Rejoinder: on the duality of missing data and causal effects

Constantine E. Frangakis, Donald B. Rubin, Ming-Wen An, and Ellen MacKenzie

We thank the editors for the opportunity to have the topic of this paper discussed. We also thank the discussants for their generally interesting comments.

1. Xie and Murphy

Xie and Murphy (XM) describe our problem using only observed-data representations, and then discuss some additional practical issues.

Physics versus pure empiricism. XM present a reduced form of our problem using only the resultant observed data. Motivated from this representation, XM (and also Robins, Rotnitzky, and Vansteelandt - RRV) suggest that one does not need to invoke potential outcomes and principal strata. We disagree: Potential outcomes and principal strata are essential in order to *formulate* the problem and goal, to state explicit assumptions (such as ignorable treatment assignment), and to *devise possible designs* to address the problem. For instance, with only notation for observed data, it is not possible even to define the meaning of a causal effect. That meaning was central in our approach – to regard a value as truly missing (i.e., not observed but observable), only if there exists, in principle, an intervention that would have caused it to be observed. The reader can appreciate the need for potential outcomes also from XM's own writing: when they comment outside of our specific problem, they too invoke potential outcomes and principal strata (see their discussion after question (d) in "practical considerations", where their R_i is defined as the difference of never jointly observable potential outcomes: $S_i(1) - S_i(0)$).

More generally, a representation in terms of potential outcomes and principal strata is required if one is to describe the theoretical, physical underlying system of the problem. Many analogies regarding such physical versus purely empirical representations can be drawn. For example, man went to the moon based on Newton's theoretical, *physical* (albeit not quite correct) model of nature's laws. That voyage would not have been possible if Newton had not persisted in seeking a physical model, but instead had proposed – and if we had accepted as appropriate – some non-differentiable step function (e.g., based on a CART - tree diagram) that would stop after "explaining" empirically only his discrete, few observations.

In summary, postulating a theory in terms of its underlying physics has been, and will continue to be more beneficial than mere explanation in terms of observed data, because a physical system is actually more parsimonious and thus more generalizable, and hence more powerful for predicting other observable events.

On practical considerations. A researcher needs to consider the thoughtful questions (a)-(d) that XM raise, and address them based on the ability to obtain data on factors approximately satisfying our assumptions. An example is question (c): if the prevention factor z is known to be effective, why does the decision maker not administer the most effective level of z to all ? The answer involves obstacles external to the decision maker. Taking, for example, the time to transport an injured patient to the hospital, and adjusting for severity of injury and knowledge that time is important, considerable variation in time can still exist because of other factors: how promptly the injury victim was first spotted and reported; how close the nearest help was; availablity of fast transport at the time; and traffic and other problems encountered by the transport mode. This comment also addresses RRV's point on ethical considerations: variation in such obstacle factors cannot be generally viewed as ethical or not, because these obstacles are rarely in the control of the ethically charged decision maker for z. Of course, Z_i is assigned by the decision maker so as to maximize the *anticipated* likelihood of survival, but this likelihood is only conditional on what the decision maker knows, and so after we condition on that knowledge, we can effectively assume ignorability.

Regarding XM's discussion of more general principal strata, certainly the meaning of the strata $S_i(z)$ depend on the meaning of the prevention factor z, but this is not a complication of principal strata, but a consequence of meanings changing with problems. Within a problem, though, the meaning of $S_i(z)$ *does not* depend on the assignment mechanism for the actual levels Z_i.

XM wonder about the distinction between the covariates we denoted as X and the input factor A, stating that "X precedes both A and Z". This is not generally correct. Some covariate values are determined prior to both A and Z, such as age or gender, but other covariate values, and often those used to ensure ignorability, are determined prior to Z but after A. In our example, X was the severity of injury as judged by the medical personnel *after* the injury occurred, whereas the input variable A was a disability whose value is determined *before the injury*, but only recorded at the interview after the injury.

More important, as we have emphasized in the paper, there is a clear *scientific* distinction between the critical covariates X and the input A: the covariates X used to ensure ignorability need only be those that were involved in the decision maker's informed choice to administer or not the prevention factor (for example, X can often leave out factors causing variation in z such as the obstacle factors just described). The key fact that makes it *easier to record* X than A is this: If the decision maker for z is a person *other* than the injured victim, we can, in principle, talk to that decision maker (whether or not the victim eventually dies) and ask for the value of *all* those variables X that the decision maker used for the assignment of the prevention factor. We cannot do the same for A because, by definition, its accurate measurement depends on the victim's ability to be interviewed, which is impossible if the victim dies.

2. Ten Have

Ten Have commented on the role of an exclusion restriction and the role of covariates, and has indicated numerous directions for possible fruitful extensions to our methods.

On ignorability and exclusion. Ten Have wonders about our assumption of ignorability, that is, $(A, P) \perp\!\!\!\perp Z \mid X$, and its relation to exclusion restrictions typically made in settings of noncompliance. Because the factor A is, by design, an input factor that preceeds the prevention factor z, the value of A cannot be changed (for any person) by changing the level of z. If we had allowed potential outcomes for A under $z = 0, 1$, i.e., $A_i(0), A_i(1)$, then the exclusion restriction $A_i(0) = A_i(1)$ would have been a *consequence* of the temporal ordering of the design, and not an assumption, and that is why we need not make it.

Now, given that A preceeds z, it follows that A, like any other covariate, will be balanced between levels of z after we condition on the variables that were used to make the assignment of the actual levels Z_i - hence the ignorability assumption 1. For that assumption, we disagree with Ten Have's claim that "neither ignorability assumption implies a null ITT effect of Z on A within principal strata": it is not difficult to show that the ignorability assumption 1, i.e., $(A, P) \perp\!\!\!\perp Z \mid X$, implies that $A \perp\!\!\!\perp Z \mid P, X$.

On relations to other problems - what is important in an isomorphism ?

Ten Have observes the similarity of the mathematics in our problem to the mathematics of noncompliance and Zelen's design, a comment related to the isomorphism discussion of RRV.

It is revealing to explicate the source of importance in an isomorphism. To do so, we will invoke a striking example - the isomorphism involved in Gödel's theorem of incompleteness (Gödel, 1931, see also Nagel and Newman, 2001). In brief, Gödel considers axiomatic systems, where axiomatic proofs are constructed by logically building those proofs based on the axioms. Assuming the system is consistent, Gödel then constructs a proposition that has a remarkable duality: (1) there exists no axiomatic proof that the proposition is true, yet (2) there exists a non-axiomatic, but fully valid, proof that the proposition is true, without invoking additional axioms. The relation to our discussion is this: Gödel's theorem was a completely *new* (and unexpected) result, yet it was isomorphic to another well known result, that indefinitely long sequences of natural numbers are uncountable. A somewhat liberal, but useful, explanation of the mapping, provided in the figure, is that the proofs that we can construct axiomatically using finite sentences must be countably many, yet there are uncountably many truths. **Figure 1 here**.

Thus, the critical element in any isomorphism is the intuition that leads us to *see, in the first place,* how one problem – in our case, missing data – could be solved by a design that draws power from an *appropriate* isomorphism – in our case, involving causal inference. The isomorphism here is not really about why instrumental variables work, but much deeper; it involves the relation, discussed in the final section of the target article, between the situation where a quantity can be legitimately viewed as missing, and the requirement that there should exist, at least in principle, an intervention that could have made that quantity observed.

On monotonicity and covariates. Ten Have also asks how important covariates are for identifying parameters under a fully parametric approach. Under the assumptions that allow such identifiability, covariates are as important in the parametric as in the nonparametric formulation, because in order to justify the ignorability assumption 1, we must condition on the covariates involved in the decision maker's choice to administer the prevention factor z. Of course, if monotonicity is relaxed, covariates that predict the *direction of effect of the prevention factor* are also important for narrowing the ranges of the plausible distributions, as Ten Have suggests.

Moreover, Ten Have points towards the connections to the very interesting problem of evaluating a vaccine's efficacy on the viral load for post-randomization infectees - a problem in which the use of principal stratification was initiated by Gilbert et al. (2003). The scientific structure of that problem, as Ten Have observes, is different from this one because an exclusion restriction in the vaccine problem is questionable and should not be assumed a priori.

Finally, it is rewarding to see how quickly Ten Have points to many new fruitful directions and challenges in which such ideas can be useful.

3. Robins, Rotnitzky and Vansteelandt

Robins, Rotnitzky and Vansteelandt (RRV) mainly comment on our assumptions and goal. Their comments about our assumptions are addressable. Their comments about our goal are not relevant to a researcher who interprets the meaning of our result in a scientific context.

i) Addressing RRV's points on assumptions. RRV's comments about our assumptions are addressable because they are assessable. For example, RRV say that physicians know that death can arise from hemorrhage when a thrombolytic drug is given [*versus if it is not given*] after an infarction. That is true but is not relevant to our assumption regarding *timing* of administration: Physicians also know that among infarction victims to whom a thrombolytic drug *is administered*, the sooner the drug is given, the higher the likelihood of survival; in fact there is strong evidence from randomized trials that the probability of hemorrhage is practically zero (est. at 0.2%) when the drug is given within two hours after a myocardial infarction, whereas when the drug is given later than two hours, the probability of hemorrhage is estimated at 2.5%, a relative risk of more than 10 fold (Steg et al. 2003). Thus the note of RRV is incomplete and could mislead the casual reader.

RRV also point to the debate about whether attempting to stabilize injured persons is better before or after transporting them to the hospital. This debate exists but is also not relevant to our assumptions. Obviously those who transport injured victims know whether or not they made such stabilizing attempts. Therefore, when this information is used (e.g., as a stratifier in the variables X that were used to make the decisions regarding transportation), its variation is controlled (in principle) and is no longer a concern. With analogous thought and adjustments, one can address RRVs other concerns.

ii) The importance of using scientific meaning in our goals. It is important to recall the two observations we made under partial preventability (Sec. 5.1). First, we stated that we cannot fully estimate the predictive distribution of mortality given the input factor. Second, we showed that we can still estimate the distributions of the input variable, A, for patients who are protectable by the prevention factor, which also implies we can estimate the distribution of A for patients who are "always survivors". That is, we can estimate

$$P(A_i = 1 \mid P_i = \text{ protectable }) \text{ and } P(A_i = 1 \mid P_i = \text{ always survivor }). \tag{1}$$

RRV's comments on this point (regarding the anti-bird flu drug) essentially repeat – but stop at – our first observation, without paying attention to our second observation. The strata "protectable" and "always survivors" have scientific *meanings*, and are not merely technical abstractions to be averaged over (or not) depending on some statistical goal. In most problems, the strata of "protectable" and "always survivors" are best understood as gradations of a condition in a single underlying system. For example, in the injury problem, this system determines how much injury damage a person can endure at various levels of a treatment factor: An "always survivor" is a person of higher endurance than a "protectable" person, who needs the more effective level of the treatment to survive. In RRV's example, the system determining "protectable" and "always survivor" is the immune system, and "always survivors" are those with generally stronger immune system than the "protectables".

Using the scientific meaning of the principal strata is crucial because it allows us to interpret the estimable comparison in (1) above: if "always survivors" have a higher proportion of the input factor's state $A = 1$ than the "protectables", then this implies that state $A = 1$ is associated with more robust states of the system that determines the principal strata. For example, it would imply that $A = 1$ is associated with higher endurance to injuries, in the injury example; and that $A = 1$ is associated with higher immunity to the virus, in the bird flu example. This conclusion is reachable without needing to identify fully the predictive distribution of mortality, although there is the need to think about the meaning of the principal strata.

Of course, the association between the input factor and the principal strata of protectable versus always survivor does not imply causality. RRV discuss this in length (in their discussion of the psychiatrist's hypothesis), but this issue seems to be entirely obvious. Our goal – to learn about the associations of A and mortality when A is missing under death – provides information to suggest that A is related to (and thus *may be* causally involved in) the system determining the endurance of a person to survive an injury. With this suggestion established, whether the factor A is or is not a causal agent must be addressed with a different design, and cannot be addressed simply by the notational arguments of RRV.

iii) On statistical issues. RRV also re-iterate at length our observation that our problem is related, indeed isomorphic, to other problems involving principal stratification. Although we also could have detailed many additional examples from our own work (e.g., Frangakis et al. 2004; Rubin 2006; Li and Frangakis, 2006; Jin and Rubin, 2007), we believe that is more beneficial to the reader to read our conceptual rejoinder to Ten Have on isomorphisms.

iv) The role of objectivity and sensitivity analysis. We agree that sensitivity analyses can be useful, but the question is how to conduct them. A sensitivity analysis can only be useful to the extent that the framework in which it is formulated is rich enough so that it can provide, at least partly, an objective assessment of the values or ranges of the sensitivity parameters. "Objective" here does not mean "absolutely correct", but it does means "based on assumptions that are understandable", because then, as we have seen from discussions like this (item (i)), one can clarify the ways to assess and address concerns about these assumptions. It is such objectivity that our design and methods provide.

We thank again the editors and discussants for the generally stimulating exchange of ideas, and Spyridon Kotsovilis for insightful discussions on Gödel's theorem.

Additional references.

[1] Frangakis, CE, Brookmeyer, RS, Varadhan, R, Mahboobeh, S, Vlahov, D, and Strathdee, SA. (2004). Methodology for evaluating a partially controlled longitudinal treatment using principal stratification, with application to a Needle Exchange Program. *Journal of the American Statistical Association*, **99** 239–249.

[2] Gödel, K. (1931). Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatshefte für Mathematik und Physik **38** 173-198.

[3] Jin, H, and Rubin DB. (2007). Principal Stratification for Causal Inference with Extended Partial Compliance: Application to Efron-Feldman Data. To appear in the *Journal of the American Statistical Association.*

[4] Li, F and Frangakis, CE (2006). Polydesigns in causal inference. *Biometrics* **62**, 343–351.

[5] Nagel, E. and Newman JR (2001). *Godel's proof.* New York University Press: New York.

[6] Rubin DB. (2006). Causal Inference Through Potential Outcomes and Principal Stratification: Application to Studies with Censoring Due to Death (with discussion). *Statistical Science* **21** 299-321.

[7] Steg PG, Bonnefoy E, Chabaud S, Lapostolle F, Dubien PY, Cristofini P, Leizorovicz A, Touboul P; Comparison of Angioplasty and Prehospital Thrombolysis In acute Myocardial infarction (CAPTIM) Investigators. (2003). Impact of time to treatment on mortality after prehospital fibrinolysis or primary angioplasty: data from the CAPTIM randomized clinical trial. *Circulation* **108** 2851–2856.

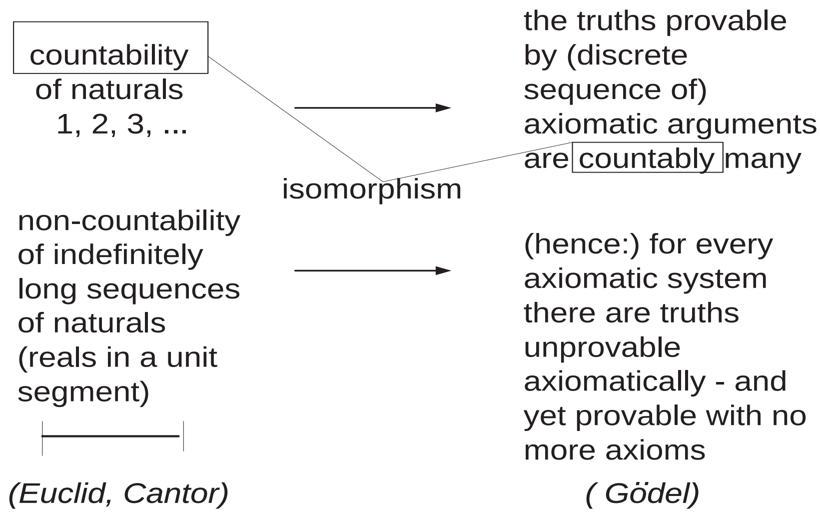

Fig. 1. A remarkable case of isomorphism.