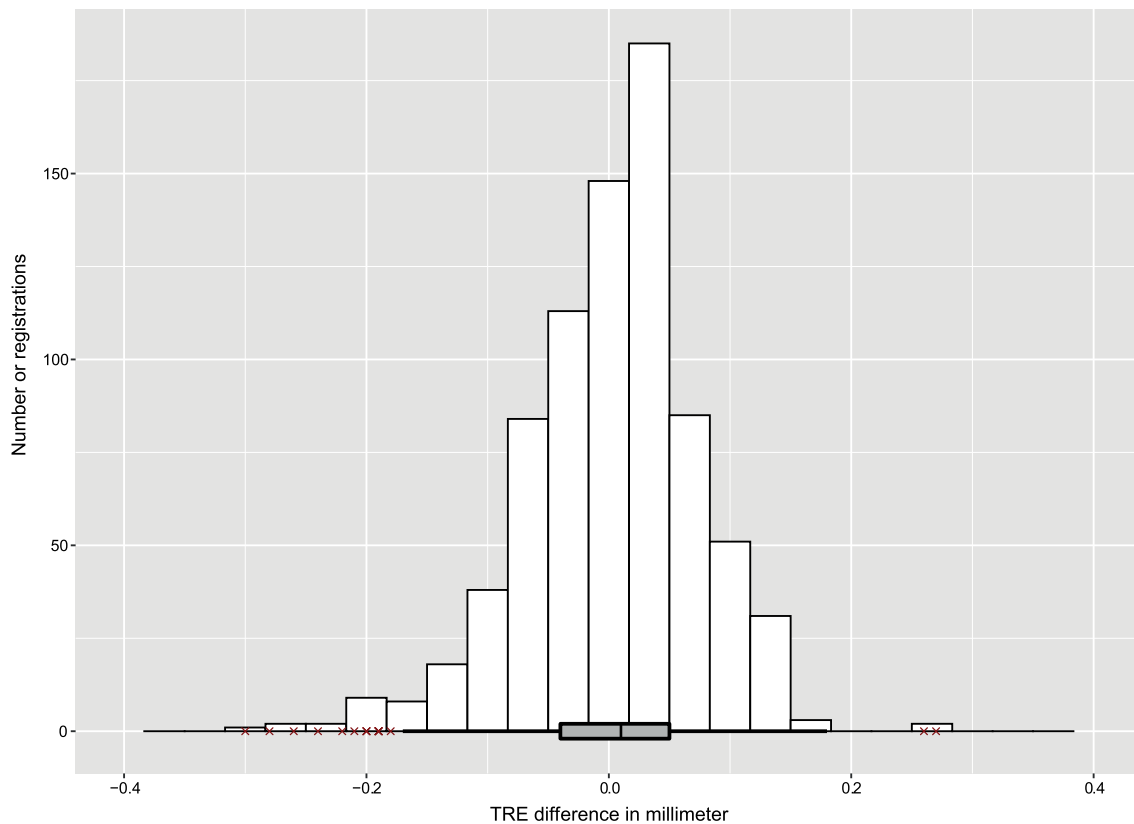
Fig. 8 The TRE difference between two points (left pedicle and center of vertebral body) within one vertebra ($n = 780$). The median TRE difference was 0.01 mm [IQR − 0.04 to 0.05 mm]. In total, 17 outliers occurred, which were in the registrations performed with rotation angle differences of 40° ($n = 3$), 56° ($n = 3$), 60° ($n = 8$) and 64° ($n = 3$), thoracic vertebrae ($n = 12$), subject A ($n = 6$), subject B ($n = 3$) and subject C ($n = 8$)

tion may be considered feasible for clinical practice if no major outliers occur and, if pedicle screws breach the pedicle wall, the breach is less than 2 mm [7, 28, 29]. When including all registrations, 95% of the TREs were below 1.16 mm but outliers existed up to 2.06 mm. Registrations with a RAD of 40° were the least accurate and responsible for all outliers. When excluding registrations with a RAD of 40° (thus keeping RADs between 56° and 90°), 95% of the TREs were below 1.02 mm and the maximum TRE was 1.28 mm. Although future studies should explore the exact boundaries of the algorithm, a RAD of 40° may be too small for safe application of 3D2D navigation. Commercially avail-

able systems often require a specific RAD of 90° for 3D2D registration [8, 13]. The practical impact may be small, but if the current algorithm remains accurate with RADs between 56° and 90°, it may provide a more flexible and safer future workflow. Registration does not require one specific angulation between the 2D images and less movement of the C-arm reduces the chance of breaking the sterile field.

The algorithm performed best in subject C and in thoracic vertebrae. Subject C had less attenuating tissue because of a smaller waist circumference and the thoracic vertebrae had less attenuating tissue than the lumbar vertebrae, which improved the quality of the fluoroscopic images [30, 31]. Also, the CT scans of the subjects with a larger waist circumference had a relatively larger voxel size. Because each CT scan had a field of view that contained the whole study subject and was not limited to the vertebral column, subject A had almost twice the voxel size of subject C. However, during pre-experimental test runs using a phantom, changing the field of view and, therefore, the relative voxel size did not alter the algorithm's accuracy. Still, if obesity becomes extreme and bone density also decreases severely, the algorithm's accuracy might compromise patient safety [8].

Bivariate analysis revealed that the algorithm was less accurate with eight skin markers than seven or nine markers, but a logical explanation is lacking. The algorithm needs a minimum of five skin markers for registration, and during the experiment, all markers were randomly applied within the same frame of 15×15 cm. If registration accuracy depends on the number of markers used, one would expect the registration accuracy to also decrease for nine markers indicating that there is a maximum number of markers for accurate registration. Future experiments should include different numbers of markers in multivariable analysis to assess if registration accuracy directly depends on the number of markers used.

In-depth analysis of translational errors showed that most errors occurred in the z-axis (toward the dorsal side of the patient). The algorithm had the lowest capture range over this axis because the study subjects were prone on the surgical table and the maximum angulation of the mobile C-arm was 45° from AP toward the lateral side of the study subject. For example, all outliers were registrations with the lowest fluoroscopic capture range (angulation of 40°). One could argue that an error in this direction is less crucial for pedicle screw placement as critical structures like the spinal cord are located medially (x-axis). However, a high registration accuracy in all axes becomes necessary if 3D2D navigation is applied for different purposes, such as navigated vertebral biopsy. Future studies may expand the latero-lateral angulations and experiment with angulation in other directions, such as cranio-caudal angulations, to optimize the registration accuracy [32].

In the current study, we only evaluated CT scans with parallel slices of 0.67 mm. The algorithm was developed to work with any CT scan vendor, so it can easily integrate in clinical practice. The algorithm would integrate even more easily if it can also cope with various scan protocols. The current algorithm interpolates the unknown values of pixels lying between slices with a known value to generate a complete 3D volume, but the present study did not assess at what slice thickness it becomes inaccurate. In the supplementary data, we explored the algorithm's accuracy using post-processed baseline CT scans with a slice thickness up to 5.0 mm. The supplementary figures indicate that the algorithm remains similarly accurate with slice thicknesses up to 2.0 mm but the exact limits should be explored using a higher number of scans (Supplement 1).

Limitations

A drawback of 3D2D registration, in general, might be the assumption that, at the time of intraoperative registration, the 3D volume of vertebrae has not altered since the baseline CT scan. Although the algorithm registered each vertebra separately and accounted for positional changes of individual vertebrae up to 5° rotation around the x-axis and translations up to 2 mm over the z-axis, it may be necessary to maintain a short interval between the baseline CT and operative treatment regarding, for instance, the 3D volume of collapsed vertebrae may change.

Another limitation was that the rotational error was not exactly calculated but estimated by the TRE difference between two points in a single vertebra representing the trajectory of a pedicle screw. The maximum difference was 0.30 mm, and apart from 17 outliers, all values were smaller than 0.19 mm, suggesting that the rotational error was small. Still, future studies need to assess the exact rotational error, for example, by evaluating the angle deviation from the planned path of a navigationally inserted pedicle screw.

Furthermore, the circumstances of the current experiments differed from normal circumstances because the subjects were frozen at the time of baseline CT and the 2D images were acquired in a thawed state. Also, the 3D2D registration was performed offline on a different moment than the 2D image acquisition. As a result, the unsuccessful registration of twelve vertebrae could not be restored but it was unlikely that these would have changed the study results. Test cases using two different 2D images to register the same vertebrae were accurate. Eventually, the navigation system will not be limited by unsuccessful registrations because the registration is executed directly in the operating room, obliging the surgeon to re-register vertebrae when initial registration fails.

Integrating the algorithm into a navigation system also generates new challenges, such as intraoperative 2D image acquisition and registration time. For example, during surgery, between the acquisitions of the two 2D images, the
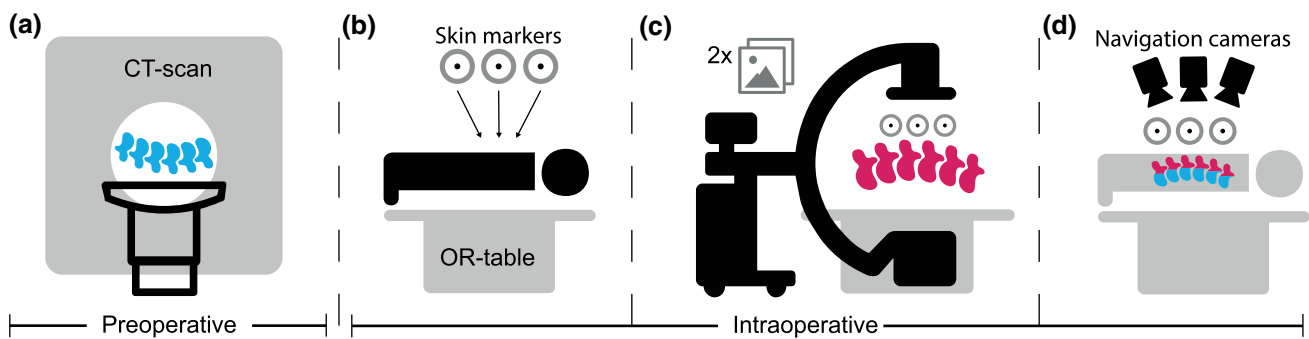
Fig. 9 Intended workflow of 3D2D navigation. **a** A preoperative CT scan is acquired (in supine position). **b** Intraoperatively, the radiopaque adhesive skin marker are applied on the back of the patient. **c** Intraoperatively, 3D2D registration is performed with the preoperative CT data, and two fluoroscopic images capturing the marker model and vertebrae. **d** After 3D2D registration, the marker model is registered to optical cameras track the marker model, thus indirectly tracking the vertebrae, to allow for intraoperative navigation

algorithm will require a patient's breath-hold (under anesthesia) so that the marker model's relative position to the vertebrae remains the same. A breath-hold under anesthesia urges the surgical team to acquire the 2D images in a short time and may be considered a complicated intervention. Another challenge for clinical implementation is keeping registration time as short as possible, preferably under a few minutes. In the current study, we did not calculate the exact registration time but the algorithm performed all registrations on a regular computer in up to a few minutes. No optimizations for limiting registration time have been implemented yet.

Intended workflow of 3D2D navigation

Spinal navigation tracks the patient's bony anatomy (such as vertebrae) and compatible surgical tools in 3D during surgery. Figure 9 shows our concept of 3D2D navigation: optical cameras indirectly track the patient's vertebrae by tracking a marker model applied to the back of the patient. The marker model consists of a pattern of hybrid adhesive skin markers (radiopaque and optical).

First, the 3D volume of vertebrae is obtained before the surgical procedure using data from a previous CT scan (without markers). Then, in the operating room, after the incision, the adhesive skin markers are applied when the patient is prepared to undergo pedicle screw insertion. The marker model is the reference to relate 3D2D registration (fluoroscopic space) to optical navigation (navigated space). The 3D2D registration algorithm registers the reconstructed 3D volume of vertebrae from the pre-existing CT scan to their intraoperative position in two fluoroscopic 2D images. The two fluoroscopic images must both fully capture the marker model and the vertebrae that must be registered so that the navigation system can relate fluoroscopic space to the navigated space. Using optical cameras, the navigation system

indirectly tracks the position of the registered vertebrae based on the marker model (Fig. 9).

Conclusion

We studied the accuracy of a new 3D2D registration algorithm based on a non-invasive skin marker model. The algorithm registered five adjacent vertebrae in the thoracic and lumbar spine, and showed a maximum target registration error of approximately 2 mm. All five adjacent vertebrae were registered with similar accuracy. To further evaluate its potential for navigated spine surgery, the algorithm may now be integrated into a complete navigation system.

Declarations

Conflict of interest MSvM, RAMW, RH and JJR are employed by Philips Healthcare. JJV is cofounder and stockholder of SentryX BV.

Ethics approval All experiments were in compliance with ethical guidelines for human cadaveric studies.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Lu VM, Alvi MA, Goyal A, Kerezoudis P, Bydon M (2018) The potential of minimally invasive surgery to treat metastatic spinal disease versus open surgery: a systematic review and meta-analysis. World Neurosurg 112:e859–e868. https://doi.org/10.1016/j.wneu.2018.01.176

2. Goldstein CL, Phillips FM, Rampersaud YR (2016) Comparative effectiveness and economic evaluations of open versus minimally invasive posterior or transforaminal lumbar interbody fusion: a systematic review. Spine (Phila Pa 1976) 41(Suppl 8):S74-89. https://doi.org/10.1097/BRS.0000000000001462

3. Marie-Hardy L, Wolff S, Frison-Roche A, Bergère A, Khalifé M, Riouallon G (2020) Minimal invasive management of early revision after minimal invasive posterior lumbar fusion. Orthop Traumatol Surg Res 106:1209–1214. https://doi.org/10.1016/j.otsr.2020.04.003

4. Yang Z, Yang Y, Zhang Y, Zhang Z, Chen Y, Shen Y, Han L, Xu D, Sun H (2015) Minimal access versus open spinal surgery in treating painful spine metastasis: a systematic review. World J Surg Oncol 13:68. https://doi.org/10.1186/s12957-015-0468-y

5. Kumar N, Malhotra R, Maharajan K, Zaw AS, Wu PH, Makandura MC, Po Liu GK, Thambiah J, Wong H-K (2017) Metastatic spine tumor surgery: a comparative study of minimally invasive approach using percutaneous pedicle screws fixation versus open approach. Clin Spine Surg 30:E1015–E1021. https://doi.org/10.1097/BSD.0000000000000400

6. Burström G, Persson O, Edström E, Elmi-Terander A (2021) Augmented reality navigation in spine surgery: a systematic review. Acta Neurochir (Wien) 163:843–852. https://doi.org/10.1007/s00701-021-04708-3

7. Perdomo-Pantoja A, Ishida W, Zygourakis C, Holmes C, Iyer RR, Cottrill E, Theodore N, Witham TF, Lo S-FL (2019) Accuracy of current techniques for placement of pedicle screws in the spine: a comprehensive systematic review and meta-analysis of 51,161 screws. World Neurosurg 126:664-678.e3. https://doi.org/10.1016/j.wneu.2019.02.217

8. Lieberman IH, Kisinde S, Hesselbacher S (2020) Robotic-assisted pedicle screw placement during spine surgery. JBJS Essent Surg Tech 10:e0020–e0020. https://doi.org/10.2106/JBJS.ST.19.00020

9. Härtl R, Lam KS, Wang J, Korge A, Kandziora F, Audigé L (2013) Worldwide survey on the use of navigation in spine surgery. World Neurosurg 79:162–172. https://doi.org/10.1016/j.wneu.2012.03.011

10. Verhey JT, Haglin JM, Verhey EM, Hartigan DE (2020) Virtual, augmented, and mixed reality applications in orthopedic surgery. Int J Med Robot 16:e2067. https://doi.org/10.1002/rcs.2067

11. Malham GM, Wells-Quinn T (2019) What should my hospital buy next?-Guidelines for the acquisition and application of imaging, navigation, and robotics for spine surgery. J Spine Surg (Hong Kong) 5:155–165. https://doi.org/10.21037/jss.2019.02.04

12. van de Kraats EB, Penney GP, Tomazevic D, van Walsum T, Niessen WJ (2005) Standardized evaluation methodology for 2-D-3-D registration. IEEE Trans Med Imaging 24:1177–1189. https://doi.org/10.1109/TMI.2005.853240

13. Avrumova F, Sivaganesan A, Alluri RK, Vaishnav A, Qureshi S, Lebl DR (2021) Workflow and efficiency of robotic-assisted navigation in spine surgery. HSS J 17:302–307. https://doi.org/10.1177/15563316211026658

14. Burström G, Nachabe R, Homan R, Hoppenbrouwers J, Holthuizen R, Persson O, Edström E, Elmi-Terander A (2020) Frameless patient tracking with adhesive optical skin markers for augmented reality surgical navigation in spine surgery. Spine (Phila Pa 1976) 45:1598–1604. https://doi.org/10.1097/BRS.0000000000003628

15. Burström G, Buerger C, Hoppenbrouwers J, Nachabe R, Lorenz C, Babic D, Homan R, Racadio JM, Grass M, Persson O, Edström E, Elmi Terander A (2019) Machine learning for automated 3-dimensional segmentation of the spine and suggested placement of pedicle screws based on intraoperative cone-beam computer tomography. J Neurosurg Spine 31:147–154. https://doi.org/10.3171/2018.12.SPINE181397

16. Markelj P, Tomaževič D, Likar B, Pernuš F (2012) A review of 3D/2D registration methods for image-guided interventions. Med Image Anal 16:642–661. https://doi.org/10.1016/j.media.2010.03.005

17. Penney GP, Weese J, Little JA, Desmedt P, Hill DL, Hawkes DJ (1998) A comparison of similarity measures for use in 2-D-3-D medical image registration. IEEE Trans Med Imaging 17:586–595. https://doi.org/10.1109/42.730403

18. Hansen N (2016) The CMA evolution strategy: a tutorial. CoRR abs/1604.0: https://doi.org/10.48550/arXiv.1604.00772

19. Umeyama S (1991) Least-squares estimation of transformation parameters between two point patterns. IEEE Trans Pattern Anal Mach Intell 13:376–380. https://doi.org/10.1109/34.88573

20. Wang P, Xu G, Cheng Y, Yu Q (2018) A simple, robust and fast method for the perspective-n-point problem. Pattern Recognit Lett 108:31–37. https://doi.org/10.1016/j.patrec.2018.02.028

21. Tomazevic D, Likar B, Pernus F (2004) "Gold standard" data for evaluation and comparison of 3D/2D registration methods. Comput Aided Surg Off J Int Soc Comput Aided Surg 9:137–144. https://doi.org/10.3109/10929080500097687

22. Kwak SK, Kim JH (2017) Statistical data preparation: management of missing values and outliers. Korean J Anesthesiol 70:407–411. https://doi.org/10.4097/kjae.2017.70.4.407

23. Lo S-FL, Otake Y, Puvanesarajah V, Wang AS, Uneri A, De Silva T, Vogt S, Kleinszig G, Elder BD, Goodwin CR, Kosztowski TA, Liauw JA, Groves M, Bydon A, Sciubba DM, Witham TF, Wolinsky J-P, Aygun N, Gokaslan ZL, Siewerdsen JH (2015) Automatic localization of target vertebrae in spine surgery: clinical evaluation

of the LevelCheck registration algorithm. Spine (Phila Pa 1976) 40:E476–E483. https://doi.org/10.1097/BRS.0000000000000814

24. De Silva T, Uneri A, Ketcha MD, Reaungamornrat S, Kleinszig G, Vogt S, Aygun N, Lo S-F, Wolinsky J-P, Siewerdsen JH (2016) 3D–2D image registration for target localization in spine surgery: investigation of similarity metrics providing robustness to content mismatch. Phys Med Biol 61:3009–3025. https://doi.org/10.1088/0031-9155/61/8/3009

25. Newell R, Esfandiari H, Anglin C, Bernard R, Street J, Hodgson AJ (2018) An intraoperative fluoroscopic method to accurately measure the post-implantation position of pedicle screws. Int J Comput Assist Radiol Surg 13:1257–1267. https://doi.org/10.1007/s11548-018-1732-8

26. Uneri A, De Silva T, Goerres J, Jacobson MW, Ketcha MD, Reaungamornrat S, Kleinszig G, Vogt S, Khanna AJ, Osgood GM, Wolinsky J-P, Siewerdsen JH (2017) Intraoperative evaluation of device placement in spine surgery using known-component 3D–2D image registration. Phys Med Biol 62:3330–3351. https://doi.org/10.1088/1361-6560/aa62c5

27. Esfandiari H, Anglin C, Guy P, Street J, Weidert S, Hodgson AJ (2019) A comparative analysis of intensity-based 2D–3D registration for intraoperative use in pedicle screw insertion surgeries. Int J Comput Assist Radiol Surg 14:1725–1739. https://doi.org/10.1007/s11548-019-02024-x

28. Kuo K-L, Su Y-F, Wu C-H, Tsai C-Y, Chang C-H, Lin C-L, Tsai T-H (2016) Assessing the intraoperative accuracy of pedicle screw placement by using a bone-mounted miniature robot system through secondary registration. PLoS ONE 11:1–11. https://doi.org/10.1371/journal.pone.0153235

29. Fayed I, Tai A, Triano M, Sayah A, Makariou E, Voyadzis J-M, Sandhu FA (2020) Robot-assisted percutaneous pedicle screw placement: evaluation of accuracy of the first 100 screws and comparison with cohort of fluoroscopy-guided screws. World Neurosurg 143:e492–e502. https://doi.org/10.1016/j.wneu.2020.07.203

30. Uppot RN, Sahani DV, Hahn PF, Kalra MK, Saini SS, Mueller PR (2006) Effect of obesity on image quality: fifteen-year longitudinal study for evaluation of dictated radiology reports. Radiology 240:435–439. https://doi.org/10.1148/radiol.2402051110

31. Martin C (2007) Optimisation in general radiography. Biomed Imaging Interv J 3:e18. https://doi.org/10.2349/biij.3.2.e18

32. Uneri A, Otake Y, Wang AS, Kleinszig G, Vogt S, Khanna AJ, Siewerdsen JH (2014) 3D–2D registration for surgical guidance: effect of projection view angles on registration accuracy. Phys Med Biol 59:271–287. https://doi.org/10.1088/0031-9155/59/2/271