4. Peckham C. *Medscape anesthesiologist lifestyle report 2017: race and ethnicity, bias and burnout* 2017. Available from: https://www.medscape.com/features/slideshow/lifestyle/2017/anesthesiology. [Accessed 7 July 2020]
5. Saha S, Beach M. Impact of physician race on patient decision-making and ratings of physicians: a randomized experiment using video vignettes. *J Gen Intern Med* 2020; **35**: 1084–91
6. Bustillo M, Gotian R. A mentoring circle supports women anaesthesiologists at every career stage. *Br J Anaesth* 2020; **124**: E190–1
7. Dobbin F , Kalev A. Why doesn't diversity training work? The challenge for industry and academia. *Anthropol Now* 2018; **10**: 48–55
8. Harvard University. *Project Implicit* 2011. Available from: https://implicit.harvard.edu/implicit/. [Accessed 4 July 2020]
9. Gotian R. *How to amplify the voice of your mentees* 2020. Available from: https://www.forbes.com/sites/ruthgotian/2020/06/15/how-to-amplify-the-voice-of-your-mentees/. [Accessed 4 July 2020]
10. Hewlett S, Marshall M, Bourgeois T. People suffer at work when they can't discuss the racial bias they face outside of it. *Harv Bus Rev* 2017. Available from: https://hbr.org/2017/07/people-suffer-at-work-when-they-cant-discuss-the-racial-bias-they-face-outside-of-it?utm_medium=email&utm_source=newsletter_weekly&utm_campaign=insider_activesubs&deliveryName=DM84811. [Accessed 4 July 2020]

doi: 10.1016/j.bja.2020.08.037
Advance Access Publication Date: 15 September 2020

Safeguarding anaesthesia research from spin

Rohan Magoon[1,2,*] and Jes Jose[1,2]

[1]Department of Cardiac Anaesthesia, Atal Bihari Vajpayee Institute of Medical Sciences (ABVIMS), Dr. Ram Manohar Lohia Hospital, New Delhi, India and [2]Department of Anaesthesia, Atal Bihari Vajpayee Institute of Medical Sciences (ABVIMS), Dr. Ram Manohar Lohia Hospital, New Delhi, India

*Corresponding author. E-mail: rohanmagoon21@gmail.com

Keywords: anesthesia research; bias; evidence-based practice; interpretation; randomised controlled trials; reproducibility; spin

Editor—Although perioperative physicians continue to debate propositions and oppositions to evidence-based practice in anaesthesia, the need for reflection upon the objectivity and accuracy of the reporting of individual RCTs is pivotal beyond any debate. The absence of robust reporting and understanding of the results emanating from these trials can substantially impact subsequent formulation of recommendations and guidelines for clinical practice. In this context, an in-depth assessment of RCTs has revealed the phenomenon of 'spin' being increasingly used in research reporting. 'Spin' in research is characterised by '…the manipulation of language to potentially mislead readers from the likely truth of the results….'[1] Given the broad latitude of language used for reporting studies, 'spin' continues to escalate albeit in a covert manner. Concerns manifest in anaesthesia and related specialities wherein misleading results of research influencing clinical practice can have far reaching implications for patients.

Despite initiatives such as checklists, structured instructions, authoring aids, and peer review, aimed at improving accuracy of research-reporting, there remain significant 'spin'ing tricks to distort the evidence for practice.[2] Spin often involves strategically achieving significance by pivoting to the secondary results, using within-group comparisons and invention of subgroups to suit the desired outcomes, which results in overinterpretation of the study results. Boutron and colleagues[3] have described various ways of misrepresentation of a study (from the title to the methods, results, and conclusion) used by authors to accomplish spin. The issues of particular relevance to clinical research are highlighted in Table 1.[3–5]

Evidence on the prevalence of spin is accumulating from diverse clinical fields.[6–9] Kinder and colleagues[7] scrutinised the abstracts of RCTs published in seven leading anaesthesia journals to delineate the extent of spin in anaesthesia-related research. They found that 23.2% abstracts displayed 'spin', with insinuation of treatment effectiveness based on a secondary outcome constituting the most common evidence of spin. With sample size of these RCTs premised on a defined primary outcome, reliance on a secondary endpoint is far from statistically robust. Another recent comprehensive analysis of 93 cardiovascular RCTs (published in six eminent high-impact cardiology and general medical journals) by Khan and colleagues[9] found that as high as 57% of abstracts and 67% of main texts revealed 'spin' in one or the other ways (mostly classified as low-level) as described in the nosology put forward by Boutron and colleagues.[3]

A major contributing concern hinges around the lack of reproducibility and transparency in research particularly when elucidation of details pertaining to methodology, protocols, analysis scripts, and raw data (in certain cases) can provide substantial justification to the conclusions limiting any chances of misinterpretation in the abbreviated forms described above. A recent description of the absence of

Table 1 Common strategies that contribute to 'spin' in research[3–5]

- Alteration in hypothesis or objectives to conform to the results
- Lack of distinction between the pre-specified and the *post hoc* analysis
- Lack of reporting of protocol deviations
- Selective focus on statistically significant results while disregarding those contradicting the hypothesis
- Interchanging 'intention-to-treat' analysis with 'per-protocol' analysis
- Results misinterpretation
  - ➢ Ignoring regression to the mean, small effect size and confounding
  - ➢ Interpreting a significant *P* value as an estimate of effect and statistical insignificance as indicative of safety or equivalence
  - ➢ Clinically ineffective but paradoxically concluded cost effectiveness in RCTs with 'doubly null' results
  - ➢ Unsupported causal language
  - ➢ Unexplored extrapolation to diverse clinically relevant settings
- Ignoring limitations of the study design and methods

analysis script availability statement in 99.2%, protocol availability statement in 96.7%, and raw data availability statements in 86.1% of publications in anaesthesia journals by Okonya and colleagues[2] speaks to the gaps in reproducibility and transparency in anaesthesia research reporting.

Incentives play an all-important in role in driving authors to resort to rhetorical publishing techniques and detract attention from the details so as to enhance the chances of high-impact publications, which in turn foster advancements in career and funding. The adage *publish or perish* bears testimony to this. A study on publication bias supports the notion that positive findings are likely to be accepted for publication in higher impact journals, thereby accounting for a natural tendency of authors to accentuate the positive aspects of their results.[10] Additional spin over and above the original report can be contributed by concomitant press releases, authors' institutional publicity, and funding agencies.[9]

A very interesting take on studying 'spin' in studies of 'spin' (SSSPIN) by Bero and colleagues[11] highlighted that the 'spin' in 'spin' studies is less prevalent than studies on other topics. The group proposes the requisite support and resources to launch effective evaluation of 'spin' in research to reduce avoidable low-quality research. In addition, 'spin' may de-escalate if due emphasis to publication of neutral results is recognised as an integral component of scientific discovery. As far as the crisis in inferential reproducibility is concerned, the concept of independent reviewers (with access to the methodology and raw data alone) formulating a version of their own discussion and conclusion, compared alongside the original version by the authors, to reduce the impact of the conflicts of interests and allegiance biases (common 'spin' motivators) constitutes a novel approach to discouraging spin, as proposed by Avidan and colleagues[12] and supported by Adam.[13]

Although the most robust editorial and peer review procedures need not always preclude the use of 'spin', development of prototypical systems for detecting spin is important to minimising the degree of 'spin'. These systems incorporate specific algorithms using text structure analysis, entity and relation extraction, classification of sentences, and assessment of semantic similarity. Moreover, context appropriateness of the interventions studied in RCTs adds inherent specificity to the results, avoiding overenthusiastic extrapolation.[14,15]

Application of rhetorical techniques in scientific publications is an area of active interest. Tuning our 'spin' detectors to the highest possible fidelity by endorsing an enhanced evaluation of the magnitude and types of 'spin' is certainly needed to safeguard and strengthen the research-evidence translational continuum as developments in the anaesthesia practice of tomorrow evolve.

Declarations of interest

The authors declare that they have no conflicts of interest.

References

1. Fletcher RH, Black B. "Spin" in scientific writing: scientific mischief and legal jeopardy. *Med Law* 2007; **26**: 511–25
2. Okonya O, Rorah D, Tritz D, Umberham B, Wiley M, Vassar M. Analysis of practices to promote reproducibility and transparency in anaesthesiology research. *Br J Anaesth* June 28 2020. https://doi.org/10.1016/j.bja.2020.03.035
3. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA* 2010; **303**: 2058–64
4. Fihn SD. Combating misrepresentation of research finding. *JAMA Netw Open* 2019; **2**, e192553
5. Raftery J, Williams HC, Clarke A, et al. 'Not clinically effective but cost-effective' — paradoxical conclusions in randomised controlled trials with 'doubly null' results: a cross-sectional study. *BMJ Open* 2020; **10**, e029596
6. Turrentine M. It's all how you "spin" it: interpretive bias in research findings in the obstetrics and gynecology literature. *Obstet Gynecol* 2017; **129**: 239–42
7. Kinder NC, Weaver MD, Wayant C, Vassar M. Presence of 'spin' in the abstracts and titles of anaesthesiology randomised controlled trials. *Br J Anaesth* 2019; **122**: e13–4
8. Ochodo EA, de Haan MC, Reitsma JB, Hooft L, Bossuyt PM, Leeflang MM. Overinterpretation and misreporting of diagnostic accuracy studies: evidence of "spin. *Radiology* 2013; **267**: 581–8
9. Khan MS, Lateef N, Siddiqi TJ, et al. Level and prevalence of spin in published cardiovascular randomised clinical trial reports with statistically nonsignificant primary outcomes, a systematic review. *JAMA Netw Open* 2019; **2**, e192622
10. Ioannidis JP, Munafo MR, Fusar-Poli P, Nosek BA, David SP. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends Cogn Sci* 2014; **18**: 235–41
11. Bero L, Chiu K, Grundy Q. The SSSPIN study-spin in studies of spin: meta-research analysis. *BMJ* 2019; **367**: 16202
12. Avidan MS, Ioannidis JPA, Mashour GA. Independent discussion sections for improving inferential reproducibility in published research. *Br J Anaesth* 2019; **122**: 413–20
13. Adam D. Reproducibility trial publishes two conclusions for one paper. *Nature* 2019; **570**: 16

14. Magoon R. Implications of practice variability: comment. *Anesthesiology* 2020; **133**: 943–4

15. Magoon R. Precision cardiac anesthesia: welcome aboard! *J Cardiothorac Vasc Anesth* 2020; **34**: 2551–2

doi: 10.1016/j.bja.2020.08.042
Advance Access Publication Date: 18 September 2020

How do anaesthetist and geriatrician perioperative frailty assessments compare?

Thomas Cope[1 *], Roisin Coary [2], Andrea Joughin [3], Philip Braude [3], David Shipway [3] and Kath Jenkins[1]

[1]Department of Anaesthesia, Southmead Hospital, Bristol, UK, [2]Department of Medicine for the Elderly, St James's Hospital, Dublin, Ireland and [3]Department of Geriatric Medicine, Southmead Hospital, Bristol, UK

*Corresponding author. E-mail: tomcope@doctors.org.uk

Keywords: clinical frailty score; emergency laparotomy; frailty assessment; geriatrics; perioperative outcome; risk assessment

Editor—Frailty is a syndrome characterised by loss of physiological reserve across multiple organ systems leading to vulnerability to homeostatic failure and organ dysfunction in the aftermath of a stressor event.[1] Frailty is associated with adverse perioperative outcomes in emergency general surgery.[2] National reports such as the National Confidential Enquiry into Patient Outcome and Death (NCEPOD) report, 'An Age Old Problem', have highlighted deficiencies in surgical pathways for this high-risk group and emphasise the need for frailty recognition as an independent marker of perioperative risk.[3]

The National Emergency Laparotomy Audit (NELA) is an initiative in England and Wales aiming to improve the care of patients undergoing emergency laparotomy through collection and publication of comparative data. In December 2018, NELA introduced a pre-admission Clinical Frailty Scale (CFS)[4] score into the dataset. In the emergency laparotomy setting, anaesthetists and surgeons are usually responsible for completing frailty assessments. It has been hypothesised that anaesthetists may be less familiar with frailty assessment than geriatricians, for whom frailty is at the core of their daily practice. In a recent survey of anaesthetist-delivered perioperative medicine services, only a fifth of respondents utilised frailty assessment tools in their clinical practice.[5]

Frailty assessment is increasingly used to guide clinical decision-making.[6] In the perioperative setting it can be used to inform decisions about appropriate levels of care, such as the need for postoperative intensive care admission. It is therefore imperative that clinicians undertaking frailty assessments are competent in the use of frailty assessment tools and understand their limitations. This study aimed to compare CFS scores assigned by anaesthetists and geriatricians and to evaluate anaesthetists' confidence in frailty assessment.

Patients presenting for emergency laparotomy between December 2018 and May 2019 at a large tertiary centre (Southmead Hospital, Bristol, UK) were assigned CFS scores preoperatively by an anaesthetist and postoperatively by a geriatrician. The geriatrician was blinded to the anaesthetist-assigned frailty score. CFS scores were assigned based on information relating to the patient's physical performance in the 2 weeks preceding admission to hospital. This information was gathered during face-to-face consultations with the patient, carer, or both by both the anaesthetist (preoperatively) and geriatrician (postoperatively).

An anonymous online survey was sent via email to all anaesthetic trainees (ST1–ST7), associate specialists, and consultants at our institution. Respondents were asked to score their confidence in assessing frailty; whether they have received formal teaching on frailty; their familiarity with frailty scoring systems, and how they thought their frailty assessment would compare with a geriatrician. The survey was designed through a team of anaesthetists and geriatricians and tested before use with colleagues not involved in its design. The frailty tools were suggested by the members of the team before the survey being sent.

Thirty-three patients were incorporated on the NELA database in the period analysed; nine were excluded, resulting in 24 patients included in the analysis. Reasons for exclusion included: no preoperative CFS score (8 patients), and one patient died after surgery before postoperative CFS.

CFS was the same in 58% of cases (n=14). The highest level of agreement between anaesthetist and geriatrician assigned CFS were in the non-frail cohorts (CFS 1–3). Anaesthetist and geriatrician assigned CFS differed in 42% (n=10); these results are displayed in Table 1. Where scores differed, anaesthetists were more likely to assign a higher frailty score than the geriatricians.

Of the 120 anaesthetists contacted, 35 responded to the survey (response rate, 25.7%), of which 70% were consultants. The mean score for confidence in assessing frailty was 4.7 out of 10 (1=not at all confident; 10=very confident); range 2–9. The 'timed up and go' test was the most recognised assessment of frailty, with 23 (65%) respondents reporting awareness of the test, and 51% (n=18) of respondents were aware of the