# The Authors Respond

***To the Editor:***

We welcome the discussion by Huitfeldt and Stensrud[1] on our recent article on generalizing study results. One assumption we listed in the set of sufficient conditions for generalizability was exchangeability between the study sample and the target population, perhaps conditional on a set of covariates $W$ (such that $S \perp Y(x) \mid W$ for $x = a, a'$).[2] Huitfeldt and Stensrud[1] state that conditional exchangeability is expected to hold only when $W$ includes all causes of the outcome whose distribution differ between the study and target populations. They then state that this condition will often not hold in randomized trials. Depending on the setting, we concur that $S \perp Y(x) \mid W$ for $x = a, a'$ may be a strong assumption. However, this assumption is analogous to the conditional exchangeability assumption between treatment arms often made in observational studies. We also note that conditional exchangeability can hold even if $W$ does not include all causes of the outcome whose distribution differ between the study and target populations. For example, consider the single world intervention graph[3] $S \leftarrow W \leftarrow U \rightarrow Y(x)$; here, conditioning on $W$ is sufficient to ensure $S \perp Y(x) \mid W$ yet $W$ is not a cause of the outcome.

The first example Huitfeldt and Stensrud[1] provide of a situation in which they assert the conditional exchangeability assumption may be problematic in fact is among the situations where generalizability is possible: a set of measured, pretreatment covariates thought to be associated with the outcome serve to selectively recruit patients into a trial. As long as some "low-risk" patients are also enrolled to satisfy the positivity assumption that $0 < P(S = 1 \mid W = w)$ for all $w$ such that $0 < P(W = w)$ and the same set of covariates $W$ are measured in the target population, because the set $W$ is known (by virtue of explicitly stated recruitment criteria), either direct standardization or inverse probability weighting could be used to generalize trial results to the target population.

The other example provided by Huitfeldt and Stensrud,[1] in which only men are recruited into a trial on the effect of homeopathy on heart disease and the target population is women, seems to conflate several issues. First, this is an example where methods for transportability rather than generalizability are required, because the study sample is not a subset of the target population. Furthermore, depending on the hypothesized causal relationship between homeopathy and heart disease, this scenario either: (1) violates the positivity assumption if sex is considered to be a covariate in $W$ or (2) is not problematic, if sex is not considered to be a cause of heart disease.[4]

Returning to the assumption of conditional exchangeability between the study sample and the target population, Huitfeldt and Stensrud[1] suggest, as was noted by Cole and Stuart,[5] that the set $W$ could be narrowed to include only covariates that are effect modifiers on a particular scale of interest, when the investigators are willing to specify a scale. This relaxation of the exchangeability assumption is potentially useful in certain situations. However, the set of covariates sufficient to generalize one effect measure (e.g., the risk ratio) under this relaxed assumption will not necessarily be the same set of covariates sufficient to generalize a different effect measure (e.g., the risk difference). Multiple effect measures are often presented in research, making a broader exchangeability assumption more relevant. Furthermore, while relaxing the exchangeability assumption to be conditional only on effect modifiers on a

Funding for this study was provided in part by the National Institutes of Health (DP2 HD084070, R01 AI100654, P30 AI50410, U01 HL121812, U01 AI103390-02, R24 AI067039). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

particular scale may relieve the investigator of measuring some components of $W$ in the target population, $W$ should still be measured in the study sample to assess whether the effect measure of interest is indeed homogeneous across those components of $W$ the investigators would like to avoid measuring in the target population.

**Catherine R. Lesko**
Department of Epidemiology
Johns Hopkins Bloomberg
School of Public Health
Baltimore, MD
clesko2@jhu.edu

**Ashley L. Buchanan**
Department of Pharmacy Practice
College of Pharmacy
University of Rhode Island
Kingston, RI

**Daniel Westreich**

**Jessie K. Edwards**
Department of Epidemiology
University of North Carolina
Chapel Hill, NC

**Michael G. Hudgens**
Department of Biostatistics
University of North Carolina
Chapel Hill, NC

**Stephen R. Cole**
Department of Epidemiology
University of North Carolina
Chapel Hill, NC

## REFERENCES

1. Huitfeldt A, Stensrud MJ. Re: Generalizing study results: a potential outcomes perspective. *Epidemiology.* 2018;29:e13–e14.
2. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing study results: a potential outcomes perspective. *Epidemiology*. 2017;28:553–561.
3. Richardson TS, Robins J. Single World Intervention Graphs (SWIGs): a unification of the counterfactual and graphical approaches to causality. Working Paper Number 128. Center for Statistics and the Social Sciences, University of Washington, 2013.
4. Rozemeijer K, le Cessie S, van Hylckama Vlieg A, et al. Exposure opportunity: the advantages of including men in analyses of female-related risk factors. *Am J Epidemiol*. 2017;185:965–973.
5. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *Am J Epidemiol*. 2010;172:107–115.