Designation: F2554 – 22

Standard Practice for Measurement of Positional Accuracy of Computer-Assisted Surgical Systems[1]

This standard is issued under the fixed designation F2554; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This document provides procedures for measurement and reporting of basic static performance of surgical navigation and/or robotic positioning devices under defined conditions. They can be performed on a subsystem (for example, tracking only) or a full computer-aided surgery system as would be used clinically. Testing a subsystem does not mean that the whole system has been tested. The functionality to be tested based on this practice is limited to the performance (accuracy in terms of bias and precision) of the system regarding point localization in space by means of a pointer. A point in space has no orientation; only multidimensional objects have orientation. Therefore, orientation of objects is not within the scope of this practice. However, in localizing a point the different orientations of the pointer can produce errors. These errors and the pointer orientation are within the scope of this practice. The aim is to provide a standardized measurement of performance variables by which end users can compare within a system (for example, with different reference elements or pointers) and between different systems (for example, from different manufacturers). Parameters to be evaluated include (based upon the features of the system being evaluated):

(1) Accuracy of a single point relative to a coordinate system.

(2) Sensitivity of tracking accuracy due to changes in pointer orientation.

(3) Relative point-to-point accuracy.

1.1.1 This method covers all configurations of the evaluated system as well as extreme placements across the measurement volume.

1.2 This practice defines a standardized reporting format, which includes definition of the coordinate systems to be used for reporting the measurements, and statistical measures (for example, mean, RMS, and maximum error).

1.3 The values stated in SI units are to be regarded as standard. No other units of measurement are included in this standard, except for angular measurements, which may be reported in terms of radians or degrees.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.5 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

E456 Terminology Relating to Quality and Statistics

E2281 Practice for Process Capability and Performance Measurement

2.2 *Other References:*[3]

ISO 10360 Geometrical Product Specifications (GPS)—Acceptance and Reverification Tests for Coordinate Measuring Machines (CMM)

3. Terminology

3.1 *Definitions:*

3.1.1 *accuracy, n*—the closeness of agreement between a measurement result and an accepted reference value. **E456**

3.1.1.1 *Discussion*—In the context of this standard, with the definitions of bias and precision (see below), it can be considered that the accuracy of a measurement of a point will be subject to some bias error and some precision error.

3.1.2 *bias, n*—the difference between the expectation of the measurement results and an accepted reference value. **E456**

[1] This practice is under the jurisdiction of ASTM Committee F04 on Medical and Surgical Materials and Devices and is the direct responsibility of Subcommittee F04.38 on Computer Assisted Orthopaedic Surgical Systems.

Current edition approved Sept. 1, 2022. Published September 2022. Originally approved in 2010. Last previous edition approved in 2018 as F2554 – 18. DOI 10.1520/F2554-22.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

[3] Available from American National Standards Institute (ANSI), 25 W. 43rd St., 4th Floor, New York, NY 10036, http://www.ansi.org.

3.1.2.1 *Discussion*—In the context of this standard, bias represents the systematic error in a set of measurements of a target reference point making their average deviate from the actual reference point with a certain magnitude and direction.

3.1.3 *calibration, n*—the pre- or intraoperative registration of an item or device to its reference element.

3.1.4 *computer-assisted surgery (CAS), n*—the use of computers to facilitate or enhance surgical procedures via the use of three-dimensional space tracking of objects.

3.1.5 *coordinate measuring machine (CMM), n*—measuring system with the means to move a stylus and capability to determine spatial coordinates on a work piece surface. **ISO 10360-1**

3.1.6 *degree of freedom (DOF), n*—set of independent displacements that specify completely the displaced or deformed position of the body or system.

3.1.7 *dynamic reference base, n*—the coordinate system of a reference element used for the tracking of other therapeutic objects.

3.1.8 *fiducial, n*—an artificial item (for example, a screw or a sphere) rigidly attached to a therapeutic object to facilitate its calibration.

3.1.9 *ground truth, n*—short name for the accepted reference value (see 3.1.1).

3.1.10 *marker, n*—a single 3-degree-of-freedom indicator on a reference element or dynamic reference base.

3.1.11 *maximum error, n*—the largest distance between any measured point and its ground truth for any trial during a testing procedure.

3.1.12 *mean, n*—of a population, u, average or expected value of a characteristic in a population; of a sample, x, sum of the observed values in the sample divided by the sample size. **E456**

3.1.13 *measurement range, n*—the interval of allowed values for a specific degree of freedom while performing the practice.

3.1.14 *measurement volume, n*—measuring range of a tracker, stated as simultaneous limits on all spatial coordinates measured by the tracker. **ISO 10360-1**

3.1.15 *navigation system, n*—a set of devices consisting of a computer, its associated software, and a tracker capturing the reference elements within the measurement volume. This system provides real-time feedback of the state of the surgical scene under operation.

3.1.16 *phantom, n*—standardized measurement object. See Appendix X1 for details regarding the design of the phantom used in this practice.

3.1.17 *pointer, n*—the device offered by the evaluated system to point and locate a position on any object including anatomical landmarks. The pointer is the whole device, including the stylus-like tip all the way to any reference element used to track it in space.

3.1.18 *precision, n*—the closeness of agreement between independent measurement results obtained under stipulated conditions. **E456**

3.1.18.1 *Discussion*—In the context of this standard, precision represents scatter of a set of measurements of a point.

3.1.19 *range, R, n*—the largest observation minus the smallest observation in a set of values or observations. **E456, E2281**

3.1.20 *reference element, n*—an artificial item composed of rigidly bound markers in a unique and asymmetrical pattern recognizable by the tracker. While being rigidly attached to a therapeutic object, the position and orientation of the reference element can be used to determine those of the therapeutic object after its calibration.

3.1.21 *registration, n*—the determination of the spatial relationship between the referential frames of two coordinate systems. This may occur between two reference elements or between the fiducials and a reference element of a therapeutic object. The registration is rigid if it consists only of rotations and translations (six degrees of freedom) and non-rigid if it also comprises scaling and/or local or global distortions (seven degrees of freedom and more).

3.1.22 *repeatability, n*—precision under repeatability conditions. **E456**

3.1.23 *reproducibility, n*—precision under reproducibility conditions. **E456**

3.1.24 *robotic positioning system, n*—use of an active mechanical (mechatronic) device to position an instrument guide at a specified location in 3D space (up to six degrees of freedom).

3.1.25 *root mean square (RMS), n*—means of estimation of the scatter of a set of values, which consists of the square root of the average of the squared values.

3.1.26 *therapeutic object, n*—a surgical item or a part of the patient.

3.1.27 *tracker, n*—a device that detects and locates fiducials and markers in its measurement volume. This can be achieved by mechanical linkage or by analyzing signals of various types (visible or infrared light, electromagnetic field, or ultrasound).

4. Summary of Practice

4.1 This practice provides recommendations for the collection, analysis, and presentation of data regarding the positional accuracy (in terms of bias and precision) of surgical navigation and robotic positioning systems under repeatable conditions.

4.2 Data to be reported consists of all measurements, their corresponding errors if applicable, their statistical analysis, the test conditions, and the system conditions.

5. Significance and Use

5.1 The purpose of this practice is to provide data that can be used for evaluation of the accuracy of different CAS systems.

5.2 The use of surgical navigation and robotic positioning systems is becoming increasingly common. In order to make informed decisions about the suitability of such systems for a given procedure, their accuracy capability needs to be evaluated under clinical application and compared to the requirements. As the performance of a whole system is constrained by

those of its subparts, a preliminary step must be to objectively characterize the accuracy of the tracking subsystem in a controlled environment under controlled conditions.

5.3 In order to make comparisons within and between systems, a standardized way of measuring and reporting accuracy is needed. Parameters such as coordinate system, units of measurement, terminology, and operational conditions must be standardized.

6. Apparatus

6.1 The system under test is considered to have at least some tracking functionality, a pointer and associated hardware, and software. If the system is provided by the manufacturer with various combinations of parts, the evaluation must be performed at least with the combination known to present the worst-case scenario in terms of accuracy. For example, the tester may use the longest pointer with the smallest reference elements.

6.2 This practice relies on a phantom. See Appendix X1 for design requirements. The phantom size and points have been designed to approximate a typical surgical site on the human body. All divots of the phantom shall be measured by a CMM (or another measurement system of similar performance traceable to NIST, FDA, EU, and ISO standards). These phantom measurements will constitute the ground truth used for the accuracy assessment. Therefore, the accuracy of the CMM must be better than that of the system being evaluated.

6.3 If the evaluated system relies on a dynamic reference base for its measurements, a reference element is attached to the phantom. This reference element and its attachment shall replicate as close as possible those used in a surgical setting.

6.4 If dedicated additional software functionality is used for assistance in performing the tasks outlined in this practice, or for statistical analysis of the measurements, this addition must not alter the way the measurements are made by the system to be used clinically.

7. Hazards

7.1 None.

8. Procedure

8.1 Set up the system to be evaluated, including the dynamic reference base, if used. The dynamic reference base shall not be repositioned relative to the phantom during the procedure. Measured points shall be expressed in the coordinate system of the phantom, to enable them to be directly compared to those made by CMM.

8.2 *Test Conditions*—The conditions used for this procedure should be as close as possible to those typically found in the surgical setting for which the system is designed, and those conditions shall be reported. They include all known factors that may influence the performance of the evaluated system, such as temperature, humidity, lighting, as well as potential sources of interference (for example, infrared noise and reflections for optical systems, large metal objects for electromagnetic systems). Some of these factors may not be replicated if their non-inclusion is reported and justified.

8.3 *System Conditions*—The system is composed of various parts and all their references and configuration shall be provided, including firmware and software versions. Any changes to the system beyond what is provided and configured by the manufacturer are to be reported and justified (for example, using third-party markers or pointer). Specific details of the phantom are also to be reported (for example, the divot dimensions). The measured points are to be acquired only through the firmware and software provided by the manufacturer.

8.4 *Phantom Placement and Registration*—In the first series of tests, place the system tracker nominally at the recommended distance from the phantom. At this location, a registration of the phantom to the dynamic reference base may be required for most systems. Any registration shall be performed as described by the manufacturer, simulating registration of a patient's anatomy in the clinical environment. Registration can only be done once to cover the sequence of steps 8.6 – 8.9, but may be repeated in between.

8.5 *Point Acquisition*—In each trial, the tester locates the individual labeled points on the phantom and acquires its position using the pointer following the system manufacturer's instructions for obtaining point data. This includes pointer orientation except for the rotation tests in 8.7.

8.6 *Test 1: Single Point Accuracy*—Treating the phantom as if it was part of the patient's anatomy, this test requires the measurement of a designated point of the phantom multiple times, to compare the positions measured versus the actual position on the phantom relative to its local coordinate system. The single point measurement is then independently performed 20 times on the central divot (#20 in Fig. X1.1). Bias is estimated by the difference between the average of the measured points and the central divot. The result is a small error vector emanating from the target reference point. Calculate the average of all the error vectors by vectorial summation then dividing the length of the resulting vector by the number of vectors. Report the average error vector and the length of the longest error vector. For the determination of precision, start by calculating the average point of all measurements, which represents the system's best estimate for the location of the target central divot. Calculate the distances of all the measurements from this average point, determine and report the RMS of these distances as well as their maximum.

8.7 *Pointer Rotation*—These steps quantify the variation in successive measurements of the same central divot (#20 Fig. X1.1) with various pointer orientations (tilt). The maximum physically possible angular range depends on the axis of rotation, but most systems have some limit to this range beyond which tracking is no longer provided. The actual range for which the system provides tracking data should be tested at no fewer than ten uniformly spaced intervals. The actual range, all angular increments, and the corresponding measurements shall be reported. The tester may use a protractor scale to measure the pointer angles or any other method. For each test in 8.7, the maximum distance between any two measurements

shall be reported, as well as the RMS of the deviations between all measurements and the average point measured in 8.6 (Test 1).

8.7.1 *Test 2*—The single point is measured under successive rotations of the pointer about its principal (shaft) axis from its nominal orientation (see left picture in Fig. X1.5). For this mode of rotation, the maximum theoretical angular range is [–180°, 180°].

8.7.2 *Test 3*—The single point is measured during successive rotations about the line between the tracker and the central divot (see middle picture in Fig. X1.5). For this mode of rotation, the maximum theoretical angular range is [–90°, 90°], limited further by various factors including the pointer leaving the divot or impacting the dynamic reference base.

8.7.3 *Test 4*—The single point is measured during successive rotations about an axis perpendicular to the line between the tracker and the central divot (see right picture in Fig. X1.5). For this mode of rotation, the maximum theoretical angular range is [–90°, 90°], limited further by various factors including the pointer leaving the divot or impacting the dynamic reference base.

8.8 *Test 5: Point-to-Point Accuracy*—The purpose of this test is to determine the accuracy of measuring the distances between various points. A minimum of 20 points are to be measured, including divot #30, with at least two points per plane of the phantom (four points for the slanted plane). The point-to-point distance shall be calculated for each possible combination of two measured points (for example, 20 measured points yield 190 distances). The errors are the differences between these calculated distances and the corresponding distances from the ground truth based on the phantom CMM measurements. The errors are to be reported, alongside the maximum error and the RMS.

8.9 *Test 6: Phantom Rotation*—During the course of a surgical operation using the system, the patient's pose may be changed and hence that of the dynamic reference base, or the tracking system may move (for example, cameras). To measure the impact of such movement on the system's accuracy, the phantom (assembly, included the reference base) shall be rotated to two extreme orientations about the vertical axis passing through the center of the phantom held horizontally (see Fig. X1.7). Those extreme orientations represent the two opposite orientation limits where tracking of the phantom remains possible. The steps described in 8.6 shall then be performed again.

8.10 *Phantom Locations*—Repeat 8.6 – 8.9 at four different locations within the measurable volume, representing its functional extremes. The four locations should be uniformly distributed at the outer boundaries of the measurable volume cross section at the furthest functional distance from the tracker (see Fig. X1.6).

9. Report

9.1 Report characteristics of the phantom including its CMM measurements and divot geometry.

9.2 Test conditions, as described in 8.2.

9.3 System conditions, as described in 8.3.

9.4 For each step in 8.6 – 8.9, the report should include:

9.4.1 The orientation of the phantom relative to the tracker as described in 8.9 and expressed in the coordinate system of the tracker.

9.4.2 The location of the phantom relative to the measurable volume as described in 8.10 and expressed in the coordinate system of the tracker.

9.4.3 The coordinates and label of all measured points with appropriate units. All measurements of points on the phantom should be expressed in a Cartesian coordinate system of the phantom with clearly described origin and axes.

9.4.4 The resulting statistical analysis, as detailed for each step in 8.6 – 8.8.

10. Precision and Bias

10.1 As with any measurement system, uncertainties and errors are present, and the purpose of this standard is to estimate them. In order for the accuracy of computer-aided orthopedic systems to be estimated in a reproducible and comparable manner among laboratories, it is essential that uniform procedures be established as specified by this standard. Sufficient data have not been produced using identical navigation or robotic systems in different laboratories to provide the precision and bias of this procedure.

10.2 The publication of this test method is intended, in part, to facilitate uniform testing and reporting of data from systems used clinically. Some data (2–4) have been published to help toward this end. Validation of this methodology may be achieved through round-robin testing.

10.3 As specified in 6.2, a CMM or other traceable measurement device of suitable precision must be used to establish a ground truth of all the locations of the points on the phantom. CMM manufacturers typically claim accuracies which exceed the expected accuracies of surgical navigation systems by several times.

11. Keywords

11.1 augmented reality assisted surgery; computer-assisted surgery; computer navigation; imageless guided surgery; infrared tracking system; magnetic tracking; optical tracking; robotic surgery

APPENDIX

(Nonmandatory Information)

X1. PHANTOM

X1.1 A phantom is used in this practice to evaluate the tracking accuracy of a system. To enable reliable comparisons between evaluated systems, all phantoms should provide the functionality described hereafter.

X1.1.1 The phantom has 47 marked divots arrayed on five different machined faces with two orientations (see Fig. X1.1 for illustration). The CAD coordinates of these points are listed in Table X1.1 and blueprints of the phantom are provided in Figs. X1.2 and X1.3.

X1.1.2 The phantom also includes an interface (for example, threaded holes) that enables attachment of the dynamic reference base. As specified in 6.3, this reference element and its attachment are specific to the evaluated system and intended to replicate those of a surgical setting. The details of this interface are not shown in the drawings and are left for the standard user to decide.

X1.2 For consistency, the Cartesian coordinate system of the phantom has point #1 set as the origin, the perpendicular lines (#1 to #18) and (#1 to #19) as the x-axis and y-axis, respectively. The z-axis, defined as the cross product of the x- and y-axes, points vertically upwards from the phantom. The coordinates of the points in CAD, the points measured by CMM, and the points measured by the evaluated system are all be expressed in this coordinate system.

X1.3 The user can construct their own phantom (or purchase one from a third party), provided each individual phantom has all its points measured with a CMM or similar measurement device. It is recommended that the phantom is made from a rigid and thermally stable material, so that its shape and thus the relationship of the 3D positions of the divots relative to each other remain stable between uses of the phantom. The detailed design of the divots can be chosen by the user to suit the shape and dimensions of the pointer tip (see Fig. X1.4 for an example of divot designed for a 1 mm diameter ball point tip).

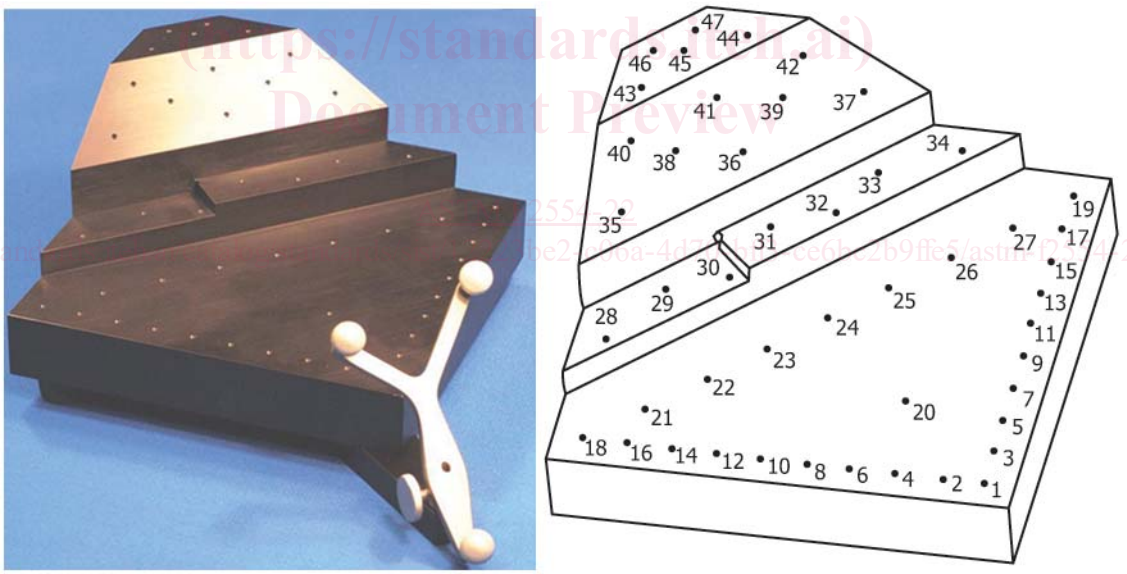

FIG. X1.1 Photo and Diagram of Phantom with Designated Numbering of Divots