

## Editorial: Sparse-data Bias—What the Savvy Reader Needs to Know

Seth S. Leopold MD, Raphaël Porcher PhD

It will surprise no one that the typical surgeon prefers surgery to statistics. But the well-read clinician must possess at least a passing familiarity with the pitfalls and shortcomings of common study designs in order not to be misled by what (s)he reads in orthopaedic journals.

---

The authors certify that neither they, nor any members of their immediate families, have any commercial associations (such as consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article.

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

The opinions expressed are those of the writers, and do not reflect the opinion or policy of *CORR*® or The Association of Bone and Joint Surgeons®.

---

S. S. Leopold, Editor-in-Chief, *Clinical Orthopaedics and Related Research*®, Philadelphia, PA, USA

R. Porcher, Center of Clinical Epidemiology, Assistance Publique-Hôpitaux de Paris, Hôtel-Dieu, Centre d'Epidémiologie Clinique; Team METHODS, Centre de Recherche en Epidémiologie et Statistiques (CRESS), Paris Descartes University; INSERM U1153, Paris, France

S.S. Leopold MD, *Clinical Orthopaedics and Related Research*®, 1600 Spruce Street, Philadelphia, PA 19013 USA, Email: sleopold@clinorthop.org

We occasionally cover the “must-know” topics on the editorial pages here, in particular focusing on newer themes as they come up, with the goal of having readers of *Clinical Orthopaedics and Related Research*® be the savviest readers in the specialty. Earlier editorials have helped readers interpret survivorship curves in the presence of competing risks (such as death and revision after reconstruction) [14], pointed readers to new tools in *CORR*® [10] that increase understanding of network meta-analyses [6, 8], and, perhaps most importantly to readers of clinical research, provided plain-language explanations about effect size and minimum clinically important differences (MCIDs) [11]. We have gone to particular lengths at *CORR* in the last several years to emphasize effect-size estimates over statistical estimates in the papers we publish, since clinicians think in terms of effect sizes and not p values when we care for patients.

On the topic of effect size, our consciousness recently has been raised to a problem called sparse-data bias, which can result in badly misleading effect-size estimates in clinical research. We credit the writers of a couple of letters to the editor in *CORR* [4, 5] for first bringing sparse-data bias to our attention, and we are glad to raise its profile further here. Although the problem was recognized nearly four decades ago [2, 3], and statistical manipulations to mitigate it have been available for quite some time [3, 7], a robust, clear approach for how to

identify and address sparse-data bias only recently has been promoted before a broad clinical audience [9].

A full description of the problem goes well beyond what we can provide here (and beyond what most readers probably want to know), though we do recommend that recent review [9] for the mathematically ambitious. Rather, here, we seek to focus attention on how to determine whether sparse-data bias may have influenced the results of an article one is reading, and what a reader should do if (s)he suspects it is present.

Some key “red flags” for sparse-data bias in a published paper are:

1. The presence of only a small number of events of interest for each clinical variable that is being studied, in particular if all (or nearly all) of the events fall into one study group and none (or nearly none) fall into the other.
2. High odds ratios, often with wide confidence intervals, that are out of line with reasonable expectations.
3. Statistical adjustments (like multivariable analysis) that adjust for confounding variables dramatically increase the estimated effect size (like the odds ratio) when one would expect the opposite to occur, as is more-often the case. Confounding variables typically increase apparent effect sizes, and so controlling for them usually pushes apparent effect-size estimates down, not up.

## Editorial

**Table 1.** Hypothetical study of admission to an ICU after total joint arthroplasty. The odds ratios are computed using usual formulas and adjusted odds ratios obtained by logistic regression.

Characteristic	Admitted to ICU (n = 10)	Not admitted to ICU (n = 90)	Odds ratio (95% CI)	Adjusted odds ratio (95% CI)
Male, number (%)	6 (60%)	27 (30%)	3.5 (0.91 to 13.4)	13.8 (1.6 to 120.0)
Current smoker, number (%)	3 (30%)	1(1%)	38.1 (3.5 to 416.4)	186.0 (9.2 to 3750)
General anesthesia, number (%)	10 (100%)	10 (11%)	Inf (0 to inf)	NA

ICU: intensive-care unit; Inf: infinity; NA: not used for adjustment.

Let's take a hypothetical example about the risk of admission to an intensive-care unit (ICU) following surgery. Imagine that charts were reviewed for 100 patients undergoing total joint arthroplasty to identify factors associated with ICU admission, and that 10 (10%) of these patients were admitted to the ICU for major complications (Table 1). Analyzing the association of sex with ICU admission does not seem problematic at first glance, but the limited number of patients admitted to the ICU makes the apparent association quite unstable. For example, if just one fewer male patient (or one more) was admitted to the ICU, that would change the odds ratio dramatically, from 3.5 to 2.3 or 5.4, respectively (not shown in the table). For the association with current smoking, the increased odds ratio for ICU admission of 38 seems unrealistically large, and the instability of the association persists; if even one more smoker had been in the non-ICU

group, the odds ratio would have decreased by half, to 19 (not shown in the table). Moreover, there is no sensible way to adjust for the association of sex by smoking status, since the only patient with a smoking history who was not transferred to the ICU could only have been either a man or a woman (in the example, she is a woman), but despite this, a logistic regression model will generate estimates. In that scenario, the sparse-data problem would go undetected, unless the reader is paying attention to the fact that these estimates are not logically sensible (since without both males and females in the analysis, it is not logical to analyze by sex). As for general anesthesia, since all patients transferred to the ICU received general anesthesia, usual methods make it impossible to estimate the odds ratio because of a zero in the denominator, resulting in an estimated odds ratio for that risk factor of infinity. Finally, one observes the two other "red-flag" signs for sparse data bias in

this example (Table 1): The adjusted odds ratios increase after adjusting for confounding variables (rather than decrease, as one would expect), and the 95% confidence intervals are extremely wide—two or three orders of magnitude wide—putting entirely unreasonable effect-size estimates within the range of possible values as defined by those confidence intervals.

The previous hypothetical may seem extreme. But in fact, a study similar to it was published in an orthopaedic journal [1] and was cited as an exemplar of sparse-data bias in the *BMJ* article mentioned earlier [9]. Looking at the risk factors that published orthopaedic study considered, we find that smoking, use of cement, and male sex had odds ratios for ICU admission of about 10, 1, and 2 prior to adjustment, which increased to 65, 56, and 4 after adjustment for confounding variables. Such a large increase after adjustment should alert the reader to a potential problem, since correcting for confounding variables

**Table 2.** Hypothetical example demonstrating complete and near-complete data separation

	Complete separation		Near-complete separation*		
	Event		Event		
Exposure	No	Yes	Exposure	No	Yes
Exposed	0	15	Exposed	0	15
Unexposed	20	0	Unexposed	18	2

\*Near-complete separation occurs when one of the four cells in a 2x2 table is zero. In cases of complete and near-complete data separation, sparse-data bias will cause usual methods of calculation to fail to estimate an accurate odds ratio and to provide meaningful CIs for other effect-size metrics (such as risk difference or relative risk).

## Editorial

usually makes effect-size estimates smaller rather than larger. Another concern should have been that the combining of those seemingly bland and very-common risk factors resulted in a risk estimate that departed radically from any sensible a priori expectation. Taken together, those three risk factors suggest that a man's odds increase for admission to the intensive-care unit after surgery were some 13,000 times greater if he ever smoked and underwent cemented arthroplasty, a finding that seems unlikely in the extreme. Unusually large confidence intervals—in that study, the risk factor of cement usage was bracketed by a confidence interval of 1.64 to 1894—also should raise a reader's antennae. As an important aside, we note that placing the value of an estimate between two and 2000 is not terribly informative.

We do not mention this to critique that article or that journal; we note that *CORR* published at least two articles last year [12, 13] where this same concern was identified after publication [4, 5]. In one of those [13], another hallmark of sparse-data bias was present: Nearly all instances of the event of interest were in one study group, while the other study group had none. “Always” and “never” are unusual in medicine, and that should be a tipoff to the careful reader. Effect-size estimates drawn from samples where there is this kind of near-complete separation of effects (Table 2) between study groups may be misleading. In the future, we will ask authors to use available

statistical techniques to try to mitigate the influence of sparse-data bias on effect-size estimates [9], and to discuss the issue in the “limitations” section of the Discussion.

We encourage readers to be attentive to this important and common problem in clinical research, and as editors, we plan to do likewise. In particular, we counsel caution in interpreting studies where some or all of the “red-flag” signs of sparse-data bias are present.

**Acknowledgment** We thank Erfan Ayubi PhD and Saeid Safiri PhD for bringing sparse-data bias to our attention with their letters to the editor.

### References

1. AbdelSalam H, Restrepo C, Tarity TD, Sangster W, Parvizi J. Predictors of intensive care unit admission after total joint arthroplasty. *J Arthroplasty*. 2012; 27:720–725.
2. Albert A, Anderson JA. On the existence of maximum likelihood estimates in logistic regression models. *Biometrika*. 1984;71:1–10.
3. Anderson JA, Richardson SC. Logistic discrimination and bias correction in maximum likelihood estimation. *Technometrics*. 1979;21:71–78.
4. Ayubi E, Safiri S. Letter to the editor: Increased risk of revision, reoperation, and implant constraint in TKA after multiligament knee surgery. *Clin Orthop Relat Res*. 2017;475:2610–2611.
5. Ayubi E, Safiri S. Letter to the editor: What injury mechanism and patterns of ligament status are associated with isolated coronoid, isolated radial head, and combined fractures? *Clin Orthop Relat Res*. [Published online ahead of print]. DOI: 10.1007/s11999-0000000000000220.
6. Chaudhry H, Foote CJ, Guyatt G, Thabane L, Furukawa TA, Petrisor B, Bhandari M. Network meta-analysis: Users' guide for surgeons part II – certainty. *Clin Orthop Relat Res*. 2015;473: 2172–2178.
7. Firth D. Bias reduction of maximum likelihood estimates [correction in: *Biometrika* 1995;82:667]. *Biometrika*. 1993; 80:27–38.
8. Foote CJ, Chaudhry H, Bhandari M, Thabane L, Furukawa TA, Petrisor B, Guyatt G. Network meta-analysis: Users' guide for surgeons part I – credibility. *Clin Orthop Relat Res*. 2015;473: 2166–2171.
9. Greenland S, Mansournia MA, Altman DG. Sparse-data bias: A problem hiding in plain sight. *BMJ*. 2016;353:i1981.
10. Leopold SS. Editorial: “Pencil and paper” research? Network meta-analysis and other study designs that do not enroll patients. *Clin Orthop Relat Res*. 2015; 473:2163–2165.
11. Leopold SS, Porcher R. Editorial: The minimum clinically important difference—The least we can do. *Clin Orthop Relat Res*. 2017;475:929–932.
12. Pancio SI, Sousa PL, Krych AJ, Abdel MP, Levy BA, Dahm DL, Stuart MJ. Increased risk of revision, reoperation, and implant constraint in TKA after multiligament knee surgery. *Clin Orthop Relat Res*. 2017;475:1618–1626.
13. Rhyou IH, Lee JH, Kim KC, Ahn KB, Moon SC, Kim HJ, Lee JH. What injury mechanism and patterns of ligament status are associated with isolated coronoid, isolated radial head, and combined fractures? *Clin Orthop Relat Res*. 2017;475: 2308–2315.
14. Wongworawat MD, Dobbs MB, Gebhardt MC, Gioe TJ, Leopold SS, Manner PA, Rimnac CM, Porcher R. Editorial: Estimating survivorship in the face of competing risks. *Clin Orthop Relat Res*. 2015;473:1173–1176.