Response to Commentary

Vehement Agreement on New Models?

Teri A. Manolio*, and Rory Collins, for the National Institutes of Health New Models Workshop Participants

* Correspondence to Dr. Teri Manolio, Director, Division of Genomic Medicine, National Human Genome Research Institute, 5635 Fishers Lane, Room 4113, MSC 9305, Rockville, MD 20892-9305 (e-mail: manolio@nih.gov).

Initially submitted October 1, 2012; accepted for publication October 8, 2012.

Dr. Bracken et al. (1) raise several issues in their interpretation of our recent commentary, “New models for large prospective studies: Is there a better way?” (2). Although the 12 points raised by Bracken et al., all of whom are described as associated with the US National Children’s Study, begin with issues related to our commentary, their later points stray into criticisms of the National Children’s Study with which we were not directly concerned. On most of the other points, however, we appear to be in full, even vehement, agreement.

We agree, for example, that harmonizing data from extant databases is difficult, and we explicitly recognized this as one of the greatest potential challenges to a nationwide US cohort study (2). Such databases can be extremely useful, however, in identifying potentially relevant outcomes for further investigation. We agree that selection bias may compromise validity and risk estimation, but only if the study population does not include people with a wide range of backgrounds and exposures. We agree that large cohort studies are inefficient for studying rare exposures and rare diseases, although they are essential for examining risk factors with modest effect sizes and avoiding potential biases of other designs.

We also agree that a probability sampling frame may be useful in generalizing results to the base population as a whole, particularly for deriving estimates of disease incidence, prevalence, or mortality, but we question the importance of a probability sampling frame for risk associations. We note that many nonprobability samples (Framingham Heart Study and the British Doctors’ Study come immediately to mind (3, 4)) have provided crucial insights into risk relations that have been widely generalizable—indeed, it is difficult to find a group to which they do not apply. We agree that simple convenience sampling is likely to lack underrepresented populations almost by definition; hence, the models we espoused critically depend upon capturing diversity in exposures and backgrounds (2), typically through oversampling or targeted assessments.

We agree that research hypotheses can be useful in designing large studies, although we would caution that hypotheses can only be as good as the assumptions on which they are based, and in long-term studies they are likely to be obsolete or downright naive by the time they can be tested. It is unclear to us how we were presumed to “appear to prefer a hypothesis-free approach” (1, p. 286), as the roles of defined hypotheses and research questions were not addressed in our commentary.

We agree on the importance of community engagement and noted the significant effort needed to keep community-based organizations involved and ensure that local concerns are addressed (2). We do question, however, the suggestion that UK Biobank’s 6%–10% enrollment rates reflect a lack of trust rather than the equally plausible explanations of a lack of time, interest, or transportation. Although we are not in a position to comment on the dissatisfaction with the role of the US government in the National Children’s Study expressed by Bracken et al., we find it ironic that this concern comes immediately after mention of “…the Framingham Heart Study, one of the most successful research cohort studies” (1, p. 288). It is often forgotten that the Framingham Heart Study was entirely directed and staffed by the US government for at least the first 2 decades of its existence (5). We also agree that the new models we have described are specifically applicable to very large cohort studies which otherwise would be unaffordable, and clearly endorsed “…the need for, and complementary nature of, a variety of designs and approaches” (2, p. 864).

Lastly, we fully agree that methods used successfully in European studies may not be directly transportable to the United States. We advocated careful piloting and critical evaluation of the feasibility of these models in the United...
States (2). However, the success of these models to date outside the United States and the near impossibility of funding similarly large cohort studies using conventional distributed designs should provoke objective and disinterested US scientists to investigate the adaptability of such models.

ACKNOWLEDGMENTS

Author affiliations: National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland (Teri A. Manolio); and Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Oxford, United Kingdom (Rory Collins).

Conflict of interest: none declared.

REFERENCES