Accuracy of Computer-Aided Techniques in Orthopaedic Surgery
How Can It Be Defined, Measured Experimentally, and Analyzed from a Clinical Perspective?

Olivier Cartiaux, PhD, Jean-Yves Jenny, MD, and Leo Joskowicz, PhD

Investigation performed at the Computer Assisted and Robotic Surgery, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Brussels, Belgium

Abstract: Surgical accuracy is multifactorial. Therefore, it is crucial to consider all influencing factors when investigating the accuracy of a surgical procedure, such as the surgeon’s experience, the assistive technologies that may be used by the surgeon, and the patient factors associated with the specific anatomical site. For in vitro preclinical investigations, accuracy should be linked to the concepts of trueness (e.g., distance from the surgical target) and precision (e.g., variability in relation to the surgical target) to gather preclinical, quantitative, objective data on the accuracy of completed surgical procedures that have been performed with assistive technologies. The clinical relevance of improvements in accuracy that have been observed experimentally may be evaluated by analyzing the impact on the risk of failure and by taking into account the level of tolerance in relation to the surgical target (e.g., the extent of the safety zone). The International Organization for Standardization (ISO) methodology enables preclinical testing of new assistive technologies to quantify improvements in accuracy and assess the benefits in terms of reducing the risk of failure and achieving surgical targets with tighter tolerances before the testing of clinical outcomes.

Context
The complexity and frequency of surgical procedures involving bone-preparation gestures (e.g., bone-cutting and the positioning of implants and prostheses) have given rise to an area of research that is of importance throughout orthopaedics. Since the 1990s, several assistive technologies have been developed to improve the accuracy of surgical procedures, with the aim of improving clinical and functional outcomes. Today, several imaging and optical navigation systems, robots, and 3-dimensional (3D)-printed patient-specific instrumentation...
Technologies are used in various fields, including knee and hip arthroplasty, spinal instrumentation, corrective osteotomies, bone tumor resections, and others. The geometrical parameters to evaluate the performance accuracy of these systems are defined according to patient-specific functional outcomes. Examples include measurement of the angular alignment of the leg in knee arthroplasty, the mechanical axis of the leg in corrective osteotomies, the safe margins in tumor resections, and the correction of scoliotic spine deformities. However, the measurement of these geometrical parameters does not permit a direct assessment of the accuracy of the actual surgical gesture in comparison with the desired gesture. Therefore, it is not yet possible to state unequivocally that improvements in the accuracy of the assisted surgical gestures substantially impact the outcomes of surgical procedures. Doubt is still cast on the real benefits of surgical-assistance technologies, hampering their integration and use in clinical routine.

In 2004, the International Society for Computer Assisted Orthopaedic Surgery and the American Society for Testing and Materials undertook the creation of a new standard for assessing the performance accuracy of surgical-assistance technologies. This standard was published in 2010 as F2554-10, using definitions of accuracy from the E177-08 standard. The F2554-10 standard was recently used to evaluate the positional accuracy of several knee systems.

ISO Methodology to Study the Accuracy of a Surgical Gesture

The ISO methodology set describes preclinical testing of new assistive technologies in order to assess improvements in the accuracy of bone preparation (e.g., bone-cutting and the positioning of implants and prostheses) that is performed using new navigation systems, robots, and PSI. The methodology is based on 3 questions: (1) How is accuracy defined? (2) How can accuracy be measured experimentally? (3) How should accuracy be analyzed from a clinical perspective?

How Is Accuracy Defined?

Objective definitions of the desired performance accuracy of surgical procedures make it possible to determine the target value for a surgical procedure and the desired level of accuracy. Target-value definitions and accuracy-measurement methods need to be stated independently of the assistive technologies that are to be used during the surgery to assist the execution of the surgical gestures.

Multifactorial Accuracy

Surgical accuracy is multifactorial. It depends on 3 key factors: (1) Accuracy depends on the level of assistance integrated into the surgical procedure; therefore, it is mostly determined by the technologies (e.g., imaging, navigation, robotics, or PSI) and by the surgical tools and implants that are used. (2) Accuracy depends on the surgeon’s experience (i.e., the number of years in practice, the surgeon’s training, etc.). (3) Accuracy depends on local difficulties specific to the intervention, such as the anatomical site (e.g., the knee, the hip, the spine, the pelvis, etc.), unique patient anatomical factors, and the surgical approach (e.g., anterior, posterior, lateral, etc.).

Trueness, Precision, and Surgical Target

The ISO 5725-1 standard for measurement methods defines the terms “accuracy,” “trueness,” and “precision.” Applying these concepts to orthopaedic surgery, accuracy can be defined as the closeness of agreement between an achieved surgical gesture and the desired surgical gesture, namely the surgical target. When applied to a series of surgical gestures executed under stipulated conditions, accuracy involves a combination of a systematic error or bias component (trueness) and a randomness component (precision). Therefore, trueness can be defined as how close the agreement is between the mean value obtained from a large series of executed surgical maneuvers and the surgical target. Precision can be defined as how close the agreement is between independent surgical gestures obtained under stipulated conditions. As illustrated in Figures 1-A and 1-B, 4 situations can be envisaged, depending on whether the data set observed has low or high variability and whether the mean is close to or far from the target value.

How Can Accuracy Be Measured Experimentally?

Experimental research protocols enable us to gather preclinical, quantitative, objective data on the accuracy of bone-preparation gestures, while taking into account the influencing factors, which include the surgeon’s experience, local difficulties associated with the anatomical site, and the technologies that may be used by the surgeons. When seeking to illustrate a method for investigating accuracy in preclinical settings, bone-tumor surgery represents a good educational example for several
reasons. This surgery is difficult, and it requires multiple 3D bone cuts around the tumor to resect it with safe margins. Its high potential for inaccuracy has motivated the development of assistive bone-cutting technologies. A large number of experiments have been carried out on synthetic bone models to investigate the effect of assistive technologies on bone-cutting accuracy. The first clinical cases involving optical navigation that were published describe tumor resection and graft-reconstruction gestures that were performed with the assistance of a direct navigation system for cutting with an oscillating saw. A publication on the integration of 3D-printing PSI technology in clinical practice is, to our knowledge, the first clinical study of patients (11 in total) with bone tumors who underwent operations in which PSI was used. Figure 2 describes the clinical use of a patient-specific surgical-guide technology in a case of pelvic bone tumor. In this clinical study, histopathological analysis of the resected specimens showed tumor-free bone-resection margins in all of the cases. Bone-cutting accuracy was computed by registering postoperative and preoperative computed tomography (CT) scans, and averaged 2.5 and −0.8 mm, respectively.

**Experimental Protocol and Factors Influencing Accuracy**

The experimental protocol that was set up to investigate bone-cutting accuracy in preclinical settings consists of 5 steps, which are illustrated in the context of bone tumors. First, the surgical target has a 10-mm safe margin (i.e., the desired surgical gesture is a bone cut located 10 mm from the tumor border). Second, success and failure are defined as a safe resection margin (bone cut made 10 mm from the tumor border) and an intralesional resection (bone cut made within the volume of the tumor), respectively. Third, the accuracy of the actual surgical procedure is defined as the error in millimeters with respect to the safe margin. For example, an accuracy of 5 mm means that the bone cut was made 15 mm from the tumor border, an accuracy of −5 mm means that the bone cut was made 5 mm from the tumor border, and an accuracy of <−10 mm means that the bone cut was made within the tumor volume. Fourth, the factors included in the research are the experience of the surgeons, the cutting technologies that were used, and any surgical difficulties (e.g., complete resection of a simulated tumor on the acetabulum using 4 cuts, including 2 iliac cuts, 1 cut in the pubis, and 1 cut in the ischium). The data set comprises 23 surgeons, each of whom carried out a tumor resection by means of 4 cuts using 3 different cutting technologies, for a total of 276 cut planes. Fifth, measurement of all of the cut planes was carried out mechanically using a mechanical measurement arm with micrometer resolution.
Impact of Assistive Technologies

The implementation of this experimental protocol enabled us to gather and analyze 276 measurements of bone-cutting error in relation to the desired 10-mm safe margin (Fig. 3). Statistical analyses show that “assistive technology” was the main factor that influenced the order of magnitude of the error in relation to the surgical target. The “surgeon’s experience” factor did not influence the order of magnitude of the error in relation to the surgical target (i.e., senior and junior surgeons carried out their bone cuts to the same performance standard for each of the 3 cutting techniques that were investigated).

The effect of the assistive technologies on the accuracy of bone-cutting in terms of trueness and precision is shown in Figure 3. The level of trueness is the same because the mean of the observed data is identical for the 3 cutting techniques (freehand, navigation, and PSI). Moreover, this level of trueness is correct because it corresponds to the surgical target (i.e., the 10-mm safe margin). The level of precision, however, is better with navigation technology, and better still when PSI is used. In fact, the variability in the observed data is substantially reduced with navigation and PSI techniques when compared with the freehand technique. This suggests that assistive technologies allow surgeons to better control the sources of variability and inaccuracy when making bone cuts.

How Should Accuracy Be Analyzed from a Clinical Perspective?

Even if improvements in accuracy in terms of precision are observed experimentally when using navigation and PSI technologies, is this enough to have confidence in the use of these technologies in clinical routine? The clinical relevance of improvements in accuracy observed experimentally may be questioned by analyzing the effect on the risk of failure as defined in the experimental protocol (i.e., the risk of intralesional resection). With the freehand-cutting technique, 5 intralesional resections were observed in 92 cases, yielding a failure rate of about 5%. This failure rate, although observed experimentally in synthetic bone models under ideal conditions (perfect tumor accessibility, no bleeding, no soft tissues, etc.), is sufficient to conclude that the risk of intralesional resection in clinical practice is certainly not negligible. With use of the navigation and PSI cutting techniques, however, no intralesional resections have been observed. The observed failure rate is therefore 0%. Consequently, this result allows...
us to draw relevant conclusions regarding the risk of intralesional resection when using navigation and PSI technologies in clinical practice.

Tolerance with Respect to the Surgical Target

The surgical target (the desired safe margin) is defined by the surgeon based on his or her experience and on the specific characteristics of the tumor that will be resected (e.g., type, volume, and site). These characteristics define the accepted tolerance in relation to the surgical target (e.g., the safety zone around the surgical target). For bone-tumor surgery, the safety zone can be defined as shown in Figure 4. The surgical target is the safe margin of 10 mm. The safety zone may be defined as extending from 5 to 15 mm, considering that the tumor is below 5 mm and there are vessels, nerves, organs, and other structures above 15 mm.

Therefore, a bone cut made at the desired safe margin of 10 mm is regarded as a success. When a bone cut is made outside of the tolerance range (i.e., outside of the safety zone), it is considered a failure. If a bone cut is made within the safety zone, it is not considered a failure; however, the farther away it is from the surgical target, the less successful it is considered to be.

The clinically relevant question in the case of bone-tumor surgery is, “What would happen if the surgical target (the desired safe margin) was smaller, and therefore had a narrower tolerance (safety zone) because it was closer to the tumor border?” If there was a smaller desired safe margin, would there still be no observed intralesional resections when using navigation and PSI cutting techniques?

Estimation of the Risk of Failure

The risk of failure is a clinically relevant concept that is linked to the concepts of “accuracy” and “surgical target.” This naturally leads clinicians and researchers to ask themselves several questions when studying the accuracy of a surgical gesture.

What is the accepted risk of failure? What is the surgical target? What is the accepted tolerance in relation to the surgical target? In other words, what will happen if the surgical gesture is carried out farther from or nearer to the target? Also, what will happen if the surgical target is modified?

Figure 5 illustrates the risk estimate of intralesional resection based on 3 parameters: the desired safe margin, the surgeon’s experience, and the cutting technique. The curves in Figure 5 were derived by statistical estimations from the 3 experimental data sets observed for a desired safe margin of 10 mm when using the freehand, navigation, and PSI cutting techniques (as shown in Fig. 3). The method to statistically estimate the proportion of risk of intralesional resection has been described by Francq and Cartiaux.

Figure 5 illustrates the clinical relevance of improvements in accuracy that have been observed experimentally with navigation and PSI techniques. For a given level of acceptable risk for intralesional resection (e.g., 10%), the desired safe margin would need to be at least 10 to 11.5 mm for the freehand-cutting technique, while the navigation and PSI techniques allow us to set safe margins between 2 and 5 mm, which are substantially smaller. In other words, for bone-tumor surgery, the effect of navigation and PSI technologies is not only the improvement of the precision of bone-cutting, but also the ability to allow for desired safe margins that are substantially smaller when compared with that of the unassisted freehand-cutting technique. The clinical consequence is that the surgeons can, for example, preserve a joint.
What New Insights Have Been Gained into Accuracy?

The 4 key observations regarding ISO methodology in preclinical settings (illustrated in Fig. 6) are listed below.

1. **Accuracy is multifactorial.** It is crucial to take into account all of the factors that may influence the accuracy of a surgical gesture, including the surgeon’s experience, assistive technologies, and local difficulties associated with the anatomical site.

2. **For in vitro investigations, accuracy should be linked to the concepts of trueness (distance from the surgical target) and precision (variability in relation to the surgical target).**

3. **The clinical relevance of improvements in accuracy that have been observed experimentally may be evaluated by analyzing the impact on the estimated risk of failure and by taking into account the level of tolerance with respect to the surgical target (the extent of the safety zone).**

4. **The ISO methodology enables preclinical testing of new assistive technologies to quantify improvements in accuracy when using assistive technologies and to assess the benefits in terms of reducing the risk of failure and achieving surgical targets with tighter tolerances.**

The ISO methodology described above presents some limitations for use in clinical settings. First, the ISO methodology requires that an accepted reference (target) value be defined. The assessment of improvements in accuracy when using navigation systems, robots, and PSI can be limited when no agreed-upon clinically adequate target has been defined. For example, with knee and hip arthroplasty, achieving neutral mechanical alignment when implanting a knee prosthesis or observing a safety zone for acetabular cup orientation recently have been shown to not correlate with clinical success. Second, clinical outcomes are multifactorial. The quality of a surgical procedure cannot be solely limited to the accuracy of the bone-preparation gestures. Many additional factors also may influence the quality of a surgical treatment, including patient demographics and comorbidities, preoperative and postoperative-care protocols, and individual adherence to these protocols. Third, the definition of success and failure of a surgical treatment is not binary and cannot rely exclusively on the execution of the bone-preparation gestures. The success of a surgical treatment also relies on the patient’s satisfaction and functional outcomes after surgery. Finally, the definition of the safety zone is intrinsic to the clinical situation. The safety zone in other types of procedures (e.g., knee and hip arthroplasty) might differ considerably from the safety zone of bone-tumor resection as described above. In general, the safety zone is related to the concept of uncertainty (i.e., not knowing the target precisely) and is independent of the concept of accuracy of the bone-preparation gestures.

In conclusion, evaluation techniques to determine accuracy will need to consider the variety of technologies and systems that can be used during or after the intervention, not to execute but to measure the accuracy of the executed bone-preparation gestures (e.g., bone-cutting and the positioning of implants and prostheses). Some image-based measurement techniques could use radiographs or CT to perform image registration and bone segmentation. However, measurements based on radiographs are known to be subject to errors of magnification as well as rotation, and if CT-based measurements are performed, we need to deal carefully with acquisition parameters such as pixel size, slice thickness, and interslice distance, among others, which could influence image quality. Additional measurement techniques could use navigation systems and robots to perform intraoperative real-time tracking and tool localization. From an engineering perspective, a commonly accepted recommendation is to minimize the measuring errors by using measurement procedures and systems with an accuracy at least an order of magnitude greater than the errors expected during the execution of the bone-preparation gestures.

---

Olivier Cartiaux, PhD1
Jean-Yves Jenny, MD2
Leo Joskowicz, PhD3

1Computer Assisted and Robotic Surgery, Institut de Recherche Expérimentale et Clinique, Université Catholique de Louvain, Brussels, Belgium
2Centre de Chirurgie Orthopédique et de la Main, Les Hôpitaux Universitaires de Strasbourg, Illkirch, France
3The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

E-mail address for O. Cartiaux: olivier.cartiaux@uclouvain.be
References


