

RESPONSE

Nicolas Wentzensen, Barbara Fetterman, Philip Castle, Mark Schiffman, Shannon Wood, Diane Tokugawa, Clara Bodelon, Nancy Poitras, Tom Lorey, Walter Kinney

Affiliations of authors: Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD (NW, MS, SW, CB); Kaiser Permanente TPMG Regional Laboratory, Berkeley, CA (BF, DT, NP, TL); Global Coalition Against Cervical Cancer, Arlington, VA (PC); Albert Einstein College of Medicine, Bronx, NY (PC); Division of Gynecologic Oncology, Kaiser Permanente Medical Care Program, Oakland, CA (WK).

Correspondence to: Nicolas Wentzensen MD, PhD, MS, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Room 7-E114, Bethesda, MD 20892-9774 (e-mail: wentzenn@mail.nih.gov).

In his correspondence, Dr. Garner discusses the implications of evaluating dual stain (DS) performance in a population of human papillomavirus (HPV)-positive women rather than a general screening population. He raises an important point about the population risk and the portability of clinical test performance estimates. We disagree, however, with his conclusion that our analysis is biased.

The restriction of our analysis to HPV-positive women is supported by the underlying natural history of cervical cancer. HPV infection is the necessary cause of almost all cervical cancers; screening is increasingly shifting to primary HPV testing (1,2). HPV-negative women are at very low risk of cervical cancer, allowing extending screening intervals (3). However, the majority of HPV infections disappear after a short period of time. The goal of triage assays is to find women at highest risk of precancer among HPV-positives (4). We can restrict the triage test to HPV-positive women and thereby focus our efforts on the approximately 10% of the screened population that includes almost all women with disease, rather than doing a second test on all women undergoing screening (Figure 1). This step-wise approach is a fundamental principle of many screening approaches with smaller populations and higher risk of disease at each step.

As outlined by Dr. Garner, we evaluated a combination of HPV testing and DS, using the logical operator “AND,” and the sequence of HPV testing followed by DS (5). It is well known that PPV and NPV are dependent on prior risk and cannot be compared between populations with different risk levels (6). Dr. Garner’s example nicely demonstrates that text book teaching about the portability of sensitivity and specificity estimates between populations does not hold up: The specificity of DS strongly differs dependent on the population to which the assay is applied.

But the example also shows differences in the cNPV between the estimates among HPV-positives and the whole screening population that is relevant for downstream management (Figure 1): The risk of CIN2+ in HPV-positive, dual stain-negative women is much higher (cNPV = 3.6%) than the risk in all

HPV-negative women (about 0.4% [3]), suggesting that a shorter follow-up than a regular screening interval is required. In contrast, the population risk among the DS-negatives in Dr. Garner’s example is 0.31%, suggesting that regular screening intervals would be sufficient. Thus, the population estimates presented by Dr. Garner cannot be used to make recommendations for management of HPV-positive women.

The two clinical performance estimates in Dr. Garner’s table address different questions. The population-wide analysis represents a programmatic perspective, combining the screening and triage steps into a single estimate. The analysis restricted to HPV-positive women represents the clinical perspective, addressing the question about who among the HPV-positives requires colposcopy and how long test intervals among DS-negatives can be. Both evaluations are valid, but the choice of analysis depends on the underlying question.

The important message is that biomarker performance estimates may not be portable between populations with very different baseline risks, and caution is warranted when using external estimates to develop guidelines and inform clinical management.

References

1. Huh WK, Ault KA, Chelmow D, et al. Use of primary high-risk human papillomavirus testing for cervical cancer screening: interim clinical guidance. *J Low Genit Tract Dis.* 2015;19(2):91–96.
2. Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE. Human papillomavirus testing in the prevention of cervical cancer. *J Natl Cancer Inst.* 2011;103(5):368–383.
3. Gage JC, Schiffman M, Katki HA, et al. Reassurance against future risk of precancer and cancer conferred by a negative human papillomavirus test. *J Natl Cancer Inst.* 2014;106(8):dju153.
4. Wentzensen N, Schiffman M, Palmer T, Arbyn M. Triage of HPV positive women in cervical cancer screening. *J Clin Virol.* 2015;S1386-6532(15)00749-0.
5. Wentzensen N, Fetterman B, Castle PE, et al. p16/Ki-67 Dual Stain Cytology for Detection of Cervical Precancer in HPV-Positive Women. *J Natl Cancer Inst.* 2015;107(12):djv257.
6. Wentzensen N, Wacholder S. From differences in means between cases and controls to risk stratification: a business plan for biomarker development. *Cancer Discov.* 2013;3(2):148–157.

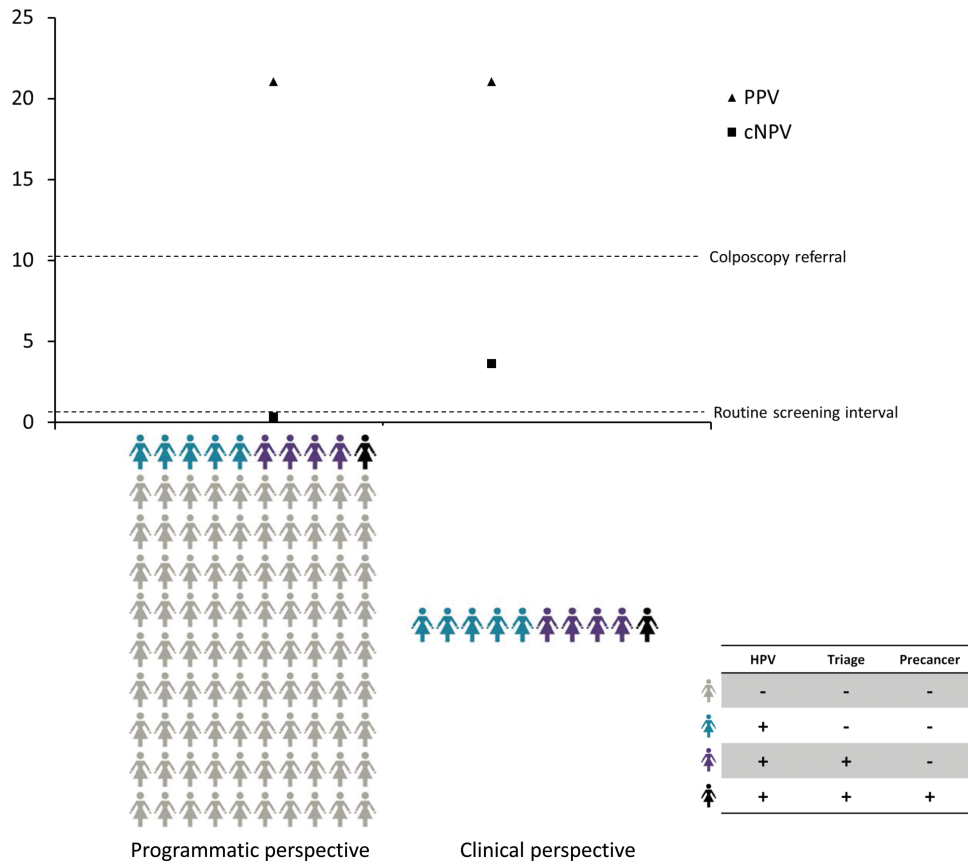


Figure 1. The figure shows an idealized scenario of a screening and triage strategy to detect cervical precancers. All women are tested for human papillomavirus (HPV) and triage testing is restricted to HPV-positives. **Gray women** represent HPV-negative women. **Blue women** are HPV-positive but negative for the triage marker and do not have precancer. **Purple women** are HPV-positive and test positive for the triage marker. The **black woman** is positive for HPV and the triage marker and is diagnosed with precancer. On the **left side**, the absolute risk estimates after testing for DS (positive predictive value and complement of the negative predictive value) are shown for the full population as described by Dr. Garner (1). On the **right side**, the risk estimates are shown for the group of HPV-positive women as published in our paper (3). Clinical management thresholds are shown in relationship to the absolute risk estimates. In this idealized scenario, women testing negative for HPV but positive for the triage marker are not shown, as they would not be identified in the sequential testing strategy. cNPV = complement of the negative predictive value; PPV = positive predictive value.