Study protocol: A systematic review and re-evaluation of clinical practice laparoscopic guidelines using GRADE (Update: 16/07/2019)

J.Leung[1]

[1] University College London Medical Student, zchaokj@ucl.ac.uk

Background:

The application of laparoscopic surgery in different fields of surgery had increased significantly over the past years around the world. In the UK, patients receiving colorectal laparoscopic surgery increased from 10% to 28.4% between 2006-2008 (1). In Sweden, a total of 537,817 procedures were performed during 1998-2014 according to the National Patient Register (2). Consequently, more guidelines were being developed for different laparoscopic procedures. The guidelines had adopted different recommendation systems. Examples of recommendation systems include GRADE, Oxford methodology and Scottish Intercollegiate Guidelines Network (SIGN)(3-5). The different recommendation systems all have their own grading criteria and this could lead to potential confusion to people who wished to use the guidelines(6). The variations between guidelines also rendered a mix in quality that would make their recommendations inconsistent(7). In our pilot study on 'Differences in the recommendation of laparoscopic clinical practice guidelines according to the recommendation system– Re-evaluation using GRADE' we found out there were a wide variety of recommendation systems for laparoscopic guidelines. There was also substantial variation between recommendations from guidelines using Grading of Recommendations Assessment, Development and Evaluation (GRADE), Oxford Methodology, Scottish Intercollegiate Guidelines Network (SIGN) and other 'bespoke' systems when re-evaluated by GRADE. However, since a purposive sampling was only conducted, a further systematic review of laparoscopic guidelines was needed to assess the extent and find out the main reasons of the problem. The Appraisal of Guidelines for Research & Evaluation instrument (AGREE-II) Tool would also be used to assess the quality of the guidelines. This was done by validating its methodological rigour, editorial independence and transparency in which the guidelines were developed.. It includes 23 criteria within 6 domains (scope and purpose; stakeholder involvement; rigour of development; clarity of presentation; applicability; editorial independence)and 2 overall assessments(overall quality of guideline and recommendation of guideline for use)(8). The AGREE-II Tool's purpose is to address the problem of variation in quality of guidelines(9). This is done by providing a methodological strategy for the development of guidelines and suggesting what information is needed or reported in the guidelines(9).

Objectives:

The primary purpose of this study is to find out if the variation in agreement of recommendation of each guideline between guidelines' authors and regrading using available information is dependent upon the quality of the guidelines .

Methods/Design:

A systematic review of laparoscopic guidelines will be conducted on Pubmed and Embase. Searches will be made using free text and controlled vocabulary terms (MeSH). Guidelines published prior to 1st March 2016 will be excluded. There will be no language restriction. Guidelines not published in English will be translated using the following two options: 1. Translation of the text using translation software e.g. Google translate 2. Translation by authors who are fluent with the language. We will only look at the first level of research from the guidelines. These guidelines will then be reclassified in a similar manner as per other English guidelines. Table 1 detailed the search strategy for Pubmed and Embase respectively. CADTH search filters will be used(10).

Data extraction will be done independently by six authors and reclassification of statements will be conducted using GRADE and appraised using the AGREE-II Tool. The first author will reclassify and appraise all the statements while the other five co-authors will do so for a portion of the statements each. One of the co-authors will also run the searches independently from the first author to make sure no guidelines are missed. The interrater reliability (concordance of strengths of recommendation between authors of this study) will be conducted using both Cohen's Kappa and percent agreement. Cohen's Kappa is similar to correlation coefficient, it can range from -1 to +1, where 0 is the agreement expected from random chance and 1 represents the prefect agreement between authors vice versa, Cohen's Kappa will be conducted as it addresses the limitations of percent agreement, which overestimates the interrater reliability by assuming the majority is always correct(11). Percent agreement is the percentage of agreement in strong and weak recommendation that data collectors give to each statement, it also will be conducted as it addresses the limitations of Cohen's kappa, which underestimates the interrater reliability by taking into account random guesses made by data extractors(11). Any differences will then be resolved by discussion. The guidelines will be appraised using three AGREE-II Tool domains which are Rigour of Development, Clarity of Presentation and Editorial Independence respectively. Then 4 possible logistic regression models will be trialled. The four models are: 1. Three domains of AGREE-II without clustering by guideline

(simple logistic regression) 2. Three domains of AGREE-II with clustering by guideline (hierarchical logistic regression)3. Rigour domain only of AGREE-II without clustering by guideline (simple logistic regression)iv.4. Rigour domain only of AGREE-II with clustering by guideline (hierarchical logistic regression). Fit statistics will be interpreted and the best model will be used to find out the association between the variation and quality of guidelines(9).

Table 1 shows the search strategy for Pubmed and Embase (10)

Pubmed (Searched from 1st– 5th March)	(((laparoscop* OR celioscop* OR coelioscop* OR abdominoscop* OR peritoneoscop*) AND (Clinical pathway[mh] OR Clinical protocol[mh] OR Consensus[mh] OR Consensus development conferences as topic[mh] OR Critical pathways[mh] OR Guidelines as topic [Mesh:NoExp] OR Practice guidelines as topic[mh] OR Health planning guidelines[mh] OR guideline[pt] OR practice guideline[pt] OR consensus development conference[pt] OR consensus development conference, NIH[pt] OR position statement*[tiab] OR policy statement*[tiab] OR practice parameter*[tiab] OR best practice*[tiab] OR standards[ti] OR guideline[ti] OR guidelines[ti] OR ((practice[tiab] OR treatment*[tiab]) AND guideline*[tiab]) OR CPG[tiab] OR CPGs[tiab] OR consensus*[tiab] OR ((critical[tiab] OR clinical[tiab] OR practice[tiab]) AND (path[tiab] OR paths[tiab] OR pathway[tiab] OR pathways[tiab] OR protocol*[tiab])) OR recommendat*[ti] OR (care[tiab] AND (standard[tiab] OR path[tiab] OR paths[tiab] OR pathway[tiab] OR pathways[tiab] OR map[tiab] OR maps[tiab] OR plan[tiab] OR plans[tiab])) OR (algorithm*[tiab] AND (screening[tiab] OR examination[tiab] OR test[tiab] OR tested[tiab] OR testing[tiab] OR assessment*[tiab] OR diagnosis[tiab] OR diagnoses[tiab] OR diagnosed[tiab] OR diagnosing[tiab])) OR (algorithm*[tiab] AND (pharmacotherap*[tiab] OR chemotherap*[tiab] OR chemotreatment*[tiab] OR therap*[tiab] OR treatment*[tiab] OR intervention*[tiab])))
Embase (Searched from 5th -10th March)	1. (laparoscop* or celioscop* or coelioscop* or abdominoscop* or peritoneoscop*).af 2. exp laparoscopic surgery/ 3. 1 or 2 4. exp clinical pathway/ 5. exp clinical protocol/ 6. exp consensus/ 7. exp consensus development conference/ 8. exp consensus development conferences as topic/ 9. critical pathways/ 10. exp practice guideline/ 11. guidelines as topic/ 12. practice guidelines as topic/ 13. health planning guidelines/ 14. (position statement* or policy statement* or practice parameter* or best practice*).ti,ab,kw. 15. (standards or guideline or guidelines).ti,kw.

	16. ((practice or treatment* or clinical) adj guideline*).ab. 17. (CPG or CPGs).ti. 18. consensus*.ti,kw. 19. consensus*.ab. /freq=2 20. ((critical or clinical or practice) adj2 (path or paths or pathway or pathways or protocol*)).ti,ab,kw. 21. recommendat*.ti,kw. 22. (care adj2 (standard or path or paths or pathway or pathways or map or maps or plan or plans)).ti,ab,kw. 23. (algorithm* adj2 (screening or examination or test or tested or testing or assessment* or diagnosis or diagnoses or diagnosed or diagnosing)).ti,ab,kw. 24. (algorithm* adj2 (pharmacotherap* or chemotherap* or chemotreatment* or therap* or treatment* or intervention*)).ti,ab,kw. 25. Or/4-24 26. 3 and 25

References:

1.   Taylor EF, Thomas JD, Whitehouse LE, Quirke P, Jayne D, Finan PJ, et al. Population-based study of laparoscopic colorectal cancer surgery 2006-2008. Br J Surg. 2013;100(4):553-60.
2.   Sundbom M, Hedberg J. Use of Laparoscopy in Gastrointestinal Surgery in Sweden 1998-2014: A Nationwide Study. Scand J Surg. 2017;106(1):34-9.
3.   Grading quality of evidence and strength of recommendations. BMJ.328(7454), p.1490.
4.   https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/
5.   Grondin S SC. Evidence-Based Medicine: Levels of Evidence and Evaluation Systems. Difficult Decisions in Thoracic Surgery. 2010:13-22.
6.   Woolf S GR, Hutchinson A, Eccles M, Grimshaw J. . Clinical guidelines: Potential benefits, limitations, and harms of clinical guidelines. BMJ. 1999;318:(7182):.
7.   Abdelsattar ZM, Reames BN, Regenbogen SE, Hendren S, Wong SL. Critical evaluation of the scientific content in clinical practice guidelines. Cancer. 2015;121(5):783-9.
8.   Hoffmann-Esser W, Siering U, Neugebauer EA, Brockhaus AC, Lampert U, Eikermann M. Guideline appraisal with AGREE II: Systematic review of the current evidence on how users handle the 2 overall assessments. PLoS One. 2017;12(3):e0174831.
9.   Brouwers M KM, Browman GP, Cluzeau F, feder G, Fervers B, Hanna S, Makarski J AGREE II: Advancing guideline development, reporting and evaluation in healthcare. Can Med Assoc J. Dec 2010.
10.   https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters
11.   McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012;22(3):276-82.