Article from:

Health Watch

May 2015 – Issue 78

When Do Cause and Effect Matter for Health Actuaries?

By Tia Goss Sawhney and Bruce Pyenson

Health actuaries are increasingly being asked to opine on whether a particular health intervention improved health or saved money. Tough questions almost always follow:

- Are the observed results based on the right metrics (a very different question than whether the metrics were calculated correctly)?
- Did the intervention actually cause the observed results, or are they merely correlated? Or perhaps positive results happened but, for some reason, did not appear in the data?
- Do the results make sense?

Actuaries who are charged with figuring out if interventions improve health or save money are stepping into territory where causality rules. Because correlation methods work so well for so much actuarial work, actuaries may not recognize situations where relying only on correlation will get them into trouble.

Fun examples of spurious correlations can be found on a popular website,[1] but the genre long predates the Internet. Statisticians have been warning us against assuming causality for a long time. A widely referenced paper from 1926 has a chart "proving" that fewer marriages in the Church of England caused a decline in the death rate.[2]

And the comedy continues. When a 2013 study by the German Institute for the Study of Labor found that more sexual activity was associated with higher wages,[3] the popular media coverage enthusiastically assumed causality. While *The Washington Post* coverage acknowledged that only correlation may be at play,[4] the science section of *Cosmopolitan* declared that "regrettable one-night-stands are actually helping us save for our European vacays."[5] Perhaps comedians, more than statisticians, are our best defense against such hubris.

Health actuaries often don't pay attention to cause and effect, and mostly they don't need to. For example, historically high medical costs in a region can "cause" high premium rates for policies in that region. The people buying insurance in that region might be sicker (in dimensions not fully reflected in risk adjustment) or the providers might be less efficient. An insurer can be successful in that region without the actuary ever figuring out the reasons for

CONTINUED ON PAGE 28

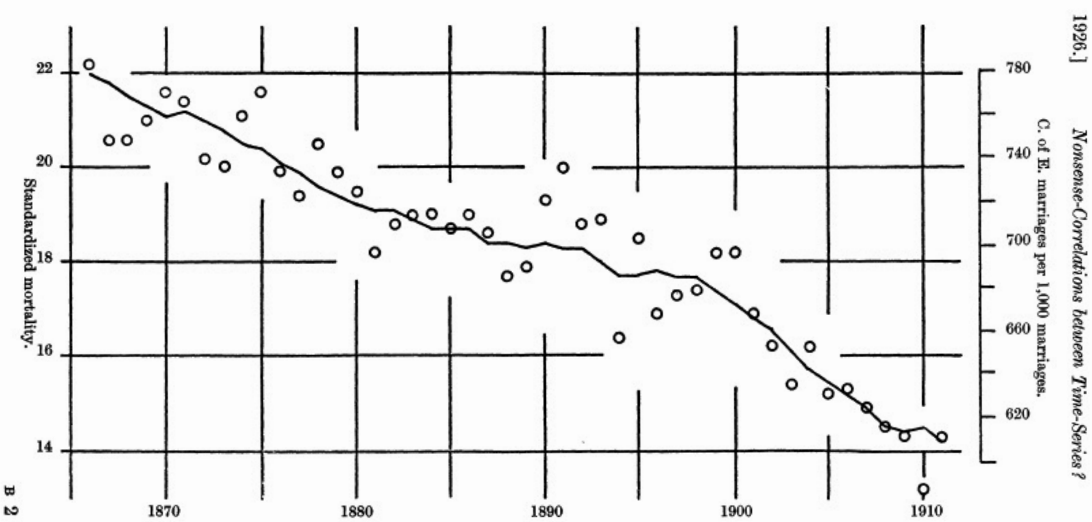

FIG. 1.—Correlation between standardized mortality per 1,000 persons in England and Wales (circles), and the proportion of Church of England marriages per 1,000 of all marriages (line), 1866–1911. $r = +0 \cdot 9512$.

Tia Goss Sawhney, DrPH, FSA, MAAA, is healthcare consultant and actuary with the New York City office of Milliman, Inc. She can be reached at *tia.sawhney@milliman.com*.

Bruce Pyenson, FSA, MAAA, is a principal and consulting actuary with the New York City office of Milliman, Inc. He can be reached at *bruce.pyenson@milliman.com*.

A good program evaluation study builds qualitative and quantitative evidence for causality while ruling out or quantifying the impact of other causes of the same outcome.

the high costs. And when it comes to agreements with providers or vendors that the actuary must evaluate, contract terms become reality—a contract could define success using metrics such as claims processing times or member satisfaction. In this case, the actuary's job is doing the calculation right, as is (hopefully) clearly defined in the contract.

Causality, however, is essential for some tasks assigned to health actuaries; for example, evaluating the potential or historical savings generated by small or large health care interventions. These evaluation studies often involve changing health care delivery, not necessarily health benefits or health insurance. An actuary may be asked to evaluate the claims of vendors selling "solutions" to health care costs. Such vendors seem to bring a simple proposition of "buy my product (or service, device, drug, IT system, network, etc.) and your company will spend less on health care." Vendors may promise to improve the health of the insurer's members (for example, disease management companies), or they may claim to help members avoid expensive services (for example, hospital discharge planning services). On a grander scale, health actuaries are asked to evaluate if big system changes in Medicaid and Medicare health care delivery and payment (such as the promulgation of patient-centered medical homes) have generated savings,[6] or if they are likely to do so.

What's an Actuary to Do?

Big system changes can draw media attention and often involve a lot of money. Should actuaries leave the evaluation of such programs to economists? We feel that actuaries with experience in payment systems, delivery systems and populations can take on the challenges of such evaluations, but they need to pay careful attention to causality.

Suppose an actuary is charged with figuring out if a program that managed "high-risk members" did indeed save money. The actuary might try to answer some of these questions:

1. Was the per capita cost trend for the population lower than budgeted?
2. Was the per capita cost trend for the managed members lower than the trend for the nonmanaged members?
3. Did hospitalizations go down for the managed members? Relative to nonmanaged members?
4. Did costs go down for managed members? Relative to nonmanaged members?

Despite the appeal of a narrow quantitative approach, focusing exclusively on answering these and similar questions can easily produce flawed results.[7]

Doing It Right

A good program evaluation study builds qualitative and quantitative evidence for causality while ruling out or quantifying the impact of other causes of the same outcome. It also looks for special circumstances that might be affecting results.

Causality can seem like peeling an onion—there is always another layer, and by the end you want to cry. For example, cigarettes cause lung cancer, but how? The tobacco industry unscrupulously used the "lack of absolute proof" to continue to promote highly addictive carcinogens for decades after there was overwhelming observational evidence proving tobacco's harms.[8] As painful and unnecessarily long as the cigarettes and lung cancer debate was, extended debate regarding causality is normative. A single study or even group of studies seldom proves causality to everyone's satisfaction. Furthermore, evaluators who acknowledge the limitations of their study have more credibility among those concerned with causation than those who don't.

Fortunately, health actuaries considering causality can find useful and accessible guidance in the epidemiology literature. For more than 100 years, public health professionals, particularly epidemiologists, have been very much concerned with establishing the direct and indirect causes of disease and health, and evaluating the effectiveness of public health interventions. Epidemiology can help actuaries avoid some obvious pitfalls, and help actuaries find out how others have tackled similar problems. Ideally, actuarial analyses that involve causation layer the actuaries' expertise in payer systems and costs upon a solid epidemiologic framework.

Fifty years ago, epidemiologist Bradford Hill proposed a framework for considering causation that is particularly applicable to health actuaries in the form of nine causality criteria.[9] These criteria are presented much like Actuarial Standards of Practice (ASOPs)—a list of issues that need to be seriously considered, not all of which are applicable all the time. The criteria are:

1. **Strong Associations.** The lung cancer rate among cigarette smokers was much higher than among nonsmokers. While strong associations are more indicative of causality, less strong associations can also be causal; for example, uncommon infectious diseases among people who are unusually vulnerable which may be hard to measure due to low sample size, not nonexistent causality.

2. **Consistency**. Associations that are replicated over time and populations are more likely to be causal. Good outcomes from a program in Hartford have more credibility if the program also works in San Diego and Birmingham.

3. **Specificity.** Single types of interventions or exposures (e.g., exposure to one chemical) and single types of outcomes (e.g., one type of cancer) make for stronger causal arguments. On the other hand, it is well known that some risk exposures (e.g., obesity) are linked to many types of illness.

4. **Temporality**. The intervention or exposure must occur before the outcome. Sometimes temporality is challenging; for example, in order to demonstrate that smoking causes lung cancer, it is not sufficient to show that more smokers die from lung cancer as they may have had the cancer before they started smoking. However, prior to lung cancer screening, there was no way to know when the cancer first appeared. This scenario is particularly relevant for potential evaluations of the negative effects of marijuana as marijuana is being used for a host of therapeutic uses, including symptom management for the terminally ill.

5. **Biological gradient**, also known as the dose-response relationship. A smoking cessation program that reaches 50 percent of smokers should have more impact on the population of smokers than one that reaches 1 percent of smokers.

6. **Plausibility**. If a relationship is believable according to current health theory, causality should be considered; absent supporting theory, causality should be questioned. That said, we are always limited by current knowledge, and the history of medicine is filled with practical revolutions that became understandable only years later. Polluted water was recognized as a cause of disease before germ theory. The process of constructing a theoretical model can help identify other variables (confounders) that may be in play. For example, while it is hard to construct a direct causal connection between compliance with cholesterol drugs and safer driving, it is not difficult to construct a causative model where an individual's tendency to comply with expectations and rules impacts both drug compliance and driving safety.[10]

7. **Coherence**. The idea that obesity can cause diabetes coheres with the historical increase in both obesity and diabetes prevalence. Widespread efforts to reduce hospital length of stay cohere with the observed reduction in length of stay.

8. **Experiment.** Experimentation, when possible, bolsters the evidence for causation. Because confounding variables and the intervention itself are better controlled, evidence from experimentation, especially randomized controlled trials, can be particularly strong. When randomized controlled trials are not possible, ethical or practical, however, we must rely on less-than-controlled and even "natural experiments." Sometimes natural experiments are advantageous, as the causality observed in randomized controlled trials may depend on conditions not generally found in the real world.

9. **Analogy**. The epidemiological evidence against cigarettes was overwhelming before the 1960s. Studies on pipe smoking excluding cigarettes were relatively rare, making popu-

CONTINUED ON **PAGE 30**

Our initial Web searches for "actuarial" and "causality" yielded thousands of hits. On inspection, these were almost all related to "causality insurance" and similar mis-entries. Perhaps today's explosion of data requires adding a 10th criterion to Bradford Hill's nine: "Humor. Computer systems can create systematic biases that have strong associations, are plausible, appear coherent, and are wonderfully misleading. Researchers should know when to enjoy a good laugh."

> lation studies of the harms of pipe smoking more difficult. However, the analogy to cigarettes was powerful evidence of causality.[11]

While these criteria seem intuitive, considering them requires considerable investigation and typically far more comprehensive analysis than what would be needed to demonstrate correlation. The successful investigator thinks big, starts with a survey of existing literature and potential theoretical models, follows with a careful analysis, and conclusion.

An overly narrow application of the Bradford Hill criteria can cause mistakes (as is the case with the ASOPs). Neither the Hill criteria nor the ASOPs should be viewed as prescriptive.[12] Suppose an intervention should specifically reduce inpatient hospital costs for a particular population, for example, an emergency room diversion program. An actuary may seem to pay homage to the specificity criteria by examining only inpatient costs. But this could miss a scenario where costs decreased across all services. Such a decline may imply other plan changes or simple regression to the mean caused the inpatient decline.

The many forces that affect health care costs make the temporality criteria particularly challenging. An intervention that happens over one year, such as a healthy diet campaign for diabetics, may be associated with a cost reduction the following year. However, if costs were already decreasing before the intervention or if nondiabetics experienced a similar cost decrease, the apparent causality between the intervention and the cost reduction might be illusory.

Selection (adverse or positive) and regression to the mean are common challenges in causality analyses, and they may be subtle. For example, a population composed of enrollees who had at least one claim in the last year is a select population. In the following year, this population will regress toward the mean of having some people without a claim and hence without cost. Likewise, a population of people who are alive at the end of a year is a select population. Over the next year some portion of the selected population will likely die and have end-of-life costs. We note that patient attribution methods widely used in accountable care organization shared savings programs are often affected by these kinds of selections, and so determining causality among attributed populations can be particularly challenging.

Actuarial Standards and Causality

Causality is rarely mentioned in the actuarial literature. Some of the mentions actually de-emphasize causality. For example, ASOP 12 "Risk Classification"[13] and the Academy's "Risk Classification Statement of Principles" explicitly make the point that risk classification systems do not need to be tied to causality.[14]

The actuary who steps into the world of causality must recognize the need to look outside the ASOPs to do competent work. The growth of provider risk sharing and accountable care organizations means actuaries will be increasingly involved in health care delivery and system changes—an environment where many professionals think about causality. Perhaps it's time to make sure actuaries can recognize and address cause-and-effect situations. We hope actuaries called upon to perform evaluation studies will thoughtfully engage in building solid evidence for or against causality and will not be satisfied with simply reporting correlation, hinting at causality, or warning about potential selection bias. With this in mind, actuaries can be satisfied by giving safer advice than the authors of Cosmo.

END NOTES

[1] Spurious correlations website: *http://www.tylervigen.com/*.

[2] Yule, G.U. Why Do We Sometimes Get Nonsense-Correlations between Time-Series?—A Study in Sampling and the Nature of Time-Series. *Journal of the Royal Statistical Society*, Vol. 89, No. 1. (January 1926), pp. 1-63. Chart reprinted with permission of The Royal Statistical Society.

[3] Drydakis, N. The Effect of Sexual Activity on Wages. Anglia Ruskin University, IZA and Scientific Centre for the Study of Discrimination, Athens. Discussion Paper No. 7529, July 2013: *http://ftp.iza.org/dp7529.pdf*.

[4] DePillis, L. The More Sex You Have, the More Money You Make. *The Washington Post*, Aug. 9, 2013: *http://www.washingtonpost.com/blogs/wonkblog/wp/2013/08/19/the-more-sex-you-have-the-more-money-you-make/*.

[5] The More Sex You Have the More Money You Make: Study Tells Us Those Regrettable One-Night-Stands Are Actually Helping Us Save for Our European Vacays. *Cosmopolitan*, Aug. 19, 2013: *http://www.cosmopolitan.com/sex-love/news/a15244/more-sex-more-money/*.

[6] Sawhney, T.G., and B. Pyenson. Enhanced Primary Care Leads to Reduced Hospital Use and Saves Costs—Or Does It? *Health Watch*, Society of Actuaries. January 2015.

[7] Fairman, K.A., and F.R. Curtiss. Still Looking for Health Outcomes in All the Wrong Places? Misinterpreted Observational Evidence, Medication Adherence Promotion, and Value-Based Insurance Design. *Journal of Managed Care Pharmacy*, Vol. 15, No. 6, July/August 2009: *http://www.amcp.org/data/jmcp/501-507.pdf*.

[8] Muggli, M.E., J.L. Forster, R.D. Hurt, et al. The Smoke You Don't See: Uncovering Tobacco Industry Scientific Strategies Aimed Against Environmental Tobacco Smoke Policies. *American Journal of Public Health*, September 2001, Vol. 91, No. 9, pp. 1419-1423. DOI: 10.2105/AJPH.91.9.1419. *http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1446797/*.

[9] Hill, B.A. The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine* 58 (1965), 295-300 *http://www.edwardtufte.com/tufte/hill*.

[10] Dormuth, C.R., A.R. Patrick, W.H. Shrank, et al. Statin Adherence and Risk of Accidents: A Cautionary Tale. *Circulation*, 2009, 119:2051-2057. DOI 10.1161/CirculationAHA.108.824151: *http://circ.ahajournals.org/content/119/15/2051.full.pdf+html*.

[11] Henley, S.J., M.J. Thun, A. Chao and E.E. Calle. Association Between Exclusive Pipe Smoking and Mortality from Cancer and Other Diseases. *Journal of the National Cancer Institute*, Vol. 96, No. 11, June 2, 2004: *http://jnci.oxfordjournals.org/content/96/11/853.full.pdf+html*.

[12] Actuarial Standards of Practice. Introductory Actuarial Standard of Practice, Actuarial Standard of Practice No. 1. Section 3.1.4, March 2013: *http://www.actuarialstandardsboard.org/pdf/asops/asop001_170.pdf*.

[13] Actuarial Standards of Practice. Risk Classification Actuarial Standard of Practice No. 12. Section 3.2.2, "Causality," December 2005, updated May 1, 2011: *http://www.actuarialstandardsboard.org/pdf/asops/asop012_132.pdf*.

[14] American Academy of Actuaries, "Risk Classification Statement of Principles": *http://www.actuarialstandardsboard.org/pdf/riskclassificationSOP.pdf*.