Manuscript type: Brief report (Methodology paper)

Preventable fraction in the context of disease progression

Authors: Bronner P. Gonçalves[1], Etsuji Suzuki[2, 3]

Affiliations:

[1] Department of Comparative Biomedical Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom

[2] Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan

[3] Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Correspondence:

Bronner P. Gonçalves, Department of Comparative Biomedical Sciences, University of Surrey, Daphne Jackson Rd, Guildford, United Kingdom (b.goncalves@surrey.ac.uk)

Etsuji Suzuki, Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8558, Japan (etsuji-s@cc.okayama-u.ac.jp)

Running head: Preventable fraction and disease progression

Conflict of interest statement: The authors declare no competing interests.

Source of Funding: ES is supported by the Japan Society for the Promotion of Science (JSPS KAKENHI Grant Numbers JP19KK0418 and JP23K09740). The funding sources had no involvement in study design; the collection, analysis, and interpretation of data; the writing of the article; and the decision to submit the article for publication.

Data and computing code availability: Not applicable.

Abstract

The relevance of the epidemiologic concept of preventable fraction to the study of the population-level impact of preventive exposures is unequivocal. Here, we discuss how the preventable fraction can be usefully understood for the class of outcomes that relate to disease progression (e.g., clinical severity given diagnosis), and, under the principal stratification framework, derive an expression for this quantity for this type of outcome. In particular, we show that, in the context of disease progression, the preventable fraction is a function of the effect on the post-diagnosis outcome in the principal stratum in the unexposed group who would have disease regardless of exposure status. This work will facilitate understanding of the contribution of principal effects to the impact of preventive exposures at the population level.

Keywords: preventable fraction, principal stratification, disease progression, disease severity, counterfactual framework

Introduction

The impact of individual-level interventions, or preventive exposures, on population-level disease burden depends on the coverage of these interventions, or correspondingly on the observed proportions of exposures, and on their effects on the risk of adverse health outcomes. This is captured by the counterfactual notion of preventable fraction, which corresponds to the relative reduction in the number of cases that would occur if, contrary to fact, all individuals in a population of interest had received the intervention; in counterfactual notation [1]:

$$\frac{\Pr(Y = 1) - \Pr(Y^1 = 1)}{\Pr(Y = 1)},$$

where $\Pr(Y = 1)$ represents the observed proportion of a clinical outcome Y in the target population and $\Pr(Y^1 = 1)$ corresponds to the outcome proportion that would have been observed had all members of the population been exposed.

In public health research, interest sometimes lies in exposures or interventions that occur, or are assigned, before disease development and that might affect disease progression or severity. For example, when studying infections, post-infection outcomes (e.g., clinical severity), that can occur after infection acquisition, are often of interest [2]. In this context, Hudgens and Halloran [3] discussed and developed methods for the quantification of vaccine effects on binary post-infection outcomes. In particular, they used the principal stratification framework [4], which involves categorisation of individuals based on their joint potential outcomes of a post-treatment variable (e.g., infection status) under different levels of exposure, to argue that comparisons of post-infection outcomes that condition on infection can only be considered causal if restricted to study participants who would be infected regardless of vaccination status – in potential outcomes terms, members of the doomed principal stratum. Another relevant example of a non-communicable disease is the study by Shepherd and colleagues [5] on finasteride and severity of prostate cancer; there, the target effect was defined for individuals who would have prostate cancer if they received or not finasteride.

In this study, for outcomes related to disease progression, we express the preventable fraction in terms of proportions of principal strata and of principal effects. Our objective was to formally assess, and to gain quantitative insights into, the relative contributions of effects on disease progression to the population-level impact of preventive exposures.

Methods

We consider the setting with perfect compliance and no loss of follow-up. For simplicity, we assume no misclassification in disease diagnosis and severity. Note that, despite application of the principal stratification framework, the target population for the preventable fraction corresponds to all individuals included in the hypothetical study. For principal stratification-related parameters, we use notation similar to that used in [2, 3]. It is important to mention here that our objective is different from those of the study by Hudgens and Halloran [3]. In particular, we did not aim to develop estimation methods or sensitivity analyses for principal causal effects.

Let the binary variable Y denote a disease progression outcome (e.g., infection-related hospitalisation), with $Y = 1$ corresponding to severe presentation, and $Y = 0$, to milder disease presentation. Further, let the binary variable A denote exposure of interest (e.g., vaccination status) (1 = exposed, 0 = unexposed). The variable S represents occurrence of disease (1 = presence of disease, 0 = absence of disease), and S_i^a represents the potential disease occurrence outcome for individual i had she or he, possibly contrary to the fact, received exposure level *a*. Under a negative monotonicity assumption of A on S ($S_i^1 \leq S_i^0$ for all i), the population can be partitioned into three principal strata based on the joint potential values of disease occurrence: doomed stratum $\{i \mid S_i^1 = 1, S_i^0 = 1\}$, preventive stratum $\{i \mid S_i^1 = 0, S_i^0 = 1\}$, and immune stratum $\{i \mid S_i^1 = 0, S_i^0 = 0\}$. The probabilities of these strata are

$$\theta_{jl} = \Pr(S^1 = j, S^0 = l).$$

Similarly, Y_i^a corresponds to the potential disease progression outcome for individual i had she or he, possibly contrary to the fact, received exposure level a. Consistent with earlier studies [3, 5], here potential disease progression or severity outcomes, Y_i^a, are not defined for individuals with $S_i^a = 0$, and consequently are not defined for at least one of the two exposure levels in the immune and preventive strata (**Table 1**). Causal effects of A on Y are thus only defined for the doomed stratum.

The parameters $\{\emptyset_{11}, \emptyset_{10}, \emptyset_{01}, \emptyset_{00}\}$ represent probabilities of response types for Y (e.g., a severe clinical presentation) in the doomed principal stratum; that is,

$$\emptyset_{mn} = \Pr(Y^1 = m, Y^0 = n \mid S^1 = 1, S^0 = 1).$$

Note that we do not assume negative monotonicity of *A* on Y; that is, we do not assume that $\emptyset_{10} = 0$. Further, the parameter γ represents the probability that an individual in the preventive principal stratum develops the disease severity outcome had she or he, possibly contrary to the fact, not been exposed, $\gamma = \Pr(Y^0 = 1 | S^1 = 0, S^0 = 1)$; this parameter is definable, despite the effect of A on Y being undefined in this principal stratum. In **Table 1**, response type probabilities (parameter sets θ and $\emptyset$, and γ) are presented for the exposed, unexposed and total populations. For simplicity, we assume no interference, which is plausible for studies on non-communicable conditions, and for infectious diseases it is not infrequent for studies estimating similar quantities to (sometimes implicitly) make this assumption; for recent examples on SARS-CoV-2 vaccination, see [6, 7].

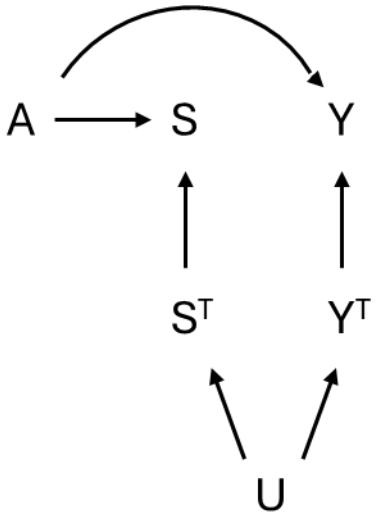
In **Figure 1**, we illustrate possible causal relations between exposure, disease and disease severity outcomes. The diagram corresponds to an extended directed acyclic graph [8, 9], which, in addition to observable variables, includes response type variables. In the diagram, an arrow is not drawn between *S* and *Y*, as *S* does not affect values of *Y*, but only whether it is definable for an individual; this is consistent with discussions in [9-11]. Note that, for simplicity, we do not show (un)measured common causes of *A*, *S*, and *Y* in **Figure 1**.

Furthermore, imputation of 0 for Y_i^a of individuals with $S_i^a = 0$ would imply a different outcome: it would correspond to "occurrence of severe disease" (defined for all individuals; see **eTable 1** for potential outcomes under this definition), rather than "severity of disease among diseased individuals". Below, unless stated otherwise, we use the first definition and assume potential outcomes in **Table 1**.

Table 1. Response types for the post-treatment variable (S) and for the outcome variable (Y) in the context discussed here. We present parameters (response type probabilities) for the exposed, unexposed and total populations; the superscripts in $\theta, \emptyset$ and γ, or their absence, indicate the population and should not be confused with potential outcomes notation.

		S^a		Probabilities of response types (S)				Y^a		Probabilities of response types (Y)		
***SY* type**	**Response type of *S***	S^1	S^0	**Exposed (A = 1)**	**Unexposed (A = 0)**	**Total population**	**Response type of *Y***	Y^1	Y^0	**Exposed (A = 1)**	**Unexposed (A = 0)**	**Total population**
1	Doomed	1	1	θ_{11}^1	θ_{11}^0	θ_{11}	Doomed	1	1	$\theta_{11}^1\emptyset_{11}^1$	$\theta_{11}^0\emptyset_{11}^0$	$\theta_{11}\emptyset_{11}$
2	Doomed	1	1				Causal	1	0	$\theta_{11}^1\emptyset_{10}^1$	$\theta_{11}^0\emptyset_{10}^0$	$\theta_{11}\emptyset_{10}$
3	Doomed	1	1				Preventive	0	1	$\theta_{11}^1\emptyset_{01}^1$	$\theta_{11}^0\emptyset_{01}^0$	$\theta_{11}\emptyset_{01}$
4	Doomed	1	1				Immune	0	0	$\theta_{11}^1\emptyset_{00}^1$	$\theta_{11}^0\emptyset_{00}^0$	$\theta_{11}\emptyset_{00}$
5	Preventive	0	1	θ_{01}^1	θ_{01}^0	θ_{01}	NA	Undefined	1	$\theta_{01}^1\gamma^1$	$\theta_{01}^0\gamma^0$	$\theta_{01}\gamma$
6	Preventive	0	1				NA	Undefined	0	$\theta_{01}^1(1-\gamma^1)$	$\theta_{01}^0(1-\gamma^0)$	$\theta_{01}(1-\gamma)$
7	Immune	0	0	θ_{00}^1	θ_{00}^0	θ_{00}	NA	Undefined	Undefined	θ_{00}^1	θ_{00}^0	θ_{00}
Total				1	1	1				1	1	1

Figure 1. Directed acyclic graph representing a possible causal structure in the disease progression setting. A, S, and Y represent exposure status, disease status, and disease progression outcome, respectively. S^T and Y^T represent response types (joint potential values under different exposure levels) of variables S and Y, respectively; in potential outcomes notation, $S^T = (S^1,\ S^0)$ and $Y^T = (Y^1,\ Y^0)$.



Results

Based on the above, under the consistency condition, the preventable fraction can be expressed as:

$$\frac{\Pr(Y=1)-\Pr(Y^1=1)}{\Pr(Y=1)}$$

$$=\frac{\{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)\}-\{\Pr(Y^1=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y=1|A=0)\Pr(A=0)+\Pr(Y=1|A=1)\Pr(A=1)}$$

$$=\frac{\{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}-\{\Pr(Y^1=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$=\frac{\Pr(A=0)\{\Pr(Y^0=1|A=0)-\Pr(Y^1=1|A=0)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$=\frac{\Pr(A=0)\{\Pr(Y^1=0,Y^0=1|A=0)-\Pr(Y^1=1,Y^0=0|A=0)+\Pr(Y^1=\text{Undefined},Y^0=1|A=0)\}}{\Pr(Y^0=1|A=0)\Pr(A=0)+\Pr(Y^1=1|A=1)\Pr(A=1)}$$

$$=\frac{(1-\pi)\left\{\theta_{11}^0(\emptyset_{01}^0-\emptyset_{10}^0)+\theta_{01}^0\gamma^0\right\}}{(1-\pi)\left\{\theta_{11}^0(\emptyset_{11}^0+\emptyset_{01}^0)+\theta_{01}^0\gamma^0\right\}+\pi\left\{\theta_{11}^1(\emptyset_{11}^1+\emptyset_{10}^1)\right\}},$$

where π represents the proportion of the exposure (i.e., $\Pr(A=1)$). In this scenario, the preventable fraction thus depends on principal stratification parameters: the proportion of the doomed principal stratum in the unexposed group multiplied by the effect of exposure on the disease severity outcome in this stratum (i.e., $\theta_{11}^0(\emptyset_{11}^0+\emptyset_{01}^0)-\theta_{11}^0(\emptyset_{11}^0+\emptyset_{10}^0)=\theta_{11}^0(\emptyset_{01}^0-\emptyset_{10}^0)$); and the proportion of the preventive principal stratum in the unexposed group (i.e., θ_{01}^0) multiplied by the probability of the disease severity outcome (under no exposure) for members of this stratum (i.e., γ^0) – this corresponds to the proportion of the *SY* response type 5 in the unexposed group (**Table 1**). Although the target population of the preventable fraction is the total population, the target of causation is not the total population but the unexposed group [1].

In *eAppendix 1*, we present an additional derivation of the preventable fraction using the alternative outcome definition mentioned in the *Methods* section. Note also that the expression above corresponds to the preventable caseload, not to a proportion [12, 13]. In e*Appendix 2*, we present the corresponding formula for the preventable proportion; under negative monotonicity of A on Y, the preventable caseload becomes equivalent to the preventable proportion [1].

We also present numerical examples (**eTable 2**): the preventable fraction varies, as expected, with changes in the proportion of the preventive stratum in the unexposed group; and it also varies with changes in the magnitude of the principal effect in the unexposed population ($\emptyset_{01}^{0} - \emptyset_{10}^{0}$).

Discussion

In sum, preventable fractions, and attributable fractions [1, 13, 14], allow quantification of population-level impact of exposures, and here we use the principal stratification framework to re-express the former. Whilst formulation of preventable fraction that does not explicitly include a post-treatment variable is useful to answer public health questions, for outcomes related to disease progression, the expression above clarifies the contribution of different components of the protection.

It is important, however, to emphasise that this formulation inherits from the principal stratification the difficulty related to the impossibility of identifying to which principal stratum an individual belongs. To translate this formal work into practice, estimation of the preventable fraction, or of its bounds, in the disease progression setting thus requires further methodological development that for example adapts and combines bounds [15, 16] or sensitivity analyses [3, 17, 18] for principal effects with assumptions on disease principal strata frequencies. Further, it is worth noting limitations of our work: we assume a time-fixed, rather than time-varying, exposure, and did not consider competing events; in many settings, investigators need to consider these factors. Finally, a question for future research, relevant given current public health situation, is on preventable fractions for post-infection outcomes that account for variation in pathogen variant-specific response type probabilities with regard to infection and severity outcomes.

References

1. Suzuki, E. and E. Yamamoto *Attributable fraction and related measures: Conceptual relations in the counterfactual framework*. Journal of Causal Inference, 2023. **11**. 20210068.
2. Halloran, M.E., I.M. Longini, and C.J. Struchiner, *Design and Analysis of Vaccine Studies*. Statistics for Biology and Health. 2010: Springer.
3. Hudgens, M.G. and M.E. Halloran, *Causal Vaccine Effects on Binary Postinfection Outcomes.* J Am Stat Assoc, 2006. **101**(473): p. 51-64.
4. Frangakis, C.E. and D.B. Rubin, *Principal stratification in causal inference.* Biometrics, 2002. **58**(1): p. 21-9.
5. Shepherd, B.E., M.W. Redman, and D.P. Ankerst, *Does Finasteride Affect the Severity of Prostate Cancer? A Causal Sensitivity Analysis.* J Am Stat Assoc, 2008. **103**(484): p. 1392-1404.
6. Consortium, H.U.C., *Undervaccination and severe COVID-19 outcomes: meta-analysis of national cohort studies in England, Northern Ireland, Scotland, and Wales.* Lancet, 2024.
7. Jia, K.M., et al., *Estimated preventable COVID-19-associated deaths due to non-vaccination in the United States.* Eur J Epidemiol, 2023. **38**(11): p. 1125-1128.
8. Suzuki, E., et al., *A counterfactual approach to bias and effect modification in terms of response types.* BMC Med Res Methodol, 2013. **13**: p. 101.
9. Suzuki, E., *Generalized causal measure: the beauty lies in its generality.* Epidemiology, 2015. **26**(4): p. 490-5.
10. Sjolander, A., *Reaction to Pearl's Critique of Principal Stratification.* The International Journal of Biostatistics, 2011. **7**(1).
11. Weuve, J., et al., *Accounting for bias due to selective attrition: the example of smoking and cognitive decline.* Epidemiology, 2012. **23**(1): p. 119-28.
12. Yamada, K. and M. Kuroki *Counterfactual-Based Prevented and Preventable Proportions*. Journal of Causal Inference, 2017. **5**. 20160020.
13. Suzuki, E., E. Yamamoto, and T. Tsuda, *On the relations between excess fraction, attributable fraction, and etiologic fraction.* Am J Epidemiol, 2012. **175**(6): p. 567-75.
14. Suzuki, E. and E. Yamamoto, *Errors in the calculation of the population attributable fraction.* Epidemiology. **in press**. 10.1097/EDE.0000000000001731.
15. Long, D.M. and M.G. Hudgens, *Sharpening bounds on principal effects with covariates.* Biometrics, 2013. **69**(4): p. 812-9.
16. Richardson, A., et al., *Nonparametric Bounds and Sensitivity Analysis of Treatment Effects.* Stat Sci, 2014. **29**(4): p. 596-618.
17. Shepherd, B.E., P.B. Gilbert, and C.T. Dupont, *Sensitivity analyses comparing time-to-event outcomes only existing in a subset selected postrandomization and relaxing monotonicity.* Biometrics, 2011. **67**(3): p. 1100-10.
18. Chiba, Y. and T.J. VanderWeele, *A simple method for principal strata effects when the outcome has been truncated due to death.* Am J Epidemiol, 2011. **173**(7): p. 745-51.

Supplemental Digital Content

Web Appendix (this includes eAppendix 1, eAppendix 2 and eTable 1)

eTable 2