

# Using GRADE to develop an evidence-based benchmark for patient safety indicators: postoperative venous thromboembolism

Bernard Burnand, Jean-Marie Januel

Institute of social and preventive medicine (IUMSP) and Cochrane Switzerland, Lausanne University Hospital, Switzerland

## RATIONALE

- Evidence-based benchmarks may help appraise the actual value of a quality indicator
- The GRADE system<sup>1</sup> may assist assessing the quality of evidence for such benchmarks
- However, GRADE is currently tailored to examine comparative designs

## OBJECTIVES

- To adapt the GRADE methodology for a rate or a proportion in the framework of a systematic review aimed at developing a benchmark for in-hospital symptomatic venous thromboembolism (VTE) in patients who had hip (TPHA) or knee (TPKA) arthroplasty<sup>2</sup>

## METHODS

- Evaluation of each component of GRADE to assess its suitability to appraise the evidence for a rate or a proportion.
- In this case study, we included randomized clinical trials (RCT) testing efficacious VTE prophylaxis and observational studies of patients receiving VTE prophylaxis. Symptomatic post-operative VTE (pulmonary embolism or deep vein thrombosis) that occurred before hospital discharge following TPHA or TPKA was the outcome.
- Two independent evaluation (BB and JMJ) of GRADE items (Table 1) were performed. Such evaluation have compared, discussed and conciliated of the two lists allowed its use in the systematic review.
- Ad hoc criteria to rate the quality of evidence were defined.

Table 1. GRADE Items at sub-group level

<b>Inconsistency of pooled estimates :</b>
- % of sub-groups that reported estimates with 3-fold less or more than the pooled mean estimate
- % of sub-groups that showed overlapping CI compared to the pooled mean estimate CI
- Across sub-groups heterogeneity: I <sup>2</sup> ≥50%
- A P-value for heterogeneity was <0.05
<b>Imprecision of pooled estimates :</b>
- % of studies with 95% CI around the estimate <10% (low imprecision)
- % of studies with 95% CI around estimate ranging from 10% to 20% (for imprecision)
- % of studies with 95% CI around estimate >20% (high imprecision)
<b>Generalizability to population of interest of pooled estimates :</b>
- % of sub-groups with no indirectness of evidence (external validity)
<b>Publication bias of pooled estimates :</b>
- Significantly asymmetric distribution of estimates across sub-groups? (based on Begg's test P-value less or more than 0.05)
- % of sub-groups with sparse data and attrition bias
- % of sub-groups with conflict of interest mentioned
<b>Limitation of pooled estimates :</b>
- % of sub-groups that presented at least one potential source of potential bias, including measurement bias, allocation concealment, and blinding

## RESULTS

- We considered GRADE elements about study design, outcome assessment, and sources of potential biases at study level and at sub-group of study level. At sub-group level criteria for evaluation are presented in Table 1
- We summarized the quality assessment of the included studies in five categories: consistency, imprecision, generalizability to the population of interest, publication bias and other limitations (allocation concealment, blinding, potential measurement bias), as exemplified for TPHA in Table 2<sup>2</sup>
- The individual sub-group and pooled estimates showed consistency, but large confidence intervals indicated lack of precision
- A potential measurement bias was present in <13% of RCT and between 67% and 75% of observational studies
- Indirectness of evidence varied largely between subgroups (80 to 93%)

Table 2. Summary of quality assessment of included studies<sup>1</sup> using GRADE methodology (adapted from Januel et al. JAMA 2012)<sup>2</sup>

# of sub-groups	Consistency	Imprecision	Generalizability to population of interest	Publication bias	Limitations
RCT – LMWH <sup>*</sup>	17% (4/23) of reported estimates were 3-fold less or more than the pooled mean estimate of the sub-group	Pooled estimate presented with 95% CI >20% of the pooled estimate mean value.	48% (11/23) had no indirectness of evidence	Begg's test: VTE : P=0.011 DVT: P=0.125 PE: P=0.004	22% (5/23) with no allocation concealment
	9% (2/23) showed overlaps CI compared to the pooled mean estimate CI	100% (23/23) presented 95% CI >20% of the estimate means value for each of them, respectively.		83% (19/23) presented sparse data	22% (5/23) with no blinding of patients/professionals
	I <sup>2</sup> was not ≥50% and P-value was <0.05			57% (13/23) presented attrition bias potentially	9% (2/23) presented potential measurement bias

\* RCT, Randomized Clinical Trial / LMWH, Low Molecular Weight Heparin

## CONCLUSION

- We propose a tentative adaptation of the GRADE system for a rate or proportion that should be further developed and assessed by the GRADE working group
- This adaptation of the GRADE system for evaluating the quality of the evidence supporting a quality benchmark, expressed as rate or a proportion was influenced by the type of evidence available, mostly RCT in this case, what may not always be the case when developing evidence-based benchmarks for quality indicators

REFERENCES 1. GRADE Working Group. www.gradeworkinggroup.org

2. Januel JM, Chen G, Ruffieux C, Quan H, Douketis JD, Crowther MA, Colin C, Ghali WA, Burnand B, for the IMECCHI Group. Symptomatic In-Hospital Deep Vein Thrombosis and Pulmonary Embolism Following Hip and Knee Arthroplasty Among Patients Receiving Recommended Prophylaxis. A Systematic Review. JAMA 2012; 307(3): 294-303.