Results For most depressed outpatients, sertraline, escitalopram or bupropion are reasonable first choices. If there is no response, the prescriber has many choices for the second trial in this algorithm because there is no clear preference based on evidence, and there are many individual patient considerations and variations in patient preference to take into account. The prescriber and patient may decide to either *switch* (to one of the above options not previously tried, or to venlafaxine, or to a nutraceutical antidepressant such as St. John's Wort or S-adenosylmethionine, or to transcranial magnetic stimulation), or to *augment* (with nutrients including l-methylfolate, or second-generation antipsychotics, or mirtazapine, or lithium or triiodothyronine). If there is no response to the second medication trial, the patient is considered to have a relatively medication-resistant depression. More recommendations follow. Comorbidities such as chronic pain, obsessive-compulsive disorder, attention-deficit hyperactivity disorder and posttraumatic stress disorder are considered.

Conclusions Utilization of this consultative tool for picking medication treatment for depression could help minimize unproductive variation in clinical care and improve clinical outcomes and produce remissions in shorter times and with fewer medication changes than with treatment as usual. Also, the algorithm encourages more cost-effective practice when generic options are recommended over expensive, brand-name products (when there is no apparent disadvantage in outcome or safety). Clinicians often overlook the large role their beliefs play in medication selection, ignoring placebo effects that contribute to their experience-based biases. These online tools provide evidence-supported ways of thinking that are available rapidly at the point of care in time to influence decision-making.

41 VALIDATION OF CROWDSOURCING FOR CITATION SCREENING IN SYSTEMATIC REVIEWS

[1,2]Nassr Nama, [3]Margaret Sampson, [4,5]Nick Barrowman, [3]Katie O'Hearn, [6]Ryan Sandarage, [4,3]Kusum Menon, [4,3]Gail Macartney, [4,3]Kimmo Murto, [4,3]Jean-Philippe Vaccani, [4,3]Sherri Katz, [4,3]Roger Zemek, [4,3]Ahmed Nasr, [4,3]James Dayre McNally. *[1]University of British Columbia, Faculty of Medicine, Vancouver, Canada; [2]BC Children's Hospital, Vancouver, Canada; [3]Children's Hospital of Eastern Ontario, Ottawa, Canada; [4]University of Ottawa, Faculty of Medicine, Ottawa, Canada; [5]Clinical Research Unit, CHEO Research Institute, Ottawa, Canada; [6]University of British Columbia, Faculty of Medicine, Vancouver, Canada*

10.1136/bmjebm-2019-EBMLive.49

Objectives Systematic reviews (SRs) are often cited as the highest level of evidence available as they involve the identification and synthesis of published studies on a topic. Unfortunately, it is increasingly challenging for small teams to complete SR procedures in a reasonable time period, given the exponential rise in the volume of primary literature. Crowdsourcing has been postulated as a potential solution. The feasibility objective of this study was to determine whether an online crowd would be willing to perform and complete abstract and full text screening. The validation objective was to assess the quality of the crowd's work, including retention of eligible citations (sensitivity) and work performed for the investigative team, defined as the percentage of citations excluded by the crowd.

Method We performed a prospective study evaluating the feasibility and validity of crowdsourcing essential components of an SR, including abstract screening, document retrieval, and full text assessment. Using the CrowdScreenSR citation screening software, 2323 articles from 6 SRs were available to an online crowd. Citations excluded by less than or equal to 75% of the crowd were moved forward for full text assessment. For the validation component, performance of the crowd was compared with citation review through the accepted, gold standard, trained expert approach.

Results Of 312 potential crowd members, 117 (37.5%) commenced abstract screening and 71 (22.8%) completed the minimum requirement of 50 citation assessments. The majority of participants were students (192/312, 61.5%). The crowd screened 16,988 abstracts (median: 8 per citation; IQR 7-8), and all citations achieved the minimum of 4 assessments after a median of 42 days (IQR 26-67). Crowd members retrieved 83.5% (774/927) of the articles that progressed to the full text phase. A total of 7604 full text assessments were completed (median: 7 per citation; IQR 3-11). Citations from all but 1 review achieved the minimum of 4 assessments after a median of 36 days (IQR 24-70). When complete crowd member agreement at both levels was required for exclusion, sensitivity was 100% (95%CI 97.9-100) and work performed was 68.3% (95%CI 66.4-70.1). Using the predefined alternative 75% exclusion threshold, sensitivity remained 100% and work performed increased to 72.9% (95%CI 71.0-74.6; $P<.001$).

Conclusions Crowdsourcing of citation screening for SRs is feasible and has reasonable sensitivity and specificity. By expediting the screening process, crowdsourcing could permit the investigative team to focus on more complex SR tasks. This requires a user-friendly online platform that allows research teams to crowdsource their reviews. Future directions should focus on assessing the application of this methodology to real life projects and determine its potential for rapid completion of systematic reviews.

42 INVESTIGATOR'S BROCHURES: DO THEY ADHERE TO BASIC PRINCIPLES OF EVIDENCE SYNTHESIS METHODS?

[1]Johannes Schwietering, [2,3]Merlin Bittlinger, [1-3]Daniel Strech. *[1]Hannover Medical School, Hannover, Germany; [2]Berlin Institute of Health, Berlin, Germany; [3]Charité Universitätsmedizin Berlin, Berlin, Germany*

10.1136/bmjebm-2019-EBMLive.50

Objectives The purpose of investigator's brochures (IB) is to compile the relevant evidence in order to enable an informed risk-benefit assessment by different reviewers including the principal investigator, a research ethics committees (REC), regulatory authorities, or data safety monitoring boards. Although a vast literature exists on the methodology of evidence synthesis for systematic reviews and meta-analyses, there is almost no literature examining the role of evidence synthesis in IBs. The primary objective of this contribution is to examine the adherence of IBs to fundamental principles of knowledge synthesis. These principles include a systematic search strategy, an evaluation of the risk of bias of the included studies and a comprehensive data collection and transparent synthesis procedure that accounts for variation in information quality.

Method We systematically examined a random sample of 30 IBs of a large sample (N = 109) from the application materials of industry-sponsored pharmaceutical clinical trials conducted between 2010–2016. IBs were obtained from three RECs of German university medical centers under data confidentiality agreements. Multiple independent examiners assessed the IBs to identify clinical trials reported in the IBs and

extracted relevant data about the reported clinical trials. For this assessment, a coding book was operationalized and specifically tailored to the context and role of IBs. The coding book included items on the reporting ofa) a search strategy to find relevant evidence b) a method of evaluating the risk of bias of included studies and c) a data collection and syntheses procedure. The coding book included a second set of items to evaluate the reporting quality of all clinical trials listed in the IB.

Results A total number of 321 clinical trials were identified in the 30 IBs. Despite the relatively similar aims of evidence reporting in IBs and systematic reviews, no IB reported a search strategy that could have demonstrated how comprehensive the reported clinical evidence is. Furthermore, no IB presented a risk of bias assessment for the reported studies. All of the included IBs reported the results of relevant clinical trials exclusively in a narrative form and did not use any method to synthesize the evidence to create a concise overview of the results. The majority of the 321 trials reported the sample size (96%, n = 308), blinding (83%, n = 266), and randomization (66%, n = 212). Specific details, however, about sample size calculation, randomization methods or blinding procedures were reported for less than 1% of the 321 studies reported in the IBs. Baseline characteristics and participant flow were reported for 10% (n = 31) of all 321 trials.

Conclusions IBs serve the important function to compile the evidence that shall justify the approval and conduct of a planned clinical trial. Our results in a random subsample of 30 IBs that reported on 321 clinical trials show that IBs do not adhere to any established principles of evidence synthesis known from systematic reviews. In sum, our findings cast doubts about the robustness and transparency of decision-making based on the evidence provided by IBs. In particular, it raises the question whether they enable the above-mentioned reviewers to conduct a meaningful risk-benefit assessment that is ethically and legally required to justify clinical trials. In the presentation, we will outline recommendations on how to improve evidence reporting in IBs.

43 GENERATING GUIDELINE RECOMMENDATIONS – EXPLORING PATIENT VALUES?

[1]Eve O'Toole, [1]Louise Murphy, [2]Leah McClimans. *[1]National Cancer Control Programme, Health Service Executive, Dublin, Ireland; [2]University College, Cork, Ireland*

10.1136/bmjebm-2019-EBMLive.51

Objectives The National Cancer Control Programme (NCCP) began developing national evidence based clinical guidelines in 2012. In 2017 we decided to generate our guideline recommendations using the GRADE process. There is clear guidance on appraising the quality of the evidence and performing bias assessments. However, the challenge begins with the integration of the best evidence, benefit and harm, patient values and resources when generating recommendations. Patients are full members of the guideline development groups; however, the question remains how do we truly capture patient values? By inviting a philosopher/medical ethicist to be a member of the GDG, integrating patient values has been explored in our update of the national clinical guideline diagnosis staging and treatment of patients with prostate cancer.

Methods A guideline development group was established with representatives from diagnostics, surgery, oncology, nursing, research, patient representatives and a philosopher/medical ethicist. The guideline questions were generated by the GDG. Literature searches, appraisal of the evidence and data extraction were performed by the NCCP research team. Generation of recommendations by the GDG follows a formal standardised protocol using an evidence to decision framework developed in NCCP. The following items are considered; the evidence, its quality its generalisability/directness, benefit and harm, patient values, resources and cost. The evidence statements are written in real time and recommendations graded strong or weak. During the first meeting to generate the guideline recommendations the philosopher challenged the group to consider the meaning behind words such as. generalisability, equity and justice. We also explored patient important values such as pain, certainty, anxiety, quality of life, trust, choice and autonomy.

Results Our evidence to decision framework has been restructured to contain a matrix to explore patient values for each option under discussion. New headings have been added with prompts to stimulate discussion with emphasis on patient important outcomes, equity, acceptability and a richer exploration on the concept of generalisation. Each evidence statement now has clear headings for quality of the evidence, benefit and harm, patient values and resources. Framing the discussion in this manner resulted in a richer conversation. For the first clinical question under review there was high level evidence, however the clinical members in the group were concerned that recommending a course of action could lead to greater inequity within our health system. This has resulted in escalation of the GDG recommendations within the health service with service redevelopment proposed.

Conclusions Involvement of the philosopher/medical ethicist led to the development of clear questions, prompts and a matrix, in the patient values domain. This resulted in the GDG having greater confidence in exploring patient values when generating guideline recommendations. This would challenge guideline developers to expand the membership of their GDG to include this expertise. Of most importance this led to a richer discussion and resulted in more consideration being given to a recommendation even in the presence of high-level evidence. As a result, there is earlier engagement with the services to enable implementation of the recommendations with fidelity to ensure we do not create further inequity within the health services.

44 OPENING THE BLACK BOX OF 'USUAL CARE' AND FINDING A BLACK HOLE: A NUMERICAL SYSTEMATIC REVIEW ON 'USUAL CARE' CONTROL GROUPS IN STROKE REHABILITATION RCTS

[1]Chiara Arienti, [2]Chiara Arienti, [2]Riccardo Buraschi, [2]Joel Pollet, [2]Massimiliano Gobbo. *[1]University of Brescia, Senago, Italy; [2]IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy*

10.1136/bmjebm-2019-EBMLive.52

Objectives Ethically, in every experimental trial new methods should be tested against the 'best' or, if not defined, the 'usual' care according to current Guidelines. To understand the latter, we focused on stroke rehabilitation as a case in point because stroke is a leading cause of disability and rehabilitation is a complex intervention whose need is increasing due to growth of chronicity and disability (World Health Organization). Our aim was to appraise stroke rehabilitation