Accuracy of baseline prevalence estimates for sample size calculations in obstetrical randomized controlled trials

OBJECTIVE: A sample size calculation is often based on disease prevalence and is a key component in planning and conducting randomized controlled trials (RCTs).[1] Overestimation of disease prevalence in the planning phase of an RCT may make it more feasible because of fewer participants needed for enrollment. However, it may lead

TABLE
Characteristics of randomized controlled trials based on accuracy of estimation of baseline primary outcome rates

Characteristic	Sample size had accurate assessment of prevalence of disease (n = 24)	Sample size had inaccurate assessment of prevalence of disease (n = 92)	*P* value
Journals			
NEJM, JAMA, and Lancet	8 (33.3)	34 (37.0)	.74
AJOG, Obstet Gynecol, and BJOG	16 (66.7)	58 (63.0)	
Centers			
Single	8 (33.3)	36 (36.6)	.58
Multiple	16 (66.7)	55 (60.4)	
Country			
United States alone	8 (33.3)	24 (26.1)	.74
United States and other countries	1 (4.2)	3 (3.3)	
Other countries	15 (62.5)	65 (70.6)	
Baseline rate derivative			
Reference	13 (54.2)	61 (67.8)	.45
Institutional rate	5 (20.8)	12 (13.3)	
No reference	6 (25.0)	17 (18.9)	
Presumed change in baseline rate			
<30%	8 (40.0)	22 (27.5)	.28
>30%	12 (60.0)	58 (72.5)	
Power			
80%–89%	16 (69.6)	68 (73.9)	.67
90%–99%	7 (30.4)	24 (26.1)	
Sample needed	440 (200–1300)	540 (242–1600)	.43
Expected lost to follow-up	5% (0%–10%)	5% (0%–10%)	.64
Individuals approached for trial			
Ineligible	286.5 (151.0–1095.0)	345.5 (45.0–1769.5)	.94
Declined participation	148.0 (37.5–823.5)	103.5 (22.0–481.0)	.52
Lost to follow-up	6 (0–14)	0 (0–15)	.30

Data are presented as number (percentage) or median (interquartile range), unless otherwise indicated.

AJOG, American Journal of Obstetrics & Gynecology; *BJOG*, British Journal of Obstetrics and Gynaecology; *JAMA*, Journal of American Medical Association; *NEJM*, New England Journal of Medicine; *Obstet Gynecol*, Obstetrics & Gynecology.

Ditter. Sample size calculations in obstetrics. Am J Obstet Gynecol 2022.

to increased risks of type II error and a trial that is not completed or noninformative.[2] When factoring in the financial costs of conducting trials and the time commitments and risks undertaken by participants, studies with erroneous sample size calculations may lack clinical usefulness and be considered unethical.[3,4]

This study aimed to ascertain whether the estimated prevalence of the outcome used for sample size calculations is similar to the actual reported prevalence (frequency of primary outcome in the placebo arm). We hypothesized that the estimated prevalence of the primary outcome differs by >10% of what is reported at the completion of the trial in most obstetrical RCTs.

STUDY DESIGN: During a 4-year period (January 2017 to December 2020), we manually identified and abstracted all obstetrical RCTs published in 6 journals (*New England Journal of Medicine*, *Journal of American Medical Association*, *Lancet*, *American Journal of Obstetrics & Gynecology*, *British Journal of Obstetrics and Gynaecology*, and *Obstetrics & Gynecology*). For each trial, we identified the estimated rate of the primary outcome (disease prevalence) used for sample size calculation and compared it to what was found at the completion of the RCT. If the prevalence of the primary outcome in the trial was at least 10% different, then the assumed rate during the study design of the trial was considered to have an inaccurate assessment of prevalence. Chi-square was used for all categorical variables, and $P<.05$ was considered significant.

RESULTS: Of 240 identified obstetrical RCTs, 116 were included in the analysis. The most common reasons for trial exclusion were noninferiority trial design and a primary outcome that did not include disease prevalence (eg, time). Of note, 24 (20.7%) of 64 (55.2%) trials had a primary outcome rate that was within 10% of the actual prevalence, and 92 (80.3%) trials had an inaccurate assessment of the prevalence of disease (Table). There was no difference in the number of individuals approached and recruited between groups. Trials with an accurate assessment of prevalence were no more likely than trials with an inaccurate assessment to rely on published references vs institutional data. Trials with an accurate baseline rate were not more likely to anticipate a change in the baseline rate of >30% for sample size and power calculations. Of the 48 trials that overestimated the prevalence of the disease by >10% and thus increased the risk of a type II error, 7 (14.6%) were positive. Moreover, 44 trials underestimated the disease prevalence by at least 10%. Of these, 16 trials (36.4%) were positive.

CONCLUSION: Approximately 80% of obstetrical RCTs have an inaccurate assessment of the prevalence of the primary outcome when calculating sample size for the trial. Further research to elucidate methods of improving disease estimates is warranted. ■

Kelly C. Ditter, BS
Ellen H. Crowe, BS
Hector Mendez-Figueroa, MD
Department of Obstetrics, Gynecology, and Reproductive Sciences
McGovern Medical School
The University of Texas Health Science Center at Houston
Houston, TX

Stephen Wagner, MD
Department of Obstetrics and Gynecology
Alpert Medical School
Brown University
101 Dudley St.
Providence, RI 02905
stephen_wagner@brown.edu

Megha Gupta, MD, MSC
Department of Obstetrics and Gynecology
Beth Israel Deaconess Medical Center
Boston, MA

Suneet P. Chauhan, MD, Hon DSc
Sean C. Blackwell, MD
Department of Obstetrics, Gynecology, and Reproductive Sciences
McGovern Medical School
The University of Texas Health Science Center at Houston
Houston, TX

The authors report no conflict of interest.

REFERENCES

1. Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. Clin Orthop Relat Res 2008;466:2282–8.
2. Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. JAMA 1994;272:122–4.
3. Bacchetti P, Wolf LE, Segal MR, McCulloch CE. Ethics and sample size. Am J Epidemiol 2005;161:105–10.
4. Speich B, von Niederhäusern B, Schur N, et al. Systematic review on costs and resource use of randomized clinical trials shows a lack of transparent and comprehensive data. J Clin Epidemiol 2018;96:1–11.

 https://doi.org/10.1016/j.ajog.2022.01.032