OXFORD

JNCI J Natl Cancer Inst (2022) 114(3): djab177

doi: 10.1093/jnci/djab177
First published online September 11, 2021
Brief Communications

Impact of Blinding on Patient-Reported Outcome Differences Between Treatment Arms in Cancer Randomized Controlled Trials

Fabio Efficace, PhD,[1,2,*] David Cella, PhD,[2] Neil K. Aaronson, PhD,[3] Melanie Calvert, PhD,[4,5] Francesco Cottone, PhD,[1] Massimo Di Maio, MD,[6] Francesco Perrone, MD,[7] Francesco Sparano, MS,[1] Eva-Maria Gamper, PhD,[8] Marco Vignetti, MD,[1] Johannes M. Giesinger, PhD[9]

[1]Italian Group for Adult Hematologic Diseases (GIMEMA), Data Center and Health Outcomes Research Unit, Rome, Italy; [2]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [3]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Amsterdam, the Netherlands; [4]Centre for Patient Reported Outcomes Research, and Birmingham Health Partners Centre for Regulatory Science and Innovation, Institute of Applied Health Research, University of Birmingham, and University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK; [5]National Institute for Health Research (NIHR) Birmingham Biomedical Research Centre, NIHR Surgical Reconstruction and Microbiology Research Centre, NIHR Applied Research Centre West Midlands, University of Birmingham, Birmingham, UK; [6]Department of Oncology, University of Turin and Mauriziano Hospital, Turin, Italy; [7]Clinical Trial Unit, Istituto Nazionale Tumori, IRCCS, Fondazione G. Pascale, Napoli, Italy; [8]Innsbruck Institute of Patient-Centered Outcome Research (IIPCOR), Innsbruck, Austria; and [9]University Hospital of Psychiatry II, Medical University of Innsbruck, Innsbruck, Austria

[*]**Correspondence to:** Fabio Efficace, PhD, Health Outcomes Research Unit, Italian Group for Adult Hematologic Diseases (GIMEMA), GIMEMA Data Center, Italy (e-mail: f.efficace@gimema.it).

Abstract

Some concerns have been raised about potential bias in patient-reported outcome (PRO) results from open-label cancer randomized controlled trials (RCTs). We investigated if open-label trials favor the experimental treatment over the standard treatment more frequently than blinded trials. We also examined if the effect of blinding differs for distal vs more proximal PROs. We assessed 538 RCTs with a PRO endpoint conducted in the most prevalent cancers, of which 366 (68.0%) were open-label, 148 (27.5%) were blinded, and 24 (4.5%) were categorized as unclear. In our multivariable logistic regression model, we did not observe a statistically significant association of the independent variable treatment concealment (blinded vs open-label) on the dependent variable measuring the proportion of trials favoring the experimental treatment (adjusted odds ratio = 1.19, 95% confidence interval = 0.79 to 1.79; 2-sided $P = .40$). This was also the case when comparing distal and proximal PROs. Our findings provide novel evidence-based data that support the validity of PRO results from open-label cancer RCTs.

Treatment concealment (ie, blinding) is considered to be an important design feature of randomized controlled trials (RCTs) that minimizes the potential influence of patients' and clinicians' expectations on possible benefits and harms associated with new treatments (1).

In its guideline document on the use of patient-reported outcome (PRO) measures in clinical trials, the US Food and Drug Administration (2) has argued that patients receiving active treatment may overestimate treatment benefit and concludes that open-label trials "are rarely adequate to support labeling claims based on PRO instruments." Concerns about the impact of study design on PROs have also been expressed by the European Medicines Agency (3).

It has also been hypothesized that bias may be more likely to occur with specific types of outcomes, such as that more distal PROs (eg, emotional or social functioning; global quality of life) could be more susceptible to open-label bias than the more proximal PROs, that is, symptoms (4). Given that many RCTs now include PROs, it is critical to better understand the relationship between study design (ie, open-label vs blinded trials) and its possible impact on PRO results.

To investigate the potential impact of blinding on PRO results in cancer RCTs, we compared the frequency of PRO results that favor the experimental treatment over the standard treatment in open-label RCTs vs RCTs in which patients were blinded to treatment. We also examined if the effect of treatment concealment differs for distal vs proximal PROs.

Data were gathered through the PROMOTION registry (promotion.gimema.it), a database containing information on published cancer RCTs with a PRO endpoint. Details have been

Received: February 19, 2021; **Revised:** July 1, 2021; **Accepted:** August 11, 2021

reported elsewhere (5). We selected from the registry RCTs published between January 2004 and February 2019 that were conducted in the 5 most prevalent cancer sites worldwide (6): breast, non-small cell lung cancer, colorectal, prostate, and gynecological cancers.

Two reviewers (EMG and JMG) independently identified those trials for which the standard and experimental treatment arm could be clearly determined and categorized these trials as "blinded" if patients were unaware of their treatment arm and as "open-label" if it was not concealed. Categorization relied primarily on information provided in trial publications. If this information was missing, we extracted it from online trial registries wherever possible; otherwise, the trial was classified as "unclear" with respect to treatment concealment. Additionally, RCTs were categorized by the reviewers as favoring the experimental or standard treatment arm, if the majority (ie, at least 1 more) of statistically significant differences of PRO endpoints, at any time point during the study period, indicated better outcomes in the respective arm. If an equal number of PRO differences favored the experimental and treatment arm, or if no statistically significant difference occurred, trials were categorized as equivalent. In case of inconsistent ratings, consensus was reached through discussion, and if necessary, a third reviewer (FE) was consulted to facilitate the reconciliation process .

For those RCTs that clearly reported information on treatment concealment, we used univariable and multivariable logistic regression models to estimate the association between reported treatment concealment (coded as blinded = 1 vs open-label = 0) and the chance of finding statistically significant differences in PROs favoring the experimental treatment (yes = 1 vs no = 0). In multivariable analysis, we adjusted for the following variables, after having checked for multicollinearity, which were included in the model based on their relevance as potential confounding factors: international study, industry funding (fully or in part), sample size, year of publication, PRO as a primary endpoint, disease stage, and quality of PRO reporting based on the CONSORT-PRO extension (7). We also performed a sensitivity analysis running the same multivariable logistic regression model, considering those trials with missing information on treatment concealment as open-label, assuming that it is rare that blinded trials are not reported as such. Also, in the subset of trials using the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 instrument (8), which allows one to distinguish type of PROs, we investigated the impact of blinding separately for proximal PROs (the QLQ-C30 symptom domains) and distal PROs (the QLQ-C30 functional scales and the global health status/quality of life scale); the QLQ-C30 financial difficulty item was not considered for this analysis. In this subanalysis, we adjusted for the same factors as in the main analysis.

We assessed 538 trials of which 366 (68.0%) were open-label, 148 (27.5%) were blinded, and 24 (4.5%) were categorized as unclear. The selection process of studies and a summary of analyses performed is reported in Supplementary Figure 1 (available online). Overall, 227 (42.2%) RCTs reported a PRO difference favoring the experimental arm (Supplementary Table 1, available online).

For the subsequent analyses, all performed by logistic regression models (2-sided Wald χ^2 test, statistical significance threshold $\alpha = .05$), we considered 514 RCTs clearly reporting information on treatment concealment. In univariable logistic regression, we did not observe a statistically significant difference between open-label and blinded trials in the corresponding proportions of those where PROs favored the experimental treatment (unadjusted odds ratio [OR] = 1.36, 95% confidence interval [CI] = 0.93 to 2.00; $P = .11$) (data not shown in tables). We note that an odds ratio above 1 indicates a higher chance that PROs favor the experimental arm in blinded RCTs. This finding was confirmed by our multivariable analyses, which found an odds ratio of 1.19 (95% CI = 0.79 to 1.79; $P = .40$) (Table 1). The multivariable sensitivity analysis considering trials with missing information on concealment as open-label showed similar results (Supplementary Table 2, available online).

In the analysis of the subset of 202 trials using the EORTC QLQ-C30, no statistically significant impact of blinding was observed for either the proximal PROs (OR = 1.08, 95% CI = 0.44 to 2.63; $P = .87$) or the distal PROs (OR = 1.22, 95% CI = 0.46 to 3.23; $P = .69$) (Table 2).

Our results are in keeping with studies in the literature based on trials across various medical fields that have reported only small or non-statistically significant effects of treatment concealment on trial results (9–11). For example, a recent meta-analysis (11) of 142 meta-analyses of 1153 trials across medical fields found no evidence that concealment of treatment is associated with estimates of treatment effects assessed by patients, health-care providers, or outcome assessors. However, there are also studies (12–14) that have found larger differences between treatment arms for subjective outcomes in open-label studies.

To date, studies investigating the association between treatment concealment and PROs have applied a superiority analysis to demonstrate such an effect. Given that a number of studies only found small effects or failed to demonstrate such an effect (even when relying on large numbers of trials), we agree with Moustgaard et al. (11) that future studies should target settings where such an effect is likely to occur rather than investigating any type of trial. In addition, we believe that it would be useful to take a noninferiority approach to the analysis of this issue—that is, to demonstrate that such an effect, if present, is weaker than a certain magnitude deemed a priori as being important. A limitation of our study is that we did not perform subgroup

Table 1. Multivariable logistic regression model for the occurrence of PRO differences favoring the experimental treatment arm (n = 514)[a]

Variable	OR (95% CI)	P[b]
Blinded (yes vs no)	1.19 (0.79 to 1.79)	.40
International (yes vs no)	0.93 (0.60 to 1.42)	.72
Industry funding, fully or in part (yes vs no)	1.33 (0.89 to 2.00)	.17
Sample size n > 200 (yes vs no)	0.68 (0.44 to 1.05)	.08
Year of publication ≥ 2014[c]	0.94 (0.63 to 1.40)	.74
PRO primary endpoint (yes vs no)	1.66 (1.05 to 2.61)	.03
Disease stage		
Metastatic/advanced vs nonmetastatic/local	0.94 (0.61 to 1.45)	.78
Both/unclear vs nonmetastatic/local	0.94 (0.58 to 1.54)	.82
CONSORT-PRO summary score[d]	1.13 (0.99 to 1.29)	.07

[a]This table represents the association of key RCT characteristics with the chance of finding statistically significant differences favoring the experimental arm. CI = confidence interval; CONSORT = CONsolidated Standards Of Reporting Trials; OR = odds ratio; PRO = patient-reported outcomes.
[b]2-sided Wald χ^2 test.
[c]This cutoff date was based on the publication of the CONSORT-PRO extension (7).
[d]The CONSORT PRO-summary score ranging from 0 to 6 was included in the model as a continuous variable. The original CONSORT-PRO extension consists of 5 items, however, for the purpose of our study, the item P6a was split in 2, to maximize information on quality of PRO reporting.

Table 2. Multivariable logistic regression model for the occurrence of PRO differences favoring the experimental treatment arm respectively in proximal and distal PROs (n = 202)[a]

Variable	Proximal PROs		Distal PROs	
	OR (95% CI)	P[b]	OR (95% CI)	P[b]
Blinded (yes vs no)	1.08 (0.44 to 2.63)	.87	1.22 (0.46 to 3.23)	.69
International (yes vs no)	0.51 (0.25 to 1.05)	.07	0.96 (0.45 to 2.07)	.92
Industry funding, fully or in part (yes vs no)	1.41 (0.70 to 2.85)	.34	1.37 (0.62 to 3.03)	.43
Sample size n > 200 (yes vs no)	1.65 (0.72 to 3.78)	.24	2.06 (0.77 to 5.50)	.15
Year of publication ≥ 2014[c]	1.17 (0.58 to 2.33)	.67	1.68 (0.80 to 3.53)	.17
PRO primary endpoint (yes vs no)	0.77 (0.31 to 1.93)	.58	1.29 (0.48 to 3.48)	.61
Disease stage				
Metastatic/advanced vs nonmetastatic/local	1.13 (0.53 to 2.39)	.75	1.27 (0.57 to 2.81)	.56
Both/unclear vs nonmetastatic/local	1.01 (0.40 to 2.51)	.99	0.32 (0.10 to 1.04)	.06
CONSORT-PRO summary score[d]	1.11 (0.88 to 1.40)	.39	1.07 (0.83 to 1.38)	.60

[a]This table represents the association of key RCT characteristics with the chance of finding statistically significant differences favoring the experimental arm in the EORTC QLQ-C30 scales, grouped as proximal PROs (fatigue, nausea and/or vomiting, pain, dyspnea, insomnia, constipation and diarrhea) and distal PROs (physical, role, emotional, cognitive, social functioning and global health status/quality of life). The EORTC QLQ-C30 financial difficulty item was not considered for this analysis. CI = confidence interval; CONSORT = CONsolidated Standards Of Reporting Trials; EORTC = European Organisation for Research and Treatment of Cancer; OR = odds ratio; PRO = patient-reported outcomes.

[b]2-sided Wald χ^2 test.

[c]This cutoff date was based on the publication of the CONSORT-PRO extension (7).

[d]The CONSORT PRO-summary score ranging from 0 to 6 was included in the model as a continuous variable. The original CONSORT-PRO extension consists of 5 items, however, for the purpose of our study, the item P6a was split in 2, to maximize information on quality of PRO reporting.

analyses to identify specific RCT settings where blinding may possibly have a more prominent impact on PROs.

In conclusion, we did not observe a statistically significant difference in the proportion of open-label vs blinded RCTs that favored the experimental treatment. This was also the case when comparing distal and proximal PROs in a subset of trials. Overall, these findings provide novel evidence-based data that support the validity of PRO results from open-label RCTs. At the same time, as even in large data sets such as ours, a lack of statistical significance does not necessarily indicate a lack of effect, we recommend additional studies that employ a noninferiority approach to the analysis.

Funding

No funding was received for this submission.

Notes

Role of the funder: Not applicable.

Disclosures: FE: consultancy for AbbVie, Amgen, Janssen, Takeda, and grants from AbbVie and Amgen (to his Institution), outside the submitted work. DC: President, FACIT.org. consulting fees from AbbVie, Amgen, Astellas, Boehringer-Ingelhein, Bristol-Myers Squibb, Daiichi-Sankyo, Dompe, Merck, Novartis, Pfizer, and Pled Pharma, outside the submitted work. MC is Director of the Birmingham Health Partners Centre for Regulatory Science and Innovation, Director of the Centre for Patient Reported Outcomes Research and is a National Institute for Health Research (NIHR) Senior Investigator. She receives funding from the NIHR Birmingham Biomedical Research Centre, the NIHR Surgical Reconstruction and Microbiology Research Centre and NIHR ARC West Midlands at the at the University of Birmingham and University Hospitals Birmingham NHS Foundation Trust, Health Data Research UK, Innovate UK (part of UK Research and Innovation), Macmillan Cancer Support, GSK and UCB Pharma. MC has received personal fees from Astellas, Takeda, Merck, Daiichi Sankyo, Glaukos, GSK, and the Patient-Centered Outcomes Research Institute (PCORI) outside the submitted work. The views expressed in this article are those of the author(s) and not necessarily those of the NIHR, or the Department of Health and Social Care. MDM: honoraria and had roles as consultant or advisor for AstraZeneca, Pfizer, Novartis, Roche, Takeda, Eisai, Janssen, Astellas; received institutional research grant by Tesaro—GlaxoSmithKline, outside the submitted work. FP: personal fees from Bayer, Ipsen, Astra Zeneca, Bristol Myers Squibb, Sandoz, Incyte, Celgene, Pierre Fabre, Janssen-Cilag and research grants (Institution) from Astra Zeneca, Bayer, Roche, Merck, Pfizer, Incyte, Sanofi, BioClin, Tesaro, outside the submitted work. MV: personal fees from Amgen, personal fees from Millennium Pharmaceuticals Inc, personal fees from Celgene, personal fees from Janssen, personal fees from Novartis, personal fees from Incyte, outside the submitted work. All other authors have no conflicts of interest to disclose.

Author contributions: Fabio Efficace: Conceptualization, Supervision, Writing—original draft, Writing—review & editing, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Validation. David Cella: Conceptualization, Methodology, Writing—review & editing, Validation. Neil K. Aaronson: Conceptualization, Methodology, Writing—review & editing, Validation. Melanie Calvert: Writing—review & editing, Methodology, Validation. Francesco Cottone: Conceptualization, Data curation, Formal Analysis, Methodology, Software, Writing—original draft, Writing—review & editing, Validation. Massimo Di Maio: Methodology, Writing—review & editing, Validation. Francesco Perrone: Methodology, Writing—review & editing, Validation. Francesco Sparano: Project administration, Investigation, Data curation, Writing—review & editing, Validation. Eva-Maria Gamper: Investigation, Writing—review & editing, Validation. Marco Vignetti: Investigation, Resources, Writing—review & editing,

Validation. Johannes M. Giesinger: Conceptualization, Investigation, Methodology, Data Curation, Writing—original draft, Writing—review & editing, Validation.

Data Availability

The data underlying this article are available upon reasonable request to the corresponding author .

References

1. Hróbjartsson A, Boutron I. Blinding in randomized clinical trials: Imposed impartiality. *Clin Pharmacol Ther*. 2011;90(5):732–736.
2. US Food and Drug Administration. *Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*. US Department of Health and Human Services Food and Drug Administration; 2009. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed July 1, 2021.
3. European Medicine Agency. *The Use of Patient-Reported Outcome (PRO) Measures in Oncology Studies. Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man*; 2016. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2016/04/WC500205159.pdf. Accessed July 1, 2021.
4. Roydhouse JK, Fiero MH, Kluetz PG. Investigating potential bias in patient-reported outcomes in open-label cancer trials. *JAMA Oncol*. 2019;5(4):457–458.
5. Giesinger JM, Efficace F, Aaronson N, et al. Past and current practice of patient-reported outcome measurement in randomized cancer clinical trials: a systematic review. *Value Health*. 2021;24(4):585–591.
6. Ferlay J, Ervik M, Lam F, et al. *Global Cancer Observatory: Cancer Today*. Lyon, France: International Agency for Research on Cancer; 2020. https://gco.iarc.fr/today. Accessed July 1, 2021.
7. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD; for the CONSORT PRO Group. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA*. 2013;309(8):814–822.
8. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*. 1993;85(5):365–376.
9. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in physical therapy trials and its association with treatment effects: a meta-epidemiological study. *Am J Phys Med Rehabil*. 2017;96(1):34–44.
10. Hartling L, Hamm MP, Fernandes RM, Dryden DM, Vandermeer B. Quantifying bias in randomized controlled trials in child health: a meta-epidemiological study. *PLoS One*. 2014;9(2):e88008.
11. Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ*. 2020;368:l6802.
12. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. *Int J Epidemiol*. 2014;43(4):1272–1283.
13. Savović J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med*. 2012;157(6):429–438.
14. Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One*. 2016;11(7):e0159267.