INSTITUTE OF PHYSICS PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **48** (2003) 2199–2216 PII: S0031-9155(03)56538-1

Neuronavigation accuracy dependence on CT and MR imaging parameters: a phantom-based study

S Poggi[1], S Pallotta[1], S Russo[1], P Gallina[2], A Torresin[3] and M Bucciolini[1]

[1] Department of Clinical Physiopathology, Medical Physics Unit, University of Florence, Florence, Italy
[2] Department of Oto-neuro-oftalmologic Surgery, Neurosurgery Unit, University of Florence, Florence, Italy
[3] Medical Physics Department, Niguarda Hospital, Milan, Italy

E-mail: marta@dfc.unifi.it (M Bucciolini)

Received 22 November 2002, in final form 9 May 2003
Published 1 July 2003
Online at stacks.iop.org/PMB/48/2199

Abstract

Clinical benefits from neuronavigation are well established. However, the complexity of its technical environment requires a careful evaluation of different types of errors. In this work, a detailed phantom study which investigates the accuracy in a neuronavigation procedure is presented. The dependence on many different imaging parameters, such as field of view, slice thickness and different kind of sequences (sequential and spiral for CT, T1-weighted and T2-weighted for MRI), is quantified. Moreover, data based on CT images are compared to those based on MR images, taking into account MRI distortion. Finally, the contributions to global accuracy coming from image acquisition, registration and navigation itself are discussed. Results demonstrate the importance of imaging accuracy. Procedures based on CT proved to be more accurate than procedures based on MRI. In the former, values from 2 to 2.5 mm are obtained for 95% fractiles of cumulative distribution of Euclidean distances between the intended target and the reached one while, in the latter, the measured values range from 3 to 4 mm. The absence of imaging distortion proved to be crucial for registration accuracy in MR-based procedures.

1. Introduction

During an open neurosurgical procedure, imaging information transposition to the operating field is becoming an important and complex step, in particular after craniotomy. Complexity is mainly due to the developments in contemporary neuroimaging, which has achieved a great level of detail, not only from a morphological point of view (CT, MRI), but also from a functional one (PET, SPECT, functional-MRI) (Dorward *et al* 1999, Vannier and Haller 1999). Neuronavigation, which has become a normal procedure in clinical settings, represents

a great help in this process, allowing a direct transposition of imaging information in the surgical field: in practice the device correlates the patient's reference system with the imaging one, and shows the image of the operating tools directly in the diagnostic images. Nowadays, image guidance advantages and clinical benefits are well established and reported.

However, the complexity of the neuronavigator's technical environment, and of image guidance in general, requires a specific knowledge of the device's limitations. In particular, the use of a localization system requires a careful definition and evaluation of different types of errors and their sources. The neurosurgeon needs to know the reliability level of his device's data from precision and bias values, choosing either to strictly follow those indications, or to proceed considering them as general hints. Unfortunately, these values are not available during surgery.

In a navigation procedure there are different sources of inaccuracies. The acquired images can show geometrical distortions, partial volume effects and can be obtained with different spatial resolution capabilities. Also the signal to noise ratio, which depends on the acquisition modality, affects the image quality. The registration process then, which translates data from the patient reference system to the imaging one, besides presenting an intrinsic inaccuracy (software algorithm), is related to the input data quality. In fact, it uses images and patient coordinates acquired with a pointer device, affected by a mechanical error (Zylka *et al* 1999, Maciunas *et al* 1994). Nowadays, the mechanical accuracy of many hand-held localization devices is submillimetric, making it comparable to the mechanical accuracy of frame-based localization systems (Kaus *et al* 1997). Its error contribution therefore has not been considered in this work.

The effect of the image quality on the registration procedure, in contrast is crucial; many registration algorithms are rigid (Hill *et al* 2001), and heavily depend on the absence of distortions in images. Indications of MR imaging distortions are present in the literature (Sumanaweera *et al* 1993, 1994, Hardy and Barnett 1998, Kondziolka *et al* 1992, Bourgeois *et al* 1999, Dipierro *et al* 1999, Debus *et al* 1996, Prott *et al* 1995, Weis *et al* 1998, Chang and Fitzpatrick 1992), but none of these results take into consideration the contribution of distortions to the specific application of neuronavigation, although this source of inaccuracy can compromise the entire navigation procedure.

This paper was aimed at evaluating the imaging accuracy and the global accuracy of the navigation procedure, which means the assessment of the system capability in allowing the user to reach the intended target. Detailed phantom studies are presented.

The main work innovations are essentially three.

The accuracy dependence on CT and MR scan parameters, such as field of view (FOV), slice thickness and specific kind of image acquisition sequence (sequential and spiral for CT images, T1-weighted and T2-weighted for MR images), has been quantified.

Since CT and MR images are both used in the clinical neuronavigation procedure, a comparison seems necessary. The differences between these two image acquisition modalities are thus discussed and correlated to MR imaging distortion. In the literature accuracy data are available only for procedures based on CT images, and their dependence on imaging parameters is not discussed.

Finally, the chosen experimental approach allows us to separately evaluate the imaging contribution and the global accuracy, while in the literature, with few exceptions (Kaus *et al* 1997, Torresin *et al* 1999), only the latter one is normally investigated (Dorward *et al* 1999, Vannier and Haller 1999, Zylka *et al* 1999, Maciunas *et al* 1994, Willems *et al* 2001).

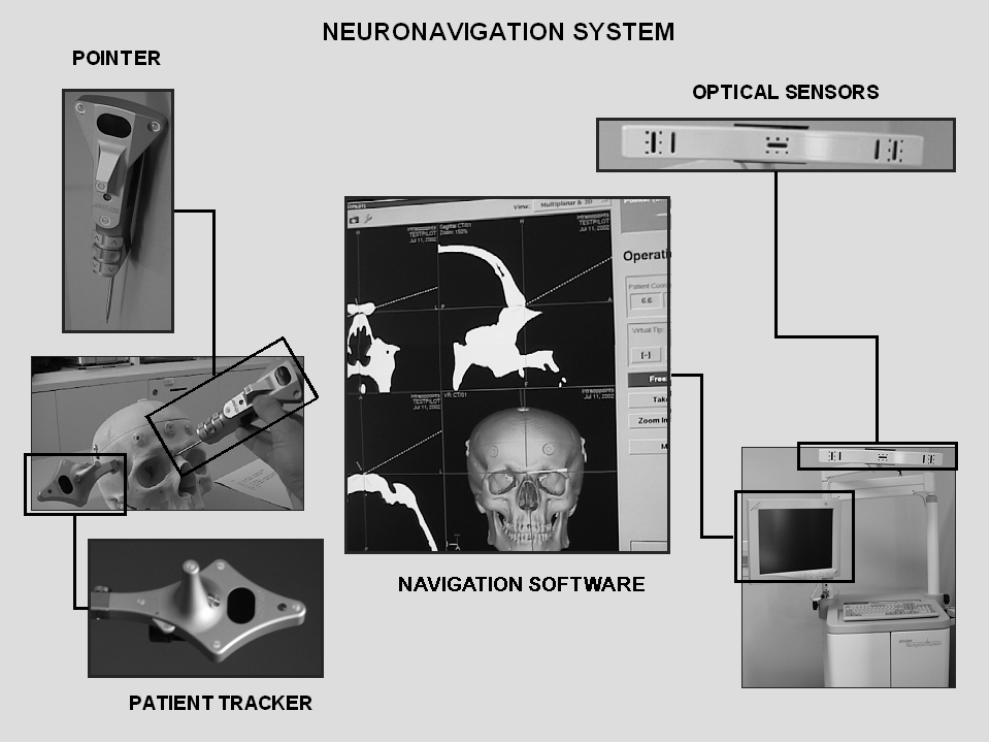

Figure 1. A schematic picture illustrating the components of the Stryker Leibinger navigation system.

2. Materials and methods

2.1. The neuronavigation system

The system used for the navigation procedure is the Stryker Leibinger navigation system. It consists of a planning software (STP 5.0), a frame on which three optical sensors are mounted, a Patient Tracker and a Pointer (figure 1). The Patient Tracker is fixed to a rigid arm, attached to the patient's head. It has five leds, whose positions are detected by the sensors. From these measurements the software calculates the centre of the patient's reference system. The Pointer is the frameless tool used by the surgeon to navigate inside the intracranial space. It is provided with four leds, whose positions are also measured by the sensors.

A neuronavigation procedure consists of different phases. A flow chart explaining the entire process is shown in figure 2. The first step is patient preparation for image acquisition. A variable number (4–10 recommended) of plastic adhesive fiducial markers (3 mm thick, 15 mm wide), specifically designed to be easily identified both on CT and MR images, are attached to the skin of the patient's head. The next step is image acquisition; during this phase the patient is not immobilized as in stereotactic image acquisition (for example, in ring frame procedures). The complete stack of images is imported in the workstation for treatment planning. Using the navigation software tools, the operator identifies each fiducial marker on the images ($\vec{m}_i^{\mathrm{Im}}$ is the vector representing the position of the centre of the ith marker in the image reference system).

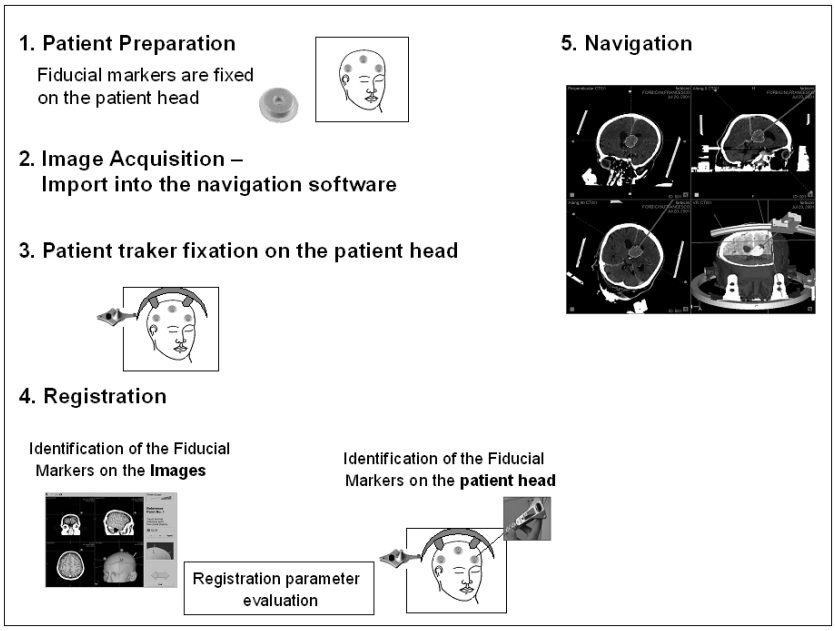

Figure 2. Flow chart showing the different phases of the neuronavigation procedure.

At this point, when the user places the tip of the Pointer in the centre of each fiducial marker, the position of each led of the Pointer is identified, and the system calculates the vector representing the position of the centre of the *i*th marker in the patient's reference system $\left(\vec{M}_i^{\mathrm{Poi}}\right)$. This procedure is repeated for the whole set of markers. Then a correlation between image and patient's reference system, which allows transformation of the coordinates of points of the image reference system into the patient's reference system, is established. This phase is called registration.

In order to perform this task, the registration algorithm (Arun *et al* 1994) calculates the best estimator of the rigid rototranslation matrix (T)

$$\left(\vec{M}_i^{\mathrm{Poi}}\right)' = T\vec{m}_i^{\mathrm{Im}}$$

minimizing the following expression:

$$\sum_i \left(\left(\vec{M}_i^{\mathrm{Poi}}\right)' - \vec{M}_i^{\mathrm{Poi}}\right)^2$$

(all the vectors and the matrix are expressed in homogeneous coordinates).

At the end of the calculation, the system shows, for each marker, the difference between the measured value $\left(\vec{M}_i^{\mathrm{Poi}}\right)$ and the calculated one $\left(\left(\vec{M}_i^{\mathrm{Poi}}\right)'\right)$ in absolute value. In the following sections, as the system does, we will refer to the mean value of these differences as the registration error.

Once the registration procedure is ended, it is possible to navigate in the intracranial space with the Pointer. The position of the Pointer's tip, as directly measured by the optical sensors, is expressed in coordinates belonging to the patient's reference system. The application of the previously calculated transformation matrix leads to the expression of the position of the Pointer's tip in image coordinates, and to visualization of it on diagnostic images. Now the tool can be guided towards the intended target along the planned trajectory.

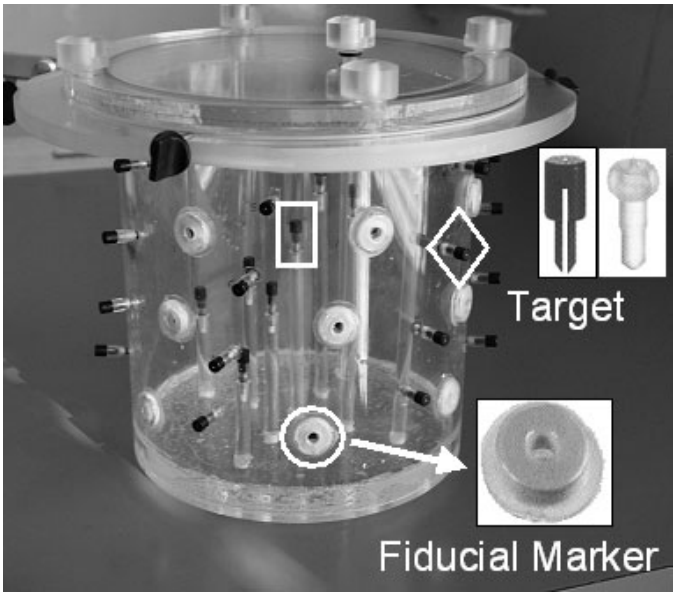

Figure 3. Phantom used for accuracy evaluation. Internal targets (in rectangular frame, cylindric for the navigation phase, spherical for image acquisitions), fiducial markers (within circle, used for the registration procedure) and external targets (in rhomboidal frame, not used in these measurements), are clearly visible. In all the acquisitions the phantom was placed with its axis along the feet–head direction.

2.2. The phantom

A plexiglas phantom developed and constructed at the Niguarda Hospital in Milan was used for the measurements (Torresin *et al* 1999) (see figure 3). It is cylindrical in shape, and one of the two circular bases can be removed in order to reach the nine internal targets, located on plexiglas columns of different heights.

On the lateral surface 16 external targets and 9 adhesive fiducial markers which can be used in the registration procedure are fixed. The coordinates of each internal target were measured with a Mitutoyo Mechanical Touch FHN906 (Quality, Novara, Torresin *et al* 1999) with an accuracy of about 0.01 mm. No measurements of the coordinates of the external targets and of the cutaneous fiducial markers were available. A planar scheme is presented in figure 4(a). In our measurements only the adhesive markers were used to perform the registration procedure.

The phantom was designed to evaluate the accuracy of the image acquisition and target localization in the navigation procedure. For this purpose different targets, to be fixed to the Plexiglas columns, were designed. Radiopaque spheres (3 mm in diameter) were used to acquire CT images, while hollow plastic spheres, filled with contrast medium, were used to acquire MR images. The filling material for MR scans was a physiological solution containing Gd (2% *Prohance*) having T1 $\approx$ 10 ms and T2 $\approx$ 16 ms. The output signal of the targets is intense in both CT and T1 weighted MR images. In T2 weighted acquisitions the spheres appear void of signal, keeping high contrast with the surrounding medium. In all cases the phantom was filled with water. The visualization tools of the neuronavigation software permit a feasible identification of the sphere centres (figure 5). Plastic cylindrical caps (1.5 mm height) with a point drawn on the top of them, were fixed to the plexiglas columns and used

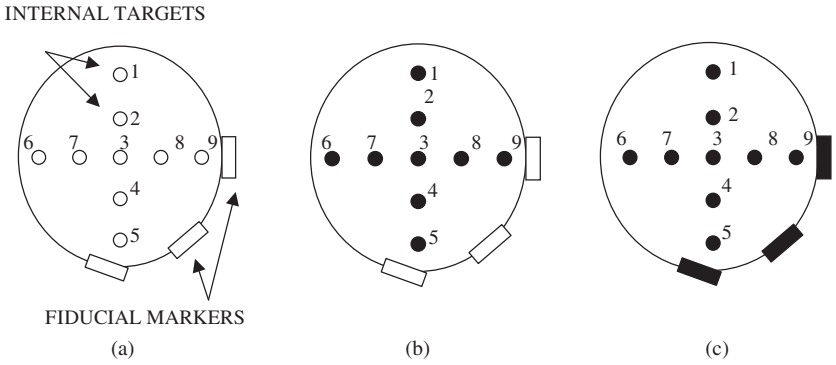

Figure 4. (a) Schematic view of the neuronavigation phantom, showing the positions of internal targets and fiducial markers. (b). In the imaging accuracy test the 1–9 internal targets are used both for the IDL registration procedure and for the accuracy evaluation. (c). In the global accuracy test the fiducial markers are used for the neuronavigator registration procedure, while the 1–9 internal targets are used for the accuracy evaluation.

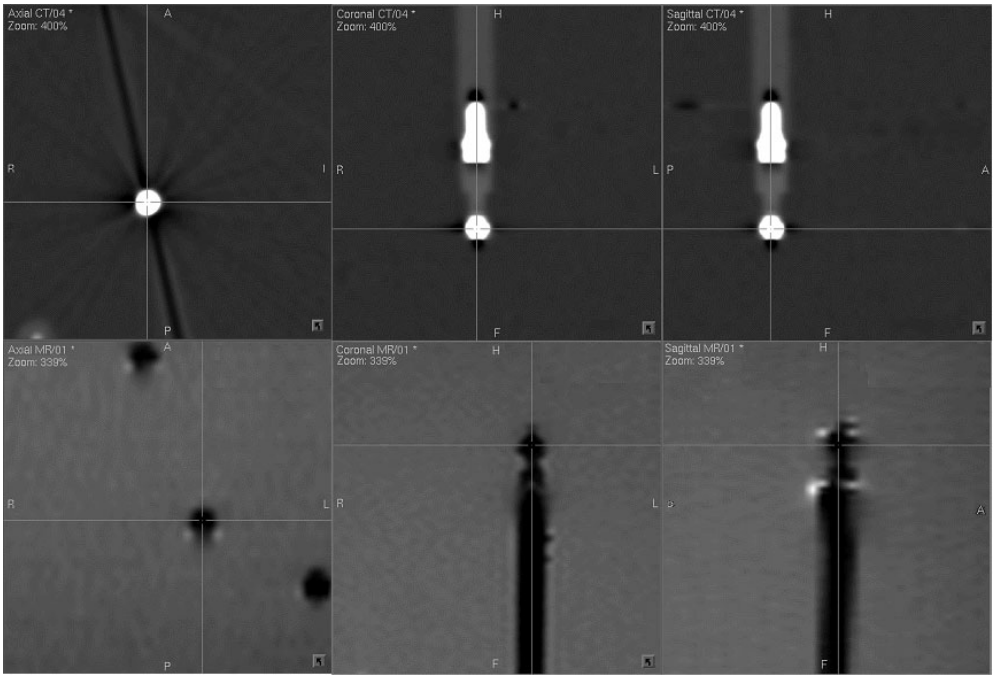

Figure 5. Detail of CT (upper) and MR (lower) phantom images where the centre of a spherical internal target is identified by the cross tool. From left to right: axial, coronal and sagittal image orientations. CT scans are sequential, MR images are T2 weighted.

as targets during the navigation procedure. The point positions on the cap coincide with the centre positions of the spheres used during CT and MRI scans. All the targets (cylindrical caps and spheres) and markers are those listed in the Leibinger catalogue.

Such a phantom allows the Pointer to be placed exactly in the centre position of the intended target, thus obtaining a direct comparison of the target coordinates measured in the

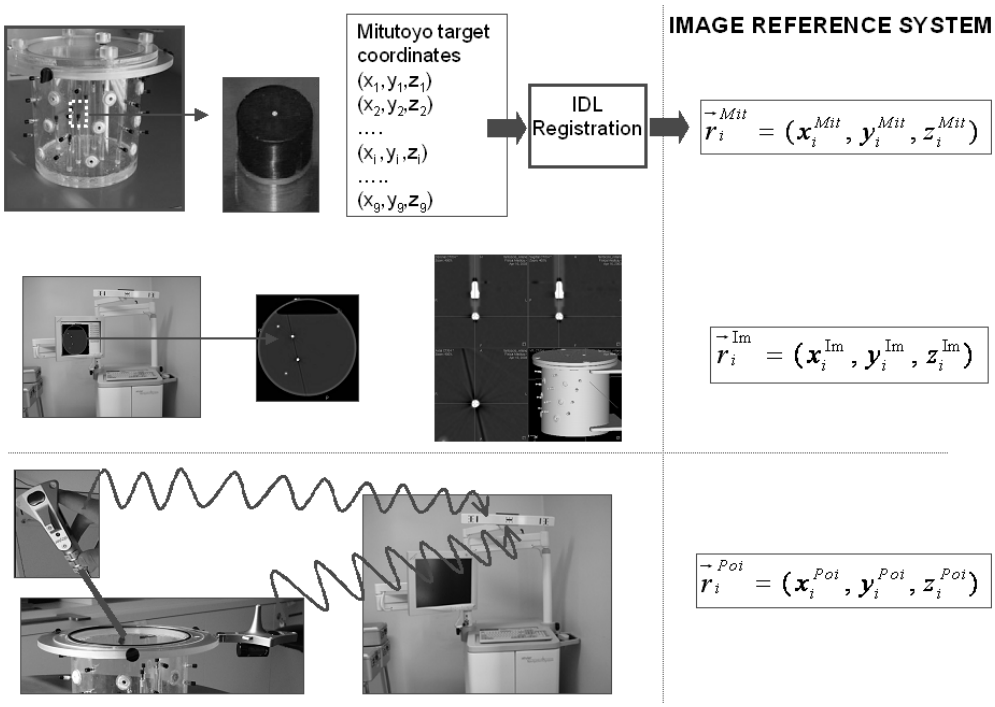

Figure 6. The three target coordinate sets: image ($\vec{r}_i^{\,\mathrm{Im}}$), pointer ($\vec{r}_i^{\,\mathrm{Poi}}$) and Mitutoyo ($\vec{r}_i^{\,\mathrm{Mit}}$); all of them are expressed in the image reference system. Because of the high accuracy of the Mitutoyo Mechanical Touch, this coordinate set is considered as the gold standard.

diagnostic images with the coordinates of the same target displayed by the neuronavigation system.

2.3. Accuracy evaluation procedure

In order to quantify the accuracy of the procedure, three coordinate sets are defined (figure 6). The first set consists of the coordinates of the internal targets measured on the images:

$$\vec{r}_i^{\,\mathrm{Im}} \rightarrow \left(x_i^{\mathrm{Im}}, y_i^{\mathrm{Im}}, z_i^{\mathrm{Im}}\right) \qquad i = 1, \ldots, 9.$$

The second one consists of the coordinates of the internal targets as they are measured by the system. When the Pointer is placed to touch the point drawn on the top of the target, the system detects the coordinates (expressed in the patient's reference system) and transforms them into the image reference system, obtaining

$$\vec{r}_i^{\,\mathrm{Poi}} \rightarrow \left(x_i^{\mathrm{Poi}}, y_i^{\mathrm{Poi}}, z_i^{\mathrm{Poi}}\right) \qquad i = 1, \ldots, 9.$$

The last one consists of the real coordinates of the internal targets in the phantom, measured with the Mitutoyo Mechanical Touch. Due to the high accuracy of these measurements, this coordinate set is considered as our gold standard. In order to compare these data with the first two sets of data (expressed in the same reference system), a rigid rototranslation has to be applied off-line. The parameters used for this transformation were calculated using a routine based on the work of Arun *et al* (1994). It was specifically developed for this purpose using the

Table 1. Protocol details of the investigated CT series. Each scan was performed with no gantry tilt or oblique plane orientation; 512 × 512 matrices were reconstructed. 120 kVp and 100 mAs were used for all series.

Sequence	Notes	FOV (mm)	Pixel size (mm)	Slice thickness (mm)
CT 1 mm	Sequential interslice gap = 0	230	0.45	1
CT 2 mm	Sequential interslice gap = 0	230	0.45	2
CT 2 mm SP	Spiral pitch = 1 length = 20 cm	230	0.45	2
CT 2 mm FOV L	Sequential interslice gap = 0	320	0.63	2
CT 4 mm	Sequential interslice gap = 0	230	0.45	4

Table 2. Protocol details of the investigated MR series. Each scan was performed with no gantry tilt or oblique plane orientation and with no interslice gap; 512 × 512 matrices were reconstructed. FOV was always 230 mm and pixel size 0.45 mm. NEX = 1 was used for all the series. The pixel bandwidth (water–fat pixel shift) was 205 Hz and the gradient mode regular. The higher SNR of the 4 mm T1 series enabled the use of a shorter TR value.

Sequence	Notes	Number of slices	Slice thickness (mm)	TR (ms)	TE (ms)
MR 1 mm T2	T2 weighted	200	1	2500	100
MR 2 mm T1	T1 weighted	100	2	1235	15
MR 2 mm T2	T2 weighted	100	2	2500	100
MR 4 mm T1	T1 weighted	50	4	597	15
MR 4 mm T2	T2 weighted	50	4	2500	100

IDL platform (Research Systems Inc., Boulder CO). The coordinates used for this calculation are the internal targets' ones (as shown in figure 4(b)).

$$\vec{r}_i^{\,\mathrm{Mit}} \rightarrow \left(x_i^{\mathrm{Mit}}, y_i^{\mathrm{Mit}}, z_i^{\mathrm{Mit}}\right) \qquad i = 1, \ldots, 9.$$

If the whole procedure was completely error free, these three coordinate sets should be coincident. Hence, comparing the three sets allows a quantification of the errors to be assigned to the described procedure. It is possible to identify the imaging accuracy and the global (navigation) accuracy, in which also the registration error and the mechanical error are considered. In this work, measurements aimed to allow an independent evaluation of the registration software algorithm and of the mechanical error associated with the neuronavigation hardware were not performed. It was chosen to estimate imaging accuracy and global accuracy and to discuss the differences, which can be ascribed to registration and mechanical inaccuracies.

2.3.1. Imaging accuracy test. In order to investigate how image acquisition parameters affect the navigation procedure accuracy, ten CT and MRI phantom studies with different scan parameters were performed.

The acquisition protocol details are reported in tables 1 and 2.

The image reproducibility was tested repeating three times selected scans (all CT acquisitions, MR 2 mm T2 and MR 2 mm T1).

A Siemens Somatom PLUS 4 Volume Zoom multislice CT scanner (8 detectors_adaptive array) was used to acquire CT images. A 1.5 T Philips Gyroscan ACS-NT 5.3 MR unit (maximum gradient 23 mT/m, rise time 0.2 ms) was used for MRI acquisitions. Standard T1-weighted and T2-weighted spin echo sequences were used. MR 1 mm T2 series was acquired with the 'half scan' technique in order to reduce acquisition time.

For all the images the X, Y and Z axes are oriented along the right–left, posterior–anterior and feet–head directions, respectively. In MR scans phase and frequency encoding directions are along the X and Y axes, respectively, the slice is selected along the Z axis.

The first set of differences, calculated out of the three coordinate sets, is

$$\left(x_i^{\mathrm{Im}} - x_i^{\mathrm{Mit}}\right) = \Delta x_i^{\mathrm{I-M}} \qquad i = 1, \ldots, 9$$

and the same for the y and z directions.

The variable $\Delta x^{\mathrm{I-M}}$ is expected to be Gaussian distributed, with a mean value equal to 0 if no systematical errors affect the whole data set. The mean values $\overline{\Delta x^{\mathrm{I-M}}}$, $\overline{\Delta y^{\mathrm{I-M}}}$, $\overline{\Delta z^{\mathrm{I-M}}}$ (biases) and the standard deviations $\sigma_X^{\mathrm{I-M}}$, $\sigma_Y^{\mathrm{I-M}}$, $\sigma_Z^{\mathrm{I-M}}$ (precisions) are used to quantify the accuracy which is only due to the imaging phase.

Starting from x_i^{Mit}, y_i^{Mit}, z_i^{Mit} $i = 1, \ldots, 9$ values, it is possible to calculate, for all the possible target couples, the distance between two different targets; in this way $d_{i,j}^{\mathrm{Mit}}$ $i, j = 1, \ldots, 9$ and $i \neq j$ are obtained. Similar calculations lead to $d_{i,j}^{\mathrm{Im}}$.

The distance between two targets measured with the two methods should be the same, unless the imaging technique introduces distortions; MRI distortions are well-known experimental findings, as referred to in the literature (Sumanaweera *et al* 1993, 1994, Hardy and Barnett 1998, Kondziolka *et al* 1992, Bourgeois *et al* 1999, Dipierro *et al* 1999, Debus *et al* 1996, Prott *et al* 1995, Weis *et al* 1998, Chang and Fitzpatrick 1992). In such cases, the distribution of

$$\Delta d_{i,j}^{\mathrm{I-M}} = d_{i,j}^{\mathrm{Im}} - d_{i,j}^{\mathrm{Mit}} \qquad i, j = 1, \ldots, 9 \quad i \neq j$$

and in particular the mean value $\overline{\Delta d^{\mathrm{I-M}}}$ and the standard deviation $\sigma_d^{\mathrm{I-M}}$ represent important parameters to quantify the described effect.

This evaluation has been performed using only internal target coordinates (as shown in figure 4(b)) because fiducial markers and external targets data were not available. Therefore, only the distortion pattern of the internal region of the phantom could be investigated. Moreover, internal targets and fiducial markers are made of different materials; since it is well known that MRI distortions are material-dependent due to magnetic susceptibility variations, the proposed evaluation is only intended as a first attempt to quantify deformations.

2.3.2. Global accuracy test. An idea of the entire neuronavigation accuracy can be drawn by the surgeon from the previously described registration error, directly calculated by the system. In a first analysis this value depends on the accuracy of imaging and registration phases but, since the registration procedure is performed by the Pointer, it also depends on the mechanical errors associated with the neuronavigation hardware.

In order to get a clearer quantification of the global accuracy of the procedure, a different test was performed. The Pointer was used to reach the internal targets (see figure 4(c)) inside the phantom.

In this way, a set of differences can be calculated:

$$\left(x_i^{\mathrm{Poi}} - x_i^{\mathrm{Mit}}\right) = \Delta x_i^{\mathrm{P-M}} \qquad i = 1, \ldots, 9.$$

As in the previous cases, the variable $\Delta x^{\mathrm{P-M}}$ is expected to be Gaussian distributed, with a mean value equal to 0 if no systematical errors affect the whole data set. Mean values $\overline{\Delta x^{\mathrm{P-M}}}$, $\overline{\Delta y^{\mathrm{P-M}}}$, $\overline{\Delta z^{\mathrm{P-M}}}$ and standard deviations $\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$ were calculated. During this test, the distance between the frame and the phantom was chosen to be about 1.5 m (as recommended by the manufacturer). This allows a minimization of the contribution of mechanical accuracy to the global error (Kaus *et al* 1997). The pointer reproducibility in

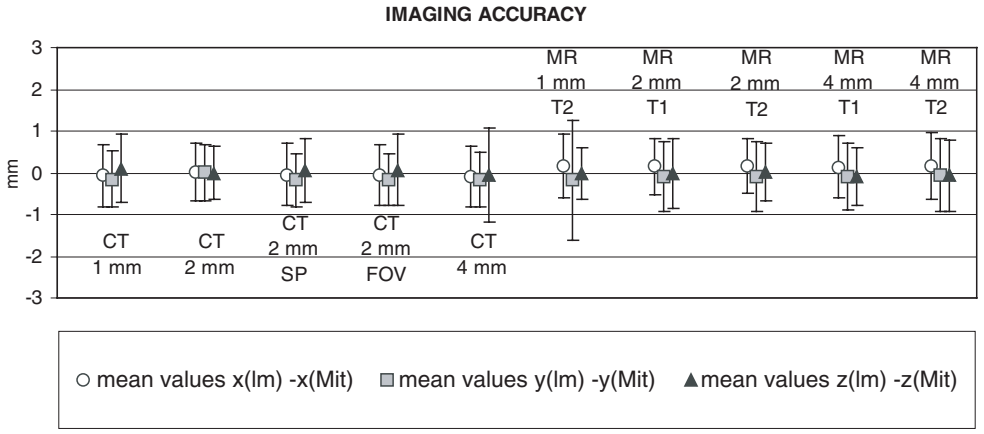

Figure 7. Mean values $\overline{\Delta x^{\mathrm{I-M}}}$, $\overline{\Delta y^{\mathrm{I-M}}}$, $\overline{\Delta z^{\mathrm{I-M}}}$; $\sigma_X^{\mathrm{I-M}}, \sigma_Y^{\mathrm{I-M}}, \sigma_Z^{\mathrm{I-M}}$ are shown as error bars. Mean values are coherent with the expected 0 value within one standard deviation. The MR 1 mm T2 series suffers the greatest random error along the y direction, probably due to artefacts, while CT 4 mm shows a decrease of precision along the z direction.

reaching targets was tested by touching the desired position (white spot on targets) several times.

In order to investigate the dependence of global accuracy on the number of fiducial markers used in the registration phase, the differences $\overline{\Delta x^{\mathrm{P-M}}}$, $\overline{\Delta y^{\mathrm{P-M}}}$, $\overline{\Delta z^{\mathrm{P-M}}}$ were calculated using three, six and nine markers. Differences among the three transformation matrices can be expected.

An overall idea of the global accuracy of the procedure can also be drawn rearranging the differences between Pointer and Mitutoyo coordinates, in order to calculate Euclidean distances $\Delta r^{\mathrm{P-M}}$:

$$\left|\vec{r}_i^{\,\mathrm{Poi}} - \vec{r}_i^{\,\mathrm{Mit}}\right| = \Delta r_i^{\mathrm{P-M}} \qquad i = 1, \ldots, 9.$$

This approach has the advantage of being near to the surgeon ideal, because it allows quantification of the distance between the intended target (identified with Mitutoyo coordinates) and the actually reached target (identified with Pointer coordinates) (Vannier and Haller 1999).

3. Results

The acquired images proved to be suitable to perform a neuronavigation procedure. Only one specific series (MR 1 mm T2) showed clear indications of artefacts. Hence, these data were only used for the imaging accuracy test.

3.1. Imaging accuracy

The tests performed to assess the image reproducibility resulted in standard deviations well below the pixel size for MR images and much lower than the pixel size in the case of CT acquisitions. The mean values $\overline{\Delta x^{\mathrm{I-M}}}$, $\overline{\Delta y^{\mathrm{I-M}}}$, $\overline{\Delta z^{\mathrm{I-M}}}$ for each imaging series are shown in figure 7, together with the corresponding $\sigma_X^{\mathrm{I-M}}$, $\sigma_Y^{\mathrm{I-M}}$ and $\sigma_Z^{\mathrm{I-M}}$ reported as error bars. Calculated biases are small; larger ones are obtained when MR data are used. Taking into specific account the values of standard deviations, one can observe that increasing the slice

Table 3. Mean values of distances between couples of target points $\overline{\Delta d_{(5)}^{\mathrm{I-M}}}$ and standard deviations $\sigma_{d(5)}^{\mathrm{I-M}}$ for the five most internal targets.

Sequence	$\overline{\Delta d_{(5)}^{\mathrm{I-M}}}$ (mm)	$\sigma_{d(5)}^{\mathrm{I-M}}$ (mm)
CT 1 mm	0.29	0.23
CT 2 mm	0.20	0.22
CT 2 mm SP	−0.03	0.20
CT 2 mm FOV L	0.01	0.33
CT 4 mm	−0.65	0.71
MR 1 mm T2	−0.80	0.81
MR 2 mm T1	−0.64	0.32
MR 2 mm T2	−0.58	0.30
MR 4 mm T1	−0.56	0.52
MR 4 mm T2	−0.88	0.53

Table 4. Mean values $\overline{\Delta x_{(5)}^{\mathrm{I-M}}}$, $\overline{\Delta y_{(5)}^{\mathrm{I-M}}}$, $\overline{\Delta z_{(5)}^{\mathrm{I-M}}}$ and standard deviations $\sigma_{X(5)}^{\mathrm{I-M}}$, $\sigma_{Y(5)}^{\mathrm{I-M}}$, $\sigma_{Z(5)}^{\mathrm{I-M}}$. Mean values calculated out of the whole CT series and out of the whole MR series are also shown. MR 1 mm T2 series is excluded from the analysis due to the discussed artefact pattern. It can be seen that the z variable is more strongly affected by the documented distortion.

Sequence	$\overline{\Delta x_{(5)}^{\mathrm{I-M}}}$	$\sigma_{X(5)}^{\mathrm{I-M}}$	$\overline{\Delta y_{(5)}^{\mathrm{I-M}}}$	$\sigma_{Y(5)}^{\mathrm{I-M}}$	$\overline{\Delta z_{(5)}^{\mathrm{I-M}}}$	$\sigma_{Z(5)}^{\mathrm{I-M}}$
CT 1 mm	0.1	0.9	0.0	0.5	0.3	0.9
CT 2 mm	0.2	0.9	0.2	0.5	0.1	0.7
CT 2 mm SP	0.1	0.9	0.0	0.5	0.1	0.7
CT 2 mm FOV L	0.1	0.9	0.0	0.5	0.2	0.9
CT 4 mm	0.1	0.8	0.0	0.5	−0.3	1.2
MR 2 mm T1	0.0	0.8	−0.1	0.9	−0.4	0.8
MR 2 mm T2	0.0	0.8	−0.1	1.0	−0.2	0.7
MR 4 mm T1	0.0	0.9	−0.2	0.9	−0.3	0.8
MR 4 mm T2	0.0	0.9	−0.1	1.1	−0.4	0.9
Mean CT	0.14		0.02		0.11	
Mean MR	−0.01		−0.17		−0.31	

thickness in CT series leads to an increase of $\sigma_Z^{\mathrm{I-M}}$ (this is only a random, not a systematic effect). Moreover, the y variable in series MR 1 mm T2 suffers the greatest random error.

The measured precisions are not extremely sensitive to slice thickness (except for CT 4 mm, as pointed out), especially in the MRI series. It can also be seen that changing the FOV from 230 × 230 mm to 320 × 320 mm does not alter the precision along the x and y directions, and choosing a spiral sequence does not introduce a lack of precision along the three directions.

Evidences of distortions in MR images were looked for. The analysis of the distances between two targets shows that the deformation effect is more evident for the five most internal ones. $\overline{\Delta d_{(5)}^{\mathrm{I-M}}}$ and $\sigma_{d(5)}^{\mathrm{I-M}}$ values, reported in table 3, represent the mean value and the standard deviation of the distribution of the $\Delta d_{i,j}^{\mathrm{I-M}}$, taking into consideration only the five most internal targets. Differences between CT-based data and MRI-based data are evident.

It is interesting to look for the direction which is most sensitive to the described distortion. In table 4 $\overline{\Delta x_{(5)}^{\mathrm{I-M}}}$, $\overline{\Delta y_{(5)}^{\mathrm{I-M}}}$, $\overline{\Delta z_{(5)}^{\mathrm{I-M}}}$ and $\sigma_{X(5)}^{\mathrm{I-M}}$, $\sigma_{Y(5)}^{\mathrm{I-M}}$, $\sigma_{Z(5)}^{\mathrm{I-M}}$ are shown. The MR 1 mm T2 series is excluded from the analysis due to the discussed artefact pattern. It can be seen that the z variable is the most strongly affected one.

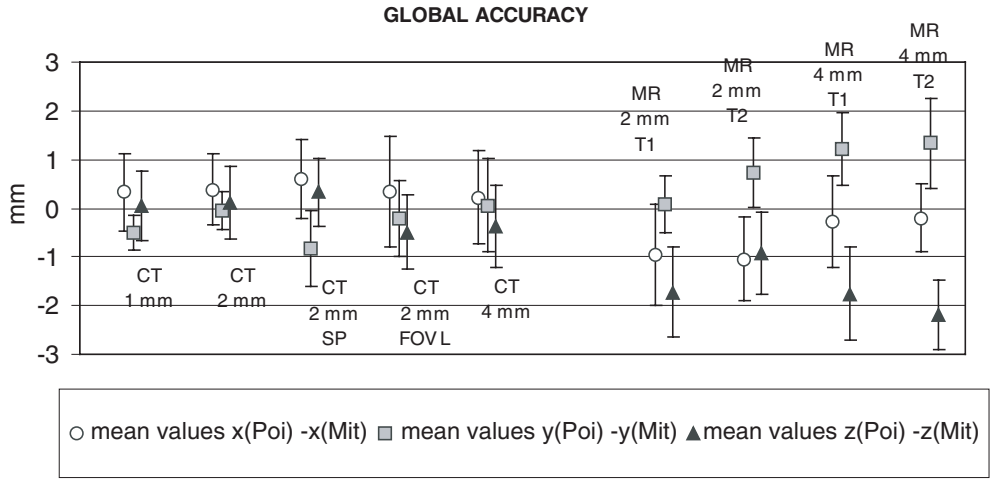

Figure 8. Mean values $\overline{\Delta x^{\mathrm{P-M}}}$, $\overline{\Delta y^{\mathrm{P-M}}}$, $\overline{\Delta z^{\mathrm{P-M}}}$; $\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$ are shown as error bars. Mean values of procedures based on MR imaging show meaningful deviations from the expected 0 value, especially in the z direction.

Table 5. Registration errors, as they are reported by the neuronavigation software.

Sequence	Registration errors (mm)
CT 1 mm	0.68
CT 2 mm	0.64
CT 2 mm SP	0.73
CT 2 mm FOV L	0.71
CT 4 mm	0.93
MR 1 mm T2	1.74
MR 2 mm T1	1.38
MR 2 mm T2	1.54
MR 4 mm T1	1.69
MR 4 mm T2	1.84

3.2. Global accuracy

The pointer reproducibility was first investigated, because it is a necessary condition for the significance of the entire procedure. Touching the desired target position several times, from different orientations, differences in coordinates are negligible.

A first idea of the overall accuracy can be drawn from the registration errors; the values obtained from a nine marker-based registration, and estimated as previously explained, are shown in table 5. If CT images are used to plan the procedure, registration errors appear smaller (mean value 0.7 mm, max. deviance 0.2 mm), while they turn out to be larger if MR images are used (mean value 1.6 mm, max. deviance 0.3 mm). The correlation between distortion and registration error is evident when analysing table 3 together with table 5: two sets of experimental points can be identified, the former with low distortions (and low registration errors) and the latter with high distortions (and high registration errors).

Differences between Pointer and Mitutoyo coordinates quantify the global errors of the navigation procedure. Biases and precisions are shown in figure 8. Data show meaningful deviations from the expected 0 value. CT series with 2 mm slices show biases and precisions which are lower than CT 2 mm SP and CT 2 mm FOV L series ones. In contrast, MRI data

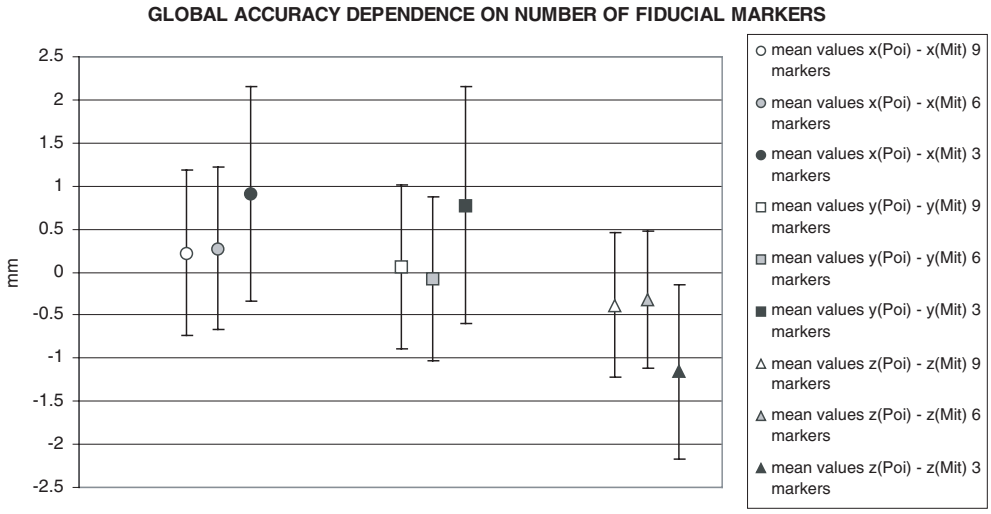

Figure 9. Effect of the number of markers used for the registration. Mean values $\overline{\Delta x^{\mathrm{P-M}}}$, $\overline{\Delta y^{\mathrm{P-M}}}$, $\overline{\Delta z^{\mathrm{P-M}}}$ and standard deviations $\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$ (error bars), are calculated out of the data obtained using nine, six or three marker-based registration. A particular series (CT 4 mm) was chosen.

Table 6. A deviation from zero of the $(\vec{r}^{\mathrm{Poi}} - \vec{r}^{\mathrm{Mit}})$ mean value indicates a systematic error. In the table the number of standard deviations ($\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$) within which the displacements along the three axes are included is reported. Meaningful deviations are found, especially in the z direction for procedures based on MR imaging.

Sequence	X interval	Y interval	Z interval
CT 1 mm	1	2	1
CT 2 mm	1	1	1
CT 2 mm SP	1	2	1
CT 2 mm FOV L	1	1	1
CT 4 mm	1	1	1
MR 2 mm T1	1	1	2
MR 2 mm T2	2	2	2
MR 4 mm T1	1	2	2
MR 4 mm T2	1	2	4

do not show clear indications of accuracy dependence on T1 and T2 weights. The increase of the FOV is responsible for a small decrease of navigation precision in the x direction, but not for the increase of the bias in that direction.

In table 6 the intervals (in units of $\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$ respectively) containing the 0 value are reported e.g., the first value of the first column is 1; it means that the 0 value is contained in the interval $\overline{\Delta x^{\mathrm{P-M}}} \pm \sigma_X^{\mathrm{P-M}}$ for the CT 1 mm series. Meaningful deviations are shown, especially in the z direction for procedures based on MR imaging.

The effect of the number of markers used for the registration was investigated. Data obtained using a nine or six marker-based registration are more accurate than those obtained using a three marker-based registration. This can be observed in figure 9, where a particular series (CT 4 mm) is chosen as an example. Mean values $\overline{\Delta x^{\mathrm{P-M}}}$, $\overline{\Delta y^{\mathrm{P-M}}}$, $\overline{\Delta z^{\mathrm{P-M}}}$ and standard deviations $\sigma_X^{\mathrm{P-M}}$, $\sigma_Y^{\mathrm{P-M}}$, $\sigma_Z^{\mathrm{P-M}}$, calculated out of the data based on three, six and nine

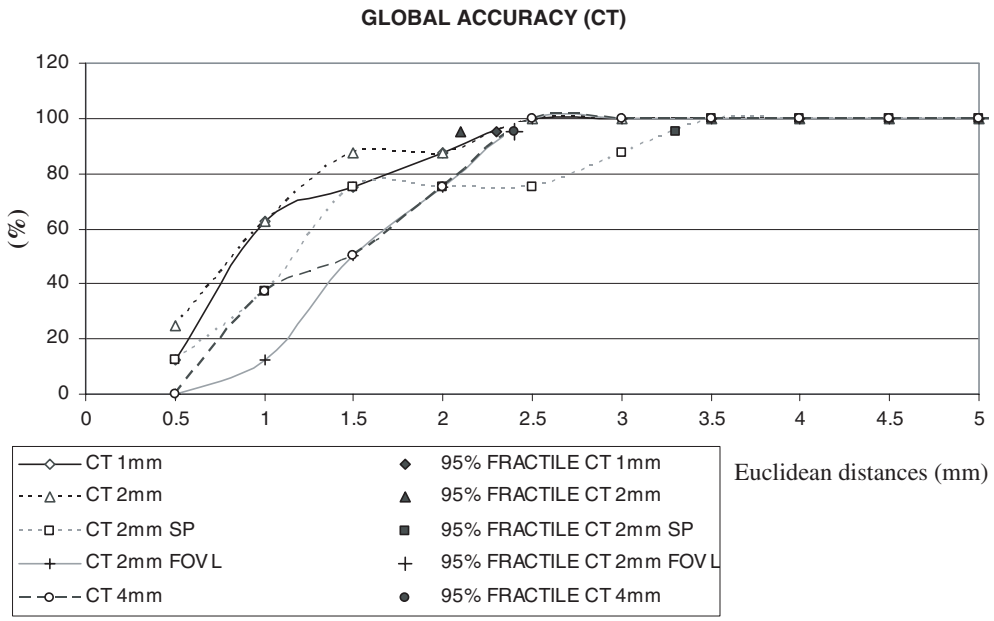

Figure 10. Cumulative distributions of the Euclidean distances $\Delta r^{\mathrm{P-M}}$ for procedures based on CT images. The 95% fractile of each distribution is also plotted. Different behaviour between CT 2 mm FOV L and CT 2 mm SP series can be observed. The 95% fractile of the CT 2 mm FOV L series coincides with that of the CT 4 mm series.

marker-based registrations, are compared. One can state that a registration based on six or nine markers approximately results in the same accuracy, while three markers result in a lower global accuracy.

Figures 10 and 11 show, for each imaging sequence, the cumulative distribution of the Euclidean distances $\Delta r^{\mathrm{P-M}}$. The 95% fractile of each distribution is also plotted. Mitutoyo coordinates are to be considered as 'intended target' coordinates, while Pointer ones are to be considered as 'reached target' ones. Different behaviour between CT 2 mm FOV L and CT 2 mm SP series can be observed. The CT 2 mm FOV L trend is quite smooth, meaning that short distances (correct targeting) and large distances (wrong targeting) are more or less equally probable; in detail, distances smaller than 2 mm were found in 75% of the cases and the 95% fractile is at 2.4 mm. In contrast, CT 2 mm SP series data show that good targeting is highly probable (distances smaller than 1.5 mm were found in 75% of the cases), while in the remaining cases wrong targeting can pass 3 mm (the 95% fractile is at 3.3 mm).

General considerations about MRI-based navigation can be drawn by analysing Euclidean distances in figure 11. Values from 3 to 4 mm are typical for 95% fractiles, and, while there is evidence that 2 mm slice thickness series are more accurate than 4 mm ones, T1 weighted and T2 weighted series show similar behaviour.

4. Discussion

4.1. Imaging accuracy

One of the most important aspects of this investigation is to compare the accuracy of a navigation procedure based on CT to that of a procedure based on MR images. It is well

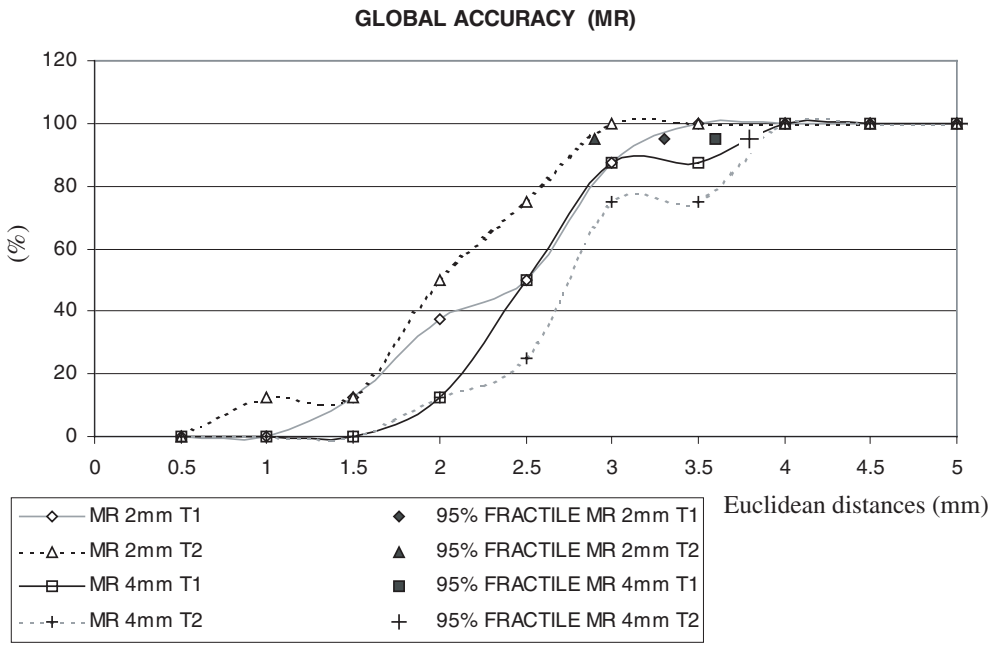

Figure 11. Cumulative distributions of the Euclidean distances $\Delta r^{\mathrm{P-M}}$ for procedures based on MR images. The 95% fractile of each distribution is also plotted. Values for 95% fractiles are larger than those obtained with procedures based on CT images.

known that MR images can be affected by many kinds of artefacts. In the worst cases (such as in our MR 1 mm T2 series, which displayed artefacts in the frequency-encoded direction, while it showed a good spatial resolution in the *xz* plane) the effect is so evident that it is necessary to discard the images.

However, imaging distortion is a misleading phenomenon; even when the image is apparently artefact-free, it can be still inaccurate and distorted, as reported in the literature (Sumanaweera *et al* 1993, 1994, Hardy and Barnett 1998, Kondziolka *et al* 1992, Bourgeois *et al* 1999, Dipierro *et al* 1999, Debus *et al* 1996, Prott *et al* 1995, Weis *et al* 1998, Chang and Fitzpatrick 1992). Additionally, MR is a low signal-to-noise ratio (SNR) technique. As an example, for 2 mm slices, the SNR found for CT sequential study is around 100 while for MR T1 and T2 weighted acquisitions it is about 13 and 28, respectively. The tests performed to assess the image reproducibility, together with the similar data dispersions found for CT and MR studies, suggest, however, that the unfavourable SNR of the MR scans does not affect the accuracy of the navigation procedure.

In many works, CT images are considered as gold standards for imaging accuracy (Sumanaweera *et al* 1993, Kondziolka *et al* 1992, Dipierro *et al* 1999, Debus *et al* 1996); in our paper this subject is studied in more detail due to the fact that, besides the coordinates of targets on CT images, also their 'real' coordinates, as measured with great accuracy by a Mitutoyo Mechanical Touch, are taken into account.

The differences between Mitutoyo and imaging coordinates (figure 7) show that using MR data results in a slightly lower accuracy than that obtained using CT data, independent of the particular chosen sequence. This effect is to be related to MRI distortions. The investigation of the distances between the nine targets confirms this hypothesis. The effect is even more evident if only the five most internal targets are considered, as shown in table 3, probably due

to a more intense deformation pattern in the internal region. A more detailed study proved the z direction to be the most sensitive one, as quantified by table 4.

This effect is heavily involved in the registration inaccuracy of MRI-based neuronavigation procedures. Therefore, global accuracy can be expected to be less accurate when MR images are used.

4.2. Global accuracy

Despite its name, the registration error depends not only on the accuracy of the registration phase itself, but also on the imaging and mechanical accuracy of the procedure. However, the last source of error is expected to act in the same way for CT- and MR-based procedures. So the far greater registration error which has been found for MR images (table 5) cannot be due to the mechanical errors of the system. Since the algorithm used by the system is a rigid one, the hypothesis that images do not introduce any distortions in marker positions is of main concern for registration accuracy. In the previously discussed imaging accuracy evaluation, we quantified MRI distortion using target coordinates. These data can only be considered as deformation guidelines; in fact, a specifically intended study of distortions for this experimental set-up should take into consideration both a wider area of the phantom (where fiducial markers are fixed), and the different materials the phantom is made of (targets and markers are made of different materials). However, general trends of distortions can be deduced from our data. Using the described deformation estimator, a correlation can be found between the registration and the presence of image distortion: imaging series with low distortions show low registration errors, while those with high distortions correlate to high registration errors (tables 3 and 5).

Our test also demonstrates that the global accuracy heavily depends on the registration accuracy; in fact results belonging to MR images show systematic errors, due to the quantified distortion effects. This idea is confirmed by figure 8 and table 6, where values based on CT images prove to be closer to the ideal 0 value than MRI-based ones.

Table 4 identifies z as the most distorted direction. Moreover, the largest biases in MR-based neuronavigations are obtained in the z direction, as shown in figure 8 and table 6. This evidence shows that the rigid point-based registration algorithm translates this larger deformation pattern into greater biases along the same direction.

Comparing the two figures displaying imaging accuracy (figure 7) and global accuracy (figure 8), different considerations can be done for CT and MR data.

CT-based global accuracy data show a greater dispersion than imaging accuracy data. This effect can be due to two different reasons. First, the global accuracy depends on the mechanical inaccuracies of the system, which, in our experimental set-up cannot be measured because it is mixed into the other error sources. On the other hand, there is an important difference between the IDL registration procedure (used for the imaging accuracy test) and the registration performed by the neuronavigation system (used for the global accuracy test): the former is based on target coordinates, while the latter is based on marker coordinates. The identification of the centre of each marker by the Pointer is less accurate due to marker size. Therefore, this error source also contributes to neuronavigation global inaccuracy. These two contributions proved to be the most important ones for CT-based procedures.

Also MR-based global accuracy data show a greater dispersion than imaging accuracy data. These data suffer from the two just mentioned error sources as well.

However, in the imaging accuracy test, MR and CT-based data do not differ as much as in the global accuracy test. The reason is related to the different sets of coordinates used

for the two registration procedures. Fiducial marker coordinates, used in the registration performed by the neuronavigation system (for the global accuracy test), can suffer from a more intense distortion pattern than target coordinates, used in the registration performed by the IDL procedure (for the imaging accuracy test). This can be due to the position of markers, which are located in the outer regions of the field of view. The registration procedure performed by the neuronavigation system proved less accurate due to its dependence on marker coordinates. This effect turns out to be responsible for the larger biases in MR-based navigations.

The best image acquisition sequences for neuronavigation can be identified from the distributions of Euclidean distances (figures 10 and 11); in this task one can look for a compromise between global accuracy and acquisition time. For CT acquisition, the best results are obtained with sequential, 1 or 2 mm scans. Spiral data show similar performances, but have the advantage of decreasing acquisition times, thus reducing motion artefacts, which can be important when imaging a patient. Due to the mentioned MRI deformation effects, the presented T1- and T2-weighted spin echo MR sequences suffer from large inaccuracies; the obtained values cannot be considered satisfying.

A possible way of solving the identified distortion problems could be the use of non-rigid, more complex registration algorithms to take into account this imaging effect. On the other hand, a different approach is to deeply investigate MRI parameters in order to find sequences which are less sensible to distortions. A preliminary study has been performed by our group in order to evaluate the results which can be obtained with gradient echo sequences, but unfortunately they proved to be extremely affected by distortions and therefore not further studied for this specific application.

Finally, specific post-processing tools capable of reducing imaging distortions could be applied and several examples are reported in the literature (Hill *et al* 2001, Sumanaweera *et al* 1993, 1994).

5. Conclusions

A detailed phantom study has been performed to measure the global accuracy in a neuronavigation procedure, based on CT and MR images. The contribution of the image acquisition phase was separately investigated. Larger data dispersion comes up when the global accuracy is compared with the imaging one. This effect outlines the significant contribution introduced by the neuronavigator registration procedure and it is more evident for the MR data. MR acquisitions are well reproducible, the standard deviations for the imaging accuracy studies are close to those observed for CT tests, nevertheless procedures based on CT images proved to be more accurate than those based on standard T1-weighted and T2-weighted MR images. Since the used registration algorithm is a rigid one, the registration accuracy is heavily dependent on the absence of image distortions. Our results show that this hypothesis appears to be verified for CT images, but not for MR images, due to distortions. One can expect this effect to cause large registration errors and, consequently, large navigation inaccuracies in procedures based on MR images.

At the moment, in our centre, with the used acquisition protocols, the precision needed for neurosurgical procedures can be reached only using CT images. In this case the extreme slice thinness is not crucial and this can be relevant from a practical point of view.

We plan to perform additional work to identify low-distortion MR sequences or to implement a post-processing technique. This could improve MR image reliability, making them suitable for neuronavigation.

Acknowledgments

This research was supported by the MURST financed project 'Radiosurgery in the minimally invasive treatment of selected brain tumours using a low-energy source. Clinical assessment, conformal dosimetry and tele-radiosurgery development'.

The authors acknowledge the invaluable help and kindness in image acquisition by Dr Cinzia Gavazzi and Francesco Placco, and English language advice by Ingunn Lovik.

References

Arun K S, Huang T S and Blostein S D 1994 Least-squares fitting of two 3-D point sets *IEEE Trans. Pattern Anal. Mach. Intell.* **9** (5)

Bourgeois G, Magnin M, Morel A, Sartoretti S, Huisman T, Tuncdogan E and Meier D 1999 Jeanmonod, accuracy of MRI-guided stereotactic thalamic functional neurosurgery *Neuroradiology* **41** 636–45

Chang H and Fitzpatrick M 1992 A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities *IEEE Trans. Med. Imaging* **11** 319–29

Debus J, Essig M, Schad L, Wenz F, Baudendistel K, Knopp M, Engenhart R and Lorenz W 1996 Functional magnetic resonance imaging in a stereotactic setup *Magn. Reson. Imaging* **14** 1007–12

Dipierro C, Francel P, Jackson T, Kamiryo T and Laws E 1999 Optimizing accuracy in magnetic resonance imaging-guided stereotaxis: a technique with validation based on the anterior commissure–posterior commissure line *J. Neurosurgery* **90** 94–100

Dorward N, Alberti O, Palmer J, Kitchen N and Thomas D 1999 Accuracy of true frameless stereotaxy; *in vivo* measurement and laboratory phantom study *J. Neurosurgery* **90** 160–8

Hardy P and Barnett G 1998 Spatial distortion in magnetic resonance imaging: impact on stereotactic localization *Textbook of Stereotactic and Functional Neurosurgery* ed P L Gildenberg and R R Tasker (New York: McGraw-Hill) ch 33 pp 271–80

Hill D L G, Batchelor P G, Holden M and Hawkes D J 2001 Medical image registration *Phys. Med. Biol.* **46** R1–R45

Kaus M, Steinmeier R, Sporer T, Ganslandt O and Fahlbusch R 1997 Technical accuracy of a neuronavigation system measured with a high-precision mechanical micromanipulator *Neurosurgery* **41** 1431–7

Kondziolka D, Dempsey P, Lunsford D, Kestle J, Dolan E, Kanal E and Tasker R 1992 A comparison between magnetic resonance imaging and computed tomography for stereotactic coordinate determination *Neurosurgery* **30** 402–7

Maciunas R, Galloway R and Latimer J 1994 The application accuracy of stereotactic frames; technical special report *Neurosurgery* **35** 682–95

Prott F, Haverkamp U, Willich N, Resch A, Stober U and Potter R 1995 Comparison of imaging accuracy at different MRI units based on phantom measurements *Radiother. Oncol.* **37** 221–4

Sumanaweera T, Adler J, Binford T and Glover G 1993 MR susceptibility misregistration correction *IEEE Trans. Med. Imaging* **12** 251–9

Sumanaweera T, Adler J, Napel S and Glover G 1994 Characterization of spatial distortion in magnetic resonance imaging and its implications for stereotactic surgery *Neurosurgery* **35** 696–704

Torresin A, Brait L, Colombo P, Minella M, Pedroli G, Cardinale F, Lo Russo G, Munari C and Scialfa G 1999 Accuracy and precision in neuronavigation using MRI and CT images *Proc. First National Congress of Italian Physics in Medicine Association (Florence, 25–26 June 1999)* pp 32–6

Vannier M W and Haller J W 1999 Navigation in diagnosis and therapy *Eur. J. Radial.* **31** 132–40

Weis J, Ericsson A, Silander H and Hemmingsson A 1998 Magnetic resonance spectroscopic imaging for visualization and correction of distortions in MRI: high precision application in neurosurgery *Magn. Reson. Imaging* **16** 1265–72

Willems P, Noordmans H J, Van der Sprenkel J W, Viergever M and Tulleken C 2001 An MKM-mounted instrument holder for frameless point-stereotactic procedures: a phantom-based accuracy evaluation *J. Neurosurgery* **95** 1067–74

Zylka W, Sabczynski J and Schmitz G 1999 A Gaussian approach for the calculation of the accuracy of stereotactic frame systems *Med. Phys.* **26** 381–91