

## The power of negative thinking

Jacques Lee, MD, MSc

SEE ALSO PAGE 327.

In this issue of *CJEM*, Frank and colleagues<sup>1</sup> report a randomized trial of magnesium administered intravenously (IV) for benign headache. They found that IV magnesium was no more effective than placebo, and had more side effects. Readers may notice that this article is unusual: It is uncommon these days to find an article reporting negative results. Journal editors shy away from negative trials, and this tendency gives rise to one of the significant problems with modern medical literature — publication bias. Emergency physicians were among the first to quantify “publication bias.” Moscatti and colleagues<sup>2</sup> found that only 15% to 16% of articles in leading emergency medicine and internal medicine journals reported negative results. Callahan and associates<sup>3</sup> further demonstrated that positive abstracts were twice as likely to be accepted to an emergency medicine annual scientific meeting, and to be subsequently published.

The second unusual aspect of Frank and colleagues’ article<sup>1</sup> is that, with only 42 subjects, it is exceedingly small compared to today’s 5-figure pharmaceutical mega-trials. So why would *CJEM* decide to publish an article that, in addition to being negative, is “small”? Especially when there is pressure on editors to reject as many articles as possible. Esteemed journals such as *JAMA* are quite proud of the fact they reject over 91% of submitted articles.<sup>4</sup> Not only do high rejection rates increase the apparent prestige and power of established journals,<sup>5</sup> they are one of the criteria used by academic and indexing agencies to assess the quality of a journal. Despite a small sample size, negative outcome, and pressures to reject, the *CJEM* editors thought that publishing this article was the right decision, both

from an ethical and scientific perspective.

Here’s why.

### Size matters, but how small is too small?

Before launching a research trial, investigators must determine how many subjects are required to answer their research question. Like all good investigators, Frank and colleagues<sup>1</sup> performed an *a priori* sample size calculation, and they used the following assumptions: 1) they would accept a 5% chance of making a type I error and a 20% chance of a type II error; 2) they were interested in detecting a 30% relative difference in successful response between IV magnesium and placebo groups (the “signal” or minimum clinically important difference [MCID]); and 3) based on previous studies, they expected 30% of patients given placebo to successfully respond (background noise). Given these assumptions, their calculated sample size was 96 patients. Since this represented less than 20% of the 500 cases of benign headache cases they treat each year, they anticipated no difficulty in reaching this target, nor did their local institutional review board. Although these expectations may have been reasonable, it comes as no surprise to those of us who have tried to enrol patients in a busy emergency department that even the most reasonable expectations are frequently disappointed. At the end of 16 months, they had only enrolled 42 subjects, and doubted they could reach their planned sample size with the resources available.

Now what?

Rather than discard the data already collected, they conducted an interim analysis that showed no apparent benefit

Assistant Professor, University of Toronto, Toronto, Ont., and Associate Researcher, Clinical Epidemiology Unit, Sunnybrook and Women’s Research Institute, Toronto, Ont.; Associate Editor, EM Advances section, *CJEM*

*Can J Emerg Med* 2004;6(5):359-61

for the intervention. Recognizing the possibility of committing a type II error, they assessed study power. With 42 subjects, their study had 53% power to detect a difference of 15 mm in visual analog scale (VAS) pain severity scores between groups, 80% power to detect a difference of 21 mm, and 98% power to detect a difference of 30 mm.

This case illustrates the problem with the traditional approach to power in the medical literature. Power is often treated if it were an “all or none” phenomena — a study either has adequate power or it doesn't. In fact, any given study has a spectrum of power, depending on the effect size deemed important. Deciding whether a study has “enough” power depends on what the reader defines as the MCID. Frank and colleagues' study tells us that if IV magnesium has any analgesic effect, it is unlikely to reduce pain severity by more than 21 mm compared to placebo.

But what is the “right” value for MCIDs, and where do the different values come from?

### **Clinically important differences in pain measurement**

Todd and coworkers<sup>6</sup> reported that a 15-mm reduction in pain severity is the smallest change that can be reliably detected, corresponding to patients' perception of “a little less pain.” This has become an accepted standard for MCIDs in emergency-based analgesic trials. However, given that many safe and effective treatments for acute migraine attacks already exist, is there a need to prove that IV magnesium has such modest analgesic effects? If IV magnesium only reduces pain to the extent that patients are able to perceive “a little less pain” (<15 mm on the VAS), would this make it a useful intervention? No. Rescue or adjuvant therapy would be needed in most or all cases. On the other hand, a 30-mm reduction in pain severity has been found to correlate with patients' perception of adequate analgesic therapy.<sup>7</sup> If IV magnesium produced such a response, it could be considered as a stand-alone treatment, and would warrant further investigation.

### **Best evidence?**

Wouldn't any self-respecting practitioner of evidence-based medicine disregard a study with only 42 subjects? In fact, No. Evidence-based medicine is not necessarily about “best evidence” but rather “best available” evidence. Frank and colleagues' study<sup>1</sup> is a meaningful addition to the existing literature base because of the poor quality of previous publications. Two uncontrolled studies and 1 pseudo-randomized, single-blinded study have shown 80% response

to IV magnesium.<sup>8–10</sup> One double-blinded randomized, placebo-controlled trial also showed impressive benefits of IV magnesium in a subgroup of 30 patients suffering migraine with aura.<sup>11</sup> Only one study that failed to show benefit of IV magnesium has been published, although magnesium was used in addition to metoclopramide in this study.<sup>12</sup> So an uncritical review of the prior literature would suggest that IV magnesium is beneficial in migraine and other benign headaches.

### **The cost of publication bias**

Why did most previous published studies suggest benefits of IV magnesium? Likely, because of publication bias. Most authors who complete a trial but do not find a statistically significant outcome simply shelve their projects.<sup>13–16</sup> In addition, journal editors and reviewers may be more likely to publish “positive” trials. Publication bias means that meta-analyses and systematic reviews will also be biased toward positive findings. In the absence of effective trial registries that track all relevant studies performed, it is impossible to know how many trials of IV magnesium were never completed, never submitted for publication, or rejected.<sup>17</sup> As a result, there is a preponderance of poor quality, positive studies published on the topic of IV magnesium in migraine.

Publication bias has the potential to harm our patients. Delays in reporting negative trials of class I antidysrhythmic conducted in the 1980s also delayed the dissemination of the knowledge that these agents tend to cause sudden death.<sup>17</sup> One author claims that this led to up as many as 75 000 unnecessary deaths per year in the US.<sup>18</sup> Although the stakes are not so high in the setting of migraine, Frank and colleagues' study<sup>1</sup> raises questions as to whether it is reasonable to enrol any more migraine patients in clinical trials of IV magnesium.

### **Ethics and the under-powered study**

One of *CJEM*'s reviewers who reviewed the paper by Frank and colleagues<sup>1</sup> questioned whether publishing a study that failed to meet its original sample-size requirement might encourage others to conduct unethical, under-powered studies. This is a serious charge, not to be taken lightly. Clearly, it is unethical to initiate a study without first doing a sample size calculation and assessing the feasibility of enrolling the required number of patients. Subjects participating in such a study would be exposed to potential risks without any possible scientific benefit accruing. But if a research team follows due diligence dur-

ing the planning of a study and is unable to meet their sample size requirements, is it ethical to discard the data collected up to that point? Any risk or inconvenience experienced by these subjects is guaranteed to have been in vain. Because of the potential to seriously bias the medical literature, there is an ethical duty to report negative results that add to the existing literature,<sup>19</sup> and a duty of journal editors to publish such results.

## Conclusion

Francis Bacon said that “the human spirit is more moved by the positive . . .”.\* In other words, humans are naturally more interested in things that work. This bias probably accounts for the preponderance of studies demonstrating benefit of IV magnesium for patients with headache. Readers of the medical literature have grown accustomed to huge megatrials that report small clinical differences, and some have developed the misconception that only large positive trials advance medical science. (The ethics of conducting overpowered trials is a topic for another editorial.) Frank and colleagues’ study illustrates that important knowledge can be gleaned from smaller negative trials. If IV magnesium is beneficial in the treatment of acute headache, the impact is minimal, and this is information we can use.

**Competing interests:** None declared.

## References

1. Frank LR, Olson CM, Shuler KB, Gharib SF. Intravenous magnesium for acute benign headache in the emergency department: a randomized double-blind placebo-controlled trial. *Can J Emerg Med* 2004;6(5):327-32.
2. Moscatti R, Jehle D, Ellis D, Fiorello A, Landi M. Positive-outcome bias: comparison of emergency medicine and general medicine literatures. *Acad Emerg Med* 1994;1(3):267-71.
3. Callahan ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *JAMA* 1998; 280(3):254-7.
4. Fontanarosa PB, Glass RM, DeAngelis CD. Thanking authors, peer reviewers, and readers — constancy in a time of change. *JAMA* 2000;283(15):2016-7.
5. Crandall R. Editorial responsibilities in manuscript review. *Behavioral Brain Sci* 1982;5:207-8.
6. Todd KH, Funk JP. The minimum clinically important difference in physician-assigned visual analog scores. *Acad Emerg Med* 1996;3:142-6.
7. Lee JS, Hobden E, Stiell IG, Wells GA. Clinically important change in the visual analog scale after adequate pain control. *Acad Emerg Med* 2003;10(10):1128-30.
8. Mauskop A, Altura BT, Cracco RQ, Altura BM. Intravenous magnesium sulphate relieves migraine attacks in patients with low serum ionized magnesium levels: a pilot study. *Clin Sci (Lond)* 1995;89(6):633-6.
9. Mauskop A, Altura BT, Cracco RQ, Altura BM. Intravenous magnesium sulfate rapidly alleviates headaches of various types. *Headache* 1996;36(3):154-60.
10. Demirkaya S, Vural O, Dora B, Topcuoglu MA. Efficacy of intravenous magnesium sulfate in the treatment of acute migraine attacks. *Headache* 2001;41(2):171-7.
11. Bigal ME, Bordini CA, Tepper SJ, Speciali JG. Intravenous magnesium sulphate in the acute treatment of migraine without aura and migraine with aura. A randomized, double-blind, placebo-controlled study. *Cephalalgia* 2002;22(5):345-53.
12. Corbo J, Esses D, Bijur PE, Iannaccone R, Gallagher EJ. Randomized clinical trial of intravenous magnesium sulfate as an adjunctive medication for emergency department treatment of migraine headache. *Ann Emerg Med* 2001;38(6):621-7.
13. Rosenthal R. The “file drawer problem” and tolerances of null results. *Psych Bull* 1979;86:638-41.
14. Dickersin K, Chan S, Chalmers TC, Sacks HS, Smith H Jr. Publication bias and clinical trials. *Control Clin Trials* 1987;8(4): 343-53.
15. Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997;9(1 Suppl):15-21.
16. Olson CM, Rennie D, Cook D, Dickersin K, Flanagan A, Hogan JW, et al. Publication bias in editorial decision making. *JAMA* 2002;287(21):2825-8.
17. Dickersin K, Rennie D. Registering clinical trials. *JAMA* 2003; 290(4):516-23.
18. Moore T. *Deadly medicine: why tens of thousands of heart patients died in America’s worst drug disaster*. New York: Simon & Shuster; 1995.
19. Chalmers I. Underreporting research is scientific misconduct. *JAMA* 1990;263(10):1405-8.

\*Thanks to Dr. Terry Klassen of the University of Alberta for this timely quotation.

**Correspondence to:** Dr. Jacques S. Lee, Emergency Medicine Research Program, Sunnybrook and Women’s College Health Sciences Centre, 2075 Bayview Ave., BG-04, Toronto ON M4N 3M5