



Masking in Pragmatic Trials: Who, What, and When to Blind

Jennifer B. Christian, PharmD, MPH, PhD, FISPE¹ ,
Emily S. Brouwer, MPH, PharmD, PhD², Cynthia J. Girman, DrPH, FISPE³,
Dimitri Bennett, MD, MPH, FISPE, FACE⁵, Kourtney J. Davis, PhD, MSPH⁶,
and Nancy A. Dreyer, PhD, MPH, FISPE¹ 

Abstract

Masking (or blinding) of treatment assignment is routinely implemented in classical randomized clinical trials (RCTs) to isolate the effect of the intervention itself and to minimize the potential for bias that could occur with traditional trials. Such biases could be introduced with the conduct, assessment of endpoints, management of conditions, analysis, and reporting when the treatment assignments are known. However, masking of treatments is not only complex but it hinders how generalizable the findings are to the “real world” clinical setting. Pragmatic RCTs (pRCTs) are intended to evaluate the effects of interventions within routine medical care, and as such, do not typically mask treatment groups; moreover, pRCTs assess comparators that are available in routine medical practice, not masked placebos. Whether pRCTs should be masked if intended for regulatory or other purposes has recently been questioned. The literature on pRCTs, while extensive, does not address how much actual benefit is gained from masking outcomes and how masking may affect the “real world” nature of a study. Here, we propose an approach to evaluate sources of bias, describe stakeholders in the conduct of pRCTs who are most likely affected, and offer a framework for considering how masking may be implemented effectively while maintaining generalizability.

Keywords

blinding, masking, pragmatic trials, randomized, clinical trials

Randomized controlled trials (RCTs) are universally accepted as the gold standard for establishing therapeutic effects of an intervention; however, the extent to which results of RCTs can be extrapolated to general populations eligible for the interventions is limited.¹ The challenge for generalizing inference from RCTs is related to the typically homogeneous population, significant control and timing of clinical assessments, and limited concomitant treatment use inherent in RCTs to isolate the effects of interventions under study. Often, these issues can be addressed by designing pragmatic randomized clinical trials (pRCTs) that are intended to be naturalistic and are embedded within routine medical care to inform practice, policy, and patient care decision making. By definition, pRCTs use randomization and are designed to measure the benefits and harms of treatments in real-world routine clinical practice to maximize generalizability, and permit assessment of the entire treatment regimen including variations in behavior and adherence.¹⁻⁴

Recognizing that personalized medicine inherently serves small markets and reliable evidence is required to guide clinical decision making, there is growing interest in making the drug development process more efficient, less costly, and more broadly useful, so as to preserve the incentive for biomedical

innovation.⁵ The 21st Century Cures Act supports this effort by directing the Food and Drug Administration (FDA) to evaluate the potential use of real-world evidence (RWE) for regulatory purposes, with the intent of reducing costs and increasing and accelerating patient access to effective medical products.⁶ In response to this Act, the FDA released a framework for evaluating RWE for use in regulatory decisions, particularly label

¹ Center for Advanced Evidence Generation, IQVIA, Durham, NC, USA

² Takeda Pharmaceuticals International Co, Global Outcomes Research and Epidemiology, Cambridge, MA, USA

³ CERobs Consulting, LLC, Chapel Hill, NC, USA

⁵ Takeda Pharmaceuticals, Epidemiology Department, Cambridge, MA, USA

⁶ Global Epidemiology, Janssen Pharmaceuticals, Titusville, NJ instead of Swarthmore, USA

Submitted 15-Dec-2018; accepted 20-Mar-2019; published online 8-Jan-2020

Corresponding Author:

Jennifer B. Christian, PharmD, MPH, PhD, FISPE, Center for Advanced Evidence Generation, IQVIA, 4820 Emperor Boulevard, Durham, NC 27703, USA.

Email: Jennifer.Christian@iqvia.com

expansions, and intends to use this framework to inform how individual supplemental applications may be evaluated and more generally to guide the advancement of FDA's RWE Program.⁷ The RWE Program involves engagement with stakeholders, the development of internal processes that bring senior leadership input into the evaluation of RWE, development of guidance documents to assist sponsors in using RWE to support their applications, and promotion of shared learning through several demonstration projects. With FDA being a strong proponent of randomized studies for causal inference, pRCTs of two or more active non-inert therapeutic groups have become a popular design for understanding the comparative benefit-risk profiles of products with equipoise in routine clinical practice. However, there is no formal guidance regarding the conduct of pragmatic trials for regulatory decision making, and to date, there has been limited discussion by regulators or in the literature regarding the role of blinding in pRCTs. However, FDA acknowledges in their recently released framework that more objective outcomes may be less prone to bias because of the lack of masking (also known as blinding) and, hence, less likely to weaken confidence in the results.

Masking of treatment assignment is routinely implemented in classical RCTs to increase the signal-to-noise ratio by minimizing the potential for bias in the conduct of the study, assessment of endpoints, management of conditions, analysis, and reporting.⁸ Some researchers and regulators have proposed applying masking of treatment assignment to pRCTs to further minimize the potential for bias and increase the validity of findings. However, masking of treatments is complex and can hinder how generalizable the findings of a trial may be to a "real world," clinical setting.^{1,9} This is particularly challenging for pRCTs where the intent is to enhance generalizability by studying outcomes that can be assessed during ordinary medical care, using practices that mimic aspects of real-world routine medical practice for both HCPs and patients and to compare the effect of the holistic treatment strategy inclusive of real world behavior and adherence. Masking treatments and their comparators in classical phase 2 and phase 3 randomized clinical trial settings is used to isolate the effects of the treatment and increase internal validity by mitigating potential biases associated with health care providers', patients', and researchers' assessment of patient outcomes that may occur with knowledge of treatment assignment.¹⁰ However, if health care providers *are* masked, this limits their ability to manage patients as they typically would, especially in terms of therapeutic decisions and medical advice. Furthermore, implementing masking affects drug supply by requiring special labeling, formulation and distribution of study treatments rather than relying on traditional approaches to refill medicines through a pharmacy. This departure from real world practice adds significant cost and complexity and limits many clinical practices from participating because of the necessary secure storage space and training required to allocate treatment to patients. Much of the extent to which costs are increased depends on whether or not the study uses commercially

available products or the clinical trial supply from the manufacturer; costs are further magnified by the duration of follow-up, treatment expenses, number of sites, and number of patients included in the study.

The literature on pRCTs, while extensive, does not address how to implement masking into a pRCT or whether or when it may not be necessary to mask treatment assignment, likely because many proponents of pRCTs believe masking detracts from the real-world nature of these studies. Here, we propose a construct for decisions during the study design process as to when masking of the treatment assignment may be desirable and specifically who or what needs to be masked; these points to consider may be particularly useful for regulatory or other authorities in evaluating the results of pRCTs in routine clinical practice settings for decision making.

Approaches to Masking in Pragmatic Randomized Trials

Treatment assignment in a classical RCT is masked for study participants, health care providers, data collectors, outcome assessors including adjudication committees, and data analysts.¹¹ The decision to mask is based on the likelihood that one or more of these groups have or may develop opinions about the effectiveness and/or safety of the intervention being investigated relative to the comparators, and these opinions may lead to systematic bias through conscious or unconscious mechanisms. In the specific scenario of a pRCT to generate evidence on comparative effectiveness and safety, the main concern when treatments are not masked is that reporting of outcomes could be differential between the new medicine and usual care medicines with which patients and HCPs have more experience (eg, higher rates of adverse event reporting for patients on a new medicine due to increased scrutiny or attribution). Recognizing that the use of placebo is not recommended in a pRCT because it does not reflect real world decision making about care and treatments and reduces generalizability, other approaches to masking can be implemented when needed in a pragmatic trial to mitigate the potential of bias.¹ We propose the following framework that offers examples on how to minimize bias for each potential concern presented, recognizing that not all outcomes are equally subject to bias and that epidemiologic techniques can be employed to estimate the impact of bias (Table 1).

When to Consider Masking in a Pragmatic Randomized Trial

The importance of masking in randomized trials, pragmatic or otherwise, varies according to the clinical question, outcome of interest, perceptions of the study treatments, and size of the anticipated treatment effect. For some pRCTs, it is impractical or may be unethical to mask patient assignment to the intervention (eg, surgical interventions or oncology treatments). However, in situations where it is ethical and feasible to mask

Table 1. Possible Ways That Bias Could Occur in pRCTs by Stakeholder and Examples of Approaches to Minimize the Impact of Bias.

Stakeholders That Could Be Masked to Treatment	Ways That Bias Can Be Introduced	Approaches to Minimize Bias
Study participants	<ul style="list-style-type: none"> – Participate (or not) in study based on treatment assignment (selection bias) – Differential reporting of symptoms¹² – Use other effective interventions or lifestyle interventions¹³ 	<ul style="list-style-type: none"> – Randomize after patient eligibility assessment and enrollment is complete to minimize selection bias and differential enrollment (a tenet of all RCTs) – Masking may not be needed if outcomes can be objectively assessed – Include and analyze those initiating therapies (new users) vs those continuing usual care therapies to minimize survival and reporting bias
Healthcare providers	<ul style="list-style-type: none"> – Differential management of patients or prescribing of cointerventions – Potential inadvertent influence reporting of subjective outcomes – Influence adherence in follow-up 	<ul style="list-style-type: none"> – Use objective outcomes where possible – Standardized approach and training on outcome ascertainment and reporting – Train on importance of follow-up visits in line with standard of care
Data collectors (study nurses or study coordinators)	<ul style="list-style-type: none"> – Potential for differential “encouragement” during performance testing 	<ul style="list-style-type: none"> – Use wearables or other objective assessments as feasible to avoid human influence, eg, sensor or device-derived data from patient by remote capture – Mask data collectors/study nurses if directly interacting with participants and/or healthcare providers if they could influence outcomes; or use third-party masked raters who are unaware of treatment assignment
Outcome assessors and/or adjudication committee	<ul style="list-style-type: none"> – Differential assessments of outcomes 	<ul style="list-style-type: none"> – Most likely to occur with more subjective outcomes – Mask assessors/committees to treatment assignment when outcomes are subjective
Data analysts	<ul style="list-style-type: none"> – Selection of time points that show the maximum or minimum effects 	<ul style="list-style-type: none"> – Finalize study protocol and statistical analysis plan at the outset of the study

the treatment intervention, the type of outcome of interest may to a large extent drive the decision for the need to mask. The potential for bias associated with knowledge of treatment assignment is higher when the measurement or interpretation of the outcome is subjective and likely to be least biased with objectively measured outcomes, such as mortality or blood test results, which are unlikely to be affected by observer bias.^{9,14,15} The increased misclassification with subjective outcomes may lead to bias away from the null in a pRCT’s effect measures that examines superiority. However, when the treatment effect anticipated is large, such biases are likely to be overwhelmed by the magnitude of the treatment effect, and less likely to substantively change the conclusions of the study, whether the study is masked or not.

In Table 2, we offer a framework that considers different types of outcomes and their qualitative likelihood to bias effect measures of a pRCT. We also provide a recommendation for masking based on the subjective nature of the outcome measure and the likelihood for misclassification. For example, when a clinical event such as all-cause mortality is the primary study outcome, the measurement of the outcome has little subjectivity. Therefore, masking is likely unnecessary in most cases in a pRCT evaluating all-cause mortality as an outcome. pRCTs that evaluate other objective clinical events, like myocardial infarction or hospitalizations, may require more clinical evaluation and interpretation than mortality as an outcome, but the degree

that knowledge of treatment assignment influences the measurement of these events is likely small and the impact, if any, can potentially be evaluated through quantitative methods.¹⁶ This small bias can be mitigated through the use of masked assessment of endpoints, without the complexity and expense of masking the entire trial. Conversely, the evaluation of outcomes that relate to a patients’ perceptions of their physical or mental state, such as pain, physical mobility, anxiety, or depression are much more likely to be subjective, require patient and clinical judgment, and may be highly variable between patients.¹⁵ In these situations, if there is a perception that the investigative treatment is more (or less) effective than the comparator arm, there is greater risk of systematic bias being introduced in a study where subjective outcomes are used. Therefore, for pRCTs that examine more subjective outcomes as primary outcomes or where such outcomes may be significantly influenced by behaviors due to knowledge of treatment assignment, masking of the patient and/or the health care provider to the treatment intervention may be helpful to minimize the potential bias, particularly if the treatment effects being studied are expected to be meaningful but small. That said, masking of treatments for the patient may require the use of a placebo, which has significant consequences on managing the patient in clinical practice and is rarely, if ever, done in pRCTs. Careful consideration should be given to using a harder primary endpoint, exploring direct measurement via remote capture technology (eg, sensor-based measurement) that

Table 2. Framework for Masking Decisions Based on Type of Outcomes in Pragmatic Randomized Trials to Support Labeling Expansion or New Indications.

Outcome Measure	Examples	Potential for Bias	Recommended Treatment Masking		Examples of Other Remediation
			Patient	Clinician	
Clinical event/clinical sign	Death, myocardial infarction, hospitalization, major adverse cardiovascular event (MACE)	Low	No	No	Central adjudication panel could be masked if perception of bias is a concern
Serious adverse events and adverse events of special interest	Serious infections, serious kidney or liver abnormalities	Low	No	No	Must be reported in an appropriate and timely manner to regulatory authorities, and rules should be applied regardless of treatment
Imaging ± (survival, clinical signs)	Bone mineral density, vertebral fractures, spleen volume	Moderately low, although some bias may occur in reading imaging based on reader perceptions of treatments	No	No	Centralized readers can easily be masked to treatment assignment
Chemistry, hematology	HbA1c, pregnancy tests, cholesterol measurements, glomerular filtration rate (GFR), neutropenia	Moderately low, particularly if measures obtained from clinical chart and not patient-reported laboratory values	No	No	
Microbiology	Sustained virologic response, plasma viral load, conversion to negative sputum	Low	No	No	
Physiological/functional measures	6-minute walk, normal sinus rhythm, forced expiratory volume in 1 second (FEV ₁)	Low	No	No	Use wearables as feasible
Clinical outcome assessments	Global clinical assessments or disease-specific rating scales such as ankylosing spondylitis scale and psoriasis severity scale	Possibly high, particularly when clinician perceives one treatment as potentially better	No	Sometimes	Simple real-world proxies may be used, or a masked clinical assessor other than treating physician may do the ratings, as appropriate for study outcome
Patient-reported outcomes by event diaries (counts of observable signs)	Counts of urinary voids, bowel movements, asthma exacerbation defined by medication / health care utilization	Low, event diaries typically assess counts of observable signs, less influenced by knowledge of treatment	No	No	
Patient-reported outcomes of emotional and health state	Perceptions of symptoms, anxiety, fear, depression, pain	High, particularly when patients perceive one treatment as better	Yes, if this is a primary endpoint	No	May be desirable to mask clinician also depending on his or her perception of treatment benefit; if interacting with patients frequently during study

avoids patient report, and reserving potentially subjective endpoints for secondary and exploratory analyses.

Discussion

Traditionally, clinical trialists have conceptualized the masking of interventions as “all or nothing” for RCTs. With the need for innovation, reduced costs, operational efficiency, and more generalizable results in clinical research, pRCTs are increasingly implemented, particularly within claims and electronic medical record distributed networks. Given the real world nature of pRCTs, we propose adopting a mentality of whether the primary endpoint will be significantly influenced by knowledge of treatment assignment, and if so, how such influence may be mitigated.

There are many ways to implement masking of interventions within pRCTs, all of which impose added burden and cost. We have proposed a construct that carefully considers different approaches to masking based on the nature of the endpoint, as well as an assessment of instances where masking may not be necessary. Masking is most desirable when outcome measures involve a great deal of judgment and where decision choices, clinical workup, and reporting may be meaningfully influenced by knowledge of the treatment assignment. Likewise, masking may also be important when primary endpoints involve perceptions reported by patients, particularly those of disease or symptom severity (eg, pain) as opposed to recording of events like incontinence episodes or occurrence of headache.¹⁵ While beyond the scope of this article, other design considerations such as ensuring that there is clinical equipoise in the study treatments being compared and requiring that all patients initiate a new treatment at the point of randomization, rather than maintain current therapies in the usual care arm, are also fundamental issues to consider when conducting a pragmatic trial and will reduce the risk of bias and improve the chances of successful implementation. We have presented several approaches to reduce reporting or assessor bias, yet still allowing for clinical decisions to be made to manage medical conditions in real world routine clinical practice.

It is important to note that pRCTs in real world settings often study multiple treatment groups where the expected relative magnitude of effect is incremental and results need to be considered holistically regarding the differentiated benefit-risk for specific patient groups with unmet needs. Regardless, it is useful to conduct sensitivity analyses that quantify the potential impact of bias on study findings.¹⁶⁻¹⁹

In all cases, the costs and complexity of masking as well as the loss of a “real world” aspect must be carefully weighed against the gains that might be achieved through mitigation of often small but unquantifiable biases that may occur with the lack of masking treatment assignments. A basic approach outlined to minimize bias and confounding at baseline is simply to randomize only after eligibility and informed consent is complete (without knowledge of treatment to be received), thus limiting selection bias at baseline but allowing routine care

subsequently throughout the study. To increase efficiency, these activities can all occur within the same visit but should be applied within a sequence of events such that the health care provider and participant are unaware of the allocated treatment until after the enrollment into the study. Regardless of whether masking is used and what approaches are implemented to mask treatment allocation, study protocols and reports should explicitly state who is masked and how masking will be or was achieved.^{20,21}

In summary, the greatest concern among strong proponents of masking is when there is an expectation on the part of the patient or clinician of how treatments will influence outcomes, and this is of greatest concern when there is subjectivity in the assessment of outcomes. Here we have presented approaches to masking that can be achieved in various ways that can avoid big increases in costs or complexity and maintain the real-world, broadly generalizable nature of pRCTs. In other cases of “harder” outcomes, masking of the clinician and patient may not be necessary, as knowledge of treatment assignment has a low likelihood to influence the results.

Authors' Note

Paper presented at: 34th International Conference on Pharmacoepidemiology & Therapeutic Risk Management, August 22-26, 2018, Prague Congress Centre, Prague, Czech Republic.


Declaration of Conflicting Interests

JC and ND are full-time employees of IQVIA; neither have declarations of potential conflicts of interests related to this article. CJG is an independent consultant for various pharmaceutical companies and owns stock in a small number of biopharmaceutical and biotechnology companies. DB is a full-time employee of Takeda Pharmaceuticals International Co. KJD is a full-time employee of Janssen and holds stock in GSK and Janssen. ESB is a full-time employee of Takeda Pharmaceuticals International Co and holds stock in Takeda Pharmaceuticals.

Funding

No financial support of the research, authorship, and/or publication of this article was declared.

ORCID iD

Jennifer B. Christian, PharmD, MPH, PhD, FISPE  <https://orcid.org/0000-0003-3696-834X>

Nancy A. Dreyer, PhD, MPH, FISPE  <https://orcid.org/0000-0003-0153-8286>

References

1. Caro JJ, Ishak KJ. Optimizing the design of pragmatic trials: key issues remain. *J Compar Effect Res.* 2012;1(4):319-327.
2. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci.* 2011;13(2):217-224.
3. Zuidgeest MGP, Goetz I, Groenwold RHH, et al; on behalf of GetReal Work Package 3. Series: pragmatic trials and real world evidence: paper 1. Introduction. *J Clin Epidemiol.* 2017;88:7-13.

4. Zuidgeest MGP, Welsing PMJ, Van Thiel GJM, et al. Series: pragmatic trials and real world evidence: paper 5. Usual care and real life comparators. *J Clin Epidemiol*. 2017;90:92-98.
5. Dreyer NA. Advancing a framework for regulatory use of real-world evidence: when real is reliable. *Therapeutic Innovation & Regulatory Science*. 2018;52(3):362-368.
6. H.R. 34 Division A – 21st Century Cures. <https://www.congress.gov/bill/114th-congress/house-bill/34>. Accessed May 21, 2018.
7. FDA Real World Evidence Guidance Website. <https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/RealWorldEvidence/UCM627769.pdf>. Accessed December 11, 2018.
8. Fransen GAJ, Van Marrewijk CJ, Mujakovic S, et al. Pragmatic trials in primary care. Methodological challenges and solutions demonstrated by the DIAMOND study. *BMC Med Res Methodol*. 2007;7:16.
9. Berger M, Daniel G, Frank K, et al. A framework for regulatory use of Real-World Evidence. https://healthpolicy.duke.edu/sites/default/files/atoms/files/rwe_white_paper_2017.09.06.pdf. Published 2017.
10. Yelland MJ, Bissett L. Can double blinding distort results in randomized controlled trials? *PM & R*. 2012;4:322-323.
11. Gotzsche PC. Masking during data analysis and writing of manuscripts. *Control Clin Trials*. 1996;17:285-290.
12. Lineberry N, Berlin JA, Mansi B, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. *BMJ*. 2016;355:i5078.
13. The Multiple Risk Factor Intervention Trial (MRFIT). A national study of primary prevention of coronary heart disease. *JAMA*. 1976;235(8):825-827.
14. Day SJ, Altman DG. Blinding in clinical trials and other studies. *BMJ*. 2000;321:504.
15. Hróbjartsson A, Emanuelsson F, Thomsen ASS, 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;43:1272-1283.
16. Lash TL, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol*. 2014;43(6):1969-1985.
17. Thabane L, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: the what, why when and how. *BMC Med Res Methodol*. 2013;13:92-103.
18. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. 2006;15:291-303.
19. Girman CJ, Faries D, Ryan P, et al. Pre-study feasibility and identifying sensitivity analyses for protocol pre-specification in comparative effectiveness research. *J Comp Eff Res*. 2014;3:259-270.
20. Moher D, Schulz KF, Altman D; for the CONSORT Group. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA*. 2001;285:15.
21. Zwarenstein M, Treweek S, Gagnier JJ, et al; for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*. 2008;337:a2390.