

The Centre for Evidence Based Medicine Outcomes Monitoring Project (COMPare)

Protocol

Ben Goldacre, Henry Drysdale, Carl Heneghan

CEBM, University of Oxford

compare-team@ebmdatalab.net

Background

Despite various efforts, academic journal reports on clinical trials continue to exhibit outcome reporting bias: pre-specified primary and secondary outcomes are left unreported, while non-pre-specified outcomes are reported as main results without any clear declaration to that effect. While there have been various publications setting out the prevalence of this problem, there has been no systematic effort to publicly identify the individual trials that have engaged in this form of research misconduct. There are good grounds to believe that this problem may become less prevalent through a process of individual accountability, and transparent access to good quality information on which individual researchers and journals are currently engaging in this practice.

We aim to implement such a monitoring system, to check whether academic journals are able to manage corrections to the record, and to reduce the prevalence of the problem. Our specific objectives are to:

1. Determine the proportion of pre-specified outcomes that are correctly reported in the top 5 general medical journals.
2. Determine the prevalence of undeclared, non-pre-specified outcomes reported in these journals.
3. Submit letters to journals setting out the discrepancies, in order to correct the record, and assess whether these are published at all, and in a timely fashion.
4. Share our results openly to increase awareness of the presence of misreported outcomes in specific trials and journals.

Methods

We will assemble a group of students and researchers able to work on the project (assessors), with one study coordinator to administer sharing out trials.

We will create an assessment template, a letter template, and a master tracker, as Google spreadsheets.

Each week the study coordinator will screen the top 5 medical journals (based on impact factor) to identify all trials, enter the brief details of these onto the tracker, and assign each trial to an assessor and second assessor.

Each assessor will carry out the steps below for each trial.

The assessment system

1. Use the assessment template to make a new assessment sheet for the trial (with study ID appended in title) and save in the study Google drive folder.
2. Check to see if there is a version of the trial protocol available that is dated before the trial start date. If so, this should be used to gather pre-specified outcomes (point 5).
3. Find the trial's registry entry, using clinicaltrials.gov or the WHO clinical trials search portal <http://apps.who.int/trialsearch/>. Copy the link to the online registry to the master tracker and assessment sheet.
4. Save in the folder for that trial:
 - a. pdf of trial paper.
 - b. Any appendices or supplementary material.
 - c. pdf of pre-trial protocol, if available.
5. Copy the pre-specified primary and secondary outcomes from either the pre-trial protocol or registry entry into the assessment sheet. Note: The trial protocol should only be used for pre-specified outcomes if it was published before the trial start date. If using the registry entry, any changes made after the trial start date will be discounted. The pre-specified outcomes will be taken from the registry entry as it was immediately before the trial start date.
6. Select 'yes', 'no', or 'unclear' from the drop-down menus next to each outcome to show if it is:
 - a. reported in main trial results table
 - b. reported somewhere in the paper other than the main results table (e.g. free text, appendix)
 - c. not reported in the paper at all.
 - d. not sufficiently defined as outcome to know whether or not it was reported. Note: If 'unclear', details should be given in a 'comment' attached to the relevant cell in the spreadsheet.

7. List all non-pre-specified outcomes reported in main results table in the assessment sheet.
8. Select 'yes' or 'no' in the box next to each of these outcomes to show if it was correctly declared as non-pre-specified in the report.
9. Enter basic results from the assessment into the 'Master Tracker' spreadsheet.

Each assessment will then be independently checked by a randomly assigned second assessor. Any discrepancies will be recorded in the master tracker as present ('yes' or 'no'), and with free text details in the next column. A solution to these discrepancies will be decided upon by discussion with the COMPare team at the weekly meeting.

Following this, the assessor will write a letter to the journal, for each trial where there are discrepancies between the registry and the publication. This will be done using a standard template letter which will explain the motivations and aims of the project, and then summarise the problems found with the trial in question. There is a specific template letter for each journal, to reflect the different correspondence requirements.

Where it adds to the content of the letter, we will also add **either** Table 1 **or** Table 2, which will be produced by the assessor.

Table 1 will be an annotated version of the trial's main results table, showing: which reported outcomes were pre-specified primary outcomes (by writing **P1** to the right of the table); which were pre-specified secondary outcomes (by writing **P2** to the right of the table); and which were not pre-specified at all (by writing **U** to the right of the table).

Table 2 will be a corrected results table for the trial that lists each pre-specified primary and secondary outcome; gives the results for each (if they are available in the trial publication, or registry results page); or reports them as 'missing' if those values are not available.

These letters will be sent for publication to the journal editor by the study team on a weekly basis. We will publish a description of the reporting errors in each trial on the COMPare website, with read-only links to the assessment and letter for that trial, in the following situations:

1. If the letter for that trial is published by the journal.
2. If the letter for that trial is rejected by the journal.
3. If four weeks have passed without the letter being published by the journal.
4. If the standard of outcome reporting was high and a letter was not required.

In these situations, we will also post our letters on PubMed Commons.

On the COMPare-trials.org website, individual trials' data will be obfuscated until conditions 1-4 above are met, but overall figures on outcome reporting will be produced that include all data from all trials coded to date.

Where responses to letters are received from trial authors or journal editors, a reply will be generated by discussion with the COMPare team.

Analysis

We will conduct an omnibus analysis of all data collected during the first 6 weeks of the project, and submit this for publication, alongside a description of our aims and methods.

Primary Outcomes:

1. Proportion of pre-specified outcomes reported.
2. Proportion of reported outcomes that are non-prespecified, and not declared as such.
3. Proportion of letters published in print.
4. Publication delay in days.

Secondary Outcomes:

1. All the above outcomes, broken down by journal.

We hope to continue monitoring outcome switching, and sending correction letters, in all five journals, until five trials in a row have no outcome switching. At this point we will switch to sampling one trial in five (by sequential allocation from a start paper determined by a randomly generated digit from 1-5). We will then return to monitoring all trials if any sampled trial has any switched outcomes.

An example of annotating a trial results table:

Table 2. Outcomes at 12 Months.

Outcome	Total No. of Assessments*		Mean Improvement in Outcome from Baseline to 12 Mo (95% CI)		Between-Group Difference in Mean Improvement (95% CI)		
			Nonsurgical-Treatment Group	Total-Knee-Replacement Group	Crude	Adjusted†	
Primary outcome							
KOOS ₄	179	193	16.0 (10.1 to 21.9)	32.5 (26.6 to 38.3)	16.5 (10.2 to 22.7)	15.8 (10.0 to 21.5)	P1
Secondary outcomes							
KOOS subscale scores							
Pain	180	194	17.2 (10.4 to 24.1)	34.8 (28.1 to 41.5)	17.6 (10.1 to 25.1)	17.1 (10.4 to 23.8)	U
Symptoms	179	194	11.4 (4.4 to 18.4)	26.4 (21.5 to 31.4)	15.0 (8.3 to 21.7)	12.7 (6.6 to 18.8)	U
Activities of daily living	180	193	17.6 (11.4 to 23.9)	30.0 (22.7 to 37.2)	12.3 (5.5 to 19.2)	12.9 (6.8 to 19.1)	U
Quality of life	180	194	17.8 (11.2 to 24.4)	38.2 (30.6 to 45.8)	20.4 (12.8 to 27.9)	20.2 (13.2 to 27.1)	U
Sports and recreation	177	193	19.3 (10.8 to 27.7)	34.5 (27.9 to 41.0)	15.2 (6.7 to 23.7)	15.6 (7.3 to 23.9)	U
Time on the timed up-and-go test (sec)	163	185	−1.2 (−1.8 to −0.6)	−2.4 (−3.1 to −1.6)	1.2 (0.4 to 1.9)	0.9 (0.2 to 1.6)	P2
Time on the 20-m walk tests (sec)	163	185	−1.0 (−1.5 to −0.4)	−2.9 (−3.8 to −1.9)	1.9 (0.9 to 2.8)	1.5 (0.7 to 2.4)	P2
EQ-5D scores							
Descriptive index	178	194	0.115 (0.063 to 0.166)	0.206 (0.141 to 0.270)	0.091 (0.026 to 0.155)	0.078 (0.023 to 0.132)	P2
Visual-analogue scale	180	193	10.2 (4.6 to 15.7)	15.0 (8.6 to 21.5)	4.9 (2.2 to 12.0)	4.4 (1.8 to 10.6)	P2
Weight (kg)‡	134	160	−2.6 (−3.9 to −1.4)	0.1 (−1.5 to 1.7)	2.8 (1.4 to 4.1)	2.8 (1.4 to 4.1)	U

Appendix 1: Updates and changes to our protocol

This protocol is principally shared as a toolkit to allow others to replicate our low-budget methods (but please get in touch and collaborate), and was developed as our internal manual for our medical student coders. This is an iterative project. Any protocol amendments, especially those relevant to data analysis, will be flagged in our publication.

1. Small modifications to specifics of “objectives” since first draft in pilot stage.
2. Addition of Appendix 2 below, during early coding, as we iteratively develop an additional taxonomy of *types* of poor reporting and poor prespecification from working with the registry entries, protocols, and publications during the project.
3. Modification of letter templates - and plans to include figures/tables - in light of journals’ submission policies and restrictions on word length and figures/tables in letters.

4. Addition of our plan for ongoing monitoring, and stopping / sampling rules for this, if person-time permits.
5. Addition of weekly meeting with team (often twice weekly due to workload).
6. Addition of rules around posting results to COMPare-trials.org (and calculating rolling summary statistics) as website developed during early phase of project.
7. Specified simple analysis for omnibus paper
8. Added team email address

Appendix 2: Correct pre-specification and correct reporting of outcomes

All primary and secondary outcomes in a clinical trial must be fully pre-specified on a clinical trials registry and in the trial protocol before the trial start date. We consider correct pre-specification of an outcome to contain details of:

1. The variable that will be measured
2. The specific method of measurement
3. The time-points at which the outcome will be reported

Some specific points regarding pre-specified outcomes:

- Where an outcome has been pre-specified at several time-points, each time-point is considered to be a separate outcome.
- Where scoring systems are used that contain several sub-scales, it must be pre-specified whether the outcome consists of the overall score or specific sub-scales.
- In general, an outcome must be unambiguously pre-specified, so that the reader can be sure if the outcome was reported or not.

Correct reporting of outcomes:

1. All pre-specified outcomes must be reported in the published trial report.
2. Primary outcomes must be clearly reported as the main results of a trial. This is usually in the main trial results table, however other clear methods of reporting, such as Kaplan-Meier plots, are also acceptable.

3. Secondary outcomes should be clearly reported as such. At a minimum, this should be in the free text results section of the trial report. Secondary outcomes must not be reported as primary outcomes.
4. Any discrepancies between the outcomes pre-specified and those reported must be clearly declared in the trial report, along with details of the timing of changes and reasons for the discrepancies.