interventions provided as 'usual care' and the terminology used, comparing high to low quality randomised-controlled trials (RCTs).

Method Systematic review of RCTs (2006–16). Inclusion criteria were as follows: stroke survivor patients, intervention, rehabilitation, control: rehabilitative 'usual care', outcome: lower limb function. We used the Cochrane 'risk of bias' tool, rating studies as low or high quality. We identified the terminology used to describe the Control Group Rehabilitation Program (CGP), performed a knowledge synthesis process and conducted a frequency analysis to sort the heterogeneity through the itemised identification of the CGP contents. Two quality groups of studies (high-low risk of bias) have been compared.

Results We included 86 publications (23% low-risk of bias). Nine per cent of articles did not describe the CGP. In the remaining 78, we identified 64 interventions: 53 were proposed once, 8 were proposed twice and 3 were proposed in 3 papers. Two interventions (gait and balance) were proposed in 52% and 51% of papers, respectively. Results did not differ in the two RCT quality groups.

Conclusions This case study on stroke rehabilitation has shown that the term 'usual care' in CGP is not adequately used: a good terminology to define interventions is missing and respecting CONSORT checklists does not avoid these problems. These results are clear, and involve authors, reviewers and both field-specific and generalist journals. Nevertheless, they should be verified by future studies in other fields, with this or other methodologies. Reporting guidelines should probably give better expert guidance on this issue.

45 FREQUENCY OF SPIN REPORTING IN RANDOMIZED CONTROLLED TRIAL PUBLISHED BY INDIAN AUTHORS

[1]Naveen Kumar, [2]Hariohm Katheresapandian, [1]Saravan Kumar. *[1]Saveetha college of physiotherapy, Chennai, India; [2]Spring physiotherapy center, Chennai, India*

10.1136/bmjebm-2019-EBMLive.53

Objectives Reports that distorts the interpretation of results is referred as spin and it has been shown to be common in randomized controlled trials. Spin can be intentional and unintentional, and studies reported its presence in several specialties. It can mislead the health care practitioner in decision making and can affect wellbeing of patient. The objectives of this study were to estimate the frequency of spin in both abstract and main text in physiotherapy clinical trials published from India.

Method Pubmed was searched for randomized controlled trial published by Indian physiotherapists between January and November 2018. Articles included were parallel group randomized controlled trial with identified primary outcome as non-significant and published by Indian authors. Abstracts and full text were retrieved. Two authors independently scored each article for presence of spin in abstract and main text result, discussion and conclusion by using pre-tested strategies defined by Boutron and colleagues. Third author was contacted in case of conflicts. The frequency of spin was estimated in all studies

Results 129 relevant articles were identified in Pubmed search. In which 32 articles met the inclusion criteria. Spin was identified in results and conclusion section of the abstract of 27 (84%) and 30(93.8%) studies, respectively. In main text results, discussion and conclusion section spin was identied in 22(68%), 26(81.3%) and 29(90.6%) studies, respectively. Spin in title was present in one study.

Conclusions Strategies of spin is frequently included in reporting results in physiotherapy related trials published from India. Readers should be aware of possibilities of reporting bias and should avoid relying on abstract as a reliable report. Peer reviewers and editors need awareness of this issue to avoid the treatment recommendation that are not supported by data.

46 CHILDREN'S FITNESS AND HEALTH: AN EPIC SCANDAL OF POOR METHODOLOGY, INAPPROPRIATE STATISTICS, QUESTIONABLE EDITORIAL PRACTICES AND A GENERATION OF MISINFORMATION

Jo Welsman, Neil Armstrong. *University of Exeter, Exeter, UK*

10.1136/bmjebm-2019-EBMLive.54

Objectives Over 30 years ago we demonstrated the poor criterion validity of a popular fitness test, the 20 m shuttle run or 'bleep' test (20mSRT). We discounted the test and assumed that others with demonstrable validity and reliability would replace its use in research. Around then, our attention was drawn to an eloquent but obscure paper by JM Tanner (1949) which detailed the fallacy of simple division by body mass to accommodate body size differences in physiological function. Tanner described how incorrect analyses led to patients having 'no more formidable disease than statistical artefact'. Aware of the significance of this paper for our own field, over the next 15 years we published numerous data and tutorial papers demonstrating appropriate methods to measure and interpret cardiorespiratory fitness (CRF) during growth. Not only is the 20mSRT not a valid estimate of measured CRF, it predicts values expressed in simple ratio with body mass.

Method Despite our efforts, the past 10 years have seen a global explosion in published research studies of children's CRF anchored in these flawed methodologies. Data from millions of children worldwide have been collated into international norms, used to examine present and predict future cardiovascular and metabolic health, and to identify *individual* children who warrant intervention to reduce their risk of future cardiovascular disease – the raising of 'clinical red flags'. Data from these studies' present patterns of temporal changes in CRF which directly conflict with rigorously collected and appropriately analysed laboratory data. The 20mSRT test is being supported by international movements as a way of monitoring physical activity levels although objective data reveal the two to be unrelated. Moreover, clinical populations of children with serious life-limiting conditions are being put through maximal laboratory exercise testing with conclusions about their health status being made upon an inappropriate statistical analysis.

Results We believe the continued use of these flawed methodologies in vast numbers of children world-wide to be ethically and morally indefensible. By way of response we have, within the past 12 months: submitted 7 original data papers based upon extensive cross- sectional and longitudinal data founded on over 2000 rigorously determined individual assessments – all of which provide details of and recommendations for statistically justified analytical methods; we have submitted 7 editorial/commentary pieces to paediatric medical, sports

medicine and physical education journals, and written 2 responses to letters from those entrenched in poor methodologies. Despite our polite, transparent, scientifically-based pleas for 'constructive, collaborative debate' we have encountered editorial bias, e.g. turned down without review; turned down despite positive reviews; appealed editorial decisions and been prevented from responding to letters commenting

Conclusions Others have attempted to diminish our contributions by employing in letters a tone of thinly disguised hostility or accusing us of evangelistic fervour whilst failing to justify their own methods. Yes, we are challenging; shifting an entire research culture, which has its roots in university teaching, is not easy – scientific rigour in aspects of our discipline plays second fiddle to practical, convenient, traditional and feasible. Although this is happening on the periphery of mainstream medical research, children's health matters and as the population becomes increasingly sedentary and overweight we urgently need to develop scientifically rigorous methods to measure and interpret CRF in health and disease. Already a generation of researchers and policy makers has been misinformed and misled by flawed data. Those of us facing these challenges need to work together to develop strategies for shifting research culture back towards defensible science.

47 COLORECTAL CANCER SCREENING – DOING NO HARM? – A SYSTEMATIC REVIEW OF THE EVIDENCE (ONGOING)

[1,2]Frederik Martiny, [1,2]Anne Katrine Lykke Bie, [1,2]Christian Jauernik, [1,2]Or Rahbek, [1,2]Sigrid Brisson Nielsen, [1,2]John Brodersen. *[1]The Section of General Practice and the Research Unit for General Practice in Copenhagen, Copenhagen, Denmark; [2]The Research Unit for General Practice in Region Zealand, Koege, Denmark*

10.1136/bmjebm-2019-EBMLive.55

Objectives Screening for cancer intends to provide benefit and it unintentionally entails risk of causing harm. The type of harm most commonly studied is the risk of complications, i.e. physical harms. The risk of physical harm due to colorectal cancer screening has been quantified in systematic reviews. However, the harms of screening tend to be underreported and inadequately studied compared to the benefits of screening. In addition, the recently published PRISMA-harms extension raise concern that systematic reviews compound simplistic definitions, inadequate measurements and poor reporting in clinical studies of the harms of medical interventions. Due to these concerns, this systematic review aims to quantify the risk of physical harms due to colorectal cancer screening, to assess the quality of the evidence and to investigate the usability of guidance in the area, including the recommendations in the PRISMA-harms extension, the GRADE approach and the ROBINS-I bias assessment tool by the Cochrane Collaboration.

Method A detailed account of the methods is available in the PROSPERO protocol (CRD42017058844). The review was conducted in line with the PRISMA harms extension and the Cochrane Handbook. Databases searched include Pubmed, Medline, Embase, Cinahl, PsycInfo and the Cochrane library. There were no restrictions concerning study design, language or date of publication. All types of studies were included if they provided data on physical harms due to screening of asymptomatic individuals in average risk of colorectal cancer by any combination of fecal occult blood testing, sigmoidoscopy and/or colonoscopy. The risk of bias was assessed per outcome via ROBINS-I, a bias assessment tool for non-randomized studies. Prior to bias assessment, the tool was extended and adjusted to increase its applicability for studies reporting harms of screening. The overall quality of the evidence was evaluated using the GRADE criteria and the reporting quality of randomized controlled trials (RCTs) was judged via the CONSORT-harms extension.

Results Analyses are ongoing. Results will be presented at the conference. Preliminary trends: In total, 89 studies were included for review, 20 RCTs and 69 non-randomized studies (NRS) of various design. The majority of studies were conducted in Europe, the United States and in Asia. Both conduct and reporting of physical harms was very heterogeneous across RCTs and NRS. Overall, the risk of physical harms seems to be underreported, narrowly defined and inadequately measured in both RCTs and NRS. There is a trend towards higher estimates of harm in well-conducted studies (low risk of bias) and in studies of ongoing colorectal cancer screening programmes compared to RCTs.

Conclusions Pending analyses, the conclusions will be presented at the conference.

48 OCTOPUS: A REVOLUTION IN SCIENTIFIC PUBLISHING

Alexandra Freeman. *University of Cambridge, Cambridge, UK*

10.1136/bmjebm-2019-EBMLive.56

Objectives Octopus is a new publishing platform designed to replace journals and papers as the means of sharing scientific knowledge and ideas. It is designed to serve the needs of science and scientists above all else: to use every digital tool possible to ensure that good scientific practice is recognised and rewarded, and that there is no longer any advantage to questionable research practices.

Method Essential features of Octopus include:

- Complete language–agnosticism: every user reads and writes in their language of choice, maximising access.
- Free open access to read and publish
- The unit of publication is not a 'paper' but instead one of 8 shorter forms:

1. Scientific problem
2. Hypothesis
3. Method/protocol
4. Results/data
5. Analysis
6. Interpretation
7. Implementation
8. Review

Each publication must be linked to at least one 'above it' in the chain, apart from a Review, which can be linked to any other publication.

- **Rating of each publication** by readers (1-5 stars) on each of three predefined criteria, chosen to represent best practice in each kind of work => work is judged on the most appropriate criteria.
- **Red flagging of publications** by readers if they suspect misconduct (scientific, ethical, legal - eg. plagiarism) or serious error (eg. statistical).
- **Every author has a page listing their contact details, affiliations, potential conflicts of interest, and publications, with their ratings** (and who rated them).