8th International Conference for EBHC Teachers and Developers
The ecosystem of evidence
Connecting generation, synthesis and translation
Taormina (Italy), 25th - 28th October 2017

Comparison of imprecision assessed by GRADE and Trial Sequential Analysis in systematic reviews

Castellini Greta, Bruschettini Matteo, Gianola Silvia, Gluud Christian, Moja Lorenzo

Speaker: **Greta CASTELLINI**
PT, MSc, PhD student
Department of Biomedical Sciences for Health, University of Milan
Unit of Clinical Epidemiology, IRCCS Orthopedic Institute Galeazzi

GRADE

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) has gained momentum as ***an internationally accepted framework*** to assess systematically and transparently ***the quality of evidence***.

The GRADE assess important sources of quality in the evidence for risks of:

- Bias (systematic errors)
- Inconsistency
- **Imprecision**
- Indirectness
- Publication bias

Trial Sequential Analysis

A frequentist method that controls for random errors of type I and type II

to reduce the uncertainty in meta-analyses results and protect against the inflation of results

The **lack of data and the repeated meta-analyses of data increase the risks of random errors**, leading to significant or neutral findings

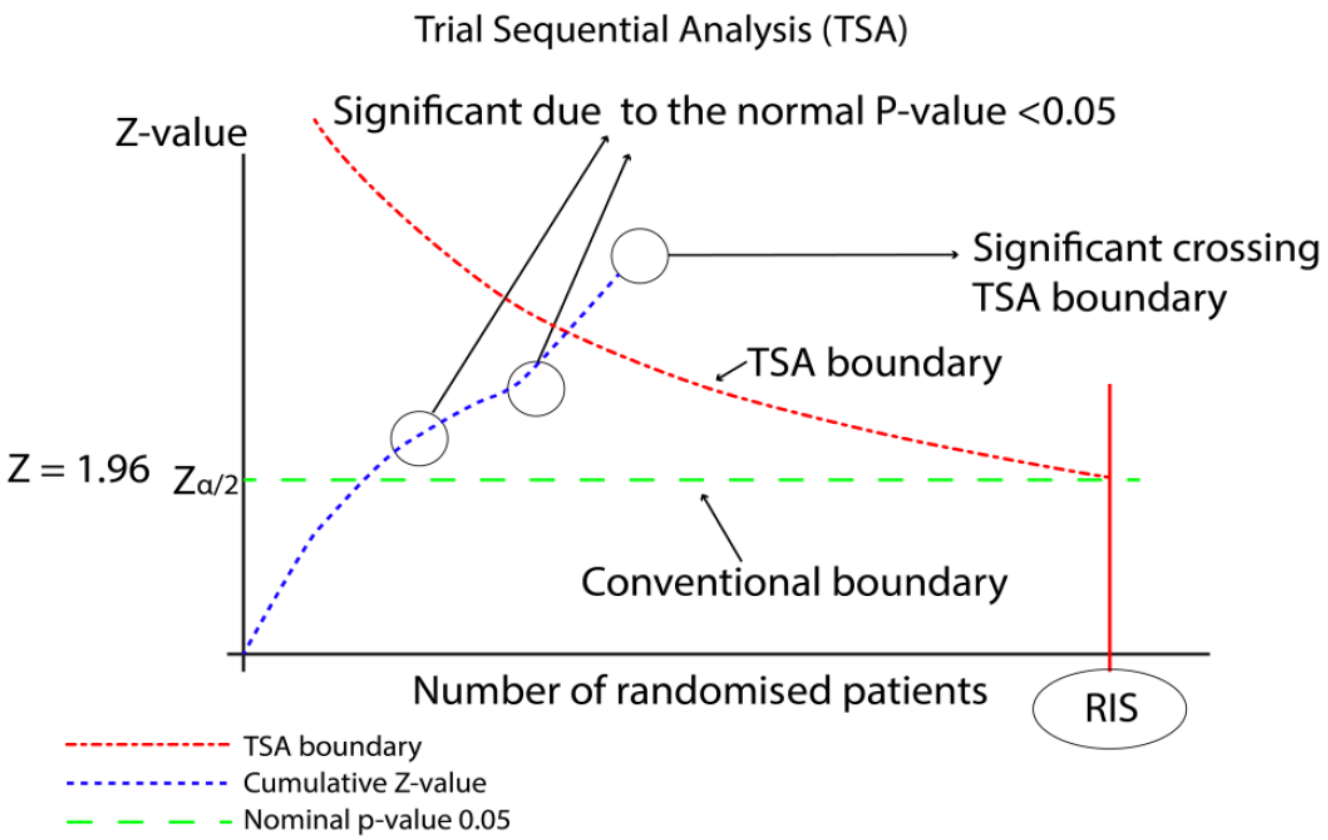
REQUIRED INFORMATION SIZE

The family-wise risk of random error increases more than 5% if accumulated data are analysed during **multiple up-dates**

TSA-ADJUSTED MONITORING BOUNDARIES

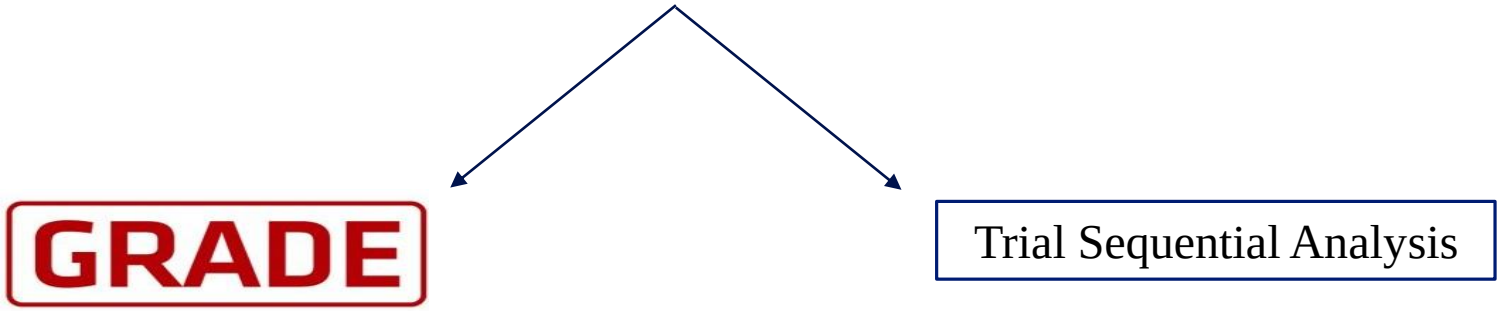
Trial Sequential Analysis

TSA Graph



Imprecision: two methods

Imprecision comprises:
Absolute sample size
Optimal or required information size
Confidence intervals of the overall effect
Critical margins of «benefit» and «harm»



TSA assesment has been advocated as a supplemet
in imprecision assessment with GRADE

Aim

- To estimate the imprecision in Cochrane systematic reviews by applying the TSA methodology

- To compare the agreement of the imprecision assessment suggested by GRADE as reported in the original Cochrane systematic reviews with generic TSA

Methods

Sample: **100** Cochrane systematic reviews

Inclusion criteria

(1) Therapeutic review assessing the effectiveness of any intervention.

(2) Limited to dichotomous outcomes.

(3) Included a meta-analysis with at least two informative randomized controlled trials.

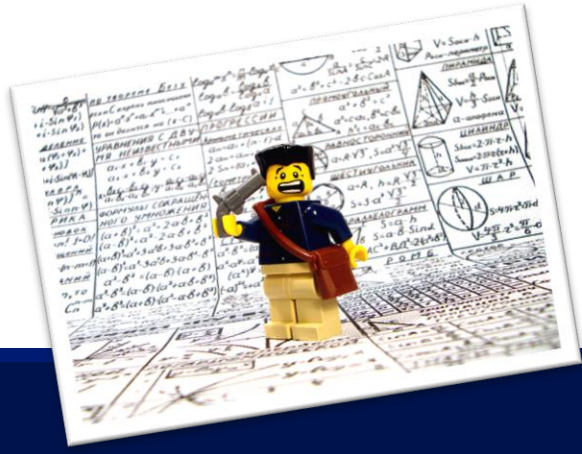
(4) SR includes Summary of Findings (SoF) table. The dichotomous outcome should be listed in the SoF table.

→ *Unit of our analysis: meta-analysis of the primary outcome*

Methods

1. We re-conducted each selected Cochrane meta-analysis using the trial data and applying the TSA method.

2. We estimated the diversity-adjusted required information size (DARIS) based on:
   a) control event proportion according to the data obtained in the meta-analysis at hand;
   b) an a priori anticipated intervention effect – i.e., risk ratio reduction suggested by GRADE' authors as default threshold of 25%;
   c) alpha of 0.05;
   d) beta of 0.20 (power of 80%);
   e) Measure of Diversity.



Methods

3. Assuming a realistic anticipated intervention effect, assessment of “imprecision” domain trough TSA is rated as follow:

If one of the boundaries for benefit, harm, or futility are crossed

↓

Not downgrading the evidence for imprecision

If none of the boundaries for benefit, harm, or futility are crossed

↓

Downgrading the evidence **two levels for** imprecision

Primary outcome: agreement of imprecision assessment between the GRADE imprecison evaluation compared with TSA (downgrade versus no downgrade)

Results

GRADE approach
48% downgraded for imprecision

TSA assessment
69% downgraded for imprecision

TSA / GRADE	Downgrade	No downgrade	Total
Downgrade	44	4	48
No downgrade	25	27	52
Total	69	31	100

The Cohen's kappa coefficient = **0.429**
The coefficient expressed a moderate strength of agreement according to the scale offered by Landis and Koch

Odds Ratio = 12.6

Limits

- The **analyses** are **restricted** to **dichotomous outcomes**

- The anticipated intervention effect was reported in only 4 of the systematic reviews. Therefore, we **chose a 25% risk relative reduction/improvement** as a realistic parameter for the outcomes

- We found difficulties in comprehending the judgements of downgrading or not downgrading in systematic reviews as they were not transparently reported

Bottom line

- We expected to find a more divergent assessment between the two systems, however, the strength of agreement between GRADE and TSA is moderate.

- Compared to GRADE assessments as conducted by Cochrane authors, TSA seems to downgrade more often for imprecision.

- Systematic reviews often do not report the anticipated clinical important intervention effect or the required information size. This information is necessary to judge the imprecision.

Thank you for your attention!

Department of Biomedical Sciences for Health, University of Milan
IRCCS Galeazzi Orthopaedic Institute, Unit of Clinical Epidemiology
Via R.Galeazzi 4, 20161 Milan
+39 02.6621.4780
greta.castellini@grupposandonato.it
greta.castellini@unimi.it