Published in final edited form as:

Epidemiology. 2013 January ; 24(1): 175–176. doi:10.1097/EDE.0b013e3182781410.

Policy-relevant proportions for direct effects

Tyler J. VanderWeele

Departments of Epidemiology and Biostatistics, Harvard School of Public Health

To the editor

Contemporary concepts from the causal inference literature have provided definitions of direct and indirect effects that can be used in models with interactions and non-linearities.^{1,2} So-called controlled direct effects (*CDE*) assess the effect of the exposure on an outcome with the intermediate fixed to a particular level. So-called natural direct effects (*NDE*) and natural indirect effects (*NIE*) allow for the decomposition of a total effect (*TE*) into direct and indirect effect components. Methods are now available to estimate these effects when they are identified.^{3–5} When both direct and indirect effects are employed, a measure of the “proportion mediated” (*PM*) is sometimes calculated as the ratio of the natural indirect effect to the total effect (e.g. $PM = NIE/TE$). This measure in some sense captures how important the pathway through the intermediate is in explaining the actual operation of the effect of the exposure on the outcome. It measures what would happen to the effect of the exposure on the outcome if we were to somehow disable the pathway from the exposure to the intermediate.^{2,6}

The use of natural direct and indirect effects is sometimes criticized because (1) they require very strong assumptions for identification^{2,6,7} and (2) they do not correspond to any particular intervention that we could actually carry out (this is because they require, for each person, fixing the mediator to the counterfactual level the person would have had in the absence of exposure).^{6,7} Such effect estimates do not correspond to actual interventions we could carry out in practice, and are thus of limited interest from a policy perspective.^{6–8}

An alternative proportion measure may be of more interest in policy settings. The controlled direct effect fixing the intermediate to level $M=m$, $CDE(m)$, captures the effect of the exposure on the outcome if the intermediate were set, possibly contrary to fact, to level m . This is an intervention we might hope to be able to carry out in practice. We might hope that by intervening on the intermediate we could block a substantial part of the effect of the exposure on the outcome. A proportion measure that could then be used, and that would be of policy relevance, would be the proportion of the effect of the exposure on the outcome that could be eliminated by intervening to set the intermediate to some fixed level m .¹ This would be $[TE - CDE(m)]/TE$. If this proportion eliminated were large we might try to implement policies to intervene on the intermediate. The proportion mediated essentially captures what would happen to the effect of the exposure if we were to somehow disable the pathway from the exposure to the intermediate (leaving it to its natural value); in contrast the proportion-eliminated measure captures what would happen to the effect of the exposure on the outcome if we were to fix the intermediate to the same fixed value $M=m$ for all persons.

Importantly, the proportion eliminated measure will not always equal the proportion mediated. Suppose that the exposure and the intermediate interacted but that the exposure

had no effect on changing the intermediate itself. In this case the indirect effect of the exposure on the outcome through the intermediate would be 0 (because the exposure does not change the intermediate) and we would have a proportion mediated of $NIE/TE=0/TE=0\%$. However, if, with interaction, the effect of the exposure were large with the intermediate but small without the intermediate, then the proportion eliminated by fixing the intermediate to 0 might be substantial. If we were to fix the intermediate to 0 for everyone, the exposure may not have much of an effect on the outcome [i.e. $CDE(m=0)$ might be quite small], and the proportion eliminated $PE=[TE - CDE(m=0)]/TE$ might be close to 100%. In the extreme case in which there is a pure interaction (so that the exposure has no effect on the outcome unless the intermediate is present) but exposure has no effect on the intermediate itself, then the proportion mediated is 0% while the proportion eliminated by fixing $m=0$ is 100%.

More generally the proportion mediated measure ($PM=NIE/TE$) and the proportion eliminated measure ($PE=[TE - CDE(m)]/TE$) may differ because, in the presence of an interaction between the exposure and the intermediate, we may have a different proportion eliminated measure for every value of m .^{1,9} Because the total effect decomposes into a natural direct effect and natural indirect effect ($TE=NIE+NDE$), we can re-express the proportion mediated as $PM=[TE-NDE]/TE$. In the technical language of the causal inference literature, the proportion eliminated and the proportion mediated may differ because the controlled direct effect may not equal the natural direct effect; the natural direct effect essentially averages over the various controlled direct effects.² The two measures will coincide when there is no interaction between the exposure and the intermediate (either at the individual level³ or at the expected population level under further assumptions⁷).

A biomedical example where these two measures do diverge is the effects of chromosome 15q25 genetic variants on lung cancer, with cigarettes smoked per day as an intermediate.¹⁰ Each variant allele on 15q25 increases the risk of lung cancer by about 1.3 fold; however, a small proportion (perhaps 5%–15%) of this effect is mediated by an increasing cigarettes per day.¹⁰ Each variant allele increases cigarettes per day by about 1, which is not enough to explain the association with lung cancer. However, there is strong interaction between these variants and cigarettes per day in their effects on lung cancer and there may, in fact, be a “pure” interaction such that the variants have no effect on lung cancer for those who do not smoke.¹⁰ The variants may operate by increasing the nicotine and toxins extracted per cigarettes smoked. In this case, the controlled direct effect if we were able to fix cigarettes per day to 0 for everyone would be $CDE(m=0)=0$ and thus the proportion of the effect eliminated would be $PE=[TE-CDE(m=0)]/TE=[TE-0]/TE=100\%$. By eliminating smoking, we eliminate all of the effect of the variants. The proportion mediated (by cigarettes per day) is small because the variants do not increase the number of cigarettes per day substantially; however, the proportion eliminated by fixing cigarettes per day to 0 is large because the variants do not seem to affect lung cancer without smoking.

The two measures, proportion eliminated and proportion mediated, have differing interpretations. The proportion eliminated is in general the more relevant policy measure. It captures how much of the effect of the exposure on the outcome we could eliminate by intervening on the intermediate. The proportion mediated captures how much of the effect of the exposure on the outcome is due to the effect of the exposure on the intermediate. It gives insight into the role of different pathways, but not necessarily on what would happen if we were to intervene on particular intermediates. The proportion-eliminated measure is attractive because it describes the estimated effect of actual policy intervention, and because it requires only the estimation of controlled direct effects, which can be identified in a broader range of circumstances than natural direct and indirect effects.^{11,12} If effect decomposition and evaluation of the operation of various pathways are of interest, natural

direct and indirect effects and the proportion-mediated measure may still be of interest. But for policy purposes, the proportion eliminated is arguably more relevant.

Acknowledgments

Supported by National Institutes of Health grant HD060696.

References

1. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992; 3:143–155. [PubMed: 1576220]
2. Pearl, J. Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*; San Francisco: Morgan Kaufmann; 2001. p. 411–420.
3. VanderWeele TJ, Vansteelandt S. Odds ratios for mediation analysis with a dichotomous outcome. *American Journal of Epidemiology*. 2010; 172:1339–1348. [PubMed: 21036955]
4. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychological Methods*. 2010; 15:309–334. [PubMed: 20954780]
5. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychological Methods*. in press.
6. Robins, JM.; Richardson, TS. Alternative graphical causal models and the identification of direct effects. In: Shrouf, P., editor. *To appear in Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*. Oxford University Press; 2010.
7. Robins, JM. Semantics of causal DAG models and the identification of direct and indirect effects. In: Green, P.; Hjort, NL.; Richardson, S., editors. *Highly structured stochastic systems*. New York: Oxford University Press; 2003. p. 70–81.
8. Hafeman DM, Schwartz S. Opening the Black Box: a motivation for the assessment of mediation. 2009; 38:838–845.
9. Kaufman JS, MacLehose RF, Kaufman S. A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation. *Epidemiologic Perspectives and Innovations*. 2004;1–4. [PubMed: 15507155]
10. VanderWeele TJ, Asomaning K, Tchetgen Tchetgen EJ, Han Y, Spitz MR, Shete S, Wu X, Gaborieau V, Wang Y, McLaughlin J, Hung RJ, Brennan P, Amos CI, Christiani DC, Lin X. Genetic variants on 15q25. 1, smoking and lung cancer: an assessment of mediation and interaction. *American Journal of Epidemiology*. 2012; 175:1013–1020. [PubMed: 22306564]
11. VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009; 20:18–26. [PubMed: 19234398]
12. Vansteelandt S. Estimating direct effects in cohort and case-control studies. *Epidemiology*. 2009; 20:851–860. [PubMed: 19806060]