The Application Accuracy of Stereotactic Frames

Robert J. Maciunas, M.D.,
Robert L. Galloway, Jr., Ph.D., Jim W. Latimer, M.D.

Departments of Neurologic Surgery (RJM, RLG, JWL) and Biomedical Engineering (RLG), Vanderbilt University Medical Center, Nashville, Tennessee

THE PURPOSE OF incorporating stereotactic methodology into neurosurgical operations is to achieve a consistently high degree of accuracy in localizing intracranial targets. Therefore, the limits of resolution for the operation are a function of the accuracy of the particular stereotactic frame system. The total clinically relevant error (application accuracy) comprises errors associated with each procedural step, including imaging, target selection, vector calculations, and the mechanical errors of stereotactic frames. To evaluate these parameters, a systematic error analysis was carried out with four commonly used stereotactic devices: the Brown-Roberts-Wells, the Cosman-Roberts-Wells, the Kelly-Goerss COMPASS (modified Todd-Wells), and the Leksell frames. Over 21,500 independent accuracy test measurements were made with 11,000 computed tomograms. The results suggest a potentially significant degree of error in the application accuracy of all stereotactic instruments, which is accentuated by but not entirely due to imaging-associated errors. Clinically encountered levels of weightbearing by stereotactic frames may have a pronounced effect on their mechanical accuracy. Both the reapplication of aiming arc assemblies and the use of phantom base units introduce independent sources of mechanical inaccuracy into stereotactic procedures. The scope of individual error values and their determining factors must be considered with every clinical use of stereotactic frame systems. (Neurosurgery 35:682–695, 1994)

Key words: Accuracy, Computed tomography, Imaging, Precision, Stereotactic, Stereotactic frame

The rationale for applying stereotactic methodology to neurosurgical procedures is to access targets accurately with a minimum of spatial error (i.e., low bias) and a high degree of reproducibility (i.e., high precision). The use of a stereotactic frame on a patient requires increased surgeon and staff training, increases the technical complexity of the surgical intervention, and may expose the patient to complications due to the application of a frame per se (11). The benefit gained by accepting these drawbacks is that of improved target localization. Clinical reliance on stereotactic instrumentation to provide accurate target localization requires that a quantitative basis exist for this degree of trust. Previous evaluations of stereotactic instruments, however, have not dealt with this issue conclusively. The limitations of previous reports include inadequate sample sizes for statistically valid analyses, failure to control for variables introduced by the imaging techniques used, artificially constrained testing conditions that minimize the mechanical errors of the instruments such as the use of target points clustered only at the center of the imaging field or the use of optimized approach trajectories, systematic failure to distinguish between mechanical "unbiasedness," mechanical precision, mechanical and application accuracy, and reliance on purely clinical criteria for a subjective judgment as to the success of localization (2, 5, 11, 14, 17, 18, 22, 31, 33). By reporting a careful analysis of the application accuracy of stereotactic frame systems, we hope to provide both the precedent and the framework to establish a format for the reporting and discussion of application accuracy claims for all neurosurgical localization systems, whether or not they use stereotactic frames.

TERMINOLOGY

Although the terms "stereotactic" and "submillimetric accuracy" are usually closely associated in the literature, the neurosurgeon must exercise considerable caution when dealing with such assumptions. To begin a considered approach to the subject of stereotactic accuracy, then, attention to the definition of terms is in order.

First, a clear distinction must be drawn between three commonly misused terms: "unbiasedness" or lack of skew, "precision," and "accuracy" (*Fig. 1*). A series of observations that tend to the true value are "unbiased" or without skew. If these observations have considerable spread, however, they lack "precision." A series of observations with little spread among them indicates precision, although they are biased or skewed if the observations tend to center at a value displaced from the true value. The term "accuracy" encompasses both unbiased-

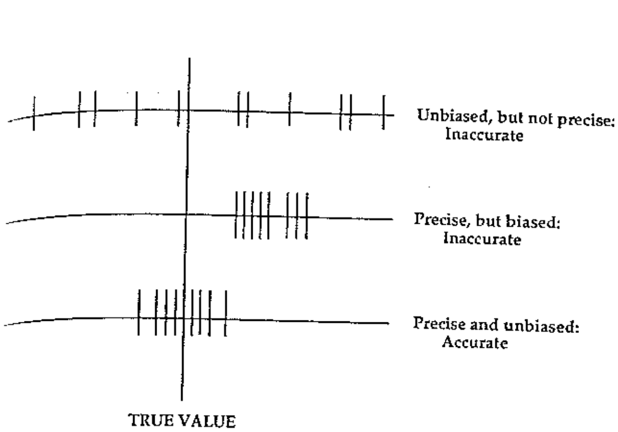

FIGURE 1. Unbiasedness, precision, and accuracy. *Top,* the wide scatter of localization attempts is imprecise. Here, their average tends to the true value; this is unbiased. This unbiased but imprecise series of localizations is inaccurate. *Middle,* the tightly clustered series of localization attempts does not tend to the true value. This is precise, with high repeatability, but there is bias to the mean. This precise but biased series of localizations is inaccurate. *Bottom,* this tightly clustered series repeatedly tends to the true value. This unbiased and precise series of localizations is accurate.

ess and precision: accurate measurements are both unbiased nd precise, whereas inaccurate measurements may be biased, mprecise, or both (13).

If a stereotactic frame demonstrates a tendency to arrive onsistently at a wrong answer, this skew is the result of echanical bias in a precise system. Because the mechanical ecision of a stereotactic frame system determines how well can return to a point, it may be thought of as, in effect, being measure of its repeatability of localization. Finer increments the verniers or better machining tolerances can be expected improve mechanical precision. It is often this number, incidentally, that is advertised for stereotactic devices. The mechanical accuracy of a stereotactic apparatus is the accuracy th which it can bring the tip of a straight probe to a given ordinate within the stereotactic coordinate system. This requires both mechanical unbiasedness and mechanical precision but does not include the effects of imaging technologies localization. The voluntarily adopted standard performance specifications for cerebral stereotactic instruments, issued the American Society for Testing and Materials, state that mechanical accuracy of a stereotactic system will be submillimetric (1).

In addition to the mechanical limitations of the stereotactic ame, the errors associated with the many steps of stereotaxis cluding imaging techniques, point selection, vector calculation, vernier settings, mechanical couplings and adjustments, all contribute to the final clinically relevant error (9, 12, 27). e purpose of this study is to go beyond the assessment of the chanical unbiasedness, precision, and accuracy of these devices in order to assess what may be termed the "application accuracy" of various stereotactic frames. This is a measure of the accuracy of these devices when used in their real-world setting. Human operator error is outside the realm of this discussion, although tedious calculations or taxing calibrations imposed by a system will make these errors more likely.

When measuring the application accuracy of stereotactic frames, a distinction must be drawn between the "mean error" and the "mean position" of localization. The application accuracy should be expressed in terms of the "mean error of localization" in order to avoid an erroneously favorable measurement from being reported. To define the mean position of localization, consider that, when a stereotactic device is used to attempt the localization of a known target, the series of such attempted localizations could be shown as a collection of points in a gaussian distribution around the true target. The mean of the positions of these points could be determined by calculating the *vector* sum of the displacements between the target and these points. If the stereotactic frame system was imprecise but without bias, then the location of any single attempt could be unacceptably distant from the true target, although the mean of these positions could turn out to be very close to the true target. Therefore, by overlooking the lack of precision in these localization attempts, one could erroneously conclude that the stereotactic frame was accurate because the mean position of localization was close to the true target. The mean error of localization is determined by calculating the mean of the sum of the *scalar* distances between the true target and the points achieved by the system. This number is based on how far from the target each attempt was, irrespective of directionality. The mean error of localization takes into account the precision of each individual localization attempt, whereas the mean position is insensitive to precision.

The fundamental flaw of defining application accuracy in terms of the mean position of localization might appear to be a "straw man proposal" because neurosurgeons would use the appropriate measurement of mean error of localization. However, that flawed definition of application accuracy has in fact been used in reporting results for surgical localization systems (D. Deckel, ISG Technologies, personal communication).

The clinical importance of recognizing this error is illustrated by this example. In attempting to target the ventral intermediate nucleus of the thalamus for the relief of parkinsonian unilateral tremor, a neurosurgeon chooses to use a stereotactic frame for surgical guidance. No neurosurgeon is interested in knowing that, if he or she were to pass the probe and make radiofrequency lesions 1000 times, the average derived from a lifetime's series of lesion positions might closely approach the ventral intermediate nucleus. On the contrary, what any neurosurgeon needs to know as he or she passes the probe is: "What is the greatest distance that this particular pass could be from the target?" This information is expressed by the mean error of localization at a confidence interval of 99.9% (a certainty of 999 cases out of 1000). In this clinical setting, a 1%

risk to the patient is sufficiently significant that it routinely requires explicit discussion during preoperative informed consent counseling. In order to be considered negligible, the risk associated with erroneous information due to the stereotactic frame should be therefore at least an order of magnitude better than 0.1%, or 1 in 1000.

MATERIALS AND METHODS

Stereotactic frames

Application accuracy was assessed in the four most commonly used stereotactic frame systems by the use of methods paralleling those used for stereotactic surgery (4, 14, 24). The frame systems tested were the Brown-Roberts-Wells (BRW; Radionics, Burlington, MA), the Cosman-Roberts-Wells (CRW; Radionics), the COMPASS (Kelly-Goerss modification of the Todd-Wells; Stereotactic Medical Systems, Seneca, NY), and the Leksell (D and G frames; Elekta Instruments, Tucker, GA). Five complete individual sets of instruments were tested for each frame type. All frames were new and were certified as being within industry standards by their respective manufacturers. Individual frames were selected to be from different manufacturing run numbers, as confirmed by their respective manufacturers. Five sets of new BRW and of CRW Phantom Base assemblies were also tested for their mechanical accuracy.

Test phantoms

Test phantoms were developed that consisted of radiodense carbon fiber targets embedded in one end of rigid acrylic plastic rods of different lengths; the nontarget ends of the rods were embedded orthogonally into a base of acrylic plastic. The carbon fiber targets were cylinders, 2 mm in diameter and 2 mm in length, conically tapering over 2 mm to a point. This tip was the target selected for this study. Twelve such target points were distributed widely throughout the entire imaging target volume enclosed by the stereotactic frame, and not just at the center of the coordinate space, where mechanical precision is likely to be optimal for a given frame. The phantom was anchored rigidly to the stereotactic frames. The weight of this type of phantom was under 1 kg (*Fig. 2*). For trials to assess the effects of clinically comparable levels of weightbearing on accuracy, test phantoms were developed that incorporated attached weights of 10 and of 25 kg.

Imaging protocols

The test phantom was affixed to the stereotactic frame, and the apparatus was imaged via Siemens Somatom DR-H-type computed tomographic (CT) scanners (Siemens, Inc., Iselin, NJ). To minimize machine-dependent biases, two individual CT scanners of this type were used for each series of tests. Both CT scanners were tested and found to be in compliance with industry standards by the manufacturer and factory service representatives before data acquisition for this study. For each stereotactic apparatus, four types of independent imaging trials were performed, each with different slice thicknesses and slice planes. Each trial was repeated for the five individual specimens of each frame system; each example was tested with five independent scan trials for each of the four different slice thickness/slice plane settings. As a result, a total of 11,000 CT scan slices (275 independent CT scan sets) were required to accomplish this study. For all images, the matrix of the image was 512 × 512 pixels. The field of view was selected to be one that was the smallest one possible that allowed the visualization of the targets and N-bar fiducials for all types of frames (magnification factor = 1.53). As a result, the inplane pixel dimensions were 0.6533 × 0.6533 mm. Standard Siemens Somatom DR-H CT software was used for data acquisition, storage, and retrieval (*Fig. 3*).

The first set of images provided an optimized situation of ideal target localization. The stereotactic frame was secured in a customized headholder such that the base ring or "zero latitudinal plane" was precisely parallel with the CT scan slices when the CT gantry tilt was zero degrees. A series of 1-mm-thick CT slices was obtained in this orientation. Each slice was hand calibrated with its respective table position carefully selected so as to cut precisely through the tip of 1 of the 12

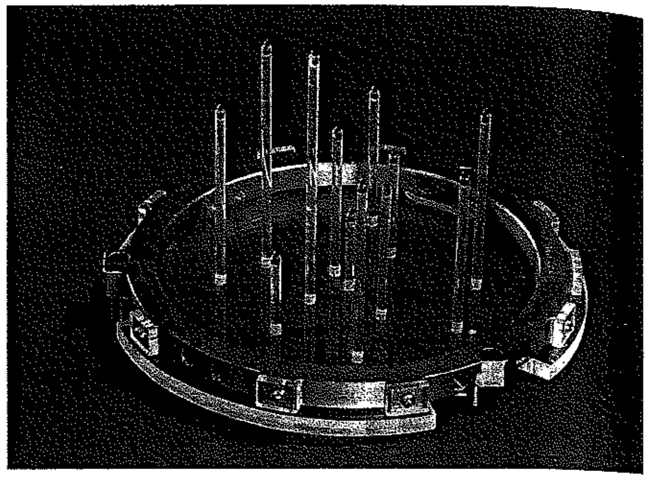

FIGURE 2. A lucite rod/carbon fiber target test phantom used in this study. This test phantom weighed 1 kg; 10- and 25-kg weighted phantoms were also used.

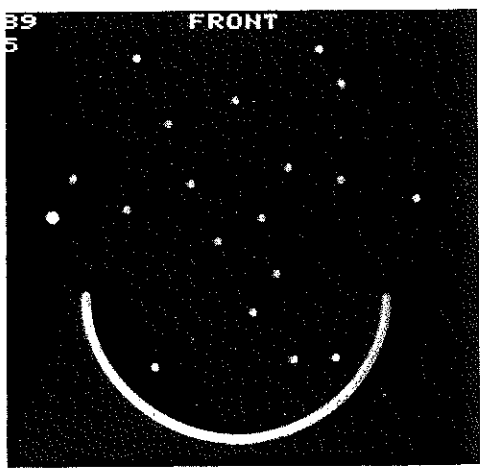

FIGURE 3. CT scan slice image of a test phantom as mounted within a BRW stereotactic frame system.

targets. This represents an optimal case in that it attempts to limit contributions by imaging-derived errors and thereby to isolate any frame-dependent errors. No other scans with thicker slice thickness were performed at these hand-calibrated table positions. Because the positioning of any thicker slice would be in effect prelocalized in the z direction by the optimized 1-mm slice and because the inplane pixel dimensions for x and y directions do not change when the slice thickness is varied, to do so would in fact be specious. The thickness of the slices would simply be distributed symmetrically about the same center of table positioning as defined by the 1-mm slices. Such prearranged constraints on the data acquisition would in fact minimize any effects of increasing the thickness of scans obtained in this way.

The second set of images assessed the effects of frame angulation upon localization in an otherwise optimized situation for ideal target localization. The stereotactic frame was positioned by varying its yaw, pitch, and roll in increments of up to 30 degrees in relation to the CT gantry. A series of 1-mm CT slices was obtained; each slice was again hand calibrated with its respective table position carefully selected so as to cut precisely through the tip of each of the 12 targets. This series was an effort to isolate the angular sensitivity of the N-bar fiducial localization algorithm from other contributions of imaging-dependent variables, in order to test the robustness of the localization algorithms provided with respect to frame orientation (6, 14, 26). This trial was not performed for the Leksell stereotactic frame system because no commercially supported software for this purpose was available at the time.

The first two sets of images, then, assumed 1) a priori knowledge of true target localization and 2) subsequent optimization by the hand calibration of selected slice acquisitions. Because this would not be possible in the clinical setting for a given patient undergoing stereotactic neurosurgery, the next two sets of images were carried out in a manner that was more directly relevant to clinical practice. The stereotactic frame was mounted in its normal position, coaxial to the bore of the CT scanner and with the base ring parallel to the orientation of the CT series. The apparatus was then scanned in the conventional manner, with slices obtained successively from the base through the highest target. Scan slices were thus generated without a priori knowledge for centering on the target tips. Instead, the third and fourth sets of images were obtained by the use of standard clinical imaging protocols with 4- and 8-mm slice thicknesses, respectively. There were no gaps between slices. There was no overlapping of slices.

The spatial locations of the fiducial marks and target points were extracted from the CT images by the use of the standard protocols supplied for each frame system. The mean of the application accuracy was plotted against the spatial position of phantom targets relative to the center of each stereotactic frame. The mean of the application accuracy is more influenced by imaging effects on target localization, whereas the standard deviation of the application accuracy is related to the repeatability as a function of the stereotactic frame's mechanical precision. The graphs were analyzed for the presence of patterns of deviation from a flat or random distribution of localization error throughout the stereotactic space.

Measurement of error

In order to control for any individual stereotactic frame's deviations from the population mean with respect to mechanical unbiasedness, precision, and accuracy, five individual specimens of each of the four frame types were tested. Inadequacies of aiming arc assembly design might be overlooked by artificially constraining the approach trajectory selected for a given target point. In order to identify any path-dependent errors, three separate entry points, and, therefore, three separate frame settings for three approach trajectories were calculated by the use of the localization algorithms supplied by the frame manufacturers. The stereotactic localization aiming arc assembly was then attached to the frame system, and a measuring probe was introduced to the calculated depth to approach each target on the test phantom. The distance resulting from the tip of the probe to the tip of the target was measured as the localization error, which defined the application accuracy. This distance was measured with digital electronic microcalipers (Mitutoyo Electronic Corporation, Tokyo, Japan) and was confirmed by two independent observers before being recorded as a result.

In summary, five CT scan sets for five trials for each of five specimens of each of four types of stereotactic frame systems were carried out with three different approach trajectories to each of 12 target points; thus, 900 measurements were made for each of the 15 imaging-defined categories (hereafter referred to as the "data cells" for *Table 1*). A total of 13,500 independent error measurements were carried out for image-guided stereotactic application accuracy tests.

Application of aiming arc assemblies

The aiming arc assembly of a stereotactic frame holds and directs the probe as it is directed to the target point. In order to assess the contribution of arc application to mechanical accuracy, the stereotactic frame system was set up so as to direct the measuring probe to touch a phantom target point. The aiming arc assembly was then uncoupled, removed, and applied again without any settings being changed. The distance from the tip of the probe to the tip of the target was measured as the error in mechanical accuracy for that stereotactic frame. The aiming arc application effects on mechanical accuracy were measured 100 times for each of the five examples of the four frame types studied; therefore, a total of 2000 independent accuracy tests were performed for this assessment.

Phantom base units

The phantom base units provided for the BRW and CRW stereotactic frame systems are only a check on the mechanical accuracy of setting up the system. With the advent of linear accelerator based radiosurgery, however, the phantom base unit has been used for adjusting the linear accelerator (19, 36). Because of this expanded role in radiosurgery, the accuracy of BRW and CRW base units was tested in relation to a calibration jig. Five phantom base units of each type were evaluated. The CRW phantom base unit is configured slightly differently than is the BRW unit, especially with regard to the support posts

TABLE 1. The Application Accuracy of Four Stereotactic Frame Systems by the Use of Four Independent Imaging Protocols[a]

CT Slice Thickness (mm)	Measurement	BRW (mm)	CRW (mm)	Compass (mm)	Leksell (mm)
1	Mean	1.9	1.8	1.2	1.7
	SD	1.0	1.1	0.6	1.0
	Min value	0.1	0.0	0.3	0.2
	Max value	5.0	4.9	3.2	4.9
	95% CI	3.6	3.6	2.2	3.4
	99.9% CI	5.0	5.2	3.1	4.8
1 canted (AE)	Mean	2.5	2.2	2.1	N/A
	SD	1.1	1.1	1.0	
	Min value	0.1	0.1	0.3	
	Max value	6.9	6.1	5.5	
	95% CI	4.3	4.0	3.8	
	99.9% CI	5.9	5.7	5.2	
4	Mean	2.7	2.6	2.5	2.6
	SD	1.3	1.5	1.2	1.4
	Min value	0.0	0.0	0.4	1.0
	Max value	7.0	7.1	6.4	7.2
	95% CI	4.8	5.1	4.5	4.9
	99.9% CI	6.7	7.2	6.2	6.9
8	Mean	6.6	6.6	5.1	5.4
	SD	3.1	3.0	2.3	2.4
	Min value	0.0	0.3	0.6	0.6
	Max value	11.2	12.0	10.9	10.2
	95% CI	11.7	11.6	8.9	9.4
	99.9% CI	16.2	15.9	12.2	12.8

[a] n = 13,500 independent accuracy measurements. CT, computed tomography; SD, standard deviation; Min, minimum; Max, maximum; N/A, not available; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells; CI, confidence interval; AE, angulation effect.

and the runners for the slide holding the target pointer. Since its introduction, this base unit has served for use with both BRW and CRW systems. Two hundred points were assessed per phantom base unit, distributed throughout the stereotactic imaging volume. A total of 2000 independent accuracy test measurements were performed.

Weightbearing

In clinical use, the stereotactic frame is called on not only to serve as the basis for fiducial registration and aiming arc assembly fixation, but also to serve as the intraoperative head support for the patient. This introduces another source of error: the weight of the patient causes sag and torque forces to be applied to the coordinate system. These effects cannot be assessed by the use of light test phantoms. Given that the weight of a patient's head and neck is approximately 10 to 20 kg and that the weight supported by a stereotactic frame in the operating room may also include some percentage of the shoulder girdle, weighted trials between 10 and 25 kg would be expected to assess adequately the effects on the application accuracy of clinically relevant load bearing by stereotactic frames. Accordingly, two similar series of experiments were performed with phantoms with 10- and 25-kg lead weights attached. The variance between target localizations obtained in five different orientations was determined with both weighted phantoms. These distances, measured as described previously, defined the effect of clinically relevant levels of weightbearing on the mechanical accuracy of stereotactic frames. Measurements were recorded 20 times for each of these five orientations at each of the two weightings, with each of the five examples of the four frame types studied; therefore, a total of 4000 independent accuracy tests were performed for this assessment.

Statistical analysis

The localization errors can be considered as distributed normally in each of the individual *x*, *y*, and *z* dimensions. The root mean square sum representing the euclidean localization error, $[x^2 + y^2 + z^2]^{1/2}$, cannot be absolutely normally distributed because no negative error values can occur. Therefore, the distribution of error is bounded on the left by zero but also on the right by the excursion of the stereotactic arc system. The distributions of these bounded error measurements were evaluated by individual histograms of cumulative distribution function and probability density function; analyzed for mean, median, symmetry, skew, and kurtosis; and studied with box plots and KS testing. The distribution of localization errors is sufficiently well described as quasinormal to be reliably predictable when analyzed in terms of a gaussian distribution. A right-tailed test was applied to determine the 95 and 99.9% confidence intervals for the data. Results were reported in terms of their mean value, standard deviation, maximum value encountered, and values for 99.9% confidence intervals.

For each cell of data reported in *Table 1*, the accuracy of each of the sample means was evaluated as an estimate of that population mean. The standard error of the mean was calculated as: $SE_M = SD / (n)^{1/2}$; where M = mean, SD = standard deviation, and n = number of sample observations. The resulting 99% confidence interval for the accuracy of the sample means was expressed as: 99% CI = M ± 3 SE_M, and the range for the confidence interval of the mean was expressed as: [(M − 3 SE_M) to (M + 3 SE_M)]; both are reported in *Table 2*.

To assess whether the individual cells of data in *Tables 1, 4, 5* and *6* and in *Figure 5* were different from each other to a statistically significant degree, the individual cells of data were compared with each other by the two-sample (unpaired) *t*-test. In all instances, the sample sizes were sufficiently large that statistically significant differences were confirmed between the groups compared ($P < 0.001$).

For each data cell in *Table 1*, the breadth of distribution was assessed by calculating the coefficient of variation, defined as: COV = SD/M, where SD = standard deviation and M = mean. These results are shown in *Table 3*. A coefficient of variation less than 1 indicated less variation around the mean and a greater central tendency for the distribution.

When these results are plotted graphically, as in *Figures 3, 4, 5*, and *7*, the correlation coefficient, *r*, was calculated for the regression line of the data and the coefficient of determination, r^2, was calculated as a descriptor of the linearity of the relationship between the two variables (7). The spatial distribution

TABLE 2. Accuracy of the Sample Means as an Estimate of the True (Population) Means for Stereotactic Frame Application Accuracies[a]

CT Slice Thickness (mm)	Measurement	BRW (mm)	CRW (mm)	COMPASS (mm)	Leksell (mm)
1	Mean ± 3 SE_M	1.9 ± 0.10	1.8 ± 0.11	1.2 ± 0.06	1.7 ± 0.10
	99% CI for the mean	1.80 to 2.00	1.69 to 1.91	1.04 to 1.26	1.60 to 1.80
1 canted	Mean ± 3 SE_M	2.5 ± 0.11	2.2 ± 0.11	2.1 ± 0.10	N/A
	99% CI for the mean	2.39 to 2.61	2.09 to 2.31	2.00 to 2.20	N/A
4	Mean ± 3 SE_M	2.7 ± 0.13	2.6 ± 0.15	2.5 ± 0.12	2.6 ± 0.14
	99% CI for the mean	2.57 to 2.83	2.45 to 2.75	2.38 to 2.62	2.46 to 2.74
8	Mean ± 3 SE_M	6.6 ± 0.31	6.6 ± 0.30	5.1 ± 0.23	5.4 ± 0.24
	99% CI for the mean	6.29 to 6.91	6.30 to 6.90	4.87 to 5.33	5.16 to 5.64

[a] The large numbers of measurements yielded small standard errors of the means and a resultant high confidence in these experimentally derived means as being representative of the true means of application accuracy for these stereotactic frame systems. SE_M, standard error of the mean; N/A, not available; CT, computed tomography; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells; CI, confidence interval.

TABLE 3. Coefficients of Variation[a]

CT Slice Thickness (mm)	BRW	CRW	COMPASS	Leksell
1	0.53	0.63	0.50	0.59
1 canted	0.43	0.50	0.48	
4	0.63	0.59	0.48	0.52
8	0.48	0.45	0.45	0.43
Average	0.52	0.54	0.48	0.51

[a] The coefficient of variation (COV) is the breadth of distribution for application accuracy in each data cell of *Table 1*, calculated by COV = SD/M, where SD = standard deviation and M = mean. A COV <1 indicates less variation about the mean and a greater central tendency for the distribution. CT, computed tomography; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells.

of localization error relative to target positions in stereotactic space was graphically displayed (*Fig. 6*) and analyzed for patterns of bias that might deviate from flat or random distributions. No consistent pattern of bias was observed. All data were tabulated with Excel spreadsheet software (version 2.01; Microsoft, Inc., Seattle, WA), statistically analyzed with StatView II software (version 1.03; Abacus Concepts, Inc., Berkeley, CA), and graphically displayed with CricketGraph software (version 1.3.2; Cricket Software, Malvern, PA).

RESULTS

A total of 21,500 independent accuracy measurements were performed in the course of this series of experiments. The results of testing for the application accuracy of the four frame types with various imaging protocols are shown in *Table 1*. The accuracy of the sample mean as an estimate of population mean was determined for each of the cells of data reported in the results of *Table 1* and is shown in *Table 2*. The extremely small standard errors of the means result in highly constrained ranges for the 99% confidence intervals of the means. We can be confident that the population mean lies within the confidence limits, and because these are well within the standard deviation for the sample data, the sample mean does not appear skewed with respect to the true population mean. Although the differences between each data cell for the four different frame types were statistically significant ($P < 0.001$), what is more remarkable is the dramatic effect on the application accuracy of the imaging protocol used during stereotactic data acquisition. A linear relationship is suggested in *Figure 4* between increasing slice thickness and worsening application accuracy ($r^2 = 0.888$; $r = 0.942$). On the other hand, not only imaging contributes to the error measured. The extrapolation of this line to the ideal case of an imaging slice thickness of zero would result in a 99.9% confidence interval of 2.28 mm for the mechanical accuracy (equivalent to application accuracy at this value). Thus, the regression line does not demonstrate a submillimetric level of ideal application accuracy, that is, mechanical accuracy for these stereotactic frame systems.

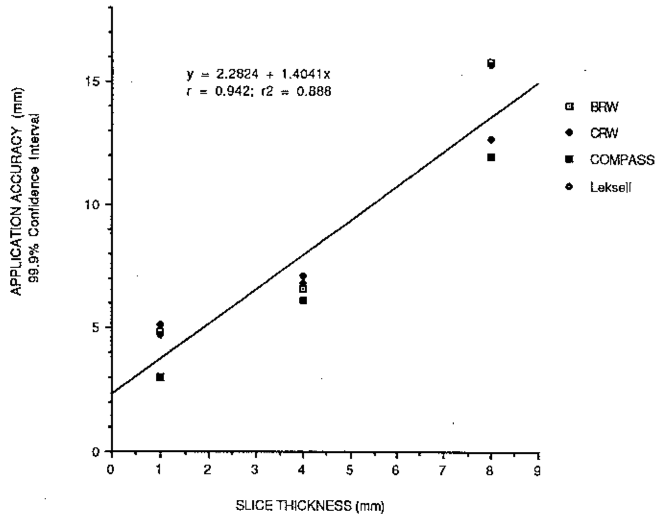

FIGURE 4. The effect of slice thickness on the application accuracy of stereotactic frame systems. n = 10,800 independent accuracy measurements. A linear relationship is observed between the slice thickness and the application accuracy ($y = 2.2824 + 1.4041x$; $r = 0.942$, $r^2 = 0.888$). When the slice thickness is zero, the application accuracy becomes equivalent to the mechanical accuracy for stereotactic frame systems; this is 2.28 mm.

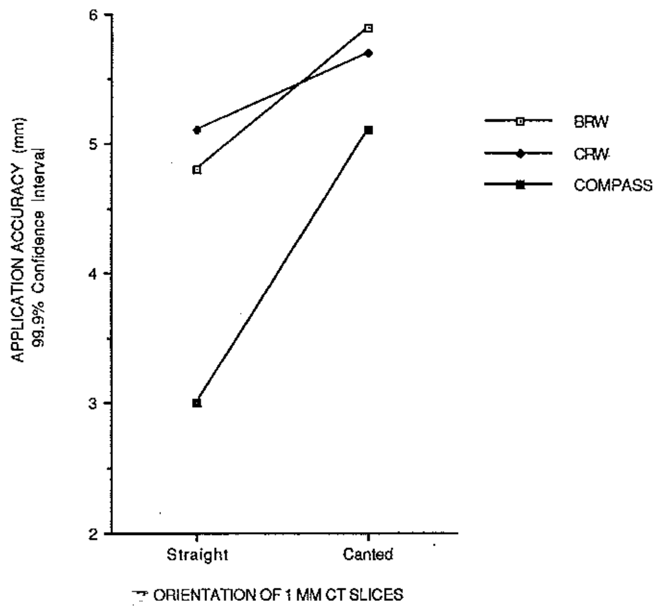

FIGURE 5. The effect on the application accuracy of frame angulation during imaging (n = 5400 independent accuracy measurements). A significant deterioration was noted for all stereotactic frame systems tested ($P < 0.001$).

The deleterious effects of frame angulation on application accuracy are shown in *Table 1* and are graphically demonstrated in *Figure 5*. This deterioration in application accuracy due to frame angulation during imaging data acquisition was significant for all three frame types tested ($P < 0.001$).

For each data cell in *Table 1*, the breadth of distribution was assessed by the calculation of the coefficient of variation. These results are shown in *Table 3*. All stereotactic frame systems demonstrated similar central tendencies, and a small range of coefficient of variation values was noted.

The spatial distribution of localization error relative to target positions in stereotactic space is shown in *Figure 6*. No consistent pattern of bias was observed that deviated from flat or random distributions.

The mechanical accuracy of the phantom base units for the BRW and the CRW stereotactic frame systems is shown in *Table 4*. Although there was a statistically significant difference between the results for the BRW and the CRW systems ($P < 0.001$), more significant is the fact that the mechanical accuracies of all of these units are not submillimetric.

The effects of the subsequent application of the aiming arc assembly (without further adjustment of any settings) on the mechanical accuracy of the four frame systems are shown in *Table 5*. The range of error introduced by this maneuver was approximately 0.5 to 1 mm, with the differences between frame types achieving statistical significance ($P < 0.001$).

No appreciable effect of weight on the mechanical accuracy of stereotactic frames was noted when the 1-kg test phantoms were used. Pronounced effects were measured when weights of 10 and then 25 kg were attached to the bases of dedicated test phantoms. This is demonstrated in *Table 6*. The mechanical accuracy of all types of stereotactic frames was adversely affected in proportion to their degree of weightbearing. Although the differences in degree between individual frame types achieved statistical significance ($P < 0.001$), the dominant result is one of universally profound deterioration. This is illustrated in *Figure 7*. Extrapolation of the regression line ($y = -1.7633 + 0.29583x$) suggests that this effect is encountered whenever weightbearing by the frame exceeds 5.85 kg.

DISCUSSION

The purpose of incorporating stereotactic methodology into neurosurgical practice is to provide an improvement in localization over that which is otherwise available. Stereotactic frame systems have proved to be highly successful in accomplishing this and are widely used as a result (11). Like any instrument, however, these stereotactic frames have their drawbacks and limitations. In order to use these devices more successfully in the operating room, a clear understanding is necessary of what these limitations are. Accordingly, the clinician requires specific information regarding what degree of accuracy can realistically be conferred on the surgical procedure by these frames. In the past, reports have indiscriminately intermixed technical standards stressing mechanical accuracy, advertising claims pointing out mechanical precision, and claims for the actual accuracy of localization during surgery. This article seeks to distinguish quantitatively between the accuracy claimed by the manufacturers and the accuracy measured by independent users.

The voluntary standard performance specifications for cerebral stereotactic instruments, as issued by the American Society for Testing and Materials, state that the mechanical accuracy of a stereotactic system shall be submillimetric (1). In addition to the mechanical limitations of the stereotactic frame, the errors associated with the many steps of stereotaxis (including imaging techniques, point selection, and vector calculation) all contribute to the final clinically relevant error (11, 12, 27). The purpose of this study is to go beyond the assessment of the mechanical accuracy of various stereotactic frames in order to assess the application accuracy of these devices. This is a measure of the accuracy of these devices when used in their real-world setting. The application accuracy includes all of the errors inherent in each step of the therapeutic process. The application accuracy is therefore the least forgiving measure of performance by stereotactic frames, because it does not ignore any of the sources of error that affect the surgical process. It should be distinguished from mechanical unbiasedness, precision, or accuracy in reports of stereotactic device performance.

In reporting the data for this analysis, care was taken to provide clinically relevant information. The error in localization was used to define the application accuracy, rather than erroneously reporting the mean position of attempted target localization. In this way, full accounting was made in the application accuracy data of the contributions of mechanical bias, imprecision, and inaccuracy. In addition to a mean, or "root mean square difference," the clinician needs to know the standard deviation and maximal encountered value. Most

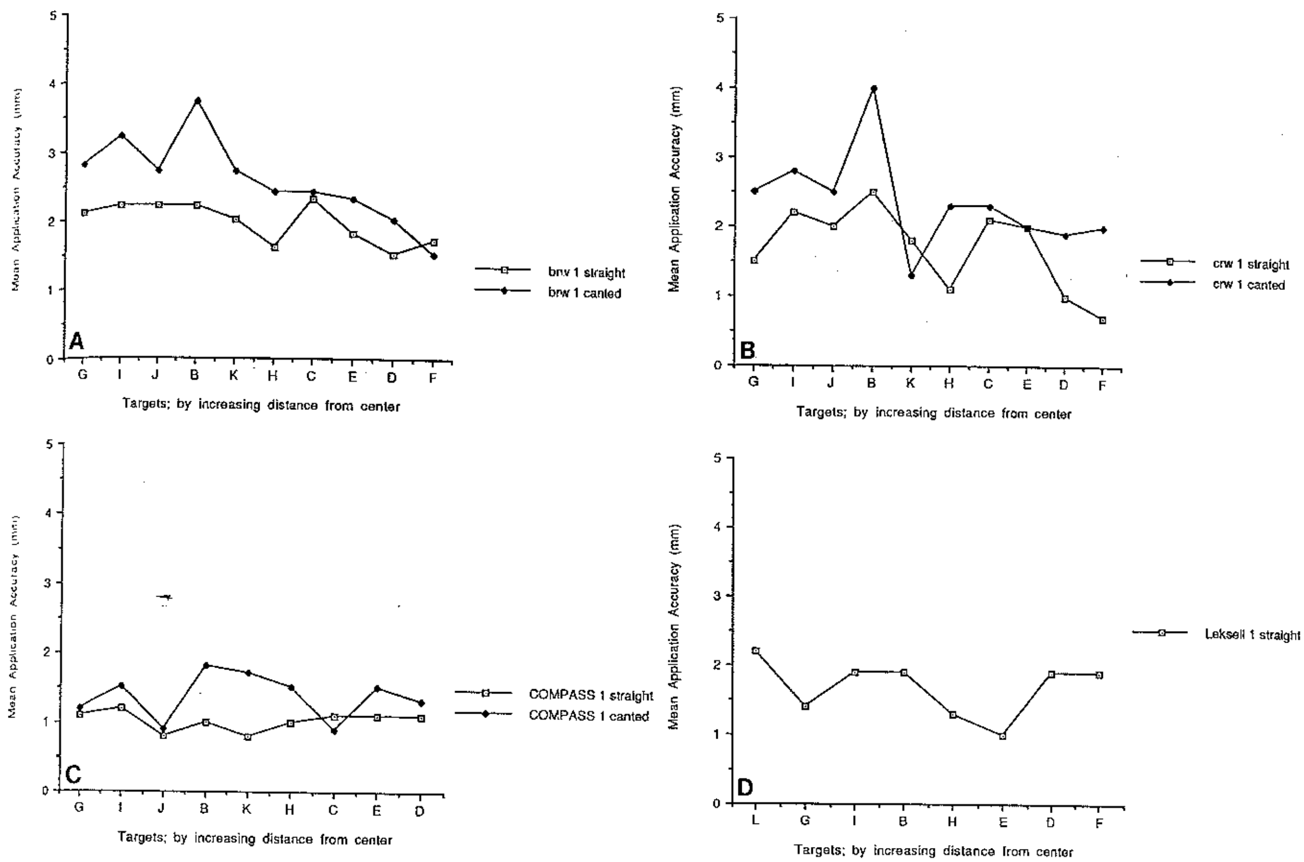

FIGURE 6. The mean application accuracy for stereotactic frame systems as related to the target's position within the stereotactic space. There was no consistent pattern of localization bias related to increasing distance from the center of the coordinate space. *A*, BRW; *B*, CRW; *C*, COMPASS; *D*, Leksell.

TABLE 4. The Mechanical Accuracy of the BRW and CRW Phantom Base Units[a]

Measurement	BRW (mm)	CRW (mm)	COMPASS	Leksell
Mean	0.6	0.7		
SD	0.4	0.4		
Max value	3.1	5.0	N/A	N/A
99.9% CI	1.8	1.9		

[a] n = 2,000 independent accuracy measurements. SD, standard deviation; Max, maximum; N/A, not available; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells; CI, confidence interval.

TABLE 5. The Mechanical Accuracy of Aiming Arc Assembly Reapplication[a]

Measurement	BRW (mm)	CRW (mm)	COMPASS (mm)	Leksell (mm)
Mean	0.26	0.32	0.20	0.35
SD	0.14	0.17	0.11	0.12
Max value	0.62	0.85	0.45	0.57
99.9% CI	0.69	0.85	0.54	0.72

[a] n = 2,000 independent accuracy measurements. SD, standard deviation; Max, maximum; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells; CI, confidence interval.

critically, what should be distinctly expressed is the 99.9% confidence interval for the error in localization. This value will always be larger, that is to say, "worse," than the mean or average or "root mean square" error. By this value alone can the neurosurgeon have sufficient information to know the likely worst case scenario before performing a given probe passage. The number of independent accuracy measurements performed during this investigation was made sufficiently large so as to provide information with a high degree of confidence. This has allowed for definitive statements to be made regarding the application accuracy of stereotactic frames.

Test phantoms

In constructing phantoms for testing this equipment, care was given to produce phantoms with a random distribution of target points throughout the stereotactic imaging volume. This distribution precluded biasing the assessment that may possibly have occurred if the target points had been limited to central, midline regions at the focus of the stereotactic system, where frame machining inaccuracies might be minimized. This phantom could test the effects on the application accuracy of more extreme excursions of the aiming arcs, such as those commonly encountered in clinical practice.

TABLE 6. The Effect of Weightbearing on the Mechanical Accuracy of Stereotactic Frames[a]

Weighting (kg)	Measurement	BRW (mm)	CRW (mm)	COMPASS (mm)	Leksell (mm)
10	Mean	0.28	0.29	0.49	0.25
	SD	0.31	0.32	0.25	0.28
	Max value	1.18	1.15	0.85	1.00
	99.9% CI	1.24	1.28	1.26	1.12
25	Mean	1.83	2.09	2.50	1.62
	SD	1.25	1.06	1.05	1.47
	Max value	6.01	5.86	4.90	5.77
	99.9% CI	5.69	5.37	5.74	6.16

[a] n = 4,000 independent accuracy measurements. SD, standard deviation; Max, maximum; BRW, Brown-Roberts-Wells; CRW, Cosman-Roberts-Wells; CI, confidence interval.

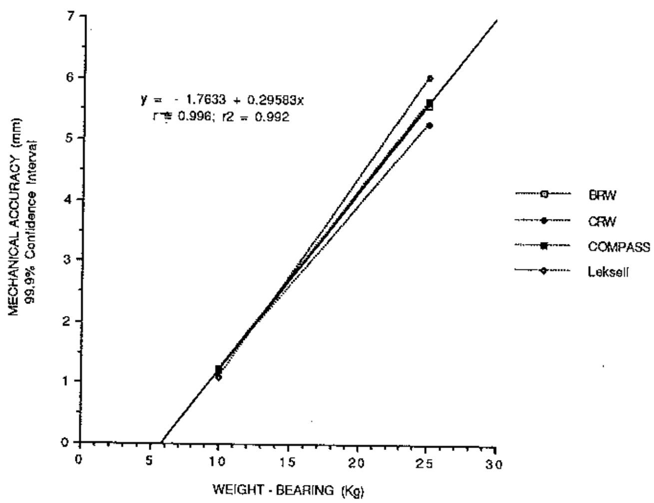

FIGURE 7. The effect of weightbearing on the mechanical accuracy of stereotactic frame systems (n = 4000 independent accuracy measurements). This effect would not be appreciable with light test phantoms ($y = -1.7633 + 0.29583x$; y = 5.8 kg when x = 0).

Imaging protocol effects

At present, the standard of care for most stereotactic localization remains dependent on CT imaging (2, 4–6, 11, 22, 26, 29). On the other hand, magnetic resonance (MR) imaging promises superior soft tissue sensitivity and resolution in comparison to CT. Currently, however, fluctuations in the spatial fidelity of MR may, at times, produce a clinically significant source of localization error. Although extremely relevant to the future of stereotaxis, issues related to the recognition and rectification of spatial distortions due to MR are outside the scope of this discussion (8). This study was therefore confined to using CT as the source of all imaging data for stereotactic localization.

The resolution of a CT scanner is related to the number of picture elements, or "pixels," that make up the digital image. The CT scanners that were used in this report provided a 512 × 512 matrix of pixels. This matrix size remains constant, regardless of the field of view encompassed by a scan image.

TABLE 7. The Relationship Between Magnification Factor, Field of View, and Pixel Size for Computed Tomography[a]

Magnification Factor	Field of View (mm)	Pixel Size (mm)
1.0	512.000	1.000
1.1	465.455	0.909
1.2	426.667	0.833
1.3	393.846	0.769
1.4	365.715	0.715
1.5	341.333	0.667
1.6	320.000	0.625
1.7	301.176	0.588
1.8	284.444	0.556
1.9	269.473	0.526
2.0	256.000	0.500
2.1	243.810	0.476
2.2	232.727	0.455
2.3	222.609	0.434
2.4	213.333	0.417
2.5	204.800	0.400
2.6	196.923	0.385
2.7	189.630	0.370
2.8	182.857	0.357
2.9	176.552	0.345
3.0	170.667	0.333

[a] Somatom DR-H CT scanner, 512 × 512 pixel matrix.

The smaller the field of view larger magnification factor, the greater the number of pixels per centimeter, therefore, the smaller is each individual pixel and the finer is the resolution of the scan image (*Table 7*). Although the term *pixel* defines the inplane measurements (x and y axes) of a picture element, each slice has a thickness in the z axis as well. The incorporation of this dimension also describes the three-dimensional volumetric picture element or "voxel." This is the building block of digital scanner–based stereotactic neurosurgery.

There is a tension between the interests of resolution that would prescribe the smallest field of view that encompasses the head and the requirement for a larger field of view than that only to encompass the head in order to visualize stereotactic frame fiducials. This tension is universal to external stereotactic localization frames. The size of each frame is slightly different (in ascending rank order, Leksell, COMPASS, BRW, and CRW). These minor differences between the frames are not sufficient to confer a significant advantage to any one system, however, and are offset by a limitation in z-positional accuracy for the smaller frames. This is because, as the dimensions of the N-bar assembly become smaller, the precision decreases for the localization of the oblique fiducial of the N-bar.

In this series of experiments, the effect of scan slice thickness and, therefore, the size of the voxel, were perhaps the most significant factors that affected the application accuracy of stereotactic localization. Increasing the thickness of slices invariably worsened the attainable accuracy of the system. This effect was universally observed, regardless of the system tested. The use of 1-mm CT slices provides better stereotactic application accuracy than does the use of 4- or 8-mm slices. The relationship between increasing slice thickness and deteriorating application accuracy could satisfactorily be described as linear. The level of localization error observed when a CT scan slice thicknesses of 8 mm is used sufficiently large to suggest that the satisfactory stereotactic localization of smaller intracranial targets should require thinner CT scan slices. Interestingly, as the magnitude of localization errors increased with slice thickness, this increase was seen not only in the z dimension, but in the x and y directions as well. For purposes of clinical application, then, increasing slice thickness led to a general deterioration of the accuracy of localization and one cannot assume that the increased localization error associated with increasing slice thickness will be confined to only the z dimension of slice thickness.

The errors of localization with stereotactic frames, as measured in this study, cannot be attributed entirely to imaging effects. This is supported by the following three considerations. First, if localization errors were entirely the result of the uncertainty of spatial position within a voxel, then the maximal localization error would be determined by the largest dimension of the voxel. Termed the "main diagonal," this is the square root of the sum $[x^2 + y^2 + z^2]$. The localization of a CT-defined target with orthogonally oriented 1-mm-thick scan slices centered on that target represents an optimized situation. In this case, the main diagonal of the voxel equals $[(0.65)^2 + (0.65)^2 + (1.0)^2]^{1/2} = 1.36$ mm. In this study, with the currently available stereotactic instrumentation, the 99.9% confidence intervals for application accuracy in this ideal case are in the range of 3.0 to 5.1 mm. This is a greater number than that attributable solely to imaging effects, being larger than the largest dimension of the voxel.

Second, this is corroborated by *Figure* 3, which plots the effect of slice thickness on the application accuracy for the four stereotactic frames examined. Although the linearity of the correlation establishes the significant contribution of imaging, note that extrapolation of the regression line intersects the axis describing application accuracy at 2.28 mm. This suggests that, even at the ideal case of zero slice thickness and, presumably, zero contribution from the imaging modality used, the 99.9% confidence interval for what is, in effect, the mechanical accuracy of stereotactic frames would not be submillimetric.

Finally, because most stereotactic frame systems are complex constructions with multiple vernier settings and other mechanical adjustments to make, ample opportunity arises for introducing error throughout the surgical procedure. Single large errors that result in dramatic misdirection are noticeable, whereas cumulative tiny errors can produce deflections that are far harder to appreciate but that are significant nonetheless.

Angulation of the stereotactic frames with respect to the plane of image acquisition led to worsened application accuracy for all three stereotactic frames tested. The algorithm for calculating the spatial position of the N-bars is theoretically independent of the orientation of the stereotactic frame with respect to the imaging plane, because the position of the center fiducial in each pair of three making up an "N" is calculated as a proportion relative to the two outer, parallel, fiducial bars (4–6). Although the requisite software for this capability is provided for the CRW, BRW, and COMPASS systems, significantly degraded spatial localization was observed when these stereotactic frames were angulated during CT data acquisition. Perhaps the commercially available software was not optimally executed, new software to localize better the centroid of each fiducial is necessary, or errors of the machining and calibration of the orientation of the three bars making up the "N" may be exaggerated in an oblique orientation as opposed to an orthogonal orientation. Whatever the reason, in order to improve spatial localization accuracy during clinical use, gantry tilts and frame angulation that would result in oblique images should be minimized.

Although the use of gaps between slices in MR imaging allows some improvement in the signal-to-noise ratio, the use of gaps coupled with relatively thinner slices for CT scanning does not confer any improvement of localization on volumetrically acquired CT data. Because no imaging advantages accrue in CT with skips between slices and the intervening patient data are omitted, stereotactic data acquisition with CT should be performed without skips between slices. The overlapping of slices is an outmoded method of improving resolution in the z axis (along the direction of slice thickness) when constrained by the inability of the CT scanner to produce sufficiently thin slices. The overlapping of thicker scan slices is, by definition, inferior in resolution to nonoverlapped scans of thinner slice thickness. Accordingly, this archaic technique is no longer used.

The proper acquisition of CT imaging data for stereotactic localization presumes appropriate attention to the fundamental details of machine calibration and maintenance by radiological technicians. To some degree, the stereotactic frame is relatively robust with regard to such errors. Such technical considerations become critical, however, when these scans are used by the emerging frameless localization systems (28, 29, 33).

Differences between frames

These trials used five new specimens of each stereotactic frame type tested, each certified as in compliance with industry standards by the manufacturer, in order to preclude any bias from being introduced through the use of a defective frame. Differences were noted between the results obtained by individual frame types, even though these differences were not the most significant determinant of application accuracy. It may be that the slight superiority of the COMPASS system was the result of a combination of a center-finding algorithm for fiducial localization and the optical encoder-controlled stepper motor-driven setting of the verniers of the arc quadrant floor stand. This is unlike the manual settings of verniers and the subjective nature of localizing the centroid of the fiducials on the other systems. What is most notable, however, is that all stereotactic frame systems behaved with remarkable similarity when the effects of imaging such as increasing slice thickness were measured. All four stereotactic frame systems demonstrated coefficients of variation of approximately 0.5, indicating similar distributions for their application accuracy. Any differences in application accuracy between the four stereotactic frame systems tested are less significant than the differences in application accuracy for each stereotactic frame between CT scans of differing slice thicknesses. Because of this uniformity of response by all of the stereotactic frames tested, further evaluations involving other less common varieties of commercially available stereotactic frames appear superfluous.

The spatial distribution of localization error relative to target positions in stereotactic space was analyzed for patterns of bias that deviated from flat or random distributions. No consistent pattern of bias was observed. Certainly, no "sweet spot" exists at the center of the stereotactic coordinate space for these frame systems.

Reapplication of aiming arc assemblies

The aiming arc assembly of a stereotactic frame holds and directs the probe as it is directed to the target point (18). The arc is mechanically coupled directly (BRW, CRW, Leksell) or indirectly (COMPASS) to the stereotactic frame base ring. Performing the mechanical coupling of these components can introduce error into the system. In some circumstances, the aiming arc assembly may need to be removed and replaced during the course of an operation, e.g., during the use of the BRW arc to define an entry point from the skull with subsequent application of the arc to direct the biopsy probe. This second arc application may have a measurable effect on the mechanical accuracy of the entire system. In doing so, errors of localization may be introduced that are on the order of 0.5 to 1 mm.

Phantom base units

Historically, a phantom target base was devised for the Riechert-Mundinger stereotactic frame because of the complexity of the required polar coordinate calculations for the aiming arc assembly to achieve a given trajectory to a target point (11, 31). This unit provided a mechanical analogue computer to solve this problem more quickly in the operating room. Later, a phantom base unit was provided for the BRW system in order to allow the intraoperative definition of entry point coordinates chosen from the patient's head and to confirm the proper set-up of the aiming arc assembly verniers (2, 17). The phantom bases provided with the BRW and CRW systems are only a check of the mechanical accuracy for setting up the system. The phantom base accuracy evaluation measures the mechanical error of the frame plus the adjustment error of the phantom base itself. It cannot and does not test the imaging aspects of the system in use, but only confirms a proper setting of verniers. Therefore, it cannot supply any information about application accuracy whatsoever. The true intraoperative position of a probe tip can only be determined by recourse to anteroposterior and lateral x-ray teleradiographs overlaid with a reference grid (22, 30).

A unique situation exists in LINAC-based stereotactic radiosurgery, where the phantom base is called on to provide a reference standard against which the mechanical accuracy of the LINAC set-up itself is tested (10, 36). In this situation, decisions are being made as to whether the therapeutic equipment used in radiosurgery is adequately calibrated, on the basis of information that is limited by the mechanical accuracy of the phantom base unit. When a discrepancy is found, it can be problematic as to whether the source of error is the LINAC set-up or the phantom base unit. Because of this, it behooves the user to test independently any such unit before relying on it as a reference standard for radiosurgical applications. It has been observed, in fact, that commercially available phantom base units for the BRW and CRW systems are often inadequate for this purpose without prior individual testing, modification, and subsequent calibration (9). The levels of mechanical accuracy demonstrated by the BRW and CRW phantom base units were insufficient to provide an absolute reference standard for radiosurgical applications. They provide adequate accuracy to serve as an assurance of proper vernier settings during surgical procedures. The CRW phantom base units are modifications of the previous BRW units and appear to be somewhat less accurate.

Weightbearing

The design of stereotactic frames has involved a compromise between the mutually exclusive goals of fiducial marker stability and maximal clinical utility. Rather than only serving to carry the fiducial markers necessary for localization, the stereotactic frame is called on to carry the load of the aiming arc assembly with its accompanying surgical instruments and also to act as the major weightbearing element for the patient's cranium, neck, and even torso during surgery. This linkage of the imaging fiducials to the weightbearing elements is suboptimal. Although the effects of this weight loading might be minimized by proper positioning and support, it nevertheless introduces an independent source of mechanical inaccuracy into the system. The deleterious effect of increasing weight on the mechanical accuracy of a stereotactic frame can easily become pronounced. It can be appreciated, for example, that the positioning of a heavy patient prone with the head sharply rotated can produce a significant source of torque and weight loading on the frame elements. This effect of weightbearing on the mechanical accuracy of stereotactic frames has not been previously addressed. The mechanical accuracy and, therefore, the application accuracy of all types of stereotactic frames tested are lessened by clinically realistic weightbearing. Weightbearing is an independent source of error that profoundly affects the mechanical accuracy of stereotactic frames in direct proportion to the load they have to support. Extrapolation of the regression line suggests that this effect has been overlooked by the practice of using test phantoms that weigh less than 5.85 kg. Because the cranium is at least this heavy, the implications of this deterioration of mechanical accuracy with weightbearing by stereotactic frames should be considered with every clinical use.

In theory, it would be optimal to uncouple the rigid intraoperative headholder mechanically from the fiducial markers used to register imaging to physical space. This would require 1) careful patient positioning during imaging and surgery; 2) the use of an independent headholder in addition to the stereotactic frame; 3) a change in the design of stereotactic frames; or 4) the development of new neurosurgical localization systems. It would appear that the first option is most directly practical at this time, whereas the last offers the greatest future potential.

Why bother with accuracy?

The application accuracy of stereotactic frames is determined by the vector summation of many independent sources of error. The results presented here might mislead some to conclude erroneously that further efforts at improving the accuracy of localization are unnecessary. The rationale for that argument is as follows: The error of magnitude × that contributed by a given aspect of the stereotactic process is not important because it is smaller than the error of magnitude ×2 due to slice thickness, etc. In effect, this argument hopes that one, smaller, error will somehow "nest" inside of another, larger, error. This is a false assumption. These two errors arising from different sources are independent events. In fact, errors arising from independent sources contribute to the total system application accuracy as a vector sum. Because all aspects of a system contribute to the total application error, any compromise in mechanical accuracy or imaging studies will invariably compromise the total application accuracy. It behooves the surgeon to minimize all sources of error that are controllable, so that the uncontrollable errors due to brain shifting, etc., are added to as small a background level of error as possible.

This caution is particularly relevant to the special case of stereotactic radiosurgery, where no intraoperative feedback exists. In radiosurgery, what is being treated directly is the imaging abnormality (10). This is indirectly correlated with the actual pathology. During the course of a radiosurgical procedure, there is no direct feedback to assure proper beam delivery to the correct anatomical location. Because of this, the mechanical accuracy of all components must be rigorously assessed to ensure sufficient accuracy. Extra precautions must be maintained for radiosurgical procedures so that systematic errors do not creep into the process and introduce bias into localization (36).

Is subvoxilar accuracy possible?

A common fallacy is that the spatial resolution of any surgical guidance system can never be better than the dimensions of the voxel, because this is the minimal element of the mosaic making up the localization data. In fact, this is not true. If the fiducials and/or the target are larger than the dimensions of a single voxel, then additional spatial information can be obtained through the calculation of their center of mass on the basis of the voxels making up their volume and the intensity of their signal within each voxel. In contrast, a fiducial or target that is smaller than a voxel cannot be resolved to an accuracy better than that of the voxel dimensions, regardless of how bright it is. Thus, the ideal fiducials and targets are as large as possible, providing the largest number of voxels from which to calculate the center of mass, regardless of how dimly it is seen above the background. This is sometimes termed *super resolution* and will likely play a significant role in future frameless stereotactic systems (28, 29, 32).

Ameliorating effects

Stereotactic frame systems have proved to be useful for improving the accuracy and precision of localization during neurosurgical operations. Despite the limitations of application accuracy delineated in this report, these devices can and do provide surgically important information. Their clinical acceptability is based on several factors. First, the application accuracy measured during a single use of a stereotactic frame will, by definition, most often approach the value of its mean. The magnitude of this number will be far smaller than the number encompassing the 99.9% confidence interval. Second, many anatomic stereotactic targets are sufficiently large that even millimetric errors of localization may go unnoticed. During biopsies, anatomic targets are usually larger than the minimal resolution of the system and are therefore somewhat forgiving of error (3, 16, 20, 21, 31). The practice of performing serial sequential stereotactic biopsies through a tumor target serves to overcome many errors of localization (11, 17, 30). Third, various forms of real-time intraoperative feedback can provide a "midcourse correction" so as to assure an ultimately successful surgical targeting. During thalamotomies, for example, electrophysiological recording feedback provides the final critical localization that is based on mechanical precision by movements along a vernier (15, 23, 25, 34, 35). Intraoperative visual inspection and frozen-section analysis of serial stereotactic biopsy specimens allow the surgeon to change the position of the probe and biopsy again as indicated (2, 11). During craniotomies, the correlation of visual feedback and stereotactic guidance provides two independent data bases from which the surgeon can choose the most useful information (19–21, 32, 33).

Caveat emptor

The application accuracy of stereotactic frame systems is less than the mechanical accuracy, precision, or unbiasedness described for these systems. Their proper clinical use, therefore, requires a mature technological understanding of the instruments available and a clear comprehension of their benefits and limitations. Clinically, the determinants of application accuracy should be considered before every use of stereotactic frames for any therapeutic intervention. The results of this study must be kept in mind by every surgeon using stereotactic instrumentation and methodology in order to tailor the imaging protocols and equipment set-up so as to achieve an appropriate confidence level for that surgical target localization. It is clearly incumbent on the neurosurgeon to determine what level of application accuracy is required for a given clinical situation and to judge whether the stereotactic localization system as configured can provide an appropriate degree of localization for that individual situation.

ACKNOWLEDGMENTS

The authors acknowledge the invaluable aid provided by C. Cobb, M.D., E. Zaccharias, M.D., A. Moore, M.D., K. Higginbotham, R.T., A. Jackson, R.T., N. Collins, R.T., T. Herron, R.T., V. Mandava, Ph.D., M. Zink, J.M. Landwehr, Ph.D., J.M.

Fitzpatrick, Ph.D., R. Larsen, Ph.D., S. Goerss, D. Canale, B. Norden, D. Deckel, D. Eisert, M.D., T.S. Chen, Ph.D., R. Taylor, S. Gadamsetty, M.S., C. Maurer, M.S., E. Longworth, A. Heflin, A.L. Failinger, M.D., and G.S. Allen, M.D., Ph.D.

Received, September 1, 1993.
Accepted, March 1, 1994.
Reprint requests: Robert J. Maciunas, M.D., Department of Neurological Surgery, T 4224 MCN, Vanderbilt University Medical Center, Nashville, TN 37232.

REFERENCES

1. American Society for Testing and Materials Committee F-4.05. Standard performance specification for cerebral stereotactic instruments, in *Annual Book of ASTM Standards, F 1266–89.* Philadelphia, American Society for Testing and Materials, 1990, pp 1–6.
2. Apuzzo MLJ, Sabshin JK: Computed tomographic guidance stereotaxis in the management of intracranial mass lesions. **Neurosurgery** 12:277–285, 1983.
3. Backlund E, von Holst H: Controlled subtotal evacuation of intracerebral hematomas by stereotactic technique. **Surg Neurol** 9:99–101, 1978.
4. Brown RA: A stereotactic head frame for use with CT body scanners. **Invest Radiol** 14:300–304, 1979.
5. Brown RA: A computed tomography–computer graphics approach to stereotaxic localization. **J Neurosurg** 50:715–720, 1979.
6. Brown RA, Roberts TS, Osborn AG: Stereotaxic frame and computer software for CT-directed neurosurgical localization. **Invest Radiol** 15:308–312, 1980.
7. Dixon WJ, Massey FJ: *Introduction to Statistical Analysis.* New York, McGraw-Hill Book Co., 1983, ed 4, p 678.
8. Dong S, Fitzpatrick JM, Maciunas RJ: Rectification of distortion in MRI for stereotaxy, in Brown J, Santago G (eds): *Computer-Based Medical Systems.* Los Alamitos, CA, IEEE Computer Society Press, 1992, pp 181–189.
9. Eisert D, Maciunas RJ, Murray MJ, Chen TS, Taylor R: Stereotactic radiosurgery: Consistency of set-up accuracy over time, in Lunsford LD (ed): *Stereotactic Radiosurgery Update.* New York, Elsevier Science Publishing Co., 1992, pp 213–217.
10. Friedman WA, Bova FJ: Linear accelerator radiosurgery for arteriovenous malformations. **J Neurosurg** 77:832–841, 1992.
11. Galloway RL, Maciunas RJ: Stereotactic neurosurgery. **CRC Crit Rev Biomed Eng** 18:181–205, 1990.
12. Galloway RL, Maciunas RJ, Latimer JW: The accuracies of four stereotactic frame systems: An independent assessment. **Biomed Instr Technol** 25:457–460, 1991.
13. Glantz SA: *Primer of Biostatistics.* New York, McGraw-Hill, 1992, ed 3, p 440.
14. Goerss S, Kelly PJ, Kall B, Alker GJ: A computed tomographic stereotactic adaptation system. **Neurosurgery** 10:375–379, 1982.
15. Guiot G, Hardy J, Albe-Fessard D: Delimitation precise des structures sous-corticales et identification des moyaux thalamiques chez l'homme par l'electrophysiologie stereotaxique. **Neurochirurgie** 5:1–18, 1962.
16. Gutin PH, Phillips PL, Wara WM, Leibel SA, Hosobuchi Y, Levin VA, Weaver KA, Lamb S: Brachytherapy of recurrent malignant brain tumors with removable high-activity iodine-125 sources. **J Neurosurg** 60:61–68, 1984.
17. Heilbrun MP, Roberts TS, Apuzzo MLJ: Preliminary experience with Brown-Roberts-Wells (BRW) computerized tomography stereotaxic guidance system. **J Neurosurg** 59:217–222, 1983.
18. Horsley V, Clarke RH: The structure and functions of the cerebellum examined by a new method. **Brain** 31:45–124, 1908.
19. Jacques S, Shelden CH, McCann GD: A computerized microstereotactic method to approach, 3-dimensionally reconstruct, remove and adjuvantly treat small CNS lesions. **Appl Neurophysiol** 43:176–182, 1980.
20. Jacques S, Shelden CH, McCann GD, Freshwater DB, Rand R: Computerized three-dimensional stereotaxic removal of small central nervous system lesions. **J Neurosurg** 53:816–820, 1980.
21. Kelly PJ, Alker GJ, Goerss S: Computer-assisted stereotactic laser mic- surgery for the treatment of intracranial neoplasms. **Neurosurgery** 10:3- 331, 1982.
22. Kelly PJ, Kall BA, Goerss B: Transposition of volumetric information deri- from computed tomography scanning into stereotactic space. **Surg Neu** 21:465–471, 1984.
23. Laitinen LV: Brain targets in surgery for Parkinson's disease. Results of survey of neurosurgeons. **J Neurosurg** 62:349–351, 1985.
24. Leksell L: A stereotaxic apparatus for intracerebral surgery. **Acta Chir Scan** 99:229–233, 1949.
25. Leksell L: The stereotactic method and radiosurgery of the brain. **Acta Chi Scand** 102:316–319, 1951.
26. Leksell L, Jernberg B: Stereotaxis and tomography. A technical note. **Act Neurochir** 52:1–7, 1980.
27. Maciunas RJ, Galloway RL, Latimer JW, Cobb C, Zaccharias E, Moore A Mandava VR: An independent application accuracy evaluation of stereotac- tic frame systems. **Stereotact Funct Neurosurg** 58:103–107, 1992.
28. Maciunas RJ, Galloway RL, Fitzpatrick JM, Mandava VR, Edwards CA Allen GS: A universal system for interactive image-directed neurosurgery **Stereotact Funct Neurosurg** 58:108–113, 1992.
29. Maurer CR, Fitzpatrick JM: A review of medical image registration, in Maciunas RJ (ed): *Interactive Image Guided Neurosurgery.* Chicago, American Association of Neurological Surgeons, 1993, pp 17–44.
30. Olivier A, Bertrand G, Peters T: Stereotactic systems and procedures for depth electrode placement: Technical aspects. **Appl Neurophysiol** 46:37–40, 1983.
31. Riechert T, Mundinger F: Indications, technique and results of the stereotactic operations upon the hypophysis using radio-isotopes. **J Nerv Ment Dis** 13:1–9, 1960.
32. Roberts DW, Strohbehn JW, Hatch JF, Murray W, Kettenberger H: A frameless stereotaxic integration of computerized tomographic imaging and the operation microscope. **J Neurosurg** 65:545–549, 1986.
33. Shelden CH, McCann G, Jacques S: Development of a computerized microstereotactic method for localization and removal of minute CNS lesions under direct 3-D vision. **J Neurosurg** 52:21–27, 1980.
34. Spiegel EA, Wycis HT, Marks M: Stereotactic apparatus for operations on the human brain. **Science** 106:349–350, 1947.
35. Spiegel EA, Wycis HT, Freed H: Stereoencephalotomy. Thalamotomy and related procedures. **JAMA** 148:446–451, 1952.
36. Winston KR, Lutz W: Linear accelerator as a neurosurgical tool for stereotactic radiosurgery. **Neurosurgery** 22:454–464, 1988.

COMMENTS

The detailed and labor-intensive study by Maciunas et al. addresses an important issue in the field of stereotactic surgery. It is too easily assumed by the surgeon that the coordinates indicated by the calculations to determine a stereotactic target and the numbers on the scales indicating how the apparatus is set invariably direct the surgeon to the intended point. Such accuracy is not always so.

As the authors indicate, there are several types of accuracy to be considered and there are several factors contributing to each type of accuracy. The performance standards that the authors cite in their Reference 1 were devised before the advent of computed tomographic or magnetic resonance imaging and consequently concern the use of stereotactic instruments with x-ray guidance. The standards refer to mechanical accuracy as the ability of a specific apparatus to reach reliably a specific point in space. Anatomical accuracy is the reliability with which a given anatomical target can be hit and may vary relative to the position of the reference atlas position of the structure. The further it is from the reference structure, the more likely inaccuracies are to be introduced because of the anatomical variability of one brain from another. The authors also introduce the concept of application accuracy, the final accuracy imparted by the entire visualization, calculation, and surgical systems working together.

The issues of mechanical and application accuracy are being addressed, perhaps overcritically, by the authors, whereas the concept of anatomical accuracy (which is addressed in their Reference 1) is not only not specifically addressed, but comments about the localization of VIM improperly assume absolute anatomical accuracy. Anatomical accuracy is far less than application accuracy. Final localization of the anatomical structures is dependent, wherever possible, on the physiological identification of the target structure.

The point is made several times that several samples from separate manufacturing runs of each device were tested, but there is no comment as to the comparison between each sample. If there were a large variability between samples, it would imply lax manufacturing standards. If there were an identical inaccuracy between samples, it would imply a design peculiarity. The inaccuracy imparted by weightbearing was measured, but no comment was made (nor was tested) on the distortion of the apparatus imparted by the torque of the application of the device to the patient's skull, which may be significant.

Although there is a brief comment about the need for the neurosurgeon to determine the accuracy required for each procedure, there could be more guidance about the range of inaccuracy that would be acceptable. There is obviously a much greater need for accuracy when aiming toward a small target in a critical area, such as the brain stem, as compared with an ill-defined large target in a relatively silent area, such as a glioma in the nondominant frontal pole. Indeed, even with the most sophisticated imaging available, it is not possible to define accurately the ideal line of resection of such an infiltrating tumor. As much critical application accuracy as possible is required in the former situation. Localization need not (and cannot) be as precise in the latter case. It is the responsibility of the neurosurgeon, however, to recognize the accuracy of the system in use and to evaluate whether it is accurate enough for the intended purpose.

Of great significance is the demonstration of the relationship between accuracy and slice thickness. Not only does a thicker slice impart inaccuracy to the localization system, but the center of the location of a lesion smaller than one slice thickness cannot be defined. The authors have exerted substantial effort in defining the accuracy of several popular systems and in alerting the reader to be cautious that the accuracy of the system is sufficient for the use to which it is put.

Philip L. Gildenberg
Houston, Texas

This article should alert those surgeons doing stereotactic surgery to the need for being aware at all times of the several factors that influence the accuracy of the system they are using. It can also benefit the several stereotactic instrument manufacturers by alerting them to the "repeatability" problems discussed in the article. The failure of the authors to note the serial numbers of the instruments they claim to have tested prevents any of the manufacturers to rebut their accuracy/inaccuracy claims and/or to take any action against the authors.

The title of the article is misleading; it is not a study of the mechanical accuracy of the several frames tested. It is instead a study of the total system inaccuracies due to scan data reduction errors, operator errors in determining the central pixel of localizer fiducial areas, and the improper setting of aiming arc scales. Actually, in the Leksell instrument, the Cosman-Roberts-Wells arc, and the COMPASS instrument, being arc-centered systems, the angular settings have nothing to do with the accuracy of getting to the target point if the linear scales have been properly set and the proper length probe is used. Claims of being able to measure distances in space to 0.01 mm with an electronic digital microcaliper are sheer nonsense. A distance of 0.01 mm is equal to 0.0004 inches, a dimension that is difficult to get with a high-quality micrometer measuring a hardened block having parallel surfaces!

The problem of head weight shifting the targets at the operating table can usually be neutralized by supporting the head in the scanner in approximately the same manner as it will be supported on the operating table. Thus, there will be very little shift of targets from those derived at the scanner console. It appears that these authors were led down the primrose path by the statistical specialists who can prove almost any point they wish if they have enough test data.

Trent H. Wells, Jr.
Coulterville, California

This article addresses the critically important issue of stereotactic frame accuracy, about which there has long been insufficient understanding, incomplete information, or both. The authors have rigorously and exhaustively documented the performance of four commonly used stereotactic frames and have made a contribution both for the data and the model of analysis. With regard to the data, this study finally provides a quantitative analysis of the accuracy one may expect from a stereotactic procedure, including not just the mechanical error of a frame (the presumed source of the often quoted but vaguely documented "less than 1 mm") but also the errors associated with imaging, registration, and whatever additional unrecognized processes may influence the final localization. The comparative accuracies reported for these particular frame systems are of relatively less interest than the objective demonstration of a considerably greater inaccuracy for any of these systems than most people have been traditionally taught. The fact that true error in any given case may well exceed several millimeters is important to know, both for that procedure and for more valid benchmarking of newer stereotactic methodologies. With respect to the model of analysis, the article sets an example for all in its rigor of investigation and adherence to engineering standards that must be respected.

David W. Roberts
Lebanon, New Hampshire