J Shoulder Elbow Surg (2018) 27, 276–281

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

# Rounded data have a high potential for false comparative statistical results as shown with elbow arc of motion

**Vilijam Zdravkovic, MD, Bernhard Jost, MD***

*Department of Orthopaedics and Traumatology, Kantonsspital St. Gallen, St. Gallen, Switzerland*

**Background and Hypothesis:** Visually measured range-of-motion (ROM) data are usually rounded to the nearest 5° interval and then recorded. Rounding might significantly influence the outcome of statistical tests.
**Methods:** We performed numerical simulation of *t* test application on 2 datasets, as typically reported for the elbow flexion-extension arc of motion. The test was performed on exact data and then repeated on the same data rounded to the nearest 5° interval. The simulation input parameters were as follows: difference in means (1°-30°), standard deviation (1°-30°), and number of cases (15, 30, 60, and 120). Diverging results were counted to find the rate of failure.
**Results:** Depending on the given difference in means, the given standard deviation, and the number of cases, the failure rate of the *t* test after rounding reached up to 40%.
**Discussion and Conclusion:** The accuracy of statistical tests performed on rounded ROM data is limited because of loss of information after rounding to the nearest 5° interval. This affects parametric and nonparametric tests, as well as paired and unpaired tests. In the future, authors should specify how ROM has been measured and recorded, explicitly addressing rounding. Furthermore, to test a zero hypothesis on rounded ROM data, authors should apply our *P* value (α) correction.
**Level of evidence:** Basic Science Study; Statistics and Measurement Error


**Keywords:** Range of motion; measurement error; arc of motion; elbow; statistics; rounding error; clinical research

Pain and reduced range of motion (ROM) are probably the most important symptoms of affected joints. Patients sometimes notice they are impeded from working and/or performing sports or even disturbed in their activities of daily living because of painful, reduced ROM. If major joints are involved, reduced ROM can be significantly disabling.[5,7] When orthopedic surgeons clinically detect reduced ROM, the aim of treatment is usually also to restore (nearly) normal ROM. In that respect, ROM is an important outcome measure, regularly reported in scientific publications in the field of orthopedic surgery. One of the main goals of many treatments is to restore, improve, or at least maintain ROM.

A standard, classic method to measure ROM (eg, elbow flexion and extension) is to use a goniometer and read the angle on the engraved scale with 1° precision (Fig. 1). As with

Institutional review board or ethical committee approval was not required for this basic science study.

*Reprint requests: Bernhard Jost, MD, Department of Orthopaedics and Traumatology, Kantonsspital St. Gallen, Rorschacherstrasse 95, CH-9007 St. Gallen, Switzerland.

E-mail address: bernhard.jost@kssg.ch (B. Jost).

1058-2746/$ - see front matter © 2017 Journal of Shoulder and Elbow Surgery Board of Trustees. All rights reserved.
https://doi.org/10.1016/j.jse.2017.10.033

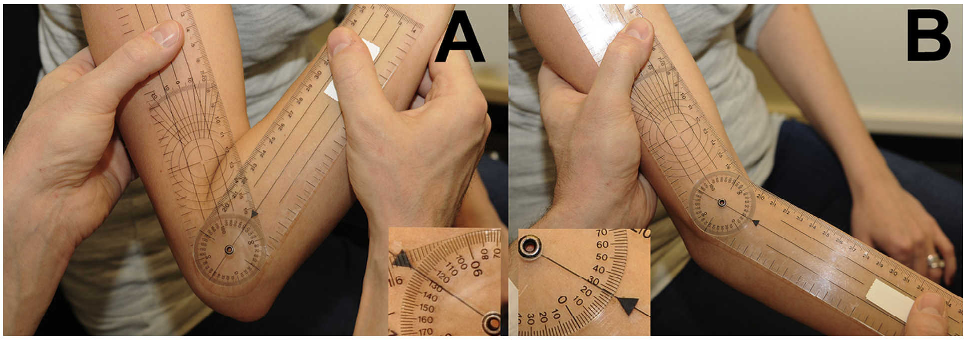


**Figure 1** Goniometric measurement of elbow extension (**A**) and elbow flexion (**B**). The difference after rounding is usually recorded as the flexion-extension arc.

**Table I** Hypothetical case

| | Preoperative | | Postoperative | |
|---|---|---|---|---|
| | Exact value (1° precision) | Rounded (at 5° interval) | Exact value (1° precision) | Rounded (at 5° interval) |
| Flexion, ° | 127 | 125 | 128 | 130 |
| Extension deficit, ° | 28 | 30 | 27 | 25 |
| Arc (1° precision) | 99 | | 101 (difference of 2° vs preoperative value) | |
| Arc (rounded at 5°) | | 95 | | 105 (difference of 10° vs preoperative value) |

Two-sided rounding at 5° intervals may potentially lead to overestimation of the difference between the preoperative and postoperative flexion-extension arc. In this hypothetical case, the 2° difference would certainly be regarded as negligible; however, the 10° difference in rounded values might appear to be a considerable—even clinically relevant—change.[1,2]

any analogue measurement, a certain margin of error is possible, depending on many factors, such as parallax, obesity of the patient, or experience of the person measuring the ROM. Because the measurement is not precise at the level of 1°, the usual practice is to round the values to 5° steps up or down. For practical purposes, those values seem precise enough to monitor the rehabilitation of a single patient during follow-up visits. It seems that rounding appears natural or logical so that authors usually do not report in the Methods section whether the data were rounded or not. However, when an arc of motion (eg, elbow flexion-extension) is being measured, the rounding on both sides could, possibly, introduce an error that appears to be clinically relevant (a hypothetical case is shown in Table I).

This error, henceforth referred to as "rounding error," is a mathematical error introduced by convention and is not a natural physical phenomenon like a measurement error. Such rounded ROM values, containing rounding errors, are used as study endpoints for research purposes as well. If one treatment option shows significantly better postoperative ROM, it will usually be considered better than the alternative. The usual method is to perform unpaired or paired *t* tests[1] on numerical ROM data. However, it is conceivable that rounding errors affect the mean values, difference in means, standard deviation, and consequently, the results of *t* tests or other parametric or nonparametric tests. To our knowledge, there are no reports about the potential effects of rounding errors on the interpretation and validity of inferential (comparative) statistics or zero-hypothesis tests in orthopedic surgery.

The purpose of this study was to explore the effect of rounding errors on the validity of statistical tests applied in testing the difference in ROM data rounded to the next 5° interval. In particular, we studied which input data (sample size, difference in means, standard deviation) might render a *t* test unreliable if applied to the elbow flexion-extension arc of motion when calculated from rounded data.

## Methods

We performed numerical simulations of the *t* test using R (R: A language and environment for statistical computing; R Foundation for Statistical Computing, Vienna, Austria; http://www.R-project.org/). Using random-number sequences, we generated normally distributed data samples for elbow arc of motion as usually reported.

We compared the outcomes of *t* tests performed on the same data samples before and after rounding to the next 5° interval. The *t* test was considered correct when the *P* value before rounding was in accordance with the *P* value after rounding (both significant or both nonsignificant at $\alpha$ level = .05). A failure of the *t* test was defined when there was no agreement between the *P* values before and after rounding. The simulation flowchart is presented in Figure 2. The