

Altmetrics in the biomedical literature: appropriateness of statistical methods

Valeria Scotti^{1,3}, Annalisi De Silvestri^{2,3}, Carmine Tinelli^{2,3}, Clinical
Moreno Curti^{1,3}, Luigia Scudeller^{2,3}

¹ *Scientific Documentation Service, IRCCS Policlinic San Matteo Foundation, Pavia, Italy*

² *Clinical Epidemiology and Biometry Unit, IRCCS Policlinic San Matteo Foundation, Pavia, Italy*

³ *ALL Metrics Team IRCCS Policlinic San Matteo Foundation, Pavia, Italy*

Corresponding author: Valeria Scotti, v.scotti@smatteo.pv.it

Abstract

Aims

Alternative metrics have recently gained much attention in the biomedical community, with increasing related publications. Formally, any new metric needs external validation against an accepted “gold-standard”; however, there is not a universally accepted method for validation. We aimed at investigating whether the statistical methods applied to validate the new metrics were correctly applied, with a focus on the biomedical community.

Methods

We performed a systematic review on Pubmed, ArXiv e-prints archive and DPLB database. Eligibility criteria for full-text retrieval were focus on alternative metrics in abstract/title and a quantitative approach. From selected papers, we extracted: publication type, citation, publication year, scientific aims formally declared, main results presented in abstract, type of statistical methods, overall judgement on appropriateness of statistical .

Results

We retrieved 200 records from ArXiv, 38 from Pubmed, and 13 from DPLB, for an overall total of 251; 185 were excluded from screening abstract (161 focus on social media rather than on alternative metrics, 24 narrative review articles or viewpoints). We retrieved 42 full text: publication year of the first report was 2005, the second publication 2011, 2 in 2012, 6 in 2013 and 16 each in 2014 and 2015. Of the 42 full-text, 4 were abstract and 3 were unpublished with no quantitative approach, 1 was a duplicate publication, 15 were not relevant.

Therefore, only 19 (4.2% of the original sample, and 24% of the full-texts retrieved) formally assessed validity of the new metrics (1 in 2012, 4 in 2013, 9 in 2014, and 5 in 2015), published by three different research groups (sometimes overlapping); 10 were published in journals with IF, 7 in journals without IF and 2 in ArXiv repository. Specific research aims were specified in all 19. Comparison with traditional indicators was performed in 12, while comparison between different altmetrics was done in 7. Only in 4 papers data were highlighted in the abstract. The statistical methods applied were purely descriptive in 2, correlation methods in 9, regression methods in 3, dimensional methods in 4, network analysis in 2, and meta-analysis of correlation in 1 (NB: total does not add up to 19 because most papers used more than one method). The median quality score was 7 (IQR 6-8, min-max range 6-9).

Discussion

Most literature on alternative metrics is represented by introductory papers, descriptive works, or narrative reviews. Original research on altmetrics represented only a minority of all publications, and was conducted by a small number of research groups, with specific interest in Altmetrics. The statistical methods were generally appropriated and the overall quality of statistical methods was fair. In conclusion, to support the uptake of altmetric indicators in the biomedical community the formal validation of alternative metrics need to be expanded.

Key words: Publication Statistics, Correlation, Validity , Altmetrics, Biomedical

Background and aims

Alternative metrics have recently gained much attention in the biomedical community, with an increasing number of related publications¹. Formally, any new metric needs external validation against an accepted “gold-standard”; however, there is not a universally accepted method for validation. Also, like in any new field, the methodology applied in the early stages might be suboptimal. Besides, most alternative metrics tools are currently a sort of “black box” and the underlying algorithms calculating the scores are not made public. We aimed at investigating whether the statistical methods applied to validate the new metrics have been correctly applied in the development literature, with a focus on the biomedical community.

Methods

A systematic review of the literature was performed on October 2015. The search was constructed and carried out by a librarian experienced on Alternative Metrics theme, using Pubmed, ArXiv e-prints archive (<http://arxiv.org/>) and DPLB database (<http://dblp.uni-trier.de/>). The following search strategy was used on PubMed: *Altmetric**, *Altmetrics* OR “*alternative metrics*”.

On Arxiv and DPLB only the general term Altmetrics was used, to find similar e-prints abstracts. No language or years restrictions were applied given the novelty of the topic. Eligibility criteria for full-text retrieval were focus on alternative metrics in the abstract or title with a quantitative approach. Selection was performed independently by two researchers (VS and ADS); if at least one researcher suggested full-text retrieval, it was retrieved.

From the selected papers, we extracted the following information: publication type, citation, publication year, scientific aims formally declared, main results presented in abstract, type of statistical methods, overall judgement on appropriateness of statistical methods (on a 0-10 scale, by an experienced statistician).

Results

We retrieved 200 records from ArXiv, 38 from Pubmed, and 13 from DPLB, for an overall total of 251. Of these 185 were excluded from screening abstract (161 focus on social media rather than on alternative metrics, 24 narrative review articles or viewpoints).

We retrieved 42 full text: publication year of the first report was 2005, the second publication 2011, 2 in 2012, 6 in 2013 and 16 each in 2014 and 2015. Of the 42 full-text:

- 4 were abstract,
- 3 were unpublished with no quantitative approach,
- 1 was a duplicate publication,
- 15 were not relevant.

Therefore, only 19 (4.2% of the original sample, and 24% of the full-texts retrieved) formally assessed validity of the new metrics (1 in 2012, 4 in 2013, 9 in 2014, and 5 in 2015-see Figure 1.), published by three different research groups (sometimes overlapping); 10 were published in journals with IF, 7 in journals without IF and 2 in ArXiv repository.

Specific research aims were specified in all 19. Comparison with traditional indicators was performed in 12, while comparison between different altmetrics was done in 7. Only in 4 papers data were highlighted in the abstract.

The statistical methods applied were:

- purely descriptive in 2,
- correlation methods in 9,
- regression methods in 3,
- dimensional methods in 4,
- network analysis in 2,
- meta-analysis of correlation in 1

(NB: total does not add up to 19 because most papers used more than one method).
The median quality score was 7 (IQR 6-8, min-max range 6-9).

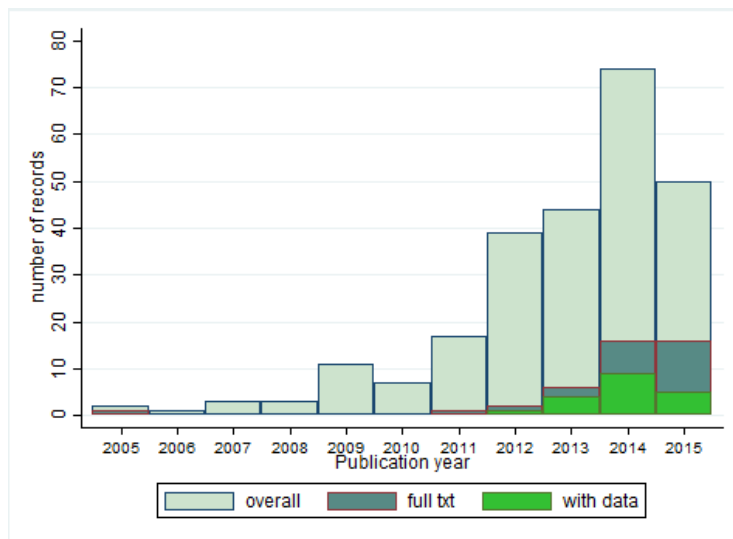


Figure 1: Correlation between publication year and number of records

Discussion

Most literature on alternative metrics is represented by introductory papers, descriptive works, or narrative reviews. The first report is 10-year old, and the number of publications increased with time.

In particular, ArXiv is heavily used by researchers who publish articles on altmetrics, while PubMed (which more easily reaches a biomedical audience) is less represented.

Original research on altmetrics represented only a minority of all publications, and was conducted by a small number of research groups, with specific interest in Altmetrics. The statistical methods were generally appropriated, though in 20% of papers only descriptive statistics were used. The overall quality of statistical methods was fair.

In conclusion, in order to support the uptake of altmetrics indicators in the biomedical community, for research evaluation or other purposes, in the future the formal validation of alternative metrics need to be expanded. Inclusion of a statistician in any research groups is strongly suggested.

REFERENCES

1. Priem J, Groth P, Taraborelli D. The Altmetrics Collection. PLoS ONE. 2012;7(11):e48753.