RECRUS Research Newsletter
HPUPM HOSPITAL PENGAJAR UPM

Appraisals in Meta-Journal Hour 1

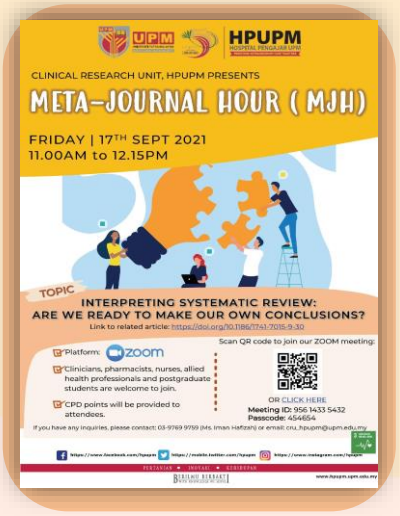
October 8, 2021
Vol. 1 Issues 8
Page 43

First published on 8 October 2021 by Nurul Iman Hafizah, Aazifah Ilham and Boon-How Chew

The Paper: Interpreting systematic reviews: are we ready to make our own conclusions? A cross-sectional study. BMC Med 2011 Mar 30;9:30. doi: 10.1186/1741-7015-9-30.

Why was this study conducted?

Appraisal of clinical evidence is an essential attribute in the practice of evidence-based medicine (EBM) for healthcare practitioners (HCPs). This includes the understanding of the evidence presented without being misled by the authors' conclusions. Despite there have been many efforts to promote the practice of EBM such as workshops and integration of EBM modules during undergraduate studies, it is still unclear on how ready the HCPs in practising EBM independently such as in interpreting clinical evidence in systematic reviews (SRs) which are highest ranked clinical evidence. Thus, this study aimed to assess how well a group of Malaysian hospital practitioners and medical students derived their own conclusions from SRs and to what extent these were influenced by their prior beliefs and the direction of the study results.



How was it done?

This cross-sectional study was conducted by a group of Malaysian clinician-researchers who were reported to be EBM teachers in a private medical school. There were two groups of participants: (1) Hospital HCPs (n=95) including **44 clinicians** and **51 allied health staffs** (AHS) consisted of nurses, pharmacist, nutritionist, physiotherapists and laboratory technicians recruited midway of an EBM course in June 2008, and (2) **Medical students** (n=35) recruited during the final six months of undergraduate medical training in November 2008. The participants were given **four trimmed Cochrane SR abstracts** consisting of background, objective and main results without the authors' conclusions. For each article, the authors highlighted one treatment outcome combination and included the corresponding forest plot with the abstract. Among the four SR abstracts, two chosen SRs demonstrated beneficial intervention ("positive"), and two others did not ("negative"). The four selected SR abstracts were **projected sequentially on a screen** for the HCPs to view, whilst the medical students received **printed copies** of the trimmed SRs to read. **Five minutes** were allocated for participants to read each trimmed SR abstracts. The participants were given an answer sheet and they were required to choose a conclusion from six options comprising different combinations of the **direction of effect** and the **strength of the evidence**:

Watch the video recording on:

YouTube UNIT PENYELIDIKAN KLINIKAL HPUPM/ UPM

Or click [HERE] and don't forget to subscribe to our channel!

A.	It is clearly beneficial and should be recommended as a possible treatment
B.	It is clearly non-beneficial and should not be recommended as a possible treatment
C.	It appears to be beneficial from limited evidence, more studies are needed to confirm the findings
D.	It appears to be non-beneficial from limited evidence, more studies are needed to confirm the findings
E.	There is insufficient evidence to comment on whether this intervention is beneficial or not
F.	I do not understand the results presented

The participants' chosen conclusions were then compared with the authors' predetermined best conclusions. The participants were also asked regarding their **prior beliefs about the intervention**. Cross-tabulations using **chi-square tests** were utilized to compare the differences of the categorical variables in this study.

UNIT PENYELIDIKAN KLINIKAL
Hospital Pengajar Universiti Putra Malaysia
Persiaran Mardi - UPM
43400 Serdang
Selangor Darul Ehsan
MALAYSIA

Untuk maklumat lanjut, sila lawati: **penyelidikan@hpupm**
03-9769 9763 / 9762 / 9761 / 9759
cru_hpupm@upm.edu.my

3 GOOD HEALTH AND WELL-BEING

What was the finding?

Overall, only **30.1%** of participants were able to identify the **most appropriate conclusion (correct direction of effect and strength of evidence)**. More students (48.2%) than practitioners (22.2%) chose the most appropriate conclusions ($P < 0.001$). Looking at the four SRs separately, higher proportions of participants identified the most appropriate conclusions in SR1 and SR3, that is, the **positive SRs** (39.2% and 36.4%, respectively), compared to SR2 and SR4, that is, the negative SRs (25.0% and 20.2%, respectively). Fewer than one-half (47%) correctly identified the direction of effect against their prior beliefs. "Positive" SRs were more likely than "negative" SRs to change the participants' beliefs about the effect of the intervention (RR 1.8, 95% CI 1.3 to 2.6) and "convert" those who were previously unsure by making them choose the appropriate direction of effect (RR 1.9, 95% CI 1.3 to 2.8).

How much can we take out from this research/paper?

Quite a good sample size of participants as a whole and fit the purpose of the hypothesis testing. A proper sample size estimation might be difficult without earlier studies in the similar setting but it could help in gauging a most economical sample size. Except the result on ability to interpret the most appropriate conclusions was stratified according to the different categories of participants, other results will have to be taken as indicating the whole group. On this note, the different approaches of the trimmed SRs in HCPs and medical students could exert differential effect on their performance. This added to the other uncertainties during the application such as discussion among the participants, referencing and Googling for references, etc.

The use of the 4 trimmed abstracts was unreal and its applicability to infer the same skill when reading other SRs is questionable when the full-texts are available. This might have caused the poorer performance on judging the strength of evidence. The same goes to the limit of time in assessing the abstracts. However unreal was the assessment methods, it might be the most feasible and sufficient method in judging the participants ability in interpreting the directions of effect, strengths of evidence and in reading forest plots for the treatment effectiveness. Additionally, the proficiency of participants on the (technical) language of the measures (answer sheet), relevancy of the study or the review topic to the clinical practice of some of the participants might be putting off and reduced their performance due to different backgrounds of practice.

Some of the constructs were not of sufficient clarity in their meaning and timing of measure such as the prior belief of the treatment/intervention effect, 'had seen or heard of the SRs' and 'perceptions of the value of the authors' conclusions'. These might cause them to be subjected to much personal experience of the participants and a poor measure across different participants.

The study was conducted among those who attended EBM workshops and by a mere logical deduction, the proportions of appropriate interpretation among other staffs could be lower as previous exposures and skills of EBM were not being assessed adequately. Another important question this study raises would be how the EBM workshops should be conducted to improve this skill, or what different workshop is needed. These questions will require an experimental study with a control group of people who never attended the EBM workshop or a qualitative study (approach?) to understand those who were skillful and those less skillful in interpreting SRs in order to understand the real reasons and possible causes from where interventions could improve on this. The skills in interpreting other types of clinical evidence are not included in the study [1].

References

1. Munn Z, Stern C, Aromataris E, et al. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol. 2018 Jan 10;18(1):5. doi: 10.1186/s12874-017-046.