Effect of Time to Diagnostic Testing for Breast, Cervical, and Colorectal Cancer Screening Abnormalities on Screening Efficacy: A Modeling Study

Carolyn M. Rutter[1], Jane J. Kim[2], Reinier G.S. Meester[3], Brian L. Sprague[4], Emily A. Burger[2], Ann G. Zauber[5], Mehmet Ali Ergun[6], Nicole G. Campos[2], Chyke A. Doubeni[7], Amy Trentham-Dietz[8], Stephen Sy[2], Oguzhan Alagoz[6], Natasha Stout[2], Iris Lansdorp-Vogelaar[3], Douglas A. Corley[9], Anna N. A. Tosteson[10]

[1] RAND Corporation, Santa Monica, CA

[2] Center for Health Decision Science, Department of Health Policy and Management, Harvard T. H. Chan School of Public Health, Boston MA

[3] Department of Public Health, Erasmus MC University Medical Center, Rotterdam

[4] Departments of Surgery and Radiology, University of Vermont Cancer Center, Burlington VT

[5] Department of Epidemiology & Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY

[6] Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison WI

[7] Department of Family Medicine and Community Health, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA

[8] Department of Population Health Sciences and Carbone Cancer Center, School of Medicine and Public Health, University of Wisconsin-Madison, Madison WI

[9] Kaiser Permanente Division of Research, Oakland, CA

[10] Departments of Medicine, Community and Family Medicine, and The Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine at Dartmouth, and Norris Cotton Cancer Center, Lebanon, NH

Funding Sources: NCI PROSPR: U54 CA163261 (Rutter); U54 CA164336 (Kim, Campos, Sy); U54 CA163262 (Meester, Doubeni, Corley, Lansdorp-Vogelaar, Zauber); U54 CA163303 (Algoz, Ergun, Sprague, Stout, Trentham-Dietz); NCI CISNET: U01 CA199335 (Rutter, Meester, Zauber, Lansdorp-Vogelaar); U01 CA199218 (Algoz, Ergun, Sprague, Stout, Trentham-Dietz); U54 CA163307 (Tosteson); P30 CA014520 (Trentham-Dietz); P30 CA008748 (Zauber).

Acknowledgement: This study was conducted as part of the NCI-funded consortium Population-based Research Optimizing Screening through Personalized Regimens (PROSPR). The overall aim of PROSPR is to conduct multi-site, coordinated, transdisciplinary research to evaluate and improve cancer screening processes. The ten PROSPR Research Centers reflect the diversity of US delivery system organizations.

Text: 4000 word limit, current word count: 3480

Abstract

Background: Patients who receive an abnormal cancer screening result require follow-up for diagnostic testing, but the time to follow-up varies across patients and practices.

Methods: We used a simulation study to estimate the change in lifetime screening benefits when time to follow-up for breast, cervical, and colorectal cancers was increased. Estimates were based on four independently developed microsimulation models that each simulated the life course of adults eligible for breast (women aged 50-74 years), cervical (women aged 21-65 years), or colorectal (adults aged 50-75 years) cancer screening. We assumed screening based on biennial mammography for breast cancer, triennial Papanicolaou testing for cervical cancer, and annual fecal immunochemical testing for colorectal cancer. For each cancer type, we simulated diagnostic testing immediately and at 3, 6 and 12 months after an abnormal screening exam.

Results: We found declines in screening benefit with longer times to diagnostic testing, particularly for breast cancer screening. Compared to immediate diagnostic testing, testing at 3 months resulted in reduced screening benefit, with fewer undiscounted life years gained per 1000 screened (breast: 17.3%, cervical: 0.8%, colorectal: 2.0% and 2.7% (from two colorectal cancer models), fewer cancers prevented (cervical: 1.4% fewer, colorectal: 0.5% and 1.7% fewer, respectively) and, for breast and colorectal cancer, a less favorable stage distribution.

Conclusions: Longer times to diagnostic testing after an abnormal screening test can decrease screening effectiveness, but the impact varies substantially by cancer type.

Impact: Understanding the impact of time to diagnostic testing on screening effectiveness can help inform quality improvement efforts.

Abstract word count: 248 (limit 250)

Introduction

Cancer screening is a multi-step process of care that often requires patients to navigate multiple health facilities and specialties[1,2], requiring coordination and potentially lengthy times to complete a screening episode. Tosteson and colleagues (2015) demonstrated variability in the time to initial follow-up of screen-positive results across patients, cancer types, and health care systems[3], but the clinical implications of this substantial variation are unknown. While the vast majority of women with an abnormal screening mammogram took the next step toward diagnosis within weeks of the screening mammogram, 4.4% did not complete the next step in a timely manner[3] (within 3 months for an incomplete exam or an exam with suspicious findings; within 9 months for an exam with probably benign findings). Among women screened for cervical cancer using a Pap test, 45% with an abnormal result (atypical squamous cells of undetermined significance [ASC-US] with HPV-positive result or worse) received recommended diagnostic assessment with either repeat testing or colposcopy/biopsy within 3 months.[3] Among adults screened for colorectal cancer using a fecal test, 68% with a positive result indicating detection of occult blood received diagnostic assessment (colonoscopy) within 3 months.[3]

Longer time to clinical follow-up and subsequent diagnostic testing theoretically leads to an increased risk for disease progression, but it is unclear if the observed variability in time to diagnostic testing has meaningful effects on clinical outcomes and how effects might differ by cancer type. Data needed to directly estimate the effects of time from an abnormal screening result to diagnostic assessment on clinical outcomes are scarce. A randomized study would be unethical and observational studies examining the effect of time to diagnostic assessment on short-term outcomes, such as cancers detected at screening and cancer stage at

detection, require a large screened population and may be biased because of non-random selection of time to diagnostic testing. For example, abnormal screening tests may be more rapidly assessed when patients have a family history of cancer, develop symptoms or have other risk factors.

Microsimulation models can be used to simulate outcomes that may not be feasible to study with randomized controlled trials and are not subject to selection bias[4]. In this paper, we use microsimulation models to estimate the impact of the length of time from an abnormal screening test to diagnostic evaluation on long-term health outcomes across three cancers.

Methods

We used four independently developed models to simulate detection of and death from breast, cervical, and colorectal cancers. Each of these models is supported by the National Cancer Institute's Cancer Intervention and Surveillance Modeling Network (CISNET).[5]

Models

Each model describes the disease processes in the absence of screening or any other intervention. In all models, individuals begin in a disease-free state and potentially transition to preclinical (asymptomatic) cancer. From the preclinical cancer state, individuals may transition to a clinical (symptomatic) cancer state. The breast cancer model allows women to transition from the preclinical cancer state back to the disease-free state, due to regression of ductal carcinoma in situ (DCIS) and some small invasive cancers. Cervical and colorectal cancer models include one or more precursor lesion states that precede preclinical cancer. The

cervical cancer model allows women to transition from precursor states back to the disease-free state, while the colorectal cancer models do not allow regression of precursor lesions.

For each cancer, we simulated outcomes with and without screening and under different screening assumptions. In the absence of screening, cancer is detected when an individual transitions into the (symptomatic) clinical cancer state. Screening can benefit patients through detection of preclinical cancer, which may result in detection at an earlier and more treatable stage than clinically detected cancer. For cancers with a precursor lesion, screening may benefit patients by detecting and removing lesions that might otherwise progress to symptomatic cancer, thereby reducing cancer incidence. In the remainder of this section we outline each of the four models. Detailed model descriptions are available online at https://resources.cisnet.cancer.gov/registry.

The **University of Wisconsin Breast Cancer Simulation Model** simulates breast cancer incidence and mortality according to estrogen receptor (ER) and human epidermal growth factor receptor 2 (HER2)-specific status in the US population over time[6,7] . The probability that a woman will develop breast cancer is informed by an age-period-cohort model reflecting breast cancer incidence in the absence of screening and accounts for population trends in risk factors and screening utilization[8]. Disease progression varies randomly across individuals.

The breast cancer model simulates mammographic sensitivity (the probability of detecting cancer when it is present) as a function of lesion size, and was calibrated to accurately predict mammography performance data by age, screening round, and breast density[9]. For example, sensitivity ranges from 0.94 for women 65 years and older with fatty (BI-RADS 1) breasts to 0.75 for women from 40 to 49 years old with dense (BI-RADS 4)

breasts; specificity ranges from 0.95 for women 65 years and older with fatty breasts to 0.85 for women 40 to 49 years old with heterogeneously dense (BI-RADS 3) breasts. Breast biopsy is assumed to diagnose breast cancer with perfect accuracy. Women diagnosed with breast cancer are assumed to receive the most effective treatment currently available based on their simulated age, stage, and ER/HER2 status at diagnosis. Treatment effectiveness is based on clinical trials and is modeled as an increase in the proportion cured relative to ER/HER2-specific survival in the absence of therapy[10].

The **Harvard cervical cancer model** simulates human papillomavirus (HPV)-induced cervical carcinogenesis[11,12]. For each woman, the model simulates the risk of HPV infection, development and regression of precursor lesions (i.e., cervical intraepithelial neoplasia, grade 2 (CIN2) or grade 3 (CIN3)), and onset and progression of invasive cancer (i.e., local, regional, and distant disease). Precursor lesions may be associated with either non-oncogenic or oncogenic HPV; only those associated with oncogenic infections can progress to cancer. The probability of transition between precursor and cancer states can vary by age, HPV type, duration of infection or lesion status, and a woman's history of prior HPV infection and CIN treatment.

The cervical cancer model simulates outcomes from Pap smear-based screening. Pap test sensitivity to detect high-grade cervical intraepithelial neoplasia (CIN2) or more advanced lesions is 0.70; specificity is 0.91. Women with Pap results of low-grade squamous intraepithelial lesions (LSIL) or worse are referred to undergo diagnostic colposcopy/biopsy, while women with ASC-US undergo HPV testing and all HPV-positive women are referred to diagnostic colposcopy/biopsy. Colposcopy/biopsy is assumed to identify preclinical cervical

cancer and precursor lesions with perfect accuracy, and women are assumed to undergo treatment immediately upon diagnosis of CIN2 or more advanced disease.

We used two colorectal cancer models: **ColoRectal Cancer Simulated Population model for Incidence and Natural history (CRC-SPIN)**[13] and **MIcrosimulation SCreening ANalysis-ColoRectal Cancer (MISCAN-Colon)**[14]. Both models describe the natural history of colorectal cancer based on the adenoma-carcinoma process[15,16]. Individuals begin in a disease-free state and may progressively transition to an adenoma state, a preclinical cancer state, and a clinically detected cancer state. Disease progression depends systematically on age and sex, and varies randomly across individuals. The two models have different assumptions about the time from adenoma initiation to preclinical cancer ('dwell time') and the time from preclinical cancer initiation to symptomatic detection ('sojourn time')[17]. Compared to CRC-SPIN, the MISCAN-Colon model assumes a shorter mean dwell time and a longer mean sojourn time.

Both colorectal cancer models simulate outcomes from fecal immunochemical testing (FIT). The person-level specificity of FIT (for individuals with neither adenomas nor preclinical cancer) is 0.964; person-level sensitivity, based on the most advanced lesion, is 0.076 for 1mm to 9mm adenomas, 0.238 for adenomas that are 10mm or larger, and 0.738 for preclinical cancers[18]. Both models use person-level FIT specificity directly. The CRC-SPIN model uses person-level FIT sensitivity directly, simulating positive FIT results based on an individual's most advanced lesion. The MISCAN model uses lesion-specific sensitivity calibrated to match person-level sensitivity. Both colorectal cancer models simulate imperfect colonoscopy, so that screen-detected lesions may be missed at diagnostic evaluation. Survival

after colorectal cancer diagnosis is based on estimates from SEER data for individuals diagnosed in 2003, the most recent year with longer-term survival information[19].

Simulated Cohorts

For each cancer type, we simulated an unscreened cohort and then simulated screening regimens under four scenarios: immediate diagnostic testing of abnormal screening result, and diagnostic testing 3, 6 and 12 months after an abnormal screening result. Immediate diagnostic testing after a positive screening test, while unrealistic, reflects the maximum benefit achievable from recommended screening. We standardized simulated cohorts so that screening is initiated in 2015 across all cancers because of the potential impact of year of detection on both overall and cancer-specific mortality. Therefore, the breast and colorectal cancer models simulated individuals born in 1965, and the cervical cancer model simulated individuals born in 1994.

Simulated screening regimens

Screening was simulated according to regimens that are consistent with current United States Preventive Services Task Force (USPSTF) recommendations (Table 1)[20-22]. To isolate the effect of time to diagnostic testing, we simulated perfect adherence to screening regimens, and varied only the time between a positive result and subsequent diagnostic testing to rule-in or rule-out disease.

Outcomes

We calculated three measures of lifetime screening benefit, from the age of screening initiation through age 100 years: 1) cancers prevented, relative to no screening; 2) cancer stage distribution at diagnosis; and 3) undiscounted life years gained (LYG), relative to no screening. LYG were estimated as the difference in the simulated number of person-years for screened versus unscreened cohorts. We also estimated two relative risks (RRs) comparing outcomes with diagnostic testing at 3, 6, and 12 months relative to immediate diagnostic testing: the lifetime RR of cancer diagnosis and the RR of cancer diagnosis at a late stage (distant or regional, or Stage 3 or 4) among those diagnosed with cancer.

Results

Screening increased the incidence of detected breast cancer, and decreased the incidence of cervical and colorectal cancers (Table 2). The breast cancer model predicted that screening would increase lifetime cancer incidence by 2.6 per 100 persons screened, with slightly smaller increases in cancer incidence as length of time to diagnostic testing increased. The cervical and colorectal cancer models predicted that screening would decrease lifetime cancer incidence, with small decrements in benefit as length of time to diagnostic testing increased. Compared to immediate diagnostic testing, testing at 3 months would result in 1.4% fewer cervical cancers prevented, and 1.7% (CRC-SPIN) and 0.5% (MISCAN-Colon) fewer colorectal cancers prevented. The RR of a lifetime cancer diagnosis among those screened with diagnostic testing at 3 months relative to immediate diagnosis was 1.00 for breast cancer, 1.11 for cervical cancer, and 1.01 (CRC-SPIN) or 1.02 (MISCAN-Colon) for colorectal cancer; at 12 months the RRs were 0.98 for breast, 1.36 for cervical cancer, and 1.05 for colorectal cancer (both CRC-SPIN and MISCAN-Colon).

All models predicted that screening would reduce the proportion of late stage cancers relative to no screening (Table 3). For the most part, longer times to diagnostic testing resulted in increases in diagnosis of late stage disease. For breast cancer, screening with immediate diagnostic testing detected 24% more cancers overall relative to no screening, specifically, more DCIS (24.5% of cancers diagnosed versus 5.3%, respectively) and less late stage disease (26.7% versus 46.6%). Cervical and colorectal cancer screening with immediate diagnostic testing resulted in declines in both cancer incidence and the percent of cancers detected at a late stage. Among individuals diagnosed with cancer, the RR of a late stage detection with diagnostic testing at 3 months relative to immediate testing was 1.08 for breast cancer, 0.99 for cervical cancer and 1.03 for colorectal cancer (both CRC-SPIN and MISCAN-Colon); the RR of late stage diagnosis with testing at 12 months relative to immediate testing was 1.26 (breast), 0.98 (cervical), 1.12 (CRC-SPIN) and 1.11 (MISCAN-Colon).

LYG incorporates the impact of time to screening on mortality through cancer prevention and stage shift. All four models predict decrements in LYG with lengthening time to diagnostic testing, relative to the potential maximum with immediate testing (Table 4), with the largest losses for breast cancer. With a three-month time to diagnostic testing, the estimated loss of potential LYG was equal to 17.3% for breast cancer, 0.8% for cervical cancer, and 2.0% (CRC-SPIN) and 2.7% (MISCAN-Colon) for colorectal cancer.

Discussion

Theoretically, screening benefit is maximized when diagnostic testing occurs immediately, on the same day as the receipt of a positive screening test. We examined the relative benefit of screening when diagnostic testing occurred at 3, 6, and 12 months after an

abnormal screening result compared to this theoretical maximum. Longer time to diagnostic testing reduced the benefit of screening, resulting in higher lifetime cancer incidence (cervical and colorectal), later stage at diagnosis (breast and colorectal), and fewer LYG (all cancers).

The effects of time to diagnostic testing on reduction in screening benefit varied across cancers, in ways that are consistent with differences in their natural history and the action of screening. Breast cancer screening focuses on early detection of malignant lesions[20,23]; our results indicate that failure to promptly diagnose and treat malignant breast lesions would have substantial impacts on screening effectiveness. Longer time to diagnostic assessment of abnormal mammograms resulted in more late stage disease and fewer life years gained via screening. Across the three cancers we simulated, breast cancer demonstrated the largest loss in LYG as the time to diagnostic testing increased.

In contrast to breast cancer screening, cervical and colorectal cancer screening can result in detection and removal of precursor lesions. Longer time to diagnostic assessment reduced cervical and colorectal cancer screening effectiveness, but with modest effects compared to breast cancer. Both cervical and colorectal cancers are characterized by slowly progressing precursor lesions, providing a long window of opportunity for disease prevention.

Even with longer time to diagnostic testing, our study suggests that screening detected most disease at an earlier and more treatable stage. Longer times to diagnostic testing had the smallest impact on disease stage distribution for cervical cancer. Cervical cancer arises from precursor lesions that may occur relatively early in life, and some lesions may regress. The effect of time to diagnostic assessment on stage at cervical cancer detection was counter-intuitive; the proportion of late stage cervical cancers declined with longer time to assessment. Over this 12-month timeframe, the cervical model simulated transition of more precursor

lesions into early stage disease than transition of preclinical cancers from early to late stage disease. Importantly, the absolute incidence of late stage cervical disease increased with longer times to diagnostic testing.

Inclusion of two colorectal cancer models provides an opportunity to compare results across models with different assumptions. The CRC-SPIN model simulates a longer adenoma dwell time on average than the MISCAN-Colon model, and predicted a slightly smaller effect of time to diagnostic testing on cancer incidence than the MISCAN-Colon model. The MISCAN-Colon model simulates a longer sojourn time than the CRC-SPIN model, and predicted slightly smaller effect of time to diagnostic testing on cancer stage than the CRC-SPIN model. Despite these differences in model specifications, the overall results were similar across the two models, and were consistent with results of a previous modeling study.[24]

The breast cancer model demonstrated that in the absence of a precursor lesion, early detection is key and time to diagnostic testing can have a large impact on screening effectiveness. The cervical cancer model demonstrated that the screening regimens are more robust when the disease trajectory is long and precursor lesions may regress.

While we predicted that longer times to diagnostic testing could meaningfully worsen the efficacy of breast cancer screening, the times to testing we simulated are not common in clinical practice; diagnostic testing generally occurs within weeks of a screening mammography suggestive of a malignancy (i.e., BIRADS 4 or 5)[3]. This is likely due to the federal Mammography Quality Standards Act[25], which regulates timely reporting of mammography screening results in plain language[26]. Our simulation study supports the rapid follow-up of abnormal mammograms, but also highlights trade-offs that are inherent to breast cancer screening. Because some DCIS is indolent, screening could result in detection and

treatment of cancers that would have caused no harm. While shorter time to follow-up resulted in more DCIS, we do not assume that diagnosis of DCIS constitutes over-diagnosis. Other collaborative modeling analyses of screening programs have shown that over-diagnosis rates are relatively insensitive to screening frequency[27], indicating that relatively short delays in diagnosis do not affect overall rates of over-diagnosis. This is consistent with our findings that the overall breast cancer incidence did not vary with time to diagnosis.

Longer time to diagnostic evaluation of abnormal tests reduces the effectiveness and therefore the cost-effectiveness of screening. Our findings are important in light of existing policy and recent findings describing variation in time to follow-up of positive screening tests[3,28]. Federal policies regulating breast cancer screening encourage timely result reporting, which may facilitate timely follow-up. There are no such regulations for either cervical or colorectal cancer screening. Instead, the quality of cervical and colorectal cancer screening is largely monitored voluntarily, through reporting of Healthcare Effectiveness Data and Information Set (HEDIS) measures, which focus on receipt of screening tests, but not follow-up of abnormal screening[29]. Our analyses suggest that new policies focused on timeliness of follow-up could improve the effectiveness of colorectal cancer screening, in particular[30]. Further research examining the barriers to diagnostic testing at both the individual and health system level can help inform strategies for improved follow-up of positive tests. For example, diagnostic assessment of positive abnormal screening results (via colonoscopy) is improved when healthcare systems directly notify gastroenterology providers about referral for follow-up[31].

It is important to keep in mind the limitations of our work. Our results are based on simulation modeling, which requires assumptions about disease processes, health practices,

and patient behaviors. Dwell time and disease progression during the preclinical detectable period drive screening effectiveness and the impact of delays on screening outcomes[32]. However, dwell time assumptions that are built into simulation models cannot be changed in isolation because models incorporate multiple assumptions that work together. Model calibration modifies each of these assumptions, via parameter selection, so that models produce plausible simulated outcomes. When assumptions cannot easily be modified, cross-model comparisons provide the best insight into the relationships between different disease assumptions and simulated outcomes. For example, as previously noted, the two CRC models make fairly different assumptions about dwell time, yet predict similar effects of delays in follow-up on mortality. The MISCAN modeling group examined the sensitivity of their model's results to assumptions about dwell time, finding a larger effect of time to follow up on CRC mortality when average dwell time was halved[24]. The Harvard cervical cancer modeling group has carried out probabilistic sensitivity analysis based on a sample of 50 calibrated natural history parameter sets that vary the model assumptions related to transitions among HPV and CIN states[33]. These sensitivity analyses show that the cervical cancer model predictions of screening benefit are stable over this set of parameters across a range of different screening intervals[12], suggesting that underlying assumptions would have relatively little impact on predicted model outcomes. The impact of dwell time assumptions may be different for breast cancer models than for CRC or cervical cancer models because breast cancer does not have an identified precursor lesion. Cross-model comparisons found that the Wisconsin breast cancer model (included in this paper) predicted greater benefits from increased screening frequency than four other breast cancer models, suggesting that the

estimates we present provide an upper bound on the effects of delayed breast cancer diagnosis[27].

Our analyses did not examine risk factors or patient characteristics that might impact the effect of delays on outcomes. For example, we did not examine whether delays may vary based on age, which is related to the underlying risk of cancer.

Use of models allowed us to extend existing evidence to address questions about the impact of time to diagnostic testing on long-term cancer outcomes. While each of the models used in our analysis is well established there is no guarantee of the accuracy of model assumptions, especially for parameters representing unobservable processes such as tumor growth and for predictions over longer periods. Our results could be strengthened by findings from other simulation studies or, potentially, large-scale observational studies that examine the effect of time to follow-up on shorter-term cancer outcomes.

In conclusion, our study demonstrates that relatively small delays in diagnostic follow-up of abnormal findings could reduce cancer screening effectiveness, although the impact varies by cancer type. This suggests that there is value in efforts to promptly perform follow-up of positive test results, either within individual health care systems or through federal policies related to regulation and reporting. In addition to improving the overall effectiveness of screening, reducing time to follow-up has the potential to reduce health disparities in patients who are more likely to experience delays and improve patient quality of life by avoiding systemic treatment. Reducing time to follow-up may also prove to be cost-effective by avoiding the need for expensive treatment regimens.

References

1. Zapka JG, Taplin SH, Solberg LI, Manos MM. A framework for improving the quality of cancer care: the case of breast and cervical cancer screening. Cancer Epidemiol Biomarkers Prev 2003;12:4-13.

2. Beaber EF, Kim JJ, Schapira MM, et al. Unifying Screening Processes Within the PROSPR Consortium: A Conceptual Model for Breast, Cervical, and Colorectal Cancer Screening. Jnci-J Natl Cancer I 2015;107.

3. Tosteson AN, Beaber EF, Tiro J, et al. Variation in Screening Abnormality Rates and Follow-Up of Breast, Cervical and Colorectal Cancer Screening within the PROSPR Consortium. J Gen Intern Med 2016;31:372-9.

4. Rutter CM, Zaslavsky AM, Feuer EJ. Dynamic microsimulation models for health outcomes: A review. Med Decis Making 2010;31:10-8.

5. Cancer Intervention and Surveillance Modeling Network (CISNET). (Accessed September 25, 2017, at http://cisnet.cancer.gov/.)

6. Fryback DG, Stout NK, Rosenberg MA, Trentham-Dietz A, Kuruchittham V, Remington PL. The Wisconsin Breast Cancer Epidemiology Simulation Model. J Natl Cancer Inst Monogr 2006:37-47.

7. Batina NG, Trentham-Dietz A, Gangnon RE, et al. Variation in tumor natural history contributes to racial disparities in breast cancer stage at diagnosis. Breast Cancer Res Treat 2013;138:519-28.

8. Gangnon RE, Sprague BL, Stout NK, et al. The contribution of mammography screening to breast cancer incidence trends in the United States: an updated age-period-cohort model. Cancer Epidemiol Biomarkers Prev 2015;24:905-12.

9. Mandelblatt JS, Cronin K, de Koning H, Miglioretti DL, Schechter CS, Stout N. Collaborative modeling of US breast cancer screening strategies. Rockville, MD: Agency for Healthcare Research and Quality 2015.

10. Early Breast Cancer Trialists' Collaborative G, Peto R, Davies C, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet (London, England) 2012;379:432-44.

11. Campos NG, Burger EA, Sy S, et al. An updated natural history model of cervical cancer: derivation of model parameters. Am J Epidemiol 2014;180:545-55.

12. Kim JJ, Campos NG, Sy S, et al. Inefficiencies and High-Value Improvements in U.S. Cervical Cancer Screening Practice: A Cost-Effectiveness Analysis. Ann Intern Med 2015;163:589-97.

13. Rutter CM, Savarino JE. An evidence-based microsimulation model for colorectal cancer. Cancer Epidemiol Biomarkers Prev 2010;19:1992-2002.

14. van Hees D, Habbemma J, Meester R, Lansdorp-Vogelaar I, van Ballegooijen M, AG Z. Should colorectal cancer screening be considered in elderly persons without previous screening? A cost-effectiveness analysis. Ann Intern Med 2014;160:750-9.

15. Muto T, Bussey HJ, Morson BC. The evolution of cancer of the colon and rectum. Cancer 1975;36:2251-70.

16. Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525-32.

17. Rutter CM, Knudsen AB, Marsh TL, et al. Validation of Models Used to Inform Colorectal Cancer Screening Guidelines: Accuracy and Implications. Med Decis Making 2016.

18. Imperiale TF, Ransohoff DF, Itzkowitz SH, et al. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014;370:1287-97.

19. Rutter CM, Johnson EA, Feuer EJ, Knudsen AB, Kuntz KM, Schrag D. Secular Trends in Colon and Rectal Cancer Relative Survival. J Natl Cancer Inst 2013.

20. Siu AL, Force USPST. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Annals of internal medicine 2016;164:279-96.

21. Moyer VA, Force USPST. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2012;156:880-91, W312.

22. Force USPST, Bibbins-Domingo K, Grossman DC, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 2016;315:2564-75.

23. Oeffinger KC, Fontham ETH, Etzioni R, et al. Breast Cancer Screening for Women at Average Risk: 2015 Guideline Update From the American Cancer Society. Jama 2015;314:1599-614.

24. Meester RG, Zauber AG, Doubeni CA, et al. Consequences of Increasing Time to Colonoscopy Examination After Positive Result From Fecal Colorectal Cancer Screening Test. Clin Gastroenterol Hepatol 2016.

25. Compliance Guidance: The Mammography Quality Standards Act Final Regulations Document #1; availability. Food and Drug Administration, HHS. Notice. Fed Regist 1999;64:13590-1.

26. Monsees BS. The Mammography Quality Standards Act. An overview of the regulations and guidance. Radiol Clin North Am 2000;38:759-72.

27. Mandelblatt JS, Stout NK, Schechter CB, et al. Collaborative Modeling of the Benefits and Harms Associated With Different U.S. Breast Cancer Screening Strategies. Ann Intern Med 2016;164:215-25.

28. Chubak J, Garcia MP, Burnett-Hartman AN, et al. Time to Colonoscopy after Positive Fecal Blood Test in Four U.S. Health Care Systems. Cancer Epidemiol Biomarkers Prev 2016;25:344-50.

29. National Committee for Quality Assurance (NCQA). HEDIS and Performance Measurement. 2016. (Accessed 07-28-2016, at http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2016.)

30. Corley DA, Haas JS, Kobrin S. Reducing Variation in the "Standard of Care" for Cancer Screening: Recommendations From the PROSPR Consortium. JAMA 2016;315:2067-8.

31. Partin MR, Burgess DJ, Burgess JF, et al. Organizational Predictors of Colonoscopy Follow-up for Positive Fecal Occult Blood Test Results: An Observational Study. Cancer Epidem Biomar 2015;24:422-34.

32. van Ballegooijen M, Rutter CM, Knudsen AB, et al. Clarifying differences in natural history between models of screening: The case of colorectal cancer. Med Decis Making 2011;31:540-9.

33. Burger EA, Kim JJ, Sy S, Castle PE. Age of acquiring causal human papillomavirus (HPV) infections: Leveraging simulation models to explore the natural history of HPV-induced cervical cancer. Clinical Infectious Diseases 2017:cix475.

Table 1. Screening regimens that were simulated for each cancer site by age group and periodicity

	Breast	**Cervical**	**Colorectal**
Test	digital mammogram	Papanicolaou (Pap)	fecal immunochemical test (FIT)
Interval	2 years	3 years	1 year
Screening Ages	50-74 years	21-65* years	50-75 years

* Screening ends at age 65 for women who have not had abnormal screening results over the past 10 years; otherwise, screening continues past age 65 until there are three consecutive normal results.

Author Manuscript Published OnlineFirst on November 17, 2017; DOI: 10.1158/1055-9965.EPI-17-0378
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Table 2. Predicted effect of screening and time to diagnostic testing of screening abnormalities on lifetime cancer incidence per 100 screened.

		Change in Cancer Incidence with Screening							
Model	Unscreened	Immediate Diagnosis		3 months		6months		12months	
Breast	11.0	2.6	(24.0%)	2.6	(23.5%)	2.5	(22.9%)	2.4	(21.7%)
Cervical	2.2	-1.9	(-88.9%)	-1.9	(-87.6%)	-1.9	(-86.6%)	-1.8	(-84.8%)
CRC-SPIN	7.2	-5.3	(-73.9%)	-5.3	(-73.6%)	-5.3	(-73.3%)	-5.2	(-72.5%)
MISCAN-Colon	6.6	-3.1	(-46.9%)	-3.1	(-46.1%)	-3.0	(-45.4%)	-2.9	(-44.2%)

Author Manuscript Published OnlineFirst on November 17, 2017; DOI: 10.1158/1055-9965.EPI-17-0378
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Table 3. Predicted effect of screening and time to diagnostic assessment of abnormal screening tests on cancer stage distribution at detection (percent in each stage), among those diagnosed with cancer (shown as the incidence per 100 screened). Small variations in numbers are shown in Tables 3 and 4 reflect rounding.

		Time to Diagnosis			
Model	Unscreened	Immediate	3 months	6 months	12 months
Breast					
Incidence per 100	11.00	13.65	13.58	13.52	13.39
% DCIS	5.3	24.5	21.8	19.2	17.0
% Local	48.1	48.9	49.5	50.1	49.5
% Regional	39.7	23.1	25.0	26.8	29.2
% Distant	7.0	3.5	3.7	3.9	4.2
Cervical					
Incidence per 100	2.18	0.24	0.27	0.29	0.33
% Local	51.3	67.1	67.5	67.6	67.9
% Regional	38.3	26.7	26.5	26.5	26.3
% Distant	10.4	6.2	6.0	5.9	5.8
CRC-SPIN					
Incidence per 100	7.17	1.87	1.89	1.92	1.97
% Stage 1	17.8	39.4	37.9	36.1	32.6
% Stage 2	36.2	28.5	29.1	30.0	31.3
% Stage 3	26.7	19.5	20.0	20.6	21.7
% Stage 4	19.2	12.6	13.1	13.4	14.4
MISCAN-Colon					
Incidence per 100	6.63	3.52	3.58	3.62	3.70
% Stage 1	17.6	51.7	49.9	48.6	46.0
% Stage 2	33.8	25.8	26.9	27.6	29.1
% Stage 3	24.1	13.3	13.8	14.3	15.0
% Stage 4	24.5	9.1	9.4	9.5	9.9

Author Manuscript Published OnlineFirst on November 17, 2017; DOI: 10.1158/1055-9965.EPI-17-0378
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

Table 4. Predicted life years gained (LYG) per 1,000 screened by time to diagnostic testing and decrement in LYG relative to immediate diagnostic testing

	LYG per 1,000 screened (and Percent Change) Relative to Immediate Diagnosis						
Model	Immediate	3months		6 months		12 months	
Breast	101.6	84.1	(-17.3%)	66.1	(-34.9%)	41.0	(-59.6%)
Cervical	281.7	279.5	(-0.8%)	277.6	(-1.4%)	273.8	(-2.8%)
CRC-SPIN	249.8	244.8	(-2.0%)	239.7	(-4.1%)	230.2	(-7.8%)
MISCAN-Colon	233.7	227.3	(-2.7%)	222.7	(-4.7%)	213.7	(-8.6%)

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Effect of Time to Diagnostic Testing for Breast, Cervical, and Colorectal Cancer Screening Abnormalities on Screening Efficacy: A Modeling Study

Carolyn M. Rutter, Jane J. Kim, Reinier G.S. Meester, et al.

Cancer Epidemiol Biomarkers Prev Published OnlineFirst November 17, 2017.