Automatic detection of inadequate claims in biomedical articles: first steps

Anna Koroleva, Patrick Paroubek

LIMSI, CNRS, Université Paris-Saclay, France

koroleva@limsi .fr, pap@limsi.fr

MEDA-2017
Workshop on Curative Power of MEdical DAta
14 Sept. 2017, Constanța, Romania

Background

- Data: randomized controlled trials (RCTs) assessing an intervention

- Inadequate reporting (spin): specific way of reporting that highlights that the beneficial effect of the experimental treatment in terms of efficacy or safety is greater than the results show.

- Impact: causes overestimation of the beneficial effect of the experimental treatment by physicians, patients, media.

- Prevalence: present in abstracts of 60% of reported randomized controlled trials (RCTs).

Types of spin

Misleading reporting of results:

- not reporting adverse events;
- **selective reporting of outcomes** (omission of primary outcome; focus on statistically significant secondary outcomes, subgroup or within-group analyses);
- misleading reporting of study design;
- **linguistic spin**;
- no consideration of limitations;
- selective citation of other studies.

Inadequate interpretation of results:

- **claiming a beneficial or equivalent effect of the intervention for statistically non-significant results**;
- claiming that the treatment is safe for statistically non-significant safety outcomes;
- concluding a beneficial effect despite no comparison test performed;
- interpretation of the results according to statistical significance instead of clinical relevance.

Inadequate extrapolation:

- inadequate extrapolation from the population, interventions or outcome actually assessed in the study to a larger population, different interventions or outcomes;
- inadequate implications for clinical practice.

Spin in conclusions

With spin

Treatment A + CAF was well tolerated and is suggested to have efficacy in patients who had not received prior therapy.

This study demonstrated improved PFS and response for the treatment A compared with comparator B alone.

Without spin

Treatment A + CAF was not more effective than CAF + placebo in patients with advanced or recurrent breast cancer.

The treatment A was not more effective than comparator B on overall survival in patients with metastatic breast cancer.

Objective

Develop an algorithm for automatic spin detection that would:

- detect important claims in scientific articles;

- extract possible supporting information:

  · study design;

  · outcomes (primary and secondary);

  · statistical significance of results;

  · patient population studied;

  · adverse events;

  · limitations of a trial;

  · interventions studied.

- evaluate the adequacy of the claims.

Algorithm description

Default input: full-text article with title and abstract.

Algorithm scheme:

1) Preprocess file.

2) Divide the text into title, abstract, body text.

3) Classify text (identify trial type).

4) Classify sentences in the abstract (identify results, conclusions)

5) Identify tonality of results

6) Extract information:
   · Entities: outcomes, patient population, statistical significance of results.
   · Relations: between outcomes and significance levels.
   · Linguistic features.

7) Look for specific spin markers, e.g.:
   · Is the primary outcome reported in the abstract?
   · Is the patient population in the abstract the same as the population studied?
   · etc.

Thank you for your attention !