Commentary: Best Practice for What?

SAM HARPER and NICHOLAS B. KING

McGill University

Routine collection, analysis, publication, and interpretation of health data is a key part of public health surveillance (Brookmeyer and Stroup 2004), and the foundation for timely and effective public health interventions. With the emergence of a substantial literature documenting social inequalities in health across a wide range of diseases and environments, and national and international ethical concerns about potential health inequities (WHO Commission on Social Determinants of Health 2008), there is a pressing need to extend monitoring and surveillance systems to health inequalities.

In their stimulating article, Frank and Haw (2011) proposed a set of “best practices” to encourage better monitoring of health inequalities. We heartily agree with the spirit of this effort, particularly the importance of both reasonable completeness and accuracy in reporting and “statistically appropriate” analysis (though there may, of course, be reasonable disagreements about what constitutes statistical appropriateness). But some elements of Frank and Haw’s list of criteria—particularly the ambiguous concepts of “clear relevance,” “sensitivity to policy,” and “avoidance of reverse causation”—appear to conflate two aspects of surveillance that we feel should remain separate: measuring inequalities as part of routine surveillance and interpreting monitoring data as evidence of causation. We suggest that some of their criteria may be useful for interpreting data on health inequalities, but should not be used for setting the objectives of a monitoring system.

One of Frank and Haw’s criteria is “clear relevance to known social determinants of health”—that is, health outcomes that “based on current knowledge . . . reflect life-course stage-specific or cumulative exposures associated with socioeconomic position” (2011, 661). While...
we agree that monitoring health indicators with known socioeconomic gradients is useful, this criterion would be hard to apply in practice. First, “current knowledge” often changes rapidly and rarely reflects a consensus regarding causality (Canning and Bowser 2010; Chandra and Vogl 2010). Limiting outcome monitoring to only those indicators that demonstrate unambiguously “clear relevance” could produce a very short list. Second, ongoing monitoring may be the first way of determining the emergence of socioeconomic gradients for some outcomes, since it is often hard to predict in advance where gradients will arise. For example, if we had adopted Frank and Haw’s criteria in the mid-twentieth century, when some studies suggested weak, nonexistent, or even positive socioeconomic gradients in ischemic heart disease (González, Artalejo, and del Rey Calero 1998; Stamler, Kjelsberg, and Hall 1960), would the absence of a “strong or consistent” relationship to socioeconomic position have led us to de-prioritize or eliminate this outcome? It is also unclear whether Frank and Haw recommend that this criterion be used to determine whether an indicator should be included at all, prioritized (put at the “top of the list”), or simply added to the list of routinely monitored indicators. We support the last use but disagree strongly with the former two. Rather than selecting or prioritizing routinely monitored health outcomes according to their presumed etiology, routine monitoring should provide the data for subsequent investigation of causal effects.

Frank and Haw’s criterion of “reversibility and sensitivity to intervention” also illustrates a potential conflation of routine monitoring and causal interpretation. Choosing to exclusively monitor or prioritize those outcomes currently thought to be “sensitive to policy” risks making surveillance prisoner to transient policy objectives, which are often arbitrary and subject to change. Moreover, it is difficult to predict in advance which outcomes may be responsive to policy or medical innovation. For example, changes in mortality inequalities following the discovery of highly active anti-retroviral therapy for HIV/AIDS (King, Kaufman, and Harper 2010; Rubin, Colen, and Link 2010) or the “Back-to-Sleep” campaigns for infant mortality (Blair et al. 2006; Pickett, Luo, and Lauderdale 2005) would have been difficult to predict in advance.

We have similar concerns about the criterion of “avoiding reverse causation.” Frank and Haw worry that decision makers may falsely attribute the cause of a particular health gradient to social factors rather
than to downward health-related selection. While we agree that this is a valid concern, the proper solution is better education of decision makers, not disqualification of any data that may be misinterpreted. Moreover, monitoring data are unlikely to allow such a fine-grained causal analysis in any case. Their solution—to measure socioeconomic position before the onset of disease—would not necessarily solve this problem since health-related selection could still play a causal role in the observed patterning. This also presumes the existence of an unambiguous consensus about the role of health-related selection for each health indicator, so that for those conditions deemed susceptible to reverse causation, socioeconomic position should be measured before disease onset, but presumably not for those conditions unaffected by reverse causation. This would seem difficult to apply systematically in practice.

Our concerns regarding Frank and Haw’s (2011) assumptions about a clear and unambiguous scientific basis for assessing causation of health inequalities are illustrated by the exchange in this issue of The Milbank Quarterly between the authors and McCartney and colleagues (2013, respectively). Where Frank and Haw find indicators lacking sensitivity to policy change, McCartney and colleagues cite numerous examples of rapid changes in inequalities and suggest that the problem may lie in the policy response rather than the indicators themselves. Where Frank and Haw see reverse causation as a problem for monitoring inequalities in alcohol-related conditions, McCartney and colleagues argue that the evidence for this assertion is weak and unlikely to explain current patterns in Scotland. Furthermore, Frank and Haw find the mental health indicator unresponsive to interventions and associated with small differences by SEP, but McCartney and colleagues provide evidence of its response to intervention and disagree that the magnitude of differences is “unpromising” as an indicator of inequalities.

This exchange illustrates the fact that “current knowledge” of the causal role (or lack thereof) of social determinants of health rarely reflects a consensus. Indeed, such an exchange is possible because current monitoring does not depend on any prior assessment of causality. For this reason, we think that Frank and Haw’s criteria would be most helpful in guiding the interpretation of data on inequalities, rather than assessing which outcomes should be included or prioritized in routine monitoring. Moreover, despite our reservations about some of their criteria, we think that the exchange between Frank and Haw and McCartney and colleagues is productive and is precisely the kind of dynamic discussion
about patterns of, and explanations for, health inequalities that solid monitoring should make possible. And, we would add, a journal like The Milbank Quarterly is an ideal venue.

References


Rubin, M.S., C.G. Colen, and B.G. Link. 2010. Examination of Inequalities in HIV/AIDS Mortality in the United States from a


**Acknowledgments:** This work was supported by Canadian Institutes of Health Research Operating Grant No. 107530. Sam Harper was supported by a Chercheur-boursier Junior 1 salary award from Fonds de recherche Santé Québec.