

## METHODOLOGY FOR THE ENDOCRINOLOGIST

**Basic aspects of confounding adjustment**Rolf H H Groenwold<sup>1,2</sup> and Olaf M Dekkers<sup>1,3</sup>

<sup>1</sup>Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands, <sup>2</sup>Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands, and <sup>3</sup>Department of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands

Correspondence  
should be addressed  
to R H H Groenwold  
**Email**  
R.H.H.Groenwold@lumc.nl

**Abstract**

The results of observational studies of causal effects are potentially biased due to confounding. Various methods have been proposed to control for confounding in observational studies. Eight basic aspects of confounding adjustment are described, with a focus on correction for confounding through covariate adjustment using regression analysis. These aspects should be considered when planning an observational study of causal effects or when assessing the validity of the results of such a study.

*European Journal of  
Endocrinology*  
(2020) **182**, E5–E7

**Introduction**

Observational studies of potential risk factors for diseases or observational studies into the effects of medical treatments often aim at making causal claims about the risk factor (or treatment) of interest. Is an elevated growth hormone (GH) level a risk factor for mortality? Does medical reduction of GH levels reduce this risk? These are both typical causal questions. To attribute observed differences in the outcome to different levels of the exposure (e.g. high vs low GH levels), these groups should differ only in their exposure levels and should be comparable regarding other characteristics affecting the outcome. In case of absence of such comparability, an observed relation (or lack thereof) may be incorrectly attributed to the exposure and the association is confounded. Here, we describe eight basic principles of confounding adjustment that should be considered when planning or when assessing an observational study of causal effects. This list of principles is not exhaustive, nor do we pretend that all details of the listed aspects are covered. We refer to the references for further reading on confounding adjustment. We use the potential causal relation between GH levels, its treatment, and mortality to exemplify.

Various methods have been proposed to control for confounding in observational studies, ranging from matching to covariate adjustment, propensity score

methods, and instrumental variable analysis. We refer to the literature for more details on these methods (1). Hereafter, we focus on covariate adjustment, meaning the correction for confounding by adding confounders as covariates in a regression model, because that method is most commonly used in clinical research (2).

**Basic aspects****1. Common cause**

Confounders are variables that are related to exposure and, independent of exposure, also affect the outcome. If such variables are ignored, this may lead to incorrect estimates of the exposure-outcome relation. More formally, confounders are common causes of exposure and outcome (3). For example, in acromegaly, tumour size is a common cause of GH levels and mortality, as larger tumours have higher GH levels and also a higher mortality risk (higher risk of ACTH-deficiency). When adjusting for confounding, neither mediators of the exposure effect, nor consequences of exposure and outcome should be adjusted for (4). While adjusting for tumour size (confounder) is appropriate, adjusting for glucose levels is not, as these are affected by

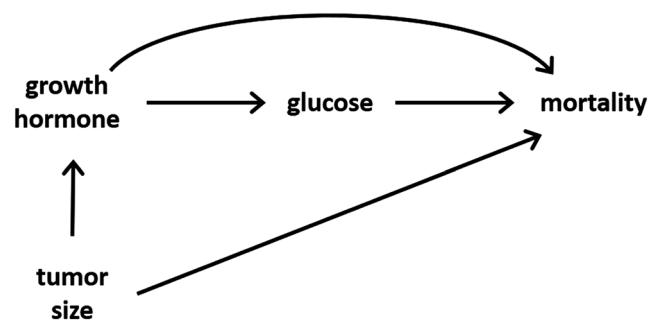
GH levels. Also, adjustment for variables that are predictive of exposure, yet do not affect the outcome (so-called instruments or instrumental variables), is inappropriate, because it may increase the uncertainty in effect estimates (i.e. wider confidence intervals) and amplify any bias due to residual confounding (see 7.) (5). In case of uncertainty about whether a variable is, for example, a mediator or instrument, causal diagrams (see point 2.) or the timing of measurements (see 6.) may be helpful. Regardless, clinical expertise is indispensable for identifying confounders.

## 2. Identification of confounders

Diagrams representing relations between different variables can be helpful to distinguish between confounders, mediators, and consequences of exposure and outcome. An example is given in Fig. 1. Such diagrams help to guide discussions among researchers about whether variables should be considered as confounders (6). From Fig. 1 it can be seen that tumour volume is a confounder of the association between GH levels and mortality in acromegaly, whereas glucose is a mediator and not a confounder. Formal rules can be applied to so-called directed acyclic graphs (DAGs) to find a minimal set of confounders that is sufficient to control for confounding (4).

## 3. Sample size

Formal sample size calculations for observational studies with confounding adjustment are rare. In addition to considerations about the distributions of exposure and outcome, also the number of confounders and their confounding effect need to be considered (7). When



**Figure 1**

Diagram representing the causal relation between growth hormone and mortality in acromegaly. Tumor size is a confounder of the relation between growth hormone levels and mortality. The effect of growth hormone on mortality is (partly) mediated by glucose level, which is a mediator of the relation of interest.

the number of confounding variables is relatively large compared to the number of events (or non-events, whichever is smallest), this may lead to instable statistical models. In that case, the sample size should be increased (if possible), the number of confounders reduced (potentially leading to residual confounding, see 7.), or alternative confounding adjustment methods could be considered. Particularly propensity score methods are a reasonable alternative when the proportion of events is relatively small, while the proportions of both exposed and unexposed are larger. For example, in an observational study assessing the effect of GH lowering drugs on mortality in acromegaly, the number of potential confounders will be large, while the number of events is likely low, thus suggesting that propensity score methods might be a viable option for confounding adjustment.

## 4. Confounder selection

In general, it is advised not to use data-driven methods for the selection of confounders, notably univariate preselection (based on observed confounder-exposure relations) or selection by the change-in-estimate criterion (8). The main reason is that data-driven methods generally cannot distinguish between confounders, mediators, instruments, and common effects of exposure and outcome (see 1). The other disadvantage is that, although the magnitude of the confounding effect of a single variable may be limited, the sum of several confounders with weak confounding effects may well be substantial (9). These confounders, however, will be dismissed by a change-in-estimate approach.

## 5. No unique set of confounders

Each research question and each exposure-outcome relation has its own set of confounders, depending on clinical context and the design of the study (10). The direction and magnitude of the observed confounder-outcome relation may be informative about the apparent validity or quality of the data. For example, if larger tumour size appears to lower risk of mortality, it may indicate that the tumour size measurement is of low quality. However, observed confounder-outcome relations do not necessarily have a causal interpretation, because the relations between each confounder and the outcome may have a different set of confounding variables than the exposure-outcome relation (11).

## 6. Timing of measurements

Many variables – including potential confounders – change over time, in which case the timing of measurements

becomes relevant. Variables that are measured before the start of exposure (or before exposure is measured) could qualify as a confounder, whereas variables measured after exposure started could be affected by exposure and thus be mediators (4). For example, when studying the effect of medical treatment in acromegaly, baseline GH levels can be considered a confounder, while GH levels after treatment will be affected by medication use and should, in general, not be adjusted for. Particularly, in studies of time-dependent exposures (for example when newly occurring pituitary deficiencies are considered), adjustment for confounding becomes challenging (3).

## 7. Residual confounding

Residual confounding due to unmeasured or inaccurately measured confounders is not controlled by conventional methods to handle confounding, including propensity score methods. If, for example, data on ACTH-deficiency are (partially) lacking, adjustment for this variable remains imperfect. Also, incorrectly modelling or categorizing continuous confounders may result in residual confounding (12), which underlines the need to check modelling assumptions and consider alternative approaches to model continuous variables (13). Unfortunately, the potential impact of residual confounding is hardly ever discussed (14, 15), even though various forms of sensitivity analysis have been proposed to guide quantitative discussions about residual confounding (16).

## 8. Other sources of bias

Apart from confounding, there may be other sources of bias in observational studies, including measurement error and missing data. Additional or other methods may be required to address these sources of bias.

### Declaration of interest

O M D is a Deputy Editor for European Journal of Endocrinology. He was not involved in the review or editorial process for this paper, on which he is listed as an author. R G reports no conflicts of interest.

### Funding

This work was supported by grants from the Netherlands Organization for Scientific Research (ZonMW-Vidi project 917.16.430) and the LUMC.

## References

- Klungle OH, Martens EP, Psaty BM, Grobbee DE, Sullivan SD, Stricker BH, Leufkens HG & de Boer A. Methods to assess intended effects of drug treatment in observational studies are reviewed. *Journal of Clinical Epidemiology* 2004 **57** 1223–1231. (<https://doi.org/10.1016/j.jclinepi.2004.03.011>)
- Groenwold RH, Van Deursen AM, Hoes AW & Hak E. Poor quality of reporting confounding bias in observational intervention studies: a systematic review. *Annals of Epidemiology* 2008 **18** 746–751. (<https://doi.org/10.1016/j.annepidem.2008.05.007>)
- Hernan MA & Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2020.
- VanderWeele TJ. Principles of confounder selection. *European Journal of Epidemiology* 2019 **34** 211–219. (<https://doi.org/10.1007/s10654-019-00494-6>)
- Pearl J. Invited commentary: understanding bias amplification. *American Journal of Epidemiology* 2011 **174** 1223–1227; discussion 1228. (<https://doi.org/10.1093/aje/kwr352>)
- Shrier I & Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology* 2008 **8** 70. (<https://doi.org/10.1186/1471-2288-8-70>)
- Hsieh FY, Bloch DA & Larsen MD. A simple method of sample size calculation for linear and logistic regression. *Statistics in Medicine* 1998 **17** 1623–1634. ([https://doi.org/10.1002/\(sici\)1097-0258\(19980730\)17:14<1623::aid-sim871>3.0.co;2-s](https://doi.org/10.1002/(sici)1097-0258(19980730)17:14<1623::aid-sim871>3.0.co;2-s))
- Witte J & Didelez V. Covariate selection strategies for causal inference: classification and comparison. *Biometrical Journal. Biometrische Zeitschrift* 2019 **61** 1270–1289. (<https://doi.org/10.1002/bimj.201700294>)
- Groenwold RH, Sterne JA, Lawlor DA, Moons KG, Hoes AW & Tilling K. Sensitivity analysis for the effects of multiple unmeasured confounders. *Annals of Epidemiology* 2016 **26** 605–611. (<https://doi.org/10.1016/j.annepidem.2016.07.009>)
- Ferguson KD, McCann M, Katikireddi SV, Thomson H, Green MJ, Smith DJ & Lewsey JD. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. *International Journal of Epidemiology* 2019 dyz220 In press. (<https://doi.org/10.1093/ije/dyz220>)
- Westreich D & Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology* 2013 **177** 292–298. (<https://doi.org/10.1093/aje/kws412>)
- Groenwold RH, Klungel OH, Altman DG, van der Graaf Y, Hoes AW, Moons KG & PROTECT WP2 (Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium, Work Programme 2 [Framework for pharmacoepidemiology studies]). Adjustment for continuous confounders: an example of how to prevent residual confounding. *Canadian Medical Association Journal* 2013 **185** 401–406. (<https://doi.org/10.1503/cmaj.120592>)
- Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, Dunkler D, Harrell Jr FE, Royston P & Heinze G. State-of-the-art in selection of variables and functional forms in multivariable analysis – outstanding issues. *arXiv Preprint* 2019. (available at: [arXiv:1907.00786](https://arxiv.org/abs/1907.00786))
- Hemkens LG, Ewald H, Naudet F, Ladanie A, Shaw JG, Sajeev G & Ioannidis JPA. Interpretation of epidemiologic studies very often lacked adequate consideration of confounding. *Journal of Clinical Epidemiology* 2018 **93** 94–102. (<https://doi.org/10.1016/j.jclinepi.2017.09.013>)
- Pouwels KB, Widyakusuma NN, Groenwold RH & Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. *Journal of Clinical Epidemiology* 2016 **69** 217–224. (<https://doi.org/10.1016/j.jclinepi.2015.08.009>)
- VanderWeele TJ & Ding P. Sensitivity analysis in observational research: introducing the E-value. *Annals of Internal Medicine* 2017 **167** 268–274. (<https://doi.org/10.7326/M16-2607>)

Received 27 January 2020

Accepted 17 February 2020