ORIGINAL ARTICLE

Comparative study of application accuracy of two frameless neuronavigation systems: experimental error assessment quantifying registration methods and clinically influencing factors

**Dimitrios Paraskevopoulos · Andreas Unterberg ·
Roland Metzner · Jens Dreyhaupt · Georg Eggers ·
Christian Rainer Wirtz**

Abstract This study aimed at comparing the accuracy of two commercial neuronavigation systems. Error assessment and quantification of clinical factors and surface registration, often resulting in decreased accuracy, were intended. Active (Stryker Navigation) and passive (VectorVision Sky, Brain-LAB) neuronavigation systems were tested with an anthropomorphic phantom with a deformable layer, simulating skin and soft tissue. True coordinates measured by computer numerical control were compared with coordinates on image data and during navigation, to calculate software and system accuracy respectively. Comparison of image and navigation coordinates was used to evaluate navigation accuracy. Both systems achieved an overall accuracy of <1.5 mm. Stryker achieved better software accuracy, whereas BrainLAB better system and navigation accuracy. Factors with conspicuous influence ($P<0.01$) were imaging, instrument replacement, sterile cover drape and geometry of instruments. Precision data indicated by the systems did not reflect measured accuracy in general. Surface matching resulted in no improvement of accuracy, confirming former studies. Laser registration showed no differences compared to conventional pointers. Differences between the two systems were limited. Surface registration may improve inaccurate point-based registrations but does not in general affect overall accuracy. Accuracy feedback by the systems does not always match with true target accuracy and requires critical evaluation from the surgeon.

Keywords Registration · Surface matching · Neuronavigation · Frameless stereotaxy · Image-guided surgery · Accuracy

D. Paraskevopoulos · A. Unterberg · C. R. Wirtz
Department of Neurosurgery, University Hospital Heidelberg,
Ruprecht-Karls University Heidelberg,
Heidelberg, Germany

R. Metzner · J. Dreyhaupt
Institute for Medical Biometry and Informatics, University
Hospital Heidelberg, Ruprecht-Karls University Heidelberg,
Heidelberg, Germany

G. Eggers
Department of Maxillofacial Surgery, University Hospital
Heidelberg, Ruprecht-Karls University Heidelberg,
Heidelberg, Germany

C. R. Wirtz
Department of Neurosurgery, BKH-Günzburg, University Ulm,
Ulm, Germany

D. Paraskevopoulos (✉)
Nathanail 1,
54644 Thessaloniki, Greece
e-mail: dimitris.parask@gmail.com

Introduction

Optical neuronavigation systems have achieved widespread acceptance, and manufacturers claim reliability and accuracy. Nevertheless, the technical complexity of a navigational environment requires experimental evaluation of different types of errors, knowledge of registration techniques, awareness of the device's limitations and reliability of accuracy feedback. Surgeons often use the square root of the mean squared deviation of registration (RMS), which is calculated and displayed by most systems, as accuracy feedback. However, some authors do not regard the registration error calculated this way as a trustworthy measure of accuracy [14,

41]. The initiative for this study arose from the observation that despite good values given from the system, such as mean distance and RMS, decreased accuracy may occur after surface registration.

In the literature, there are only few experimental phantom studies with the Stryker (Stryker Navigation, Leibinger, Freiburg, Germany) and the BrainLAB VectorVision (Brain-LAB, Feldkirchen-Munich, Germany) systems. In this study, we seek to evaluate and compare the accuracy of these two commercially available systems and quantify clinically relevant influencing factors. Most related studies are evaluations of registration accuracy [8] or evaluate navigation accuracy on the surface alone but not of deep lying target points [41]. Accuracy data on BrainLAB are already reported [1, 8, 9, 20, 31, 40]; however most authors refer to an estimation of clinical accuracy rather than conducting experimental quantitative studies. Accuracy evaluations of the Stryker system are rare [34]. A comparative study of these two systems, which represent two different types of systems, active and passive, has not been published so far. The development of a phantom, simulating skin and soft tissue motility and viscoelastic deformation properties is an innovation, allowing testing experimentally the influence of surface registration and fiducial movement, as well as the influence of soft tissue behaviour when touched by a pointing instrument and during image acquisition. To the best of our knowledge, there is no other literature report of an anthropomorphic head phantom with similar properties of soft tissue simulation for the specific research question [46].

Materials and methods

Stryker® (Stryker Navigation, Leibinger, Freiburg, Germany) is an active optical system consisting of a workstation with planning software, an infrared camera frame with optical sensors, a Patient Tracker and a Pointer with light emitting diodes (LED) (Fig. 1) [34]. VectorVision® Sky (BrainLAB, Feldkirchen-Munich, Germany) is a passive reflective marker system (Fig. 2), consisting of a navigation and a planning workstation, an infrared camera and a touch screen monitor, both attached to the ceiling, a reference star with three reflecting spheres, a Pointer with two reflecting spheres and an alternative laser instrument (z-touch) for surface registration.

A new phantom was developed for the study of surface registration accuracy and error sources. A transparent cranial model was fitted with eight inner target bushings from the Leibinger Cranial Marker System (Fig. 3), which could get fitted with radiopaque spheres for CT, hollow spheres with gadolinium for MRI, or titanium cylinder caps with a mark for navigation measurements. Two additional cylindrical markers were screwed into the lateral surface as

superficial targets. Interchangeable markers with central pivots corresponding exactly to the centres of the imaging markers improve the accurate placement of navigation probes [45–47]. In general, artificial targets allow for a more accurate and precise target definition [2, 28, 46]. Through repeated casting, an outer layer was formed. A silicon gel (Sylgard 527 A & B, Dow Corning, UK) was chosen, which has already been used as a soft tissue simulant for brain [6], breast tissue [3] and—in combination with gelatine—for traumatic deformations [11]. Its simulation advantages are viscoelastic behaviour [6], temperature resistance, allowing for controlled movement, and good magnetic properties similar to the ones of human tissue in T1- and T2-weighted MRI sequences [3]. On top of this layer, a transparent skin-friendly material with good MRI signal from gelatine and glycerine was cast (ProsMaster KI, Kerling International GmbH, Backnang, Germany). The use of a multilayer model affects contact-pressure distribution and agrees with in vivo data regarding deformation of the skin surface [10].
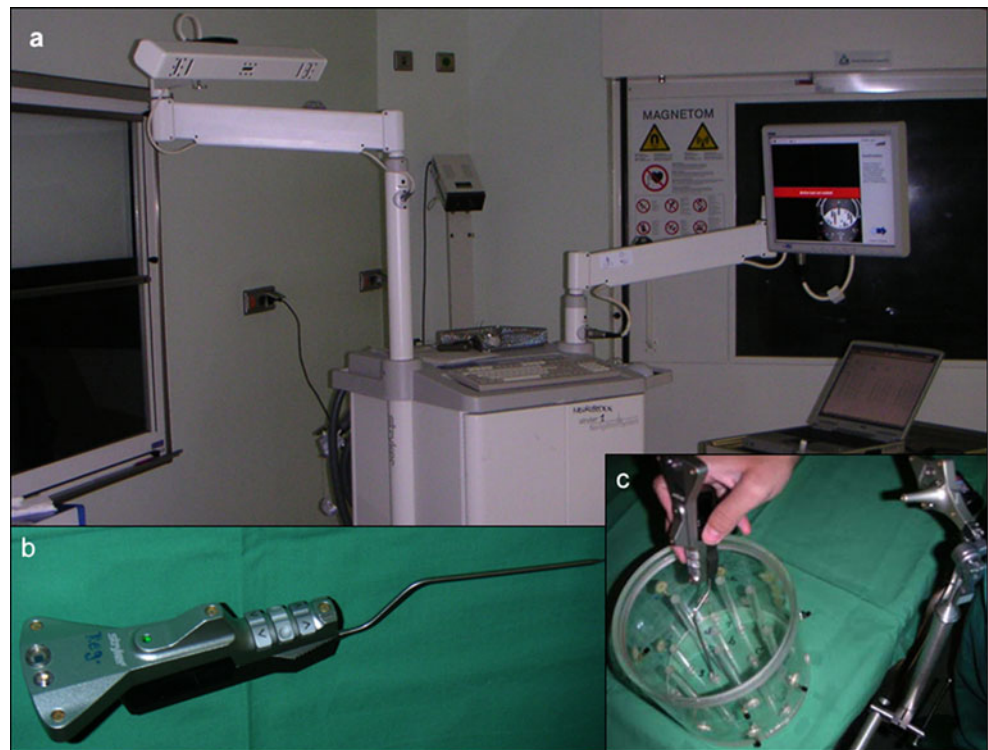
Imaging protocols

The phantom was scanned with a helical cone beam CT (Volume Zoom, Siemens) with no gantry tilt, slice thickness of 2.0 mm recalculated to 1.0 mm, 512×512 matrix and in-plane pixel size of 0.43×0.43 mm^2, as well as with an MR tomograph (Symphony 1.5 Tesla, Siemens) with a T1-weighted magnetisation prepared rapid gradient echo (MPRAGE) sequence with 1.3 mm slice thickness, 0.5 mm squared pixel size and 512×512 matrix. For testing imaging accuracy, additional scans with standard navigation protocols were obtained: a Picker Edge 1.5 T MR tomograph (T1 weighted flash-3D sequence, 135 slices, 1.3 mm slice thickness, 256×256 matrix, 1.0 mm squared pixel size), a Siemens Magnetom Open 0.2 T MR tomograph (T1-weighted flash 3D, 96 slices, TE 12.0 ms, TR 34 ms, field of view 250×250, 2.0 mm slice thickness, 256×256 matrix) and a Siemens Trio 3 T MR tomograph (T1 weighted MPRAGE, 1.3 mm slice thickness, 512×512 matrix, 0.5 mm squared pixel size) .

Definition of accuracy terms

(a) *Software accuracy*: coordinates measured on image data with the system software compared to coordinates measured with a coordinate high precision machine [computer numerical control (CNC)] (Fig. 4).
(b) *Imaging accuracy*: software accuracy subclassified with respect to the imaging modalities (Fig. 4).
(c) *System (or digitising-) accuracy*: navigation coordinates compared to CNC coordinates. It represents the accuracy of system components with which the digitising unit achieves calculation of a point's space

coordinates and includes precision of stereoscopic cameras and pointer instruments (Fig. 4).

(d) *Navigation accuracy*: coordinates from the navigation procedure compared to coordinates on image data sets. It represents the accuracy of the localisation of target points after registration and represents the practical overall accuracy, as it includes the digitising component, the software and imaging modality, the registration procedure as well as various intraoperative factors. Including all these possible error sources, a larger deviation than for the other types of accuracy was expected (Fig. 4).

Types of measurements conducted

(a) Computer numerical control machines, representing high precision standard, were used to obtain true coordinates of target points as a reference set. The phantom was measured (Fig. 5) with an accuracy of 5 μm (Global Image 9158, Brown & Sharpe, SN: 00072, QS Engineering, Bad Friedrichshall, Germany).

(b) Target point coordinates, localised with the software of each system on image data sets (ten repetitions).

(c) Navigation coordinates were measured with both systems in a clinical setup (Figs. 1 and 2). With

Fig. 2 BrainLAB VectorVision Sky in experimental setup: **a** standard Plexiglas phantom with the reference star up close on the left, **b** the BrainLAB monitor and camera attached to the OR ceiling (view through an intraoperative CT gantry)
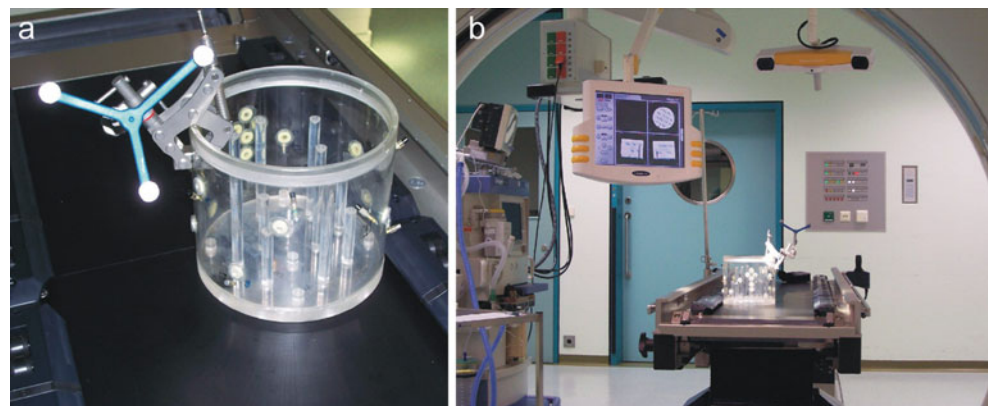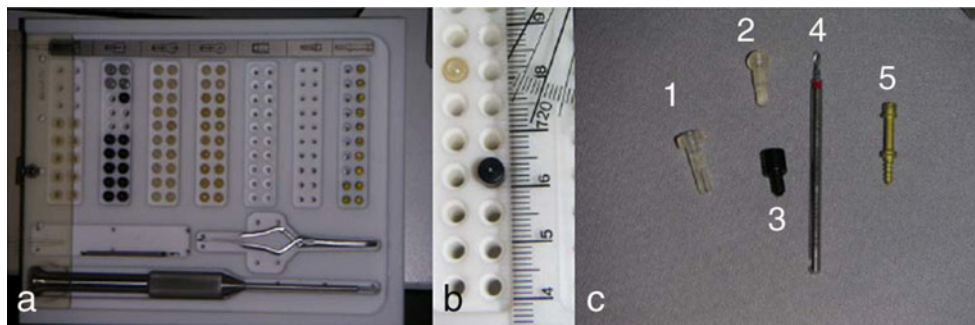
Fig. 3 Cranial Marker System set from Leibinger (Freiburg, Germany) used for the targets of the newly developed phantom: **a** the whole set with screw driver, forceps, drill bit and (*from left to right*) bushings, titanium cylinders with target point (*black*), hollow spheres for contrast medium, radiopaque spheres and screw extensions (note the schematic drawing on top of each element), **b** hollow spheres and black titanium markers in magnification with their dimensions, **c** bushing (*1*), hollow sphere for contrast medium (*2*), titanium marker (*3*), drill bit (*4*) and screw extension (*5*)

Stryker 6 to 9, fiducials in an asymmetric non-coplanar pattern surrounding the lesion were used for registration [34, 39]. The camera was warmed up and positioning was checked, so that the phantom, patient tracker and pointer were well inside the camera field in a nearly ideal distance (1.5–1.8 m). After registration, the RMS was recorded; RMS more than 2.5 mm was not accepted. Success of every registration was controlled visually by pointing at well-defined structures. Each target was touched with the pointer tip, and

Fig. 4 Flow chart showing types of measurements (image data set coordinates, computer numerical control coordinates and navigation coordinates), comparisons and transformation as well as types of accuracy (software, system and navigation accuracy)
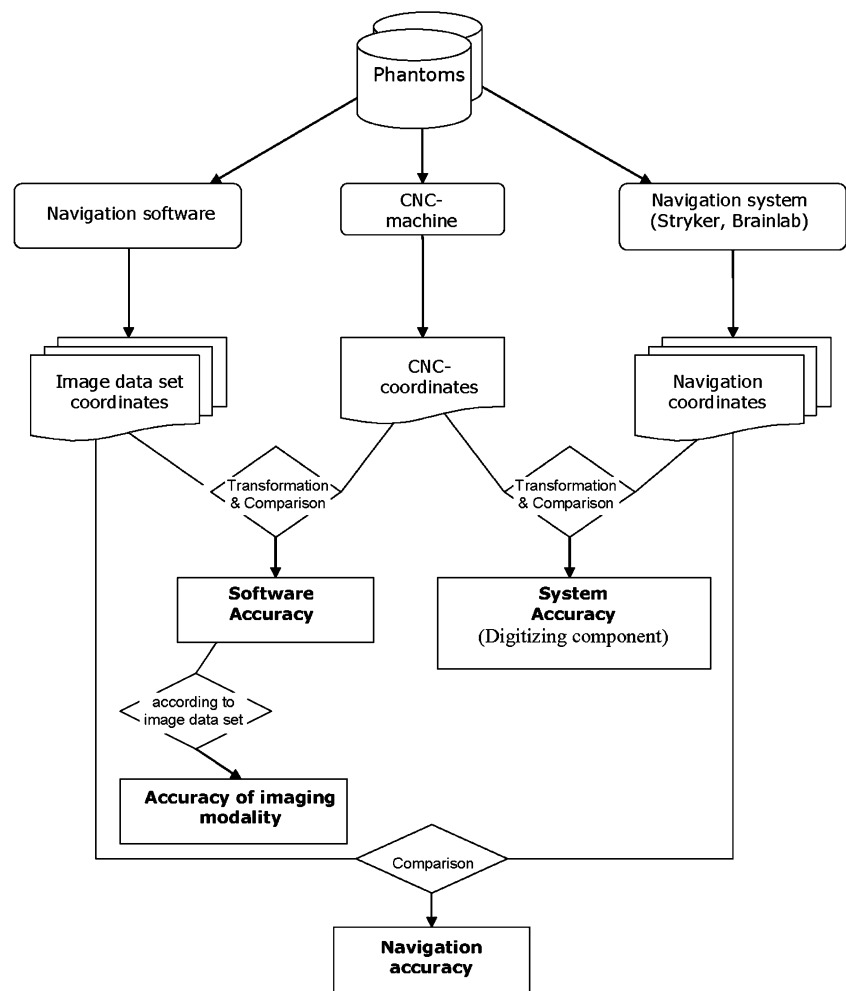
Fig. 5 The new phantom and its true target coordinates (*x*, *y*, *z* coordinates in millimeters) as measured by a Computer Numerical Control high precision machine: the main view (**a**) over the skull model and the small lateral view (**b**) show the deformable layer, adhesive fiducials (*rings*), two external target points (*yellow dots*) and openings towards the internal targets (*red dots*)
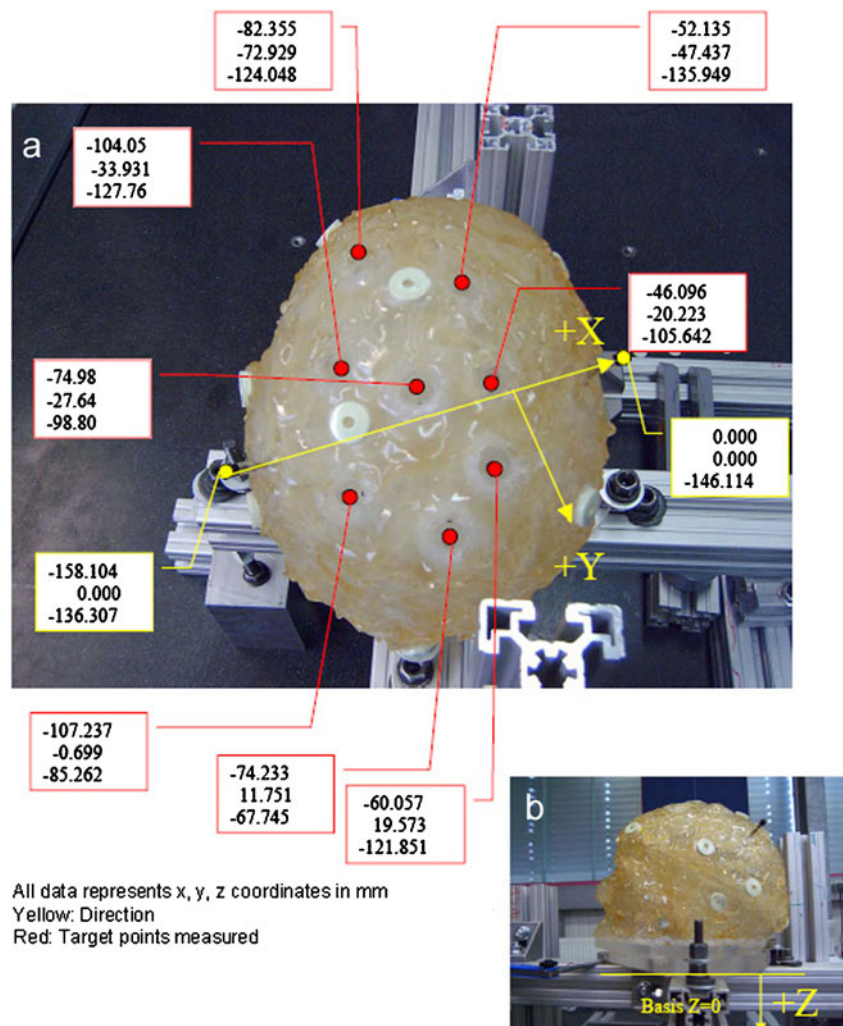
with the "freeze" function, coordinates were acquired. For each measurement series, the procedure was repeated ten times to avoid systematic error factors, and each coordinate measurement was repeated three times to avoid random errors when pointing a target point, which was taken into account during statistical processing. To examine factors such as replacement of non-sterile instruments with sterile ones after registration or the use of a sterile drape over the patient tracker to avoid replacement, extra measurement series were performed.

Slight modifications were necessary with Brain-LAB due to the user interface. Target points were defined as "labelled points" and uniquely numbered. Patient data was loaded, and a new standard registration began with the pointer. Seven fiducials were used, which is the maximum number accepted by the system. An asymmetric non-coplanar wide placed pattern was obtained. After registration and visual checking of accuracy at well-known structures, such as

the nasion, all target points were touched with the pointer and acquired as "intraoperative points," as the system does not directly display coordinates. The "pointer to target" distance in millimetres was recorded for every intraoperative point and its closest labelled point, which represents the deviation between image data set coordinates and navigation coordinates. After each series three log files were copied, which contained all necessary information, including coordinates of target points and acquired points as well as a precision estimation of the registration. The same strategy for repeated measurements was followed.

(d) Surface matching measurements were conducted with both systems. With Stryker these were carried out in analogy to the accuracy measurements. Two identical CT image data sets and two identical Symphony MRI data sets were used, differing only in the fixation during image acquisition. A control measurement of all target points with fiducial registration was conducted in order to obtain a reference. At least 30

surface points were acquired. The system suggests the acquisition of points from well-defined structures, such as the nose and the periorbital areas. Numbers of acquired and accepted surface points as well as "mean distance" were documented. Three head positions were examined: a normal supine position with asymmetrically widely distributed markers, a lateral position with almost linear placement of markers and a park-bench right position with markers occipitally, temporally and frontally. Registration of surface points on the nose or not was varied as a parameter. For each fiducial registration, all target points were measured with three to six consecutive surface registrations, while varying the examined parameters such as position, marked structures etc. Fiducial only and fiducial plus surface registration were connected variables, which were taken into account for the statistical evaluation.

Surface registration measurements with BrainLAB were conducted analogous to the corresponding accuracy measurements. Due to system limitations, the surface registration was not performed on top of the fiducial registration, but instead of it. The same four image data sets as with Stryker were imported. As being independent of previous fiducial registrations, double measurement series were necessary, i.e. extra fiducial registration as control. Surface points were acquired with a laser instrument (z-touch) plus an additional series with the mechanical pointer as a control. In addition, a series for testing the influence of different head positions and the corresponding acquired surface points was conducted. For this purpose, z-touch and pointer were combined, as in clinical routine. After registration, all target points were denoted with the pointer, and their coordinates were acquired as intraoperative points.

For each measurement series, the procedure was repeated ten times, and each coordinate measurement was repeated three times to avoid random errors when pointing a target point.

Data transformation and statistics

Coordinate systems of data sets had to be matched, to permit a mathematical comparison. This was achieved through computed matching with the help of a well established fitting algorithm (Bevington's "least squares fit") [5, 48]. Residuals were calculated with a corresponding algorithm in R-project (R version 2.6.1). Transformation was only necessary for comparing CNC and navigation coordinates and CNC and image data coordinates. For the comparison of navigation with image data coordinates, no

transformation was necessary, since both refer to the same coordinate system. Means and standard deviations were calculated as measures of accuracy and variance. Mixed effects regression models were used to account for the repeated measures and the hierarchical structure of the complex study design. These were used to examine the influence of parameters of the experimental setting on overall navigation accuracy as well as on software and system accuracy. Nested mixed effects regression models were compared to each other with the likelihood ratio test. All P values in this work refer to the comparison of nested regression models, unless explicitly mentioned. Due to the exploratory nature of the study, no adjustment was made for multiple testing. All tests were performed at a significance level of 0.05. Test results could not be interpreted in a confirmatory sense. Statistical analysis was performed using R-project (R version 2.6.1).

Results

A mean deviation for Stryker's software of 0.38 ± 0.24 mm was calculated. Respective values for BrainLAB's software were 0.44 ± 0.28 mm (Table 1). The difference was conspicuous ($P < 0.001$).

Examination of image data set accuracy showed a difference between CT and MR (Table 2). CT was more accurate with a mean Euclidean deviation of 0.28 ± 0.11 mm. All MR sequences were less accurate than CT (all $P < 0.01$). The worst accuracy was obtained with Open-MR (0.2 T intraoperative MR system). Corresponding analysis with BrainLAB showed similar results (Table 2).

Regarding accuracy of the digitising component, data showed a difference in favour of BrainLAB (Table 1). A mean Euclidean deviation of 0.72 ± 0.38 mm was calculated for Stryker as opposed to 0.33 ± 0.19 mm for BrainLAB ($P < 0.01$).

A larger deviation was expected for navigation accuracy, as it represents the overall application accuracy, including all possible error sources. This expectation was confirmed experimentally, yielding a mean Euclidean deviation of 1.45 ± 0.63 mm for Stryker and of 1.27 ± 0.53 mm for BrainLAB (Table 1). The difference of 0.18 mm was conspicuous (95% confidence interval: 0.09; 0.27, $P < 0.01$) but not constant in all three directions in favour of BrainLAB. Maximal outliers for both systems in all directions were of comparable magnitude.

Surface matching series resulted in deviations of 2.51 ± 1.49 mm for Stryker and 2.61 ± 1.56 mm for BrainLAB; these data show no significant difference (Table 1). Markedly larger deviations were observed from outliers for both systems with surface matching.

Table 1 Software, digitising (or system) and overall navigation accuracy (in italics) of the tested systems comparatively

System	Software accuracy	Digitising accuracy	Navigation accuracy	Navigation accuracy (surface registration)
Stryker® Leibinger	0.38±0.24	0.72±0.38	*1.45±0.63*	2.51±1.49
BrainLAB VectorVision®	0.44±0.28	0.33±0.19	*1.27±0.53*	2.61±1.56

The fourth column shows overall Navigation accuracy after surface registration. Mean Euclidean deviations plus/minus one standard deviation. All values are in millimetres

Influencing factors on fiducial registration

– No difference was discovered between Stryker's long or short, sterile or non-sterile pointer. Nevertheless, one of the non-sterile long pointers worsened accuracy by 0.55 mm (95% confidence interval: 0.10; 0.99; $P=0.02$). Although the specific pointer had a visible geometrical deformation, the validation procedure was successful. In Stryker, touching a predefined point on the tracker with the pointer tip performs this. In BrainLAB, such a procedure is replaced by the position of the pointer in the instrument case, where the two pins in predefined positions allow a visual check of intact geometry.

– Replacement of non-sterile instruments after registration with Stryker proved to decrease system accuracy by 0.11 mm marginally (95% confidence interval, 0.01; 0.20, $P=0.03$) and navigation accuracy conspicuous by 1.13 mm (95% confidence interval, 0.52; 1.76, $P<0.01$). This is possibly due to relative movement of the patient tracker in relation to the patient's head during replacement. An alternative is the use of a sterile drape to cover the tracker. This method also increased the observed deviations as compared to no sterile drape. System accuracy was worsened by 0.17 mm (95% confidence interval, 0.11; 0.23, $P<0.01$) and navigation accuracy by 0.69 mm (95% confidence interval: 0.28; 1.11, $P<0.01$).

– Results from regression analysis showed no relation between precision data given by the two systems (RMS and worst fiducial deviation displayed by Stryker, log-file

data in BrainLAB) and digitising or navigation accuracy. The "pointer to target" distance calculated by the BrainLAB system, however, was related with navigation accuracy ($P<0.01$).

Influencing factors on surface matching

– A conspicuous difference was detected for both systems in favour of CT, regarding software accuracy. The difference was 0.25 mm (95% confidence interval, 0.21; 0.28, $P<0.01$) for Stryker and 0.28 mm (95% confidence interval, 0.26; 0.30, $P<0.01$) for BrainLAB. In the Stryker system, this difference did not affect system- or navigation accuracy, i.e. imaging modality had no influence on the overall navigation accuracy. Nevertheless, the data show a difference for BrainLAB, where MR worsened the navigation accuracy by 0.71 mm (95% confidence interval, 0.28; 1.14, $P<0.01$).

– After surface matching, based on an existing fiducial registration (with Stryker), the data show no differences between the three tested head positions. In a park-bench right position of phantom, only with fiducial registration, a conspicuous difference was found, decreasing navigation accuracy by 1.67 mm (95% confidence interval, 1.39; 1.95, $P<0.01$). This effect disappeared after the additional surface match, thus proving helpful in improving an inaccurate fiducial registration in this setting.

– With Stryker, no influence on the mean deviation after surface matching was found regarding localisation of

Table 2 Mean deviation in millimeters±standard error and 95% confidence intervals (CI) of image data examined with the software from Stryker–Leibinger and the respective software from BrainLAB

Image data	Stryker			BrainLAB		
	Mean deviation±standard error	95% CI	*P* value	Mean deviation±standard error	95% CI	*P* value
CT	0.28±0.02	0.23, 0.32	–	0.25±0.03	0.19, 0.31	–
Edge MR	0.38±0.02	0.34, 0.43	<0.01	0.45±0.02	0.41, 0.49	<0.01
Open MR	0.68±0.03	0.63, 0.74	<0.01	0.74±0.02	0.69, 0.78	<0.01
Symphony MR	0.34±0.03	0.29, 0.39	<0.01	0.33±0.02	0.29, 0.38	<0.01
Trio MR	0.49±0.03	0.42, 0.56	<0.01	0.43±0.02	0.39, 0.48	<0.01

P values (regression model analysis) are in comparison to CT, which proved as the most accurate data set

CI confidence intervals

registered surface points or use of well-defined structures (e.g. nose), despite the subjective impression during measurements, that acquisition of points periorbitally and on the nose improved registration [28]. Acquisition of points on the nose as a separate factor could not be evaluated with BrainLAB, as they were almost always necessary for a successful registration; in such cases, a message would be shown from the system, requiring points from well-defined anatomical structures. Registration of periorbital points improved navigation accuracy with BrainLAB, decreasing the mean Euclidean deviation by 1.10 mm (95% confidence interval, −1.67; −0.53, $P<0.01$).

– The RMS and "mean distance" values from Stryker after combined fiducial and surface registration showed no influence on navigation accuracy, indicating that precision feedback does not always relate to true accuracy. Nevertheless, navigation accuracy was influenced by the number of acquired surface points ($P=0.01$) and the percentage of accepted points for registration ($P<0.01$), the product of both yields the total number of surface points actually being employed for registration. Feedback from BrainLAB, regarding achieved accuracy, can only be obtained retrospectively from a system log file. This value after surface registration proved a strong relation with navigation accuracy ($P<0.01$).

– Finally, a comparison between pointer, z-touch and combination of the two during registration revealed no differences regarding digitising or navigation accuracy.

Surface matching vs. fiducial registration

The digitising accuracy of both systems was marginally worsened with surface matching. Data showed conspicuous differences but were restricted in absolute value, thus undermining their clinical relevance (Table 3). A comparison of navigation accuracy with Stryker showed no difference between fiducial based and surface matching ($P=0.1910$). The corresponding difference with BrainLAB (Table 3) was significant in favour of fiducial registration ($P<0.01$). Mean Euclidean deviations of the two systems were of similar magnitude (Stryker 2.51±1.49 mm; BrainLAB 2.61± 1.56 mm). Analysis of accuracy in spatial directions revealed a greater variance and more outliers in all directions with both systems after surface matching.

Discussion

In neuronavigation, knowledge of achieved accuracy and system limitations is crucial. Error analysis includes fiducial localisation error, deviation of fiducials after

registration and deviation of points of interest other than the fiducials after registration [44]. Despite the fact that these errors represent vectors, they are given as values representing only their length, i.e. as "root mean square" (RMS) of the vector components. Since deviation of target points other than the fiducials [15], e.g. points in the vicinity of a deeply situated lesion, is hard to measure exactly with the navigation system, the surgeon must rely on a statistical estimation of this error, based on the known fiducial localisation accuracy or the error fit of the fiducials (RMS) [44]. Registration error calculated this way is nevertheless regarded by some authors as a non-trustworthy measure of registration accuracy [14, 41] and of real accuracy [39, 49]. After all, target registration error is considered anisotropic and may vary significantly in different areas of the head [43]. Different definitions [7, 25–27, 39, 44] and doubtful suitability of the systems' feedback for the actual application accuracy make it necessary to apply methods of quantitative analysis with externally measured coordinates.

Quantitative feedback regarding registration precision given did not show any influence on accuracy, agreeing with previous studies [12, 39]. The RMS relies only on the geometric alignment of the fiducial markers and is not the accuracy that can be expected by the surgeon when approaching a target [1, 19, 46]. Disagreements between RMS and localisation accuracy are caused by the fact that the registration markers are only rarely evenly spread across the whole registration volume; only in this case— and if the object is spherical—the RMS may provide a good measure of the accuracy across the whole volume [12]. This raises the question how transparent systems are towards the surgeon, regarding the way they estimate achieved accuracy. Visualisation of quantified accuracy estimators on the image data in a 3D manner might prove useful in the future [32].

Imaging modalities

Advanced CT imaging accuracy was confirmed in this study, which could be explained due to MR distortion [24, 39, 50], being in accordance with previous studies [12, 34, 50]. However, no differences on overall navigation accuracy were detected. In the literature, conflicting results are reported in favour of CT [12, 34, 50] or MRI [18, 37]. The intraoperative MR system presented the most inaccurate results, which was expected from a low field system with accordingly low signal/noise ratio.

Comparison of the systems

In this study, the term navigation accuracy represents global accuracy, including all possible error sources, while system

Table 3 Effect of surface matching compared to fiducial registration on digitising and navigation accuracy

System	Accuracy type	Difference (mm)	95% confidence intervals	P value
Stryker® Leibinger	Digitising accuracy	−0.14	−0.16, −0.12	<0.01
	Navigation accuracy	0.06	−0.03, 0.15	0.1910
BrainLAB VectorVision®	Digitising accuracy	−0.05	−0.07, −0.02	<0.01
	Navigation accuracy	−0.89	−1.01, −0.78	<0.01

Differences (in millimeters) were conspicuous in favour of fiducial registration regarding digitising accuracy for both systems. Regarding navigation accuracy, only the difference with the BrainLAB system was conspicuous in favour of fiducial registration

accuracy coincides with global accuracy in analogue studies [34]. Standardised reporting of error assessment, including the methodology used, may assure comparability of different accuracy reports [46].

Conspicuous differences detected in favour of BrainLAB regarding system and navigation accuracy are of limited clinical relevance, due to the submillimetric absolute values. The magnitude of the differences may be statistically significant for some measures, but in the context of the greater error seen in a clinical setting, these differences appear relatively unimportant.

Accuracy results of this work for Stryker are comparable to literature. A direct comparison is not feasible due to different imaging protocols and accuracy definitions. Poggi et al. performed a phantom study with the Stryker–Leibinger system [34]. They defined as global accuracy the difference in a reference set of coordinates acquired with a coordinate high precision machine with pointer coordinates after transformation. Influence of mechanical accuracy was not taken into consideration. Gellrich et al. [17] reported an overall accuracy of 1 mm for Stryker, but this was only the registration error reported and not the total experimental accuracy.

Regarding BrainLAB, results of this work are in accordance with previous studies. Values of 1.45 ± 0.99 mm are reported for phantom studies and 4.05 ± 3.62 mm for clinical studies [31]. Reports of registration error vary between 1.6 mm [8] and 2.1 mm [40]. Gumprecht et al. gave a target localisation error of 4.0 ± 1.4 mm, but it was an examination of a previous model with different methodology [20]. In vivo accuracy below 2 mm has been reported for the retrosigmoid approach using anatomical landmarks [9].

Influencing factors

Intact instrument geometry proved a necessity for accuracy. Validating the Stryker pointer with the patient tracker could sometimes be successful even with a deformed geometry, which could prove dangerous in clinical practice. This raises the question of whether validation thresholds should be re-evaluated, in order to find a balance between safe, strict validation on the one side and time-consuming validation failures and repeats on the other side.

Instrument replacement after non-sterile registration decreased accuracy conspicuously with Stryker, due to the change in relative position of the patient tracker in relation to the patient head (phantom) by the replacing procedure. The corresponding arm is robust, well fixed and has a specific interface for the patient tracker; nevertheless, a slight deviation cannot always be avoided. The alternative of covering the tracker with a sterile drape decreased accuracy as well. Refraction of infrared rays from the LEDs through the drape could be the reason. In passive systems like BrainLAB, problems are reported from the interruption of the visual line by the microscope drape [20].

A "park-bench right" position tested decreased navigation accuracy conspicuously. This effect disappeared after additional surface matching, indicating that surface registration may enable the system in this case to improve an inaccurate fiducial registration.

Fiducial vs. surface registration

Comparison of fiducial vs. surface registration showed a small difference in system accuracy of both systems in favour of fiducials, as well as a greater variance and more outliers with surface registration in all spatial directions. In general, algorithms for surface matching represent an effort to limit the influence of movable skin; nevertheless, they may lead to even greater registration errors [18, 22, 41], possibly due to sampling errors owing to uneven grouping of scalp points chosen for registration [18].

In daily clinical practice, registration with anatomic landmarks and/or surface matching is more practical. Insignificant loss of registration accuracy using natural anatomic landmarks compared to skin adhesive fiducials has been reported [50]. Nevertheless, fiducial registration alone remains the gold standard [22] of all non-invasive methods [33], being more accurate than the combination of anatomic landmarks and surface registration [18, 38]. Surface matching may produce larger deviations when added to a fiducial registration, and though it may improve accuracy after landmark registration, this combination fails

to reach the level of fiducials alone [22]. In other studies, all three registration techniques (fiducials, anatomic landmarks and combined anatomic landmarks and surface matching) provided comparable deviations [49]. Mascott et al. reported of the superiority of implanted fiducials, but no major differences between adhesive fiducials, surface matching or anatomic landmark registration [28]. In recent studies, distribution templates for fiducials are provided, and a configuration with eight fiducials optimised over the whole head is proposed [42, 43].

Phantom critic

The deformable outer layer simulation provides an additional error source, playing a role not only during surface matching but also due to fiducial displacement. Skin displacements can cause problems during imaging, head fixation, sampling of fiducial markers or surface points for registration with the pointer [22, 38, 50]. Furthermore, rigid phantoms cannot simulate the contribution of soft tissue properties to MR errors [39]. The anthropomorphic form of the phantom allows the study of accuracy in relation to the position of target lesions, since larger deviations are often reported for posterior localisations [22, 36]. The disadvantage of non-anthropomorphic phantoms is the lack of correct anatomical simulation [4, 16, 23, 29, 30, 34]. Although the target volume may be similar to the real anatomical target volume, a substantial difference is seen in the registration process [46]. Anthropomorphic skull phantoms represent the bony anatomy, but it is very difficult to correctly simulate soft tissues [1, 12, 13, 21, 35]. Furthermore, the fact that the real head is not spherical and has a moveable scalp layer gives rise to areas with a real error substantially deviating from the RMS [12]. Skin flexibility and elasticity have been connected to problems in surface algorithms; pressure deformity during image acquisition and lack of skin turgor with presence of wrinkles in adults may all contribute to decreased accuracy using surface-fit registration [41].

The present phantom represents a novel effort towards realistic simulation of human-like anatomy and possible inherent inaccuracy factors. Such models are necessary for conducting high-quality phantom studies, which are essential to assure patient safety before clinical application, as well as for standardised evaluation and quality assurance of a constantly evolving technology [46]. Naturally, patient studies still remain essential, as it would be misleading to extrapolate phantom accuracy to true intraoperative conditions.

Conclusions

Experimental study of two optical systems provided evidence for the same level of accuracy. Passive frame-

works need not necessarily be less accurate than active LED systems. Overall navigation accuracy of both systems was definitely acceptable (< 1.5 mm), while surface registration failed to show better results than fiducial registration. The newly developed phantom allows the experimental study of surface registration and represents a novelty towards non-rigid phantoms.

Accuracy feedback by the navigation systems should be considered with caution, as it does not always correlate well with real localisation accuracy and is therefore to be seen only as an averaged estimation for the quality of the registration procedure. Checking accuracy visually directly after registration with the help of anatomical landmarks on the surface ought to be repeated during the operation by verifying with well-defined deeper structures.

Distinguishing different types of accuracy experimentally may allow identification of error sources, and consideration of all potential influencing factors may minimise errors in navigation procedures.

References

1. Abbasi HR, Hariri S, Martin D, Shahidi R (2001) A comparative statistical analysis of neuronavigation systems in a clinical setting. Stud Health Technol Inform 81:11–17
2. Alp MS, Dujovny M, Misra M, Charbel FT, Ausman JI (1998) Head registration techniques for image-guided surgery. Neurol Res 20:31–37
3. Azar FS, Metaxas DN, Miller RT, Schnall MD (2000) Methods for predicting mechanical deformations in the breast during clinical breast biopsy. 26th IEEE Annual Northeast Bioengineering Conference, University of Connecticut, Storrs, CT, USA
4. Benardete EA, Leonard MA, Weiner HL (2001) Comparison of frameless stereotactic systems: accuracy, precision, and applications. Neurosurgery 49:1409–1415, discussion 1415–1406
5. Bevington PR (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York
6. Brands DWA (2002) Predicting brain mechanics during closed head impact—numerical and contitutive aspects. Dynamics and biomechanics. Technische Universiteit Eindhoven, Eindhoven
7. Cartellieri M, Kremser J, Vorbeck F (2001) Comparison of different 3D navigation systems by a clinical "user". Eur Arch Otorhinolaryngol 258:38–41
8. Castilla JM, Martin V, Fernandez-Arconada O, Delgado P, Rodriguez-Salazar A (2003) First steps in neuronavigation. Neurocirugia 14:398–408 (in Spanish)
9. da Silva EB, Jr LAG, Milano JB, da Silva LF, Jr CRS, Ramina R (2010) Image-guided surgical planning using anatomical landmarks in the retrosigmoid approach. Acta Neurochir (Wien) 152:905–910
10. Dandekar K, Raju BI, Srinivasan MA (2003) 3-D finite-element models of human and monkey fingertips to investigate the mechanics of tactile sense. J Biomech Eng 125:682–691
11. Didieren AM, van Riet EJM, Paulissen JJM (2005) Blunt trauma. TNO Prins Maurits Laboratory. Available at http://www.pml.tno.nl/en/bk/pdf/PML03d028.pdf. Accessed 4 February 2005
12. Dorward NL, Alberti O, Palmer JD, Kitchen ND, Thomas DG (1999) Accuracy of true frameless stereotaxy: in vivo measure-

ment and laboratory phantom studies. Technical note. J Neurosurg 90:160–168

13. Eggers G, Muhling J, Marmulla R (2005) Template-based registration for image-guided maxillofacial surgery. J Oral Maxillofac Surg 63:1330–1336

14. Fitzpatrick JM, Hill DL, Shyr Y, West J, Studholme C, Maurer CR Jr (1998) Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain. IEEE Trans Med Imaging 17:571–585

15. Fitzpatrick JM, West JB, Maurer CR Jr (1998) Predicting error in rigid-body point-based registration. IEEE Trans Med Imaging 17:694–702

16. Galloway RL Jr, Maciunas RJ, Latimer JW (1991) The accuracies of four stereotactic frame systems: an independent assessment. Biomed Instrum Technol 25:457–460

17. Gellrich NC, Schramm A, Hammer B, Rojas S, Cufi D, Lagreze W, Schmelzeisen R (2002) Computer-assisted secondary reconstruction of unilateral posttraumatic orbital deformity. Plast Reconstr Surg 110:1417–1429

18. Golfinos J, Fitzpatrick B, Smith L, RF S (1995) Clinical use of a frameless stereotactic arm: results of 325 cases. J Neurosurg 83:197–205

19. Grunert P, Darabi K, Espinosa J, Filippi R (2003) Computer-aided navigation in neurosurgery. Neurosurg Rev 26:73–99. discussion 100–101

20. Gumprecht HK, Widenka DC, Lumenta CB (1999) BrainLab VectorVision Neuronavigation System: technology and clinical experiences in 131 cases. Neurosurgery 44:97–104. discussion 104-105

21. Hauser R, Westermann B, Probst R (1997) Noninvasive tracking of patient's head movements during computer-assisted intranasal microscopic surgery. Laryngoscope 107:491–499

22. Helm PA, Eckel TS (1998) Accuracy of registration methods in frameless stereotaxis. Comput Aided Surg 3:51–56

23. Krempien R, Hassfeld S, Kozak J, Tuemmler HP, Dauber S, Treiber M, Debus J, Harms W (2004) Frameless image guidance improves accuracy in three-dimensional interstitial brachytherapy needle placement. Int J Radiat Oncol Biol Phys 60:1645–1651

24. Maciunas RJ, Fitzpatrick JM, Gadamsetty S, Maurer CR Jr (1996) A universal method for geometric correction of magnetic resonance images for stereotactic neurosurgery. Stereotact Funct Neurosurg 66:137–140

25. Maciunas RJ, Galloway RL Jr, Fitzpatrick JM, Mandava VR, Edwards CA, Allen GS (1992) A universal system for interactive image-directed neurosurgery. Stereotact Funct Neurosurg 58:108–113

26. Maciunas RJ, Galloway RL Jr, Latimer J, Cobb C, Zaccharias E, Moore A, Mandava VR (1992) An independent application accuracy evaluation of stereotactic frame systems. Stereotact Funct Neurosurg 58:103–107

27. Maciunas RJ, Galloway RL Jr, Latimer JW (1994) The application accuracy of stereotactic frames. Neurosurgery 35:682–694. discussion 694–685

28. Mascott CR, Sol JC, Bousquet P, Lagarrigue J, Lazorthes Y, Lauwers-Cances V (2006) Quantification of true in vivo (application) accuracy in cranial image-guided surgery: influence of mode of patient registration. Neurosurgery 59:ONS146–ONS156. discussion ONS146-156

29. Maurer CR Jr, Fitzpatrick JM, Wang MY, Galloway RL Jr, Maciunas RJ, Allen GS (1997) Registration of head volume images using implantable fiducial markers. IEEE Trans Med Imaging 16:447–462

30. Moriarty TM, Quinones-Hinojosa A, Larson PS, Alexander E 3rd, Gleason PL, Schwartz RB, Jolesz FA, Black PM (2000) Frameless stereotactic neurosurgery using intraoperative magnetic resonance imaging: stereotactic brain biopsy. Neurosurgery 47:1138–1145. discussion 1145–1136

31. Muacevic A, Uhl E, Steiger HJ, Reulen HJ (2000) Accuracy and clinical applicability of a passive marker based frameless neuronavigation system. J Clin Neurosci 7:414–418

32. Oberhammer P, Eisenmann U, Metzner R, Paraskevopoulos D, Wirtz CR, Dickhaus H (2007) Optimization and quantification of accuracy for rigid point based registration for computer aided surgery. In: Buzug TM, Holz D, Bongartz J, Kohl-Bareis M, Hartmann U, Weber S (eds) Advances in medical engineering, vol 114. Springer, Berlin, pp 230–235

33. Pfisterer WK, Papadopoulos S, Drumm DA, Smith K, Preul MC (2008) Fiducial versus nonfiducial neuronavigation registration assessment and considerations of accuracy. Neurosurgery 62:201–207, discussion 207-208

34. Poggi S, Pallotta S, Russo S, Gallina P, Torresin A, Bucciolini M (2003) Neuronavigation accuracy dependence on CT and MR imaging parameters: a phantom-based study. Phys Med Biol 48:2199–2216

35. Quinones-Hinojosa A, Ware ML, Sanai N, McDermott MW (2006) Assessment of image guided accuracy in a skull model: comparison of frameless stereotaxy techniques vs. frame-based localization. J Neurooncol 76:65–70

36. Raabe A, Krishnan R, Wolff R, Hermann E, Zimmermann M, Seifert V (2002) Laser surface scanning for patient registration in intracranial image-guided surgery. Neurosurgery 50:797–801, discussion 802-793

37. Roessler K, Ungersboeck K, Aichholzer M, Dietrich W, Czech T, Heimberger K, Matula C, Koos WT (1998) Image-guided neurosurgery comparing a pointer device system with a navigating microscope: a retrospective analysis of 208 cases. Minim Invasive Neurosurg 41:53–57

38. Sipos EP, Tebo SA, Zinreich SJ, Long DM, Brem H (1996) In vivo accuracy testing and clinical experience with the ISG Viewing Wand. Neurosurgery 39:194–202, discussion 202-194

39. Steinmeier R, Rachinger J, Kaus M, Ganslandt O, Huk W, Fahlbusch R (2000) Factors influencing the application accuracy of neuronavigation systems. Stereotact Funct Neurosurg 75:188–202

40. Tirakotai W, Riegel T, Sure U, Bozinov O, Hellwig D, Bertalanffy H (2004) Clinical application of neuro-navigation in a series of single burr-hole procedures. Zentralbl Neurochir 65:57–64

41. Villalobos H, Germano IM (1999) Clinical evaluation of multi-modality registration in frameless stereotaxy. Comput Aided Surg 4:45–49

42. Wang M, Song Z (2010) Distribution templates of the fiducial points in image-guided neurosurgery. Neurosurgery 66:143–150, discussion 150–141

43. Wang M, Song Z (2010) Guidelines for the placement of fiducial points in image-guided neurosurgery. Int J Med Robot 6(2):142–149

44. West JB, Fitzpatrick JM, Toms SA, Maurer CR Jr, Maciunas RJ (2001) Fiducial point placement and the accuracy of point-based, rigid body registration. Neurosurgery 48:810–816

45. Widmann G, Eisner W, Kovacs P, Fiegele T, Ortler M, Lang TB, Stoffner R, Bale R (2008) Accuracy and clinical use of a novel aiming device for frameless stereotactic brain biopsy. Minim Invasive Neurosurg 51:361–369

46. Widmann G, Stoffner R, Sieb M, Bale R (2009) Target registration and target positioning errors in computer-assisted neurosurgery: proposal for a standardized reporting of error assessment. Int J Med Robot 5:355–365

47. Widmann G, Widmann R, Widmann E, Jaschke W, Bale RJ (2005) In vitro accuracy of a novel registration and targeting technique for image-guided template production. Clin Oral Implants Res 16:502–508

48. Wirtz CR (2000) Bildunterstuetzte Operationsverfahren in der Neurochirurgie. Neuronavigation und Intraoperative Magnetreso-

nanztomographie. Neurochirurgische Universitaetsklinik. Ruprecht-Karls-Universitaet Heidelberg

49. Woerdeman PA, Willems PW, Noordmans HJ, Tulleken CA, van der Sprenkel JW (2007) Application accuracy in frameless image-guided neurosurgery: a comparison study of three patient-to-image registration methods. J Neurosurg 106:1012–1016
50. Wolfsberger S, Rossler K, Regatschnig R, Ungersbock K (2002) Anatomical landmarks for image registration in frameless stereo-tactic neuronavigation. Neurosurg Rev 25:68–72

Comment

Fumio Yamaguchi, Tokyo, Japan

The authors described the differences between two navigation systems regarding their accuracy. Essentially, this report is very important to elucidate the real accuracy of those systems that we use in daily surgeries. In addition, this is good feedback to these companies manufacturing better systems in the future. In this study, the development of a phantom-simulating skin and soft tissue is one of the authors' efforts and appreciated. I agree on the conclusion the authors elucidated that surface registration may improve inaccurate point-based registrations. Furthermore, it is good to know that surface registration does not improve the accuracy with good point-based registration.

Even though the overall accuracy of neuronavigation systems is <1.5mm, we should always pay attention to brain shift and the other factors which are fluctuating during surgery in real clinical settings. In addition, we should remember the influences of the pressure of head rest of MRI apparatus and the gravity to soft tissue of scalp and face those changes depending on patient's position. The important thing is that we should know the characteristics of image guidance systems and rely on a surgeon's final judgment.