

The Health Checkup

Was It Ever Effective? Could It Be Effective?

Harold C. Sox, MD

IN THIS ISSUE OF *JAMA*, THE *JAMA* CLINICAL EVIDENCE Synopsis¹ by Krogsbøll et al² summarizes a systematic review of 9 randomized trials of a general health check. A general health check is defined as health care motivated by the need to assess general health and prevent future illness rather than to attend to symptoms. Screening tests are performed and counseling may occur during a general health check. The Cochrane systematic review³ on which this Clinical Evidence Synopsis is based concluded that a general health check was not associated with an increase or decrease in total mortality.

The US Medicare program has only recently supported screening and prevention, starting with the Omnibus Budget Reconciliation Act of 1990. Medicare Part B now offers a one-time “welcome to Medicare” preventive visit and yearly “wellness” visits. The purpose of both is to assess risks and prevent disease. These visits are not subject to a co-payment, and the physician may bill Medicare for them. Part B of Medicare also covers screening, counseling services, and immunizations that are recommended by the US Preventive Services Task Force. More recently, the Affordable Care Act stipulated that newly insured adults and children are eligible for a large number of proven preventive services without requiring a co-payment or coinsurance. Just as preventive care is gaining traction with payers, the Cochrane systematic review appears to challenge these insurance reforms. Or does it?

A good way to think about the null main result of the Cochrane systematic review is to analyze internal validity (the causal relationship inferred from the data is the only plausible explanation for the results) and external validity (whether the results apply to current medical practice).

The meta-analysis found that general health checks were associated with a 0.99 relative risk of death from any cause compared with no general health check, with a narrow 95% confidence interval.³ When interpreting a null result, the main concern is a type II error: the failure to detect a true difference. A type II error is usually due to a small, deviant sample of outcome events; this does not seem to be a prob-

lem with the 9 trials, which included 11 940 deaths over 9 years of follow-up.

The main concern about using this result to conclude that the health check is ineffective is the possibility that most of the 11 940 deaths were from causes that the health checks could not have prevented. The leading causes of death in people younger than 65 years—similar to enrollees in these trials—are unintentional injury, heart disease, cancer, suicide, and homicide, which account for 60% of deaths up to age 65.⁴ The health check trials screened for heart disease risk factors but not the other common causes of death and did not provide counseling about avoiding violent death. Moreover, the annual incidence of the specific diseases addressed by the health check is very small in healthy individuals. For example, the annual incidence of the target cancers in the randomized trials of screening for colorectal and breast cancer is approximately 1 in 1000.^{5,6} In the health check trials, the number of healthy individuals who died from a condition targeted by the screening intervention and could therefore have benefited from early detection was probably quite small relative to those who died of causes that the screening intervention did not target effectively, such as cancer and unintentional injury.

A second possible explanation for a false-negative result is that poor execution of the trials biased the results toward the null. This explanation seems unlikely. In most of the trials, participants were identified from population lists and randomized before they were contacted about enrolling. This method avoids bias in randomization. The uptake of the intervention was 80% to 90%; whether nonadherent study participants differed from the enrolled population is not known. All-cause mortality, the primary outcome in the health check trials, is an unambiguous outcome that reflects harms and benefits, avoids bias in ascertaining subjective outcomes from unblinded participants, and uses public records of vital status. By selecting all-cause mortality as the primary end point, the trialists helped ensure an unbiased assessment of the number of outcomes, but they increased the chance of a type II error.

The general health check trials have limited application to contemporary practice because the researchers used

Author Affiliation: Department of Medicine and the Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire.

Corresponding Author: Harold C. Sox, MD, 31 Faraway Ln, West Lebanon, NH 03784 (hsox@comcast.net).

See also p 2489.

screening tests for which the effects on outcomes were, and still are, largely unknown. Seven of the 9 trials were performed before 1975, well before randomized trials of cancer screening tests showed that sigmoidoscopy, fecal occult blood testing, chest CT scans, and mammography reduced mortality from the target cancer. Of the 9 trials included in the systematic review, 1 included sigmoidoscopy, 1 included mammography, and 1 included cervical cytology. Consequently, these interventions would have benefited far too few individuals to affect total mortality as measured in the systematic review. The most commonly performed screening studies were measurement of blood pressure, serum cholesterol levels, and height and weight, as well as performance of electrocardiography, spirometry, and urinalysis. These screening procedures have not been subjected to a randomized trial testing their effects on mortality. The 9 trials used screening tests whose effects on mortality are still unknown and failed to use screening tests now known to be effective.

The design of the systems to deliver screening services also affects applicability to contemporary primary care. Most of the studies took place in a special unit designed to perform efficient delivery of screening and counseling services to thousands of study participants. With few exceptions, the delivery of the service took place in parallel with primary care, not in primary care. This arrangement may have been necessary for an efficient community-based screening program, but was probably a barrier to coordinating follow-up care and perhaps to effective counseling. Most of the interventions were delivered once, not periodically. These health checks are not what the architects of Medicare policy and the authors of the Affordable Care Act had in mind.

What are the lessons from this body of evidence? Most important, the findings do not necessarily apply to contemporary practice and Medicare policy, for 2 reasons. First, the systematic review describes trials of 1-time multiphasic testing, not annual disease prevention visits with continuity provided by a primary care clinician. Second, the trials tested a weak package of screening tests. With the exception of a few trials, the health checks did not include cancer screening tests that are now known to reduce cancer-specific mortality. In contrast, the health check trials included screening tests that target heart disease. The intervention in all 9

trials included blood pressure and serum cholesterol measurement, and 7 trials included a resting electrocardiogram. Eight trials measured cardiovascular mortality. These trials constitute a test of the package of these 3 screening tests, and were associated with a 1.03 relative risk (95% CI, 0.91-1.17) of death from heart disease.

This review also raises questions about whether all-cause mortality is an appropriate primary outcome for a diverse set of screening tests. The main disadvantage is dilution of the effects of effective interventions by deaths from causes like violence that screening interventions do not target. Measuring cause-specific mortality will help identify the effective components of the general health check.

Improving the health of a population will require integration of the public health system and the medical care system to achieve disease prevention, early detection of disease, and effective treatment across the lifespan. The results of the Cochrane systematic review summarized in this JAMA Clinical Evidence Synopsis² should not discourage clinical researchers from designing and testing a large-scale community-based preventive intervention that is closely aligned with the current medical care system. The investigators who implemented the programs studied in these nine trials thought on a grand scale. That type of thinking is needed to plan and develop systems to optimize population health in the 21st century.

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

REFERENCES

1. McDermott MM, Livingston EH. Introducing JAMA Clinical Evidence Synopsis: from systematic reviews to clinical practice. *JAMA*. 2013;309:89.
2. Krogsbøll LT, Jørgensen KJ, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *JAMA*. 2013;309(23):2489-2490.
3. Krogsbøll LT, Jørgensen KJ, Grønhøj Larsen C, Gøtzsche PC. General health checks in adults for reducing morbidity and mortality from disease. *Cochrane Database Syst Rev*. 2012;10:CD009009. doi:10.1002/14651858.CD009009.pub.2.
4. Injury prevention and control: data and statistics (WISQARS). Centers for Disease Control and Prevention. <http://www.cdc.gov/injury/wisqars/index.html>. Accessed May 31, 2013.
5. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med*. 2000;343(22):1603-1607.
6. Humphrey LL, Helfand M, Chan BKS, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2002;137(5 pt 1):347-360.