

But how many people died?

Health outcomes in perspective

BEFORE WE DISPENSE ADVICE about staying healthy, we should know the effect of whatever we are recommending—be it diet, supplements, chemoprevention, or screening—on all meaningful outcomes, including overall mortality, quality of life, harms, inconveniences, and cost. Even though looking at all these outcomes may seem self-evidently wise, many research studies do not do it, and health care providers do not do it enough.

How would looking at all the outcomes change our opinion of health practices?

■ COMPARING GRAPEFRUIT AND PEACHES

A 2013 study linked eating berries with lower rates of myocardial infarction in women,¹ another found that people who ate some fruits (blackberries and grapefruit) but not others (peaches and oranges) had a lower rate of incident diabetes,² and a third linked a healthy diet to a lower incidence of pancreatic cancer.³ However, none of these studies examined all-cause mortality rates. A fourth study found that drinking green tea was associated with a lower risk of death from pneumonia in Japanese women, but not men.⁴

For the sake of argument, let us put aside concern about whether observational studies can reliably inform recommendations for clinical practice⁵ and concede that they can. The point is that studies such as those above look at some but not all meaningful outcomes, undermining the utility of their findings. If healthy people conclude that they should eat grapefruit instead of peaches, they may miss out on benefits of peaches that the study did not examine. Eating a healthy diet remains prudent, but the study linking it to a lower rate of pancreatic

doi:10.3949/ccjm.82a.14058

cancer is no tipping point, as pancreatic cancer is just one way to die. And advocating green tea to Japanese women but not men, to avoid pneumonia, would be a questionable public health strategy. Pneumonia is the sixth leading cause of death and accounts for 3.9% of disability-adjusted life-years lost,⁶ but what about the first five causes, which account for 96.1%?

These and many other studies of dietary habits of people who are well fail to consider end points that healthy people care about. Suppose that drinking more coffee would prevent all deaths from pancreatic cancer but would modestly increase cardiovascular deaths—say, by 5%. On a population level, recommending more coffee would be wrong, because it would result in far more deaths. Suppose that drinking tea decreased deaths from pneumonia—we should still not advise patients to drink tea, as we do not know whether tea's net effect is beneficial.

Some may argue that these epidemiologic studies are merely hypothesis-generating, but my colleagues and I analyzed all the nonrandomized studies published in several leading medical journals in 1 year and found that 59% made specific practice recommendations.⁵ Other studies may be misused in this fashion, even though the authors refrained from doing so.

■ CALCIUM PROTECTS BONES, BUT WHAT ABOUT THE HEART?

Narrow end points are not limited to dietary studies. Calcium supplementation with or without vitamin D has been vigorously promoted for decades⁷ to treat and prevent osteoporosis in pre- and postmenopausal women, and data confirm that these agents decrease the risk of fracture.⁸

We should know the effect of what we recommend on all meaningful outcomes

But bone health is only one end point important to women, and long-term supplementation of a mineral or vitamin with the goal of strengthening bones may have unforeseen adverse effects.

In 2010, calcium supplementation without vitamin D was linked to higher rates of myocardial infarction (with some suggestion of increased rates of all-cause death) in pooled analyses of 15 trials.⁹ In 2011, a higher risk of cardiovascular events (stroke and myocardial infarction) was found in recipients of calcium with vitamin D in a reanalysis of the Women's Health Initiative Calcium/Vitamin D Supplementation Study,¹⁰ adjusting for the widespread use of these supplements at baseline, and this was corroborated by a meta-analysis of eight other studies.¹⁰ A more recent study confirmed that supplemental calcium increases cardiovascular risk in men.¹¹

Although the increase in cardiovascular risk seems to be modest, millions of people take calcium supplements; thus, many people may be harmed. Our exuberance for bone health suggests that, at times, a single outcome can distract.

■ DOES SCREENING IMPROVE SURVIVAL?

On the whole, the evidence for screening continues to focus only on certain outcomes. With the exception of the National Lung Cancer Screening Trial,¹² to date, no cancer screening trial has shown an improvement in the overall survival rate.

In fact, a 2013 Cochrane review¹³ found that mammographic screening failed to lower the rate of death from all cancers, including breast cancer, after 10 years (relative risk [RR] 1.02, 95% confidence interval [CI] 0.95–1.10) and the rate of death from all causes after 13 years (RR 0.99, 95% CI 0.95–1.03). Although screening lowered the breast cancer mortality rate, the authors argued that we should not look at only some outcomes and concluded that “breast cancer mortality was an unreliable outcome” that was biased in favor of screening, mainly because of “differential misclassification of cause of death.”¹³

Black et al¹⁴ found that of 12 major cancer screening trials examining both disease-specific mortality and all-cause mortality, 5 had dif-

ferences in mortality rates that went in opposite directions (eg, the rate of disease-specific mortality improved while overall survival was harmed, or vice-versa), suggesting paradoxical effects. In another 2 studies, differences in all-cause mortality exceeded gains in disease-specific mortality. Thus, in 7 (58%) of the 12 trials, inconsistencies existed between rates of disease-specific mortality and all-cause mortality, prompting doubt about the conclusions of the studies.¹⁴

For some cancers, data suggest that screening increases deaths from other causes, and these extra deaths are not included in the data on disease-specific mortality. For instance, men who are screened for prostate cancer have higher rates of death from cardiovascular disease and suicide,¹⁵ which might negate the tenuous benefits of screening in terms of deaths from prostate cancer.

Studies of screening for diseases other than cancer have also focused on only some outcomes. For example, the United States Preventive Services Task Force supports screening for abdominal aortic aneurysm once with ultrasonography in men ages 65 to 75 who have ever smoked,¹⁶ but the recommendation is based on improvements in the death rate from abdominal aortic aneurysm, not in all-cause mortality.¹⁷ This, along with a declining incidence of this disease and changes in how it is treated (with endovascular repair on the rise and open surgical repair declining), has led some to question if we should continue to screen for it.¹⁸

■ CHEMOPREVENTION: NO FREE LUNCH

Finasteride

In 2013, an analysis¹⁹ that looked at all of the outcomes laid to rest 10 years of debate over the Prostate Cancer Prevention Trial, which had randomized more than 18,000 healthy men over age 55 with no signs or symptoms of prostate cancer to receive finasteride or placebo, with the end point of prostate cancer incidence. The initial results, published in 2003,²⁰ had found that the drug decreased the rate of incident prostate cancer but paradoxically increased the rate of high-grade (Gleason score ≥ 7) tumors. Whether these results were real

Significance-chasing and selective reporting are common in observational studies

or an artifact of ascertainment was debated, as was whether the adverse effects—decreases in sexual potency, libido, and ejaculation—were worth the 25% reduction in prostate cancer incidence.

Much of the debate ended with the 2013 publication, which showed that regardless of finasteride's effect on prostate cancer, overall mortality curves at 18-year follow-up were absolutely indistinguishable.¹⁹ Healthy patients hoping that finasteride will help them live longer or better can be safely told that it does neither.

Statins as primary prevention

As for statin therapy as primary prevention, the best meta-analysis to date (which meticulously excluded secondary-prevention patients after analyzing individual patient-level data) found no improvement in overall mortality despite more than 240,000 patient-years of follow-up.²¹ Because of this, and because the harms of statin therapy are being increasingly (but still poorly) documented, widespread use of statins has been questioned.²²

Proponents point to the ability of statins to reduce end points such as revascularization, stroke, and nonfatal myocardial infarction.²³ But the main question facing healthy users is whether improvement in these end points translates to longer life or better quality of life. These questions remain unresolved.

Aspirin as primary prevention

Another example of the importance of considering all the outcomes is the issue of aspirin as primary prevention.

Enthusiasm for aspirin as primary prevention has been recently reinvigorated, with data showing it can prevent colorectal cancers that overexpress cyclooxygenase-2.²⁴ But a meta-analysis of nine randomized trials of aspirin²⁵ with more than 1,000 participants found that, although aspirin decreases the rate of nonfatal myocardial infarction (odds ratio [OR] 0.80, 95% CI 0.67–0.96), it does not significantly reduce cancer mortality (OR 0.93, 95% CI 0.84–1.03), and it increases the risk of non-trivial bleeding (OR 1.31; 95% CI 1.14–1.50). Its effects on overall mortality were not statistically significant but were possibly favorable (OR 0.94, 95% CI 0.88–1.00), so this requires further study.

After broad consideration of the risks and benefits of aspirin, the US Food and Drug Administration has issued a statement that aspirin is not recommended as primary prevention.²⁶

■ WHY STUDIES LOOK ONLY AT SOME OUTCOMES

There are many reasons why researchers favor examining some outcomes over others, but there is no clear justification for ignoring overall mortality. Overall mortality should routinely be examined in large population studies of diet and supplements and in trials of medications²⁷ and cancer screening.

With regard to large observational studies, it is hard to understand why some would not include survival analyses, unless the results would fail to support the study's hypothesis. In fact, some studies do report overall survival results,²⁸ but others do not. The omission of overall survival in large data-set research should raise concerns of multiple hypothesis testing and selective reporting. Eating peaches as opposed to grapefruit may not be associated with differences in rates of all-cause mortality, myocardial infarction, pneumonia, or lung cancer, but if you look at 20 different variables, chances are that one will have a *P* value less than .05, and an investigator might be tempted to report it as statistically significant and even meaningful.

Empirical studies support this claim. One group found that for 80% of ingredients randomly selected from a cookbook, there existed Medline-indexed articles assessing cancer risk, with 65% of studies finding nominally significant differences in the risk of some type of cancer.²⁹

An excess of significant findings such as this argues that significance-chasing and selective reporting are common in this field and has led to calls for methodologic improvements, including routine falsification testing³⁰ and upfront registration of observational studies.³¹

■ WHY ALL OUTCOMES MATTER

Healthy people do not care about *some* outcomes; they care about *all* outcomes. Some patients may truly have unique priorities (quality of life vs quantity of life), but others may overestimate their risk of death from some

Healthy people
do not care
about *some*
outcomes;
they care about
all outcomes

causes and underestimate their risk from others, and practitioners have the obligation to counsel them appropriately.

For instance, a patient who watches a brother pass away from pancreatic adenocarcinoma may wish to do everything possible to avoid that illness. But often, as in this case, fear may surpass risk. The patient's risk of pancreatic cancer is no different than that in the general population: the best data show³² an odds ratio of 1.8, with a confidence interval spanning 1. As such, pancreatic cancer is still not among his five most likely causes of death.

Some patients may care about their bone mineral density or cholesterol level. But again, physicians have an obligation to direct patients' attention to all of the outcomes that should be of interest to them.

■ OBJECTIONS TO INCLUDING ALL OUTCOMES

There are important objections to the argument I am presenting here.

First, including all outcomes is expensive. For studies involving retrospective analysis of existing data, looking at overall mortality would not incur additional costs, only an additional analysis. But for prospective trials to have statistical power to detect a difference in overall mortality, larger sample sizes or longer follow-up might be needed—either of which would add to the cost.

In chemoprevention trials, the rate of incident cancer has been called the gold standard end point.³³ To design a thrifty chemoprevention study, investigators can either target a broad population and aim for incident malignancy, or target a restricted, high-risk population and aim for overall mortality. The latter is preferable because although it can inform the decisions of only some people, the former cannot inform any people, as was seen with difficulties in interpreting the Prostate Cancer Prevention Trial and trials showing reduced breast cancer incidence from tamoxifen, raloxifene, and exemestane.

In large cancer screening trials, the cost of powering the trial for overall mortality would be greater, and though a carefully selected, high-risk population can be enrolled, historically this has not been popular. In can-

cer screening, it is a mistake to contrast the costs of trials powered for overall mortality with those of lesser studies examining disease-specific death. Instead, we must consider the larger societal costs incurred by cancer screening that does not truly improve quantity or quality of life.³⁴

The recent reversal of recommendations for prostate-specific antigen testing by the United States Preventive Services Task Force³⁵ suggests that erroneous recommendations, practiced for decades, can cost society hundreds of billions of dollars but fail to improve meaningful outcomes.

The history of medicine is replete with examples of widely recommended practices and interventions that not only failed to improve the outcomes they claimed to improve, but at times increased the rate of all-cause mortality or carried harms that far outweighed benefits.^{36,37} The costs of conducting research to fully understand all outcomes are only a fraction of the costs of a practice that is widely disseminated.³⁸

A second objection to my analysis is that there is more to life than survival, and outcomes besides overall mortality are important. This is a self-evident truth. That an intervention improves the rate of overall mortality is neither necessary nor sufficient for its recommendation. Practices may improve survival but worsen quality of life to such a degree that they should not be recommended. Conversely, practices that improve quality of life should be endorsed even if they fail to prolong life.

Thus, overall mortality and quality of life must be considered together, but the end points that are favored currently (disease-specific death, incident cancer, diabetes mellitus, myocardial infarction) do not do a good job of capturing either. Disease-specific death is not meaningful to any patient if deaths from other causes are increased so that overall mortality is unchanged. Furthermore, preventing a diagnosis of cancer or diabetes may offer some psychological comfort, but well-crafted quality-of-life instruments are best suited to capture just how great that benefit is and whether it justifies the cost of such interventions, particularly if the rate of survival is unchanged.

Preventing stroke or myocardial infarction is important, but we should be cautious of

The history of medicine is replete with practices that harmed more than helped

interpreting data when decreasing the rate of these morbid events does not lead to commensurate improvements in survival. Alternatively, if morbid events are truly avoided but survival analyses are underpowered, quality-of-life measurements should demonstrate the benefit. But the end points currently used capture neither survival nor quality of life in a meaningful way.

■ WHEN ADVISING HEALTHY PEOPLE

Looking at all outcomes is important when caring for patients who are sick, but even

more so for patients who are well. We need to know an intervention has a net benefit before we recommend it to a healthy person. Overall mortality should be reported routinely in this population, particularly in settings where the cost to do so is trivial (ie, in observational studies). Designers of thrifty trials should try to include people at high risk and power the trial for definite end points, rather than being broadly inclusive and reaching disputed conclusions. Research and decision-making should look at all outcomes. Healthy people deserve no less. ■

■ REFERENCES:

- Cassidy A, Mukamal KJ, Liu L, Franz M, Eliassen AH, Rimm EB. High anthocyanin intake is associated with a reduced risk of myocardial infarction in young and middle-aged women. *Circulation* 2013; 127:188–196.
- Muraki I, Imamura F, Manson JE, et al. Fruit consumption and risk of type 2 diabetes: results from three prospective longitudinal cohort studies. *BMJ* 2013; 347:f5001.
- Arem H, Reedy J, Sampson J, et al. The Healthy Eating Index 2005 and risk for pancreatic cancer in the NIH-AARP study. *J Natl Cancer Inst* 2013; 105:1298–1305.
- Watanabe I, Kuriyama S, Kakizaki M, et al. Green tea and death from pneumonia in Japan: the Ohsaki cohort study. *Am J Clin Nutr* 2009; 90:672–679.
- Prasad V, Jorgenson J, Ioannidis JP, Cifu A. Observational studies often make clinical practice recommendations: an empirical evaluation of authors' attitudes. *J Clin Epidemiol* 2013; 66:361–366.e4.
- Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; 380:2197–2223.
- Eastell R. Treatment of postmenopausal osteoporosis. *N Engl J Med* 1998; 338:736–746.
- Looker AC. Interaction of science, consumer practices and policy: calcium and bone health as a case study. *J Nutr* 2003; 133:1987S–1991S.
- Bolland MJ, Avenell A, Baron JA, et al. Effect of calcium supplements on risk of myocardial infarction and cardiovascular events: meta-analysis. *BMJ* 2010; 341:c3691.
- Bolland MJ, Grey A, Avenell A, Gamble GD, Reid IR. Calcium supplements with or without vitamin D and risk of cardiovascular events: reanalysis of the Women's Health Initiative limited access dataset and meta-analysis. *BMJ* 2011; 342:d2040.
- Xiao Q, Murphy RA, Houston DK, Harris TB, Chow WH, Park Y. Dietary and supplemental calcium intake and cardiovascular disease mortality: the National Institutes of Health-AARP diet and health study. *JAMA Intern Med* 2013; 173:639–646.
- The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011; 365:395–409.
- Gøtzsche PC, Jørgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013 Jun 4; 6:CD001877.
- Black WC, Haggstrom DA, Welch HG. All-cause mortality in randomized trials of cancer screening. *J Natl Cancer Inst* 2002; 94:167–173.
- Fall K, Fang F, Mucci LA, et al. Immediate risk for cardiovascular events and suicide following a prostate cancer diagnosis: prospective cohort study. *PLoS Med* 2009; 6:e1000197.
- Prasad V. An unmeasured harm of screening. *Arch Intern Med* 2012; 172:1442–1443.
- Guirguis-Blake JM, Beil TL, Senger CA, Whitlock EP. Ultrasonography screening for abdominal aortic aneurysms: a systematic evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2014; 160:321–329.
- Harris R, Sheridan S, Kinsinger L. Time to rethink screening for abdominal aortic aneurysm? *Arch Intern Med* 2012; 172:1462–1463.
- Thompson IM Jr, Goodman PJ, Tangen CM, et al. Long-term survival of participants in the prostate cancer prevention trial. *N Engl J Med* 2013; 369:603–610.
- Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med* 2003; 349:216–224.
- Ray KK, Seshasai SR, Erqou S, et al. Statins and all-cause mortality in high-risk primary prevention: a meta-analysis of 11 randomized controlled trials involving 65,229 participants. *Arch Intern Med* 2010; 170:1024–1031.
- Redberg RF, Katz MH. Healthy men should not take statins. *JAMA* 2012; 307:1491–1492.
- McEvoy JW, Blumenthal RS, Blaha MJ. Statin therapy for hyperlipidemia. *JAMA* 2013; 310:1184–1185.
- Chan AT, Ogino S, Fuchs CS. Aspirin and the risk of colorectal cancer in relation to the expression of COX-2. *N Engl J Med* 2007; 356:2131–2142.
- Seshasai SR, Wijesuriya S, Sivakumaran R, et al. Effect of aspirin on vascular and nonvascular outcomes: meta-analysis of randomized controlled trials. *Arch Intern Med* 2012; 172:209–216.
- US Food and Drug Administration. Use of aspirin for primary prevention of heart attack and stroke. www.fda.gov/Drugs/ResourcesForYou/Consumers/ucm390574.htm. Accessed February 5, 2015.
- Ioannidis JP. Mega-trials for blockbuster. *JAMA* 2013; 309:239–240.
- Dunkler D, Dehghan M, Teo KK, et al; ONTARGET Investigators. Diet and kidney disease in high-risk individuals with type 2 diabetes mellitus. *JAMA Intern Med* 2013; 173:1682–1692.
- Schoenfeld JD, Ioannidis JP. Is everything we eat associated with cancer? A systematic cookbook review. *Am J Clin Nutr* 2013; 97:127–134.
- Prasad V, Jena AB. Prespecified falsification end points: can they validate true observational associations? *JAMA* 2013; 309:241–242.
- Ioannidis JPA. The importance of potential studies that have not existed and registration of observational data sets. *JAMA* 2012; 308:575–576.
- Klein AP, Brune KA, Petersen GM, et al. Prospective risk of pancreatic cancer in familial pancreatic cancer kindreds. *Cancer Res* 2004; 64:2634–2638.
- William WN Jr, Papadimitrakopoulou VA. Optimizing biomarkers and endpoints in oral cancer chemoprevention trials. *Cancer Prev Res (Phila)* 2013; 6:375–378.
- Prasad V. Powering cancer screening for overall mortality. *Ecanermedicallscience* 2013 Oct 9; 7:ed27.
- US Preventive Services Task Force. Final recommendation statement. Prostate cancer: screening. <http://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/prostate-cancer-screening>. Accessed February 5, 2015.
- Prasad V, Cifu A, Ioannidis JP. Reversals of established medical practices: evidence to abandon ship. *JAMA* 2012; 307:37–38.
- Prasad V, Vandross A, Toomey C, et al. A decade of reversal: an analysis of 146 contradicted medical practices. *Mayo Clin Proc* 2013; 88:790–798.
- Elshaug AG, Garber AM. How CER could pay for itself—insights from vertebral fracture treatments. *N Engl J Med* 2011; 364:1390–1393.

ADDRESS: Vinay Prasad, MD, National Cancer Institute, National Institutes of Health, 10 Center Drive, 10/12N226, Bethesda, MD 20892; e-mail: vinayak.prasad@nih.gov