Acta Orthopaedica ›

Volume 81, 2010 - Issue 3

Open access

904 Views | 5 CrossRef citations to date | 1 Altmetric

Listen

Letter

The need for better analysis of observational studies in orthopedics

A retrospective study of elbow fractures in children

Per-Henrik Randsborg, Einar A Sivertsen, Inge Skråmm, J rat Šaltyt Benth & Pål Gulbrandsen

Pages 377-381 | Received 21 Sep 2009, Accepted 11 Jan 2010, Published online: 08 May 2010

Cite this article https://doi.org/10.3109/17453674.2010.487243

Abstract

Background and purpose The conventional statistical methods employed in observational studies in orthopedics require the fundamental assumption that the outcomes are independent. However, fractures treated by the same surgeon cannot be regarded as being independent of each other and should be nested in the statistical analysis. If the effect on outcome of early rather than delayed surgery depends on the severity of the fracture, we have a case of interaction. This is rarely considered in orthopedic research, but could affect the conclusions drawn. The aim of this paper is to describe the concepts of multilevel modeling and interaction in orthopedics.

Patients and methods In a cohort of 112 patients with single supracondylar humerus fractures, 78 patients were examined clinically on average 4 years after surgery. The

range of motion was measured and the global satisfaction was assessed. The results were used to compare traditional least-squares regression analysis with a 2-level model with interactions.

Results We found that 25% of the variance in outcome could be attributed to between-surgeon variance. We identified an interaction between the surgeons' experience and the severity of the fractures that influenced the conclusions. The variable “number of pins” was not significant in the 2-level model ($p = 0.07$), while the ordinary least-squares analysis gave a result that was statistically significant ($p = 0.01$).

Interpretation Researchers should consider the need for a 2-level model and the presence of interactions. Standard statistical methods might lead to wrong conclusions.

Supracondylar humerus fractures are the most common elbow injuries in children who require surgery. Percutanous pinning has become the method of choice in most clinics, and severe complications are rare (Otsuka and Kasser 1997). Recent reports have suggested that delay of surgery until the next day is safe (Iyengar et al. 1999, Mehlman et al. 2001, Leet et al. 2002, Gupta et al. 2004). However, vascular injuries and compartment syndromes still occur (Ramachandran et al. 2008), and some authors recommend treating these fractures as early as possible (Walmsley et al. 2006). The statistical methods employed in these publications are t-test, the chi-square test, and simple and multiple regression analyses. These methods require the fundamental assumption that the observed outcomes are independent, which implies that none of the fractures should have been operated by the same surgeon, and the distribution of fractures among surgeons should be random. This is rare, if not impossible, in observational studies on orthopedic surgery.

In the social sciences, there has been a gradual awareness of the inability of conventional statistical methods to analyze the complexity of human interaction. Researchers in public health have repeatedly encouraged the integration of social science methods in medical research (Singer and Ryff 2000, Office of Behavioral and Social Sciences Research 2001). For instance, we should not consider a priori that there is no correlation between outcomes of patients who are operated by the same surgeon. This is not simply a matter of controlling for the experience of the surgeon. Regardless

of the ability of the surgeon, the results for fractures treated by the same hands are likely to be more similar than if they were not. Ignoring such correlations may lead to underestimation of standard errors, increasing the risk of committing a type-I error with the conclusion that a variable is statistically significant when it is not. A multilevel approach (also called hierarchical modeling) accounts for potential correlations by modeling intercepts and regression coefficients as random. The intraclass correlation coefficient (ICC) expresses the amount of dependency among observations and is calculated to decide whether a multilevel analysis is appropriate. The ICC can take values from 0 to 1. A non-zero value of ICC implies that the observations are not uncorrelated and that there is a need for multilevel modeling.

Furthermore, orthopedic researchers should consider the existence of interaction, also known as effect modification (Moyé 2006). An interaction is defined as a factor that modifies the independent factor under study. This is analytically more complex than simple confounding. A confounder has the same effect on outcome for all values of the other independent variables studied. Interactions reflect that the effect of one variable depends on the values of one or more other variables. For example, the influence of the surgeon's experience on outcome could be stronger for severe fractures than for less complicated ones. In such cases, a statistical model with interactions should be tested.

Here we describe to the orthopedic community the concept of multilevel modeling and interactions as necessary statistical tools in observational studies. We show that the conventional statistical methods currently employed in retrospective reports may yield misleading results.

Patients and methods

The study protocol was approved by the Regional Ethics Committee on October 25, 2007 (registration number 1.2007.2093). The patient population consisted of all children who underwent reduction and pinning of a displaced supracondylar humerus fracture in our institution between 1999 and 2006. The patients were identified in our computer files and the medical records were examined. Patients treated with closed reduction without pinning (n = 14) and patients treated at different institutions (n = 6) were excluded, as well as 1 patient with bilateral fractures. We included 112 supracondylar fractures in the study (Table 1). The child's age, sex, preoperative soft

time to pin removal, any complications, and postoperative status were recorded. Surgeons were classified in two groups according to their level of experience: consultants and residents. If two surgeons performed an operation together, the most experienced was chosen as being responsible for treatment of that fracture.

Table 1. Demographic and fracture characteristics of 112 supracondylar humerus fractures in 112 children

Download CSV Display Table

Radiographic analysis

Standard AP and lateral radiographs were used. The Gartland classification of the lateral film was recorded by PHR (Gartland 1959). This classification has good intra- and interobserver reliability (Barton et al. 2001).

Clinical examination

All patients were given written information about the study and were invited to attend a follow-up clinic. 3 patients had moved abroad or to a different part of the country. 9 patients were lost to follow–up, and 22 patients refused to participate, leaving 78 patients available for clinical examination on average 4.3 (1.5–9) years after admission. The surgeon's experience was not revealed to the investigators examining the patients (PHR, EAS, and IS). The cubitus angle and range of motion were measured clinically using a goniometer. The Flynn criteria (Flynn et al. 1974) were used to assess the functional and cosmetic outcome. Neurovascular status was assessed clinically and the patients completed the Quick-DASH score (www.dash.iwh.on.ca 2009) with the help of their parents. The overall satisfaction of the end-result was assessed using a VAS score where zero was the worst imaginable result and 10 was the best (equivalent to the uninjured arm).

Statistics

Standard descriptive analyses were performed. The variances in Flynn score and the Quick-DASH score were very small and the distribution was much too skewed to allow further statistical analysis. However, the VAS score could be analyzed statistically, and

was the main outcome variable we used to demonstrate the difference between the statistical methods (Table 4).

Initially, an unconditional model (one without explanatory variables) was estimated to calculate the ICC, which can be interpreted as the percentage of variation in VAS score that could be explained by the surgeon level.

A hierarchical linear modeling approach with 2 levels (patient and surgeon) was chosen to model the main outcome variable VAS score. Even though the VAS score was a highly skewed variable, a 2-level linear model (Fitzmaurice et al. 2004) was estimated. Simple analyses were then used to assess the association between the outcome variable and the main predictors. 4 variables were considered to be main predictors for the VAS score. The Gartland classification and the surgeon's level of experience (either consultant or resident) were considered to be predictor variables based on the notions that more displaced fractures have more complications and that more experienced surgeons have better results. The variable night (between midnight and 8 a.m.) or day was selected based on the tendency to perform fewer operations during the night to reduce the risk of adverse outcome (Rothschild et al. 2009). There is no consensus on the number of pins needed for adequate fixation. Some authors recommend the use of 3 pins in unstable fractures (Gordon et al. 2001, Vlahovic and Bumci 2002). We therefore included the number of pins as a predictor variable in the model. Time to surgery, preoperative neurovascular status, age, and sex were considered as confounders.

Technically, to account for interactions a new variable, which is the product of 2 (usually) independent variables, is introduced into the model. One should not try out all possible interactions since this increases the risk of spurious results. Thus, when building the multiple model, we only considered variables that we thought might produce an interaction. We decided that the surgeon's level of experience (consultant or resident) and time to surgery might have different effects on outcome depending on the severity of the fracture. We therefore included these 2 interaction terms in the original multiple model. The R^2, quantifying the proportion of outcome variation explained, was used for model reduction. Akaike's information criterion was applied for covariance model selection (Akaike 1974). Even though the hypothesis about normally distributed residuals could not be accepted, the qq-plot for residuals demonstrated an almost straight line with some small exceptions in the tails.

Results

There were no cases of permanent neurovascular injury, deep infection, or compartment syndrome. The retrospective review of the medical records of the 112 fracture patients revealed 21 postoperative complications (19%) (Table 2). Gartland 3 fractures were operated earlier and more often by consultants, and had more complications than Gartland 2 fractures.

Table 2. Postoperative complications recorded for the 112 fractures

Download CSV Display Table

11 of the 78 elbows available for clinical examination had a cubitus varus malunion. The mean VAS score was lower in patients who developed cubitus varus (7.7 vs. 9.6) than in patients who had a positive carrying angle ($p < 0.001$, Mann-Whitney test). 20 patients had a positive Quick-DASH score (Table 3).

Table 3. Clinical outcome of 78 fractures at follow-up, mean 4.3 years after surgery

Download CSV Display Table

The overall amount of variation in outcome explained by our model was 18%. By estimating the unconditional model, the ICC was calculated to be 0.25; i.e. 25% of the total variance in VAS score could be accounted for by the surgeon level. This suggests that an ordinary least-squares (OLS) regression analysis would probably yield misleading results. The simple analyses did not show any statistically significant relationships between the 4 predictors and VAS score. However, 10–26% of the explainable variation in outcome was explained by the predictors, which was sufficiently high to be able to include all considered predictors in the multiple model (Table 4).

Table 4. Differences in coefficients and corresponding p-values between an

interactions, and a two-level model with interactions. Age and one interaction (Gartland classification × time to surgery) were excluded from the model

Download CSV Display Table

In the OLS model without interactions, there was a positive correlation between number of pins and patient satisfaction (p = 0.02), while in the 2-level model with interaction terms included this correlation was not significant (p = 0.07). Neither the surgeon's experience nor the severity of the fracture influenced the VAS score significantly. The combination (interaction) of the two, with marginally significant effect on the outcome (p = 0.05), influenced the conclusions sufficiently to be included in the final model.

Although the 2-level model corrected the problem of underestimated standard errors in OLS regression, the results did not differ much in the 2 models. Only the number of pins was no longer statistically significant in the 2-level model. If we had used only a simple OLS model, we would have reported number of pins as an explanatory variable.

Discussion

Our point is not that the last word about treatment of supracondylar fractures has been said, but that simple statistical models do not suffice. We found that the clinical outcome after pinning of displaced supracondylar humerus fractures in children is good or excellent in most cases, which is consistent with previous reports (Foead et al. 2004, Devkota et al. 2008). The few complaints were mainly cosmetic, without loss of function. This is a difficult starting point when searching for factors to improve treatment, and we could only explain 18% of the variance in outcome. Even so, the use of statistical analyses that adhere more closely to reality than a simple OLS model changed conclusions. We have considered the most commonly used explanatory variables for inferior outcome and found that an interaction variable affected conclusions, and that as much as 25% of the explained variance could be attributed to the surgeon level. Hence, we argue that future observational studies in orthopedics must consider inclusion of interactions. Furthermore, since surgery is 2-level in nature (patient and surgeon), the need for a 2-level model should always be explored by calculating the intraclass correlation coefficient (ICC). This is possible even in relatively

in published papers, so that readers can consider the clinical relevance of statistically significant results. Furthermore, findings in previous studies where an OLS model without interactions has been used—and where the conclusions have had an impact on clinical practice—should be reviewed using these statistical methods.

PHR initiated and designed the study, collected data, organized the follow-up clinic, examined patients, and wrote the manuscript. EAS and IS examined patients and reviewed the manuscript. JSB built the statistical models and revised the manuscript. PG contributed to the study design and to critical revision of manuscript.

We are grateful to Laila Øian and Inger Elisabeth Lia for their invaluable help in organizing the follow-up clinic. The study was supported by grants from the Sophies Minde Foundation.

No competing interests declared.

Related Research Data

Blind Pinning of Displaced Supracondylar Fractures of the Humerus in Children
Source: The Journal of Bone and Joint Surgery (American)
Early Versus Delayed Reduction and Pinning of Type III Displaced Supracondylar Fractures of the Humerus in Children: A Comparative Study
Source: Journal of Orthopaedic Trauma
Reliability of a Modified Gartland Classification of Supracondylar Humerus Fractures
Source: Journal of Pediatric Orthopaedics
Comparison of Two Methods of Percutaneous Pin Fixation in Displaced Supracondylar Fractures of the Humerus in Children
Source: Journal of Orthopaedic Surgery
Effect of Surgical Delay on Perioperative Complications and Need for Open Reduction in Supracondylar Humerus Fractures in Children

1. Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr 1974;19(6):716-23.

Web of Science ® | Google Scholar

2. Barton KL, Kaminsky CK, Green DW, Shean CJ, Kautz SM, Skaggs DL. Reliability of a modified Gartland classification of supracondylar humerus fractures. J Pediatr Orthop 2001;21(1):27-30.

PubMed | Web of Science ® | Google Scholar

3. Devkota P, Khan JA, Acharya BM, Pradhan NM, Mainali LP, Singh M, Shrestha SK, Rajbhandari AP. Outcome of supracondylar fractures of the humerus in children treated by closed reduction and percutaneous pinning. JNMA J Nepal Med Assoc 2008;47(170):66-70.

PubMed | Google Scholar

4. Fitzmaurice GM, Laird N, Ware J. Applied longitudinal analysis. John Wiley & Sons, 2004.

Google Scholar

5. Flynn JC, Matthews JG, Benoit RL. Blind pinning of displaced supracondylar fractures of the humerus in children. Sixteen years' experience with long-term follow-up. J Bone Joint Surg (Am) 1974;56(2):263-72.

PubMed | Web of Science ® | Google Scholar

6. Foead A, Penafort R, Saw A, Sengupta S. Comparison of two methods of percutaneous pin fixation in displaced supracondylar fractures of the humerus in children. J Orthop Surg (Hong Kong) 2004;12(1):76-82.

PubMed | Google Scholar

7. Gartland JJ. Management of supracondylar fractures of the humerus in children. Surg Gynecol Obstet 1959;109(2):145-54.

PubMed | Google Scholar

8. Gordon JE, Patton CM, Luhmann SJ, Bassett GS, Schoenecker PL. Fracture stability after pinning of displaced supracondylar distal humerus fractures in children. J Pediatr Orthop 2001;21(3):313-8.

PubMed | Web of Science ® | Google Scholar

9. Gupta N, Kay RM, Leitch K, Femino JD, Tolo VT, Skaggs DL. Effect of surgical delay on perioperative complications and need for open reduction in supracondylar humerus fractures in children. J Pediatr Orthop 2004;24(3):245-8.

PubMed | Web of Science ® | Google Scholar

10. Iyengar SR, Hoffinger SA, Townsend DR. Early versus delayed reduction and pinning of type III displaced supracondylar fractures of the humerus in children: a comparative study. J Orthop Trauma 1999;13(1):51-5.

PubMed | Web of Science ® | Google Scholar

11. Leet AI, Frisancho J, Ebramzadeh E. Delayed treatment of type 3 supracondylar humerus fractures in children. J Pediatr Orthop 2002;22(2):203-7.

PubMed | Web of Science ® | Google Scholar

12. Mehlman CT, Strub WM, Roy DR, Wall EJ, Crawford AH. The effect of surgical timing on the perioperative complications of treatment of supracondylar humeral fractures in children. J Bone Joint Surg (Am) 2001;83(3):323-7.

PubMed | Google Scholar

13. Moyé LA. Statistical reasoning in medicine. The intuitive p-value primer. Springer, 2006.

Google Scholar

14. Office of Behavioral and Social Sciences Research. Towards higher levels of analysis: Progress and promise in research on social and cultural dimensions of health. 2001. NIH Publication 21-5020. Ref Type: Report

Google Scholar

15. Otsuka NY, Kasser JR. Supracondylar Fractures of the Humerus in Children. J Am Acad Orthop Surg 1997;5(1):19-26.

PubMed | Google Scholar

16. Ramachandran M, Skaggs DL, Crawford HA, Eastwood DM, Lalonde FD, Vitale MG, Do TT, Kay RM. Delaying treatment of supracondylar fractures in children: has the pendulum swung too far? J Bone Joint Surg (Br) 2008;90(9):1228-33.

PubMed | Google Scholar

17. Rothschild JM, Keohane CA, Rogers S, Gardner R, Lipsitz SR, Salzberg CA, Yu T, Yoon CS, Williams DH, Wien MF, Czeisler CA, Bates DW, Landrigan CP. Risks of complications by attending physicians after performing nighttime procedures. JAMA 2009;302(14):1565-72.

PubMed | Web of Science ® | Google Scholar

18. Singer B, Ryff C. New Horizons in Health: An Integrative Approach. National Academy Press, Wahington DC 2000.

Google Scholar

19. The Dash Outcome Meassure. http://www.dash.iwh.on.ca/translate.htm, 2009.

Google Scholar

20. Vlahovic T, Bumci I. Biomechanical evaluation of the value of osteosynthesis in supracondylar fracture of the humerus using Kirschner pins in children. Eur J Pediatr Surg 2002;12(6):410-5.

PubMed | Web of Science ® | Google Scholar

21. Walmsley PJ, Kelly MB, Robb JE, Annan IH, Porter DE. Delay increases the need for open reduction of type-III supracondylar fractures of the humerus. J Bone Joint Surg (Br) 2006;88(4):528-ß30.

PubMed | Google Scholar

Related research

People also read | Recommended articles | Cited by 5

Wrist function recovers more rapidly after volar locked plating than after external fixation but the outcomes are similar after 1 year: A randomized study of 63...

Maria K T Wilcke et al.
Acta Orthopaedica
Published online: 1 Feb 2011

Numerical investigation of mechanical effects caused by various fixation positions on a new radius intramedullary nail

A. Çelik et al.
Computer Methods in Biomechanics and Biomedical Engineering
Published online: 20 May 2013

Low infection rates after 34,361 intramedullary nail operations in 55 low- and middle-income countries: Validation of the Surgical Implant Generation Network (S...

Sven Young et al.
Acta Orthopaedica
Published online: 9 Nov 2011

Information for

Authors

R&D professionals

Editors

Librarians

Societies

Open access

Overview

Open journals

Open Select

Dove Medical Press

F1000Research

Opportunities

Reprints and e-prints

Advertising solutions

Accelerated publication

Corporate access solutions

Help and information

Help and contact

Newsroom

All journals

Books

Keep up to date

Register to receive personalised research and resources by email

Sign me up

Copyright © 2026 Informa UK Limited Privacy policy

Cookies Terms & conditions Accessibility

Registered in England & Wales No. 01072954
5 Howick Place | London | SW1P 1WG

Taylor & Francis by informa

Article contents

Related research