Annals of Public Health and Research

Commentary

What If we're wrong? Flawed Death Certification Risks Flawed Epidemiology

Peter Pressman[1], Colin Berry[2], Anthony Dayan[1], Roger A Clemens[3], Wallace Hayes A[4]

[1]The Daedalus Foundation, USA
[2]Queen Mary University of London, London
[3]School of Pharmacy & International Center for Regulatory Science, University of Southern California, USA
[4]College of Public health, University of South Florida, USA

*Corresponding author
Peter Pressman, The Daedalus Foundation, USA, Tel: 2022365462; Email: drpressvm2@gmail.com

Submitted: 05 March 2020
Accepted: 17 March 2020
Published: 18 March 2020

COMMENTARY

Much research into public health, whether epidemiological, medical or socioeconomic, is based on modeling techniques which attempt to link the occurrence of a specific disease state treated as a dependent variable with various factors explored as independent variables, usually called 'risk factors'.

If a statistically significant correlation is shown between particular sets of variables, they are often considered to have an important role in determining the prevalence of the disease in any population. This identification of risk can lead to recommendations both for the individual and the community in predicting, preventing, and treating illness, and on the consequential social and economic costs. If, however, our information about the dependent variable were unreliable the analysis and advice might be wrong.

The most common and in many countries the principal source of information about the apparent prevalence of disease and often about risk factors comes from significantly flawed data on causes of death-death certificates.

A substantial body of research over the past 40 or more years and continuing to the present has repeatedly established conventional death certification as imprecise and often erroneous. There is an extensive literature about major errors in conventional recording of the cause of death: see examples listed in Table 1.

There are many potential sources of errors in death certificates. Simple clerical and coding mistakes occur, and the Table 1 illustrates the well-established discrepancies between clinical diagnoses and autopsy findings. There will be differences between disorders diagnosed in general practice and those revealed in clinics and hospitals after more exhaustive investigation of patients. A further notable source of uncertainties and mistakes is the difficulty posed by the complex decisions faced by the physician in classifying primary and secondary factors as causes of death.

The structure of death certificates does differ between countries but in most countries, as in the UK, a definite decision is required about the (principal) "cause" of death; a decision must be made for the process of death certification to be completed. Left ventricular failure due to myocardial ischemia is often a convenient label. Although long standing changes may underlie the immediate event, it is not easy to express nuances and qualifications, in the certification process and so to take account of all the principal disorders contributing to death. Atherosclerosis is likely to have been the pathogenetic process that underlies the diagnosis of left ventricular failure, myocardial infarction, ruptured abdominal aortic aneurysm, cerebral infarction or hemorrhage, renal failure or peripheral vascular disease with infection, amputation and so on. As an immediate cause, left ventricular failure from myocardial fibrosis, as a result of atherosclerosis may be "conveniently" recorded by pathologists as a summary term, notably in coronial autopsies. Atherosclerosis is commonly the pathogenetic process that underlies a diagnosis of left ventricular failure. However, atherosclerosis is also a proximate cause of myocardial infarction and ruptured aortic aneurysm, events that may include additional pathology relating to blood pressure, thrombosis and atheromatous plaque morphology. These conditions give rise to a mode of death resembling the classification of the cause of death made above, judging by autopsy data.

From the publications cited in Table 1, it is clear that conclusions about the prevalence of disease drawn from these data require review before the information from certification alone can safely be used to calculate putative risk. The risk factors calculated will be inaccurate if the dependent variable (cause of death) is simply wrong, perhaps 50% of the time from established data. Uncritically constructed predictive models for major illnesses, the factors that contribute to their causation

Table 1: Errors in Death Certification: Some Recent Publications.

References	Country	Number of Cases	Error (major) Rate %
Corwin, Wolf, Kannel and McNamara 1982	**USA**	280	False negatives 40
			False positives 21
Nielsen, Björnsson and Jonsson 1991	Iceland	433	53
James and Bul 1995	UK	900	14.9
Myers and Farquhar 1998	Canada	229	32.9
Shojana, Burton, McDonald and Goldmann 2003	Review of English language articles in Medline 1996-2002	53 autopsy series from 11 countries	Median 23.5
			range 4.1-49.8
Madea and Rothschild 2010	Germany, Federal data for 2007	818271 deaths	~40-60 depending on cause of death, where patient died and who issued certificate
Lefeuvre, Pavillon, Aouba, Lamarche-Vedel, Fouille, Jougla and Rey 2014	France, National mortality database 2010	552271	51% imprecision depending on criteria applied
Mieno, Tanaka, Arai, Kawahara, Kuchibe, Ishikawa and Sawabe 2016	Japan	450 between 2000-2002	52
McGivern, Shulman, Carney, Shapiro and Bundock 2017	USA	601	51
		From July 2015-January 2016	
Schuppener, Olson and Brooks 2019	Wisconsin, USA	182 between 2013-2016	47
Gao, Calloway, Zhao, Brayne and Matthews 2019	UK	7762 between 2008-2016	2-21 at different times

and the appropriateness of preventative resources will be significantly flawed.

All clinicians know that death certificates, particularly in general practice, tend to be completed under time pressure and the entries will be heavily influenced by the practitioner's expectations based on past experience and beliefs, as well as by the most recent clinical findings in the patient. The immediate cause of death, the final common pathway, is invariably some form of cardiopulmonary failure, but the "primary" disease underlying the death may well be another condition, and so may be the most significant disorder contributing to the final cause of death.

Consider the potential errors that might be introduced over several million deaths per year, for example in reporting heart disease and cancer as the leading underlying causes of death in most developed countries. If heart disease were grossly over-reported, there could be a pervasive impact on the relative funding of public health research and interventions and on advice to the individual patient.

A study population identified as dying from Myocardial Infarction (MI) includes both those dying of MI and others who will have died from other causes. It is a yes/no response type of categorical dependent variable. The independent variables are contributing reasons believed to predict the risk of this cause of death, including obesity, Type II diabetes mellitus, sedentary lifestyle, elevated blood pressure, abnormally high plasma lipid levels and a history of cigarette smoking. These factors may be independent of each other, such that two of them do not necessarily predict a third and may not be associated. Those interactions can be identified and allowed for by appropriate statistical analysis. By this statistical analysis the amount of change in each variable is evaluated 'holding all the other variables constant' against outcome, in this instance fatal MI. Each increment for each variable is evaluated. An overall estimate is obtained that permits identification of the amount of variation in the dependent variable that can be explained by each independent variable, and by the entire model.

The model produced should be able to predict the probability that individuals will die of MI or not in a given period. This does not imply that the variables cause the MI (though that is often the inference drawn and a role in causation may be real). Rather, it means only that those variables appear to contribute to, or are predictive of altered risk of death from MI. Independent variables that do not explain very much, or perhaps little of the variation are discarded. Other variables are tested, but usually in a model including already identified independent variables which do explain significant amounts of variation. It is an iterative, time consuming process.

For most of the past several decades, statistical models have set either a specific cause of death, or a specific event, as the dependent variable (i.e. the outcome is dependent upon the reported levels of the independent variables). If the dependent variable is not precisely and accurately defined, there cannot be adequate qualitative and especially not quantitative confidence in what the independent variables are actually predicting. A great deal of variation may appear to be accounted for by all the independent variables identified, but variation in what?

Accuracy of death certification can be improved in carefully focused studies but they are uncommon as they require costly interventions over a period of time and are thus unsuitable for routine purposes [see, for example, the careful evaluation of population data for deaths from prostate cancer in the European Randomized Study of Prostate Cancer Screening].

Hence the question we pose-What if we're wrong? Research strongly suggests we are, indeed, wrong. These decades of painstakingly constructed models may have relied on an ill-defined or incorrectly recorded data with a fundamental weakness at their heart.

There must be marked consequences for medicine if there were no "well-defined" incident event at the point of death, and one had to rely solely on death certification to identify the fundamental cause of death. There are many cases in which the underlying causes of death may not be well understood, or may have been misidentified. What if those causes of death are also not clearly defined, or under- or over-represented, on death certificates? Again, there may be statistical confidence in the predictive model but if information about the dependent variable is unreliable, the model can only be unreliable [1].

Practitioners are under many sorts of pressure to complete death certificates, often based on limited clinical information, especially in general practice. Many death certificates are accurate, but given that there must be reasonable doubt about many recorded causes of death, there should be concern about the rationale for the direction taken by epidemiological and medical research based on those conclusions. How big might be the effect on society of decisions and actions based on possible incorrect or unfounded conclusions?

Much research effort and funding has been allocated in direct relation to the ranking of causes of death in population surveys. It seems rational to allocate resources and funding in ways commensurate with the magnitude of different causes of death in the population, but would priorities change if the categories were inflated by the errors and uncertainties of death certificates? Heart disease and cancer get the most funding because they appear to be the most important. What if they were not the leading causes of death?

Diagnostic accuracy in medicine continues to improve but the majority of those who die are not in a framework of intensive current investigation; historical data and agonal clinical examination is often all the information available to the certifier. Errors are therefore likely but raw certification data are not reliable. It is tempting to speculate why the inertia of professional medical, administrative and research subcultures and of our gigantic research establishments has given such limited attention to what we already know about the inaccuracies of cause of death reporting. Apart from utilizing and interpreting cause-of-death data with a more critical eye, it may be both timely and valuable to consider the inertia that preserves comfortably familiar but potentially misdirected funding priorities.

Without aggrandizing the prospect of 'Digital Health' one corrective pathway might be greater use of powerful computerized/electronic medical record data collection and interrogation and data analytics and another would be to get better real-time diagnostic information to physicians in their practices and clinics.

REFERENCES

1. Walter SD, de Koning HJ, Hugosson J, Talala K, Roobol MJ, Carlsson S, et al. Impact of cause of death adjudication on the results of the European prostate cancer screening trial. Br J Cancer. 2017; 116: 141-148.

Cite this article

Pressman P, Berry C, Dayan A, Clemens RA, Wallace HA (2020) What If we're wrong? Flawed Death Certification Risks Flawed Epidemiology. Ann Public Health Res 7(1): 1089.