DNA Cytometry Testing for Cervical Cancer Screening:
Approaches and Reporting Standards for New Technologies

Graham A. Colditz1 and John Crowley2

Abstract

Evaluation of new technologies requires rigorous methods to provide unbiased estimates of the performance and so inform future clinical practice. We review evidence on DNA cytometry reported earlier in this journal and point to the standards for reporting of diagnostic accuracy as a metric against which this article can be evaluated. The cross-sectional nature of the data and incomplete reporting limit the clinical utility of the study. With application of improved reporting standards for diagnostic tests and improved design and evaluation of new technologies for screening, we may better inform practices to improve clinical outcomes and population health. Clin Cancer Res; 17(22): 6971–2. ©2011 AACR.

In an earlier issue of Clinical Cancer Research, Tong and colleagues published in 2009 suggested that DNA cytometry might be beneficial as a screening strategy compared with conventional cytology in developing settings (1). For their study, the authors completed a flow diagram showing that out of 25,000 women evaluated for eligibility, 22,993 underwent randomization. Individual women were randomized to the order of testing; group 1 had cytometry followed by cytology and group 2 had cytology followed by cytometry. All women at entry had a cervical scrape with cervix brush, and smears were classified according to the Bethesda System terminology. DNA cytometric assessments were also conducted at the same initial visit. In the primary report, data were presented for all women regardless of the order of testing when rates of positive test results and of cancer were presented (in their Table 2). The primary outcome of the trial was diagnosis of cervical cancer through the screening process. Secondary outcomes included cervical intraepithelial neoplasia 1 to 3 and death. Almost all of the women received both tests, but 48 received cytologic testing only and 46 received DNA cytometry only. This use of both tests on almost all participants adds difficulty to the interpretation of the trial results. The primary presentation ignores the randomization and the order of testing, and the cancer endpoint was reported as a rate per all women receiving the specific test (40/21,471 DNA cytometry; 24/21,693 cytologic testing, \( P = 0.04 \)). The authors report 1 year follow-up of participants by telephone. It is unclear what, if any, data were collected through this follow-up.

As there is much interest in the evaluation of new technologies and in particular the use of technologies for cervical cancer screening (2), it is important to consider the standards that have been proposed for reporting and interpretation of studies evaluating these technologies. In 2003, a report addressing standards for reporting of diagnostic accuracy (STARD) encouraged investigators to more thoroughly report the details of studies to improve interpretation for researchers and for clinicians and the public (3). Although numerous threats to internal and external validity of diagnostic tests exist, the primary goal of improving reporting is to help readers detect any bias in a study. Important areas for reporting include (i) details on the participants, including details of recruitment; and (ii) the index test and the reference standard. With regard to the test methods, details of the test, criteria for positivity (including cut points), and the blinding (or not) of readers of the index test and the reference test to the results of the other test are necessary (3). For participants, one must report the beginning and end dates, the clinical and demographic characteristics, and the time frame from index test to the reference standard. Cross tabulations of the index test against the results of the reference standard are needed along with details of any missing data. Estimates of diagnostic accuracy should be reported, as well as how indeterminate results, missing data, and outliers of the index test are handled. If possible, an estimate of test reproducibility should be given.

An additional requirement of all randomized trials is that the data be reported and analyzed according to the randomized arm, referred to as the intent-to-treat principle. Within the context of the reported trial, Tong and colleagues compared 2 tests but did not use a gold standard as the comparator; thus, sensitivity of the tests cannot be
Clinical Cancer Research

Translational Relevance

Evaluation of emerging technology requires rigorous methods to gain unbiased estimates of the contribution of technology above existing standard care. Assessment of reporting on DNA cytometry identifies numerous gaps in design and reporting that may inform interpretation and guide clinical application or future studies. Applying rigorous reporting standards will improve evaluation of new technologies and speed implementation of effective advances that will improve clinical outcomes and population health.

estimated. The details of the population are somewhat lacking, and in particular the screening history (if any) of the women prior to implementing the program clouds consideration of the test as a true screening test. Women were excluded if they self reported a history of cervical cancer.

Although some of the appropriate data for reporting of diagnostic test evaluation were added through a response to correspondence (1, 4), the underlying data from the reported trial fail to provide full details expected by the STARD criteria. If we compare this report against the STARD criteria in the context of screening, we note that a diagnostic test evaluation would require that all subjects receive the gold standard or reference test. This may not be suitable in the context of screening programs when the majority of the population is free from disease. This is the classic problem of how we identify true negatives from false negatives when screening for cancer. Importantly, the goal of screening is the reduction of cancer mortality. Reducing advanced-stage disease may also be used as a primary endpoint in some settings (typically when treatment of advanced disease has little or no clinical benefit). Introduction of a new test that increases the detection of early-stage lesions makes incidence of premalignant and early-stage lesions not particularly relevant as endpoints. For cervical cancer, this is relevant, as new technology makes incidence of premalignant and early-stage lesions not particularly relevant as endpoints. Accordingly, Arbyn and colleagues recommend that evaluation of screening technologies focus on the reduction of cancer in subsequent screening rounds rather than settling for cross-sectional evaluation of test performance (2).

With regard to the intent-to-treat principle, in the data reported by randomized arm (Table 1 of ref. 1), we see that 32 of 11,832 women in the cytometry arm were found to have cancer, as opposed to 22 of 11,859 in the cytology arm, with a $P = 0.17$, quite a different outcome and conclusion.

A final concern in evaluation of screening technology is the use of cross-sectional performance versus longitudinal performance and outcomes. The trial conducted by Tong and colleagues seems to report incidence of cancer from the screening test only. This, then, is only a cross-sectional evaluation. Improving performance in a cross-sectional study does not imply that adding the test to a screening program will necessarily lead to reduction in incidence of aggressive or lethal cancers. Accordingly, Arbyn and colleagues recommend that evaluation of screening technologies focus on the reduction of cancer in subsequent screening rounds rather than settling for cross-sectional evaluation of test performance (2).

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Grant Support

G.A. Colditz is supported by P30CA091842 and the Barnes-Jewish Hospital Foundation. He is also supported in part by an American Cancer Society Cissy Hornung Clinical Research Professorship.

Received July 19, 2011; revised August 30, 2011; accepted August 30, 2011; published OnlineFirst September 22, 2011.

References
