The Impact of Different Screening Model Structures on Cervical Cancer Incidence and Mortality Predictions: The Maximum Clinical Incidence Reduction (MCLIR) Methodology

Inge M. C. M. de Kok🔗, Emily A. Burger, Steffie K. Naber, Karen Canfell, James Killen, Kate Simms, Shalini Kulasingam, Emily Groene🔗, Stephen Sy, Jane J. Kim, and Marjolein van Ballegooijen

Background. To interpret cervical cancer screening model results, we need to understand the influence of model structure and assumptions on cancer incidence and mortality predictions. Cervical cancer cases and deaths following screening can be attributed to 1) (precancerous or cancerous) disease that occurred after screening, 2) disease that was present but not screen detected, or 3) disease that was screen detected but not successfully treated. We examined the relative contributions of each of these using 4 Cancer Intervention and Surveillance Modeling Network (CISNET) models. **Methods**. The maximum clinical incidence reduction (MCLIR) method compares changes in the number of clinically detected cervical cancers and mortality among 4 scenarios: 1) no screening, 2) one-time perfect screening at age 45 that detects all existing disease and delivers perfect (i.e., 100% effective) treatment of all screen-detected disease, 3) one-time realistic-sensitivity cytological screening and perfect treatment of all screen-detected disease, and 4) one-time realistic-sensitivity cytological screening and realistic-effectiveness treatment of all screen-detected disease. **Results**. Predicted incidence reductions ranged from 55% to 74%, and mortality reduction ranged from 56% to 62% within 15 years of follow-up for scenario 4 across models. The proportion of deaths due to disease not detected by screening differed across the models (21%–35%), as did the failure of treatment (8%–16%) and disease occurring after screening (from 1%–6%). **Conclusions**. The MCLIR approach aids in the interpretation of variability across model results. We showed that the reasons why screening failed to prevent cancers and deaths differed between the models. This likely reflects uncertainty about unobservable model inputs and structures; the impact of this uncertainty on policy conclusions should be examined via comparing findings from different well-calibrated and validated model platforms.

Keywords
cervical cancer, comparative modeling, microsimulation modeling, screening

Microsimulation models are increasingly being used to estimate the long-term population effects of cervical cancer screening programs. Collaborative modeling can enhance the rigor of modeling research through the use of multiple independent models, with standardized model inputs, to answer the same research or policy

Corresponding Author
Inge M. C. M. de Kok, Erasmus MC—University Medical Center, Department of Public Health, P.O. Box 2040, Rotterdam, Zuid-Holland, 3000 CA, The Netherlands (I.dekok@erasmusmc.nl).

questions. The US National Cancer Institute–funded Cancer Intervention and Surveillance Modeling Network (CISNET) was established in 2000 to use collaborative modeling to improve our understanding of the impact of cancer prevention, screening, and treatment dissemination on population trends in cancer incidence and mortality.[1] The CISNET Cervical Cancer Working Group started in 2015 and includes 5 modeling teams. Prior to CISNET, each model was independently developed for 3 different countries (i.e., the United States, Australia, and the Netherlands) and is unique in its structure, assumptions, and methods of synthesizing data. Consequently, they are unique in how they project the impact of screening and treatment on cervical cancer incidence and mortality. Results that are similar across multiple models, despite differences in assumptions and modeling approach, enhance the credibility of the findings and are more likely to be robust than conclusions obtained from a single model; results that are different across models reveal areas of uncertainty and priorities for future research.

While country-specific factors (e.g., cost-effectiveness thresholds, disease burden) may contribute to relevant variations in policy conclusions, it is important to understand how different model structures and combinations of assumptions contribute to these differences. Detailed

Department of Public Health, Erasmus MC—University Medical Center, Rotterdam, Zuid-Holland, The Netherlands (IMCMdK, SKN, MvB); Department of Health Policy and Management, Harvard T. H. Chan School of Public Health, Boston, MA, USA (EAB, SS, JJK); Cancer Research Division, Cancer Council NSW, Sydney, Australia (KC, JK, KS); School of Public Health, University of Sydney, Sydney, Australia (KC); Prince of Wales Clinical School, University of New South Wales, Sydney, Australia (KC); School of Public Health, University of Minnesota, Minneapolis, MN, USA (SK, EG). The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: IdK, EB, SN, JK, KS, SK, EG, SS, JK, and MvB have no conflict of interests to report. KC is co–principal investigator of an unrelated investigator-initiated trial of cervical cytology and primary human papillomavirus screening in Australia ("Compass"), which is conducted and funded by the VCS Foundation, a government-funded health promotion charity. The VCS Foundation has received equipment and a funding contribution for the Compass trial from Roche Molecular Systems and Ventana, Inc. However, neither KC nor her institution on her behalf (Cancer Council NSW) receives direct funding from industry for this or any other project. The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Financial support for this study was provided entirely by a grant from the National Cancer Institute as part of the Cancer Intervention and Surveillance Modeling Network (CISNET), grant U01CA199334. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

model descriptions are necessary and contribute to model transparency. Although all inputs of a model should always be reported, assumptions on natural history parameters, such as the dwell time (i.e., the period in which the asymptomatic disease could be caught by screening) and their implications, remain difficult to assess even with comprehensive documentation.

In the absence of screening, cervical cancers will only be diagnosed as a result of clinical symptoms (i.e., clinical incidence). Clinical incident cancer cases and cancer-specific deaths that occur after screening have 3 possible origins: 1) the disease was not present at the time of screening, 2) the disease was present but not detected by screening (i.e., limitations in test sensitivity), or 3) preclinical disease (precancerous lesions and preclinical cancers) was not completely removed or recurrence of that lesion occurred. The maximum clinical incidence reduction (MCLIR) method can be used to examine and disentangle the relative contributions of these 3 origins of different screening models. The MCLIR method is able to isolate the unobservable effects of lesion onset and duration, screening test sensitivity, and precancerous lesions and cancer treatment by comparing model results before and after imposing a one-time screening intervention under varying assumptions about screening performance and treatment effectiveness. These effects are unobservable because of limited knowledge of underlying natural history processes and parameters, and they include parameters for which traditional epidemiologic studies would be impossible or unethical to conduct. Previously, the MCLIR method has been applied to 3 CISNET colorectal[2] and 6 breast cancer[3] models to delineate the effects of natural history assumptions and model structure on colorectal and breast cancer predictions. In this study, we applied the MCLIR method to understand how differences among the CISNET cervical cancer models affect predictions for screening effectiveness by projecting the clinical incidence and mortality reductions after a one-time screening test at age 45 years among women without prior screening. The results are intended to provide a greater understanding of how the CISNET cervical models depict unobservable processes (such as the progression of underlying disease and the missed detection of existing disease) and how those representations may systematically affect conclusions about screening impact on incidence and mortality.

Methods

We used 4 well-established CISNET-Cervical microsimulation models (Harvard,[4] Microsimulation Screening

Analysis [MISCAN]–Cervix,[5] Policy1-Cervix [CCNSW],[6] and University of Minnesota [UMN]–HPV Cancer [CA][7]) to project outcomes for a hypothetical cohort of individual women. This analysis was conducted prior to standardization of each model to the US population and screening performance.

Each of the 4 CISNET-Cervical models simulates the underlying natural history of human papillomavirus (HPV)–induced cervical cancer but differs with respect to the type and number of health states, HPV genotypes included, cycle length, and data sources used to parameterize the baseline model prior to model fitting (i.e., calibration) to the observed data (Box). In general, all 4 models simulate a large study population with individual life histories, in which women can acquire HPV infections, develop precancerous lesions, and, in some cases, develop cervical cancer. Death can occur from background causes or cervical cancer among those with disease. Infections can clear or progress to cervical intraepithelial lesions that may regress. Women can acquire multiple infections during a lifetime, each with its own clinical course. The models produce age-specific outputs of disease incidence and mortality over time. Screening can affect the life histories of individual women. These changes constitute the effects of the intervention and are represented by the numbers of events and the stage of cancer that is detected or prevented.

Standardized profiles of each model's structure and underlying model parameters and assumptions, with additional references, are available at http://cisnet.cancer.gov/profiles/. Additional model specifications and key parameters are provided in Table 1.

MCLIR Analysis

To illustrate the effects of model structure and assumptions about lesion development, lesion progression, screening test ability to detect lesions, and precancerous lesion treatment on cervical cancer incidence and mortality predictions, the MCLIR analysis consists of comparisons between 4 scenarios. Three scenarios involve a one-time screening test at age 45 years, and the remaining scenario of no screening serves as a comparator. The study population for each scenario is a cohort of average-risk women born in 1980 who have never been screened or diagnosed with cervical cancer prior to age 45 years. Age 45 years was selected to illustrate model differences because it is in the middle of the start and stop ages of recommended screening guidelines,[8] and there is sufficiently high incidence of cervical cancer at this age to illustrate model differences. Women were followed for 5 years (i.e., up to age 50) and 15 years (i.e., up to age 60) to capture the short- and longer-term effects of the

intervention. Model outcomes included cervical cancer incidence and mortality by age.

MCLIR Scenarios

To disentangle the impacts of the 3 possible sources of residual cancer following one-time screening, the MCLIR method compares changes in the number of clinically detected cervical cancer cases and deaths between 4 simplified scenarios (Figure 1): 1) no-screening; 2) one-time perfect cervical screening at age 45 years, which detects all existing precancerous lesions and assumes perfect treatment of all screen-detected lesions (i.e., lesion completely removed and no recurrent disease); 3) one-time realistic-sensitivity cytological screening at age 45 years, which assumes imperfect detection of precancerous lesions and perfect treatment of all screen-detected lesions; and 4) one-time realistic-sensitivity cytological screening at age 45 years and realistic-effectiveness treatment of all screen-detected lesions. Assumptions for realistic cytological screening and realistic treatment of all screen-detected lesions differed between the models (Table 1).

We calculated the attributable fraction (i.e., origin) of cervical cancer incident cases and deaths in the 4 CISNET models, 5 years and 15 years after screening at age 45 years. We did this by dividing the difference in number of cancers/deaths between the different modeled scenarios (see Figure 1) with the number of cancers in the no-screening scenario. Because in some models, women were not all screened exactly on their 45th birthday but within the 45th year of life, we excluded the first follow-up year when calculating the attributional fractions.

Results

Figure 2 shows how the attributable fractions (Table 2) are calculated, using cervical cancer mortality in the Harvard model as an example.

Incidence Reduction Predictions

Short-term (5-year) effects of screening. When we assumed realistic screening at age 45 years, we found that after 5 years, the proportion of cancer cases prevented differed between the models, ranging from 65% to 81% (Table 2). The short-term differences in prevented clinical cancer cases were primarily caused by differences in assumed test sensitivity; for example, 19% of the prevalent lesions that caused cancer within 5 years were not detected in the Policy1-Cervix model, compared to 26%

Box Key Attributes of the 4 Microsimulation Screening Models of the Cervical Working Group of Cancer Intervention and Surveillance Modeling Network (CISNET): Harvard, MISCAN (Erasmus MC Rotterdam), Policy1-Cervix (CCNSW), and UMN-HPV CA

	Microsimulation Model			
	Harvard	MISCAN	Policy1-Cervix	UMN-HPV CA
Model attributes				
Dynamic (interactive) or static (noninteractive)	Static	Static	Static	Static
Mode of analysis, simulating life histories	Individual based	Individual based	Aggregate	Individual based
Cycle length	Monthly	Continuous time	6 or 12 monthly	Annual
HPV transmission and infection				
HPV types included	HPV 16, HPV 18, HPV 31, HPV 33, HPV 45, HPV 52, HPV 58, pooled other high-risk HPV, pooled other low-risk HPV	HPV 16; HPV 18; pooled 31, 33, 45, 52, and 58; and pooled other high-risk genotypes	HPV 16; HPV 18; pooled 31, 33, 45, 52, and 58; and pooled high-risk genotypes	HPV 16; HPV 18; pooled 31, 33, 45, 52, and 58; and pooled other high-risk types
Natural immunity	Reduced probability of future type-specific infection	Reduced probability of future infection linked from dynamic model	Reduced probability of future type-specific infection from dynamic model	Reduced probability of future type-specific infection
Cervical carcinogenesis				
Health states included	Healthy, HPV, CIN 2, CIN 3, cancer (stage specific)	Healthy, HPV, CIN 1, CIN 2, CIN 3, cancer (stage specific)	Healthy, HPV, CIN 1, CIN 2, CIN 3, cancer (stage specific)	Healthy, HPV, CIN 1, CIN 2, CIN 3, cancer (stage specific)
Progression and regression rates	Age specific, HPV and lesion persistence	Age specific	Age specific	Age specific
Model calibration				
Calibrated parameters	HPV incidence, HPV and CIN progression and regression rates; HPV natural immunity; cancer symptom detection	HPV and CIN progression and regression; duration of CIN 3; cancer stages; cancer symptom detection; cytology test characteristics; background risk in nonattenders	HPV and CIN progression and regression rates; undetected asymptomatic cancer by stage; stage-specific detection for symptomatic and screen-detected disease	HPV incidence; CIN progression and regression rates; cancer symptom detection

CIN 1, cervical intraepithelial neoplasia grade 1; CIN 2, cervical intraepithelial neoplasia grade 2; CIN 3, cervical intraepithelial neoplasia grade 3; HPV, human papillomavirus.

Table 1 Assumptions for Realistic Cytological Screening and Realistic Treatment in the 4 Models

Model Parameter	Assumptions, %			
	Harvard	**MISCAN**	**Policy1-Cervix**	**UMN-HPV CA**
Test probability of at least ASC-US for				
CIN grade 1	NA	40	39.4	NA
CIN grade 2	70.0	50	64.9	72.7
CIN grade 3 or worse	70.0	75	83.9	72.7
Test probability of at least HSIL for				
CIN grade 1	NA	4	3.0	NA
CIN grade 2	17.5	18	22.5	
CIN grade 3	45.5	56	35.0	
Cervical cancer	53.9	60	35.0	
Specificity (CIN grade 1 or worse)[a]	91.0	97.6	96.2	91.9
Cure rate precancerous lesions	100	100	93.6[b]	100
10-year age-specific probability that clinical cervical cancer will lead to death (100%—10-year survival)[c]				
Local				
0	5.7	9.7	21.3	4.4
30	5.3	9.7	21.3	4.4
45	8.3	10.8	21.3	9.4
60	13.3	22.9	21.3	31.2
80	39.0	34.5	21.3	31.2
100	39.0	34.5	21.3	31.2
Regional				
0	42.9	45.5	46.7	49.8
30	38.7	45.5	46.7	49.8
45	38.0	51.1	46.7	58.9
60	43.9	55.4	46.7	70.3
80	66.7	68.7	46.7	70.3
100	66.7	68.7	46.7	70.3
Distant				
0	86.6	45.5	81.0	71.6
30	77.6	45.5	81.0	71.6
45	79.5	51.1	81.0	73.5
60	81.2	55.4	81.0	88.2
80	94.7	68.7	81.0	88.2
100	94.7	68.7	81.0	88.2
Reduction of the risk of dying of cervical cancer after screen-detected cancer per cancer stage				
Micro invasive	NA	89.4	NA	NA
Local	NA	50	15	NA
Regional/distant	NA	20	0	NA

ASC-US, atypical squamous cells of undetermined significance; CIN, cervical intraepithelial neoplasia; HSIL, high-grade squamous intraepithelial lesion; NA, not applicable.

[a]CIN grade 2 or worse for the Harvard and UMN-HPV CA models, since they do not model CIN grade 1.

[b]Recurrence in women treated for precancerous lesions is modeled explicitly with a different natural history.

[c]For the Harvard model, these are the 5-year probabilities of dying applied in the first 12 months for squamous cell carcinoma, after which, survival improves with time since diagnosis through year 20. No excess mortality is assumed after 20 years of survival.

in MISCAN, 27% in the UMN-HPV CA model, and 32% in the Harvard model.

Long-term (15-year) effects of screening. After 15 years, the proportion of cancer cases prevented ranged from 55% to 74% across the models (Table 2). The long-term differences in the percentages of prevented cancer cases were caused by differences in all 3 possible sources of

residual cancer following one-time screening: the lesion that caused cancer prior to age 60 years was not present at age 45 years in 12% of the cases in the UMN-HPV CA model, compared to 6% in MISCAN, 4% in the Policy1-Cervix, and 2% in the Harvard model. Furthermore, the percentage of the lesions present, but not detected, ranged from 22% to 38%. In the Harvard model, 5% of the precancerous lesions that were detected at age 45 years were
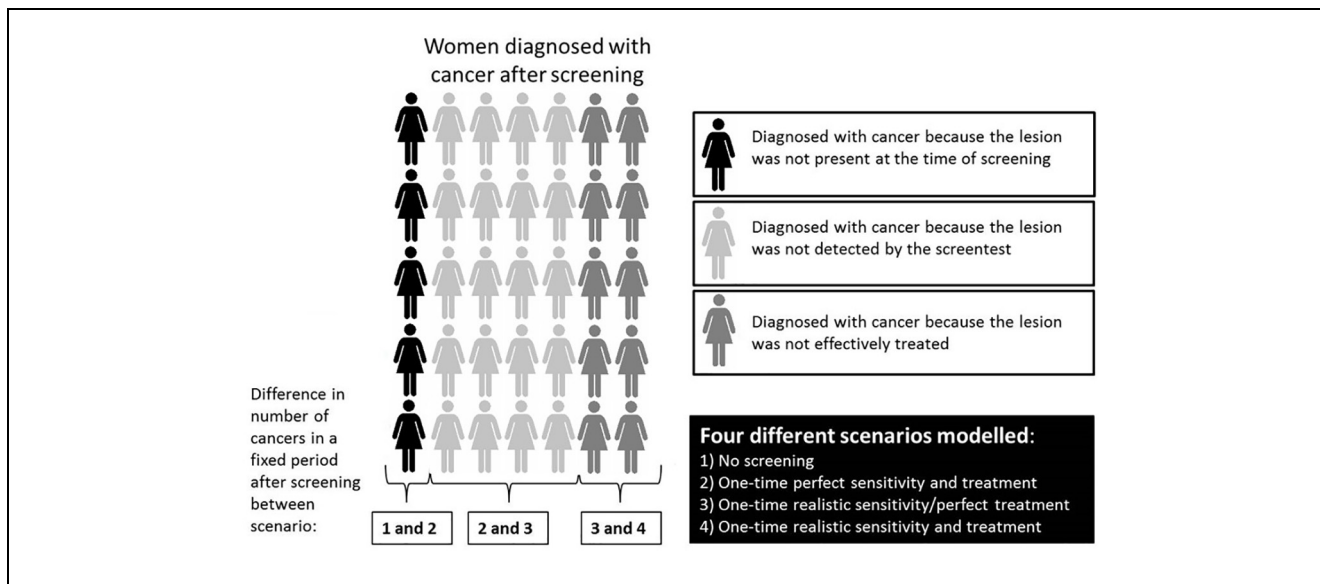
Figure 1 A schematic presentation how the impact of 3 contributors to cancer (mortality) after screening can be disentangled with modeling.
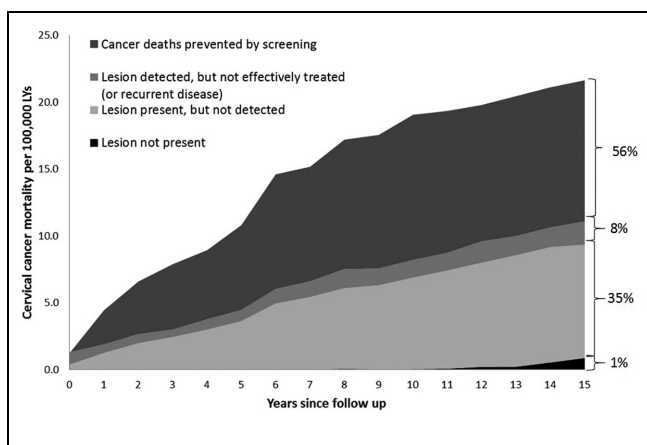

Figure 2 Illustrative example using the Harvard model: cervical cancer mortality rate, divided by explanation ("contributor") of the mortality, after 1 screening round at age 45 years. Each area in the graph represents a comparison between numbers of cancer deaths in one scenario v. another scenario (see Figure 1). The colors of the graph match the colors of the "women" in Figure 1, and the percentages refer to the attributional fractions presented in Table 2 (mortality in the Harvard model after 15 years of follow-up). The dark gray area ("cancer deaths prevented"; i.e., the difference between "no screening" and "realistic sensitivity and treatment") is the screening effect (Figure 3 and Table 3).

not treated effectively and caused cancer before the age of 60 (compared to 0% in the other models).

Screening effects over time. The proportion of clinical cancer cases prevented by realistic screening after age 45 years decreased after the screening age in all models (Figure 3). The difference between the models in the reduction of cancer as a result of screening was stable over time.

Mortality Reduction Predictions

Short-term (5-year) effects of screening. After 5 years, the proportion of deaths prevented by screening in each model ranged from 42% to 59% (Table 2). This difference was primarily caused by differences in the effectiveness of treatment (or probability of recurrent disease) of detected lesions (proportion of deaths ranging from 38% to 10%).

Long-term (15-year) effects of screening. After 15 years, the proportion of deaths prevented by screening in each model was comparable and ranged from 56% to 62% (Table 2). In all models, the most important contributor to cancer death after screening was imperfect test sensitivity: in the Harvard model, of the lesions that were present at age 45 years and resulted in a cervical cancer death before age 60 years, 35% were not detected with screening at age 45 years, compared to 21% in the Policy1-Cervix model, 27% in MISCAN, and 26% in the UMN-HPV CA model. Of the detected lesions that resulted in a cervical cancer death before age 60 years, 16% could not be

Table 2 Attributable Fraction of Cervical Cancer Incident Cases and Deaths by Model[a]

Origin	Harvard, %	MISCAN, %	Policy1-Cervix, %	UMN-HPV CA, %
Incidence				
5 years of follow-up				
Lesion not present	0	0	0	2
Lesion present but not detected	32	26	19	28
Lesion detected but not effectively treated (or recurrent disease)	3	0	0	0
Cancer cases prevented by screening	65	74	81	70
15 years of follow-up				
Lesion not present	2	6	4	12
Lesion present but not detected	38	27	22	25
Lesion detected but not effectively treated (or recurrent disease)	5	0	0	0
Cancer cases prevented by screening	55	66	74	63
Mortality				
5 years of follow-up				
Lesion not present	0	0	0	0
Lesion present but not detected	31	25	20	29
Lesion detected but not effectively treated (or recurrent disease)	10	20	38	17
Cancer deaths prevented by screening	59	55	42	53
15 years of follow-up				
Lesion not present	1	3	2	6
Lesion present but not detected	35	27	21	27
Lesion detected but not effectively treated (or recurrent disease)	8	8	16	7
Cancer deaths prevented by screening	56	62	61	60

[a]Attributable fraction (i.e., origin) of cervical cancer incident cases and deaths that occur 15 years and 5 years after screening at age 45 years in the 4 Cancer Intervention and Surveillance Modeling Network models. Each fraction represents the difference between 2 modeled scenarios (see Figure 1); the percentage cancer cases/deaths prevented is the difference between the "realistic sensitivity and treatment" scenario and "no screening."

successfully treated (or recurrent disease occurred) in the Policy1-Cervix model, compared to only 8% in MISCAN and Harvard and 7% in UMN-HPV CA.

Screening effects over time. The proportion of cancer deaths prevented by realistic screening after age 45 years increased rapidly after the screening age in all models (Figure 3); however, in the Harvard model, the maximum preventable proportion was reached more rapidly (after approximately 3 years) compared to the other models (approximately 10 years for UMN-HPV CA, 9 years for Policy1-Cervix, and 5 years for MISCAN). As a result, the attributable fractions of cervical cancer death for the different origins vary in time between the 4 different CISNET models (Table 2).

Discussion

This study is the first to apply the MCLIR method to illustrate how model structure and assumptions affect cervical cancer incidence and mortality predictions. To understand variations in model estimates of screening effects, the analysis decomposed the relative contributions of model-specific structures and assumptions regarding the preclinical duration of cervical lesions, the ability of a screening test to detect lesions, and the role of treatment of cervical lesions and their subsequent recurrence after treatment as contributors to cervical cancer incidence and mortality predictions.

The independently developed models showed similar long-term reductions in cervical cancer incidence and mortality after a one-time screening test at age 45 years. We found that these reductions were primarily caused by the probability of detecting a present lesion and by the effectiveness of treatment of detected lesions, although the attributional fraction of the 2 was different across the 4 models. Differences across models in the probability to detect a lesion can be explained by the sensitivity of the screening test per health state (which ranges from 70.0% for cervical intraepithelial neoplasia type 3 or worse in the Harvard model, compared to 83.9% in Policy1-Cervix, for example) but also by the prevalence of the different health states at the time of screening. The latter can differ across models because of differences in natural history assumptions. The fraction of cancer cases (and deaths) that occurred within 15 years after screening (i.e., were not yet present at the time of screening) also differed across the models (from 2% to 12%). This difference can be explained by assumptions on durations of the different health states. In separate CISNET analysis, we showed
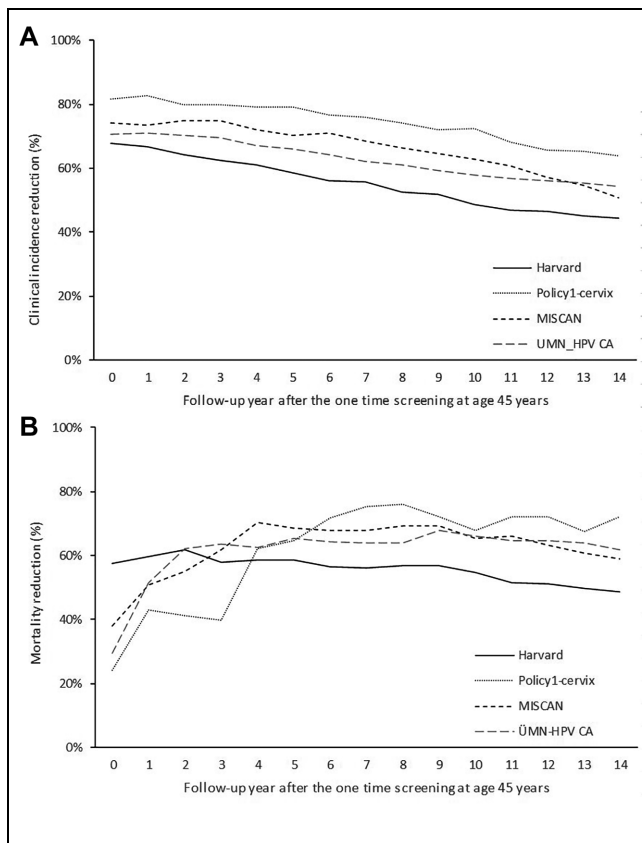
Figure 3 Screening effects in 4 different models (refers to dark gray area in Figure 2, the illustrative example): reduction in clinical incidence (A) and incidence-based mortality (B) after one-time realistic screening at age 45 years for the 4 different models.

that the median time from HPV acquisition to cancer detection ranged from 18.0 to 26.0 years across the 4 models.[9] Furthermore, we found that the short-term reductions in cervical cancer deaths were different across the models. This variation is mainly explained by differences in the effectiveness of treatment of detected cancers (attributional fraction ranged from 10% to 38%), which is the effect of the stage distributions of cancers detected without screening (clinically) and with screening, as well as the different survival probabilities per stage (Table 1).

Uncertainty on Unobservable Model Inputs and Structures

There is limited knowledge of certain model inputs and structures, because the underlying natural history processes and parameters are unobservable in the real world. First, the absolute test sensitivity of cytological screening has never been examined in a randomized clinical trial in a high-income country, and Pap characteristics have been

shown to be highly heterogeneous across settings.[10] Furthermore, since opportunistic cervical cancer screening was already introduced decades ago in most countries, including the United States,[11] information on national cervical cancer incidence and mortality before the introduction of screening is lacking in most high-income countries. As a result, the effect of cervical cancer screening on cervical cancer incidence and mortality is somewhat uncertain. Second, the preclinical duration for lesions that caused cervical cancer incident cases and deaths (i.e., dwell times) cannot be observed using epidemiological data. To estimate the preclinical durations, statistical models have been used. An evaluation of British Columbia data[12] showed that, after correcting for lead time bias, a median duration of 15 years was estimated from carcinoma in situ to invasive cancer.[13] More recently, longer durations were estimated using statistical models that included HPV.[14,15] Third, the rate of successful treatment and recurrence after treatment for detected cervical lesions depends on the grade of the lesion as well as the assumptions on loss to follow-up of diagnosed women, sensitivity of diagnostic colposcopy, and type of treatment modality. Most models assumed a 100% success rate for treatment of precancerous lesions, but some assume a certain risk of recurrence. Survival for invasive lesions is based on observed survival rates from cancer registries (Table 1).

Because it is highly unlikely that new data on these 3 issues will become available, the way to handle uncertainties in model estimates is to perform sensitivity analyses to investigate the robustness of the results.

Limitations

Usually, models are characterized by describing the modeling approach, model inputs, and assumptions. In this study, we examined model outcomes to provide insight into the mechanics of incidence and mortality predictions. Several caveats should be considered in evaluating this method. First, the effect of a single screen on cervical cancer incidence and mortality is not the same as the effect of repeated screening over decades, nor does it capture a common context for evaluations in high-income countries, which involves evaluating the impact of changes to screening (e.g., the introduction of HPV testing) in the context of long-established stable cytology programs. The results in this study are therefore not directly translatable to projections of the effects of a routine screening program on overall cervical cancer incidence and mortality. Second, we chose to present the MCLIR for a single screen at age 45 years because this is approximately in the middle of the recommended cervical screening age range (25–65 years in many

high-income countries). However, the relative contributions of the different origins may vary by age. Future work evaluating the MCLIR scenarios at different ages might provide insight into age-specific and between-model differences in lesion inception, progression, and test characteristics and the impact of these on cervical cancer incidence and mortality.

Conclusion and Recommendation

CISNET collaborative modeling predictions are increasingly used by policymakers to inform screening guidelines.[16–19] How different models arrive at their predictions of the harms and benefits of screening and treatment may be perceived as opaque due to the complexity of the models. A previous study proposed to include the projected MCLIR as a prediction measure in the description of screening models.[2] A complete description of the structure and inputs of a model is still necessary, but the knowledge about the models generated with MCLIR analyses can help explain model differences in predictions of screening effectiveness.

We showed that 4 independently developed CISNET models showed similar long-term reductions in cervical cancer incidence and mortality, but the relative contributions of the 3 possible factors varied across the models. The differences among the 4 cervical cancer screening models reflect genuine uncertainty on unobservable model inputs and structures. Further sensitivity analyses should be carried out to understand the impact of variability in modeling these uncertainties.

ORCID iDs

Inge M. C. M. de Kok (ID) https://orcid.org/0000-0002-9419-0452
Emily Groene (ID) https://orcid.org/0000-0003-1636-7061

References

1. National Cancer Institute. The Cancer Intervention and Surveillance Modeling Network. Available from: https://cisnet.cancer.gov/

2. 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(4):540–9.

3. van den Broek JJ, van Ravesteyn NT, Mandelblatt JS, et al. Comparing CISNET breast cancer models using the maximum clinical incidence reduction methodology. *Med Decis Making*. 2018;38(1, suppl):112S–25S.

4. 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(5):545–55.

5. Naber SK, de Kok IM, Matthijsse SM, van Ballegooijen M. The potential harms of primary human papillomavirus screening in over-screened women: a microsimulation study. *Cancer Causes Control*. 2016;27(4):569–81.

6. Simms KT, Smith MA, Lew JB, Kitchener HC, Castle PE, Canfell K. Will cervical screening remain cost-effective in women offered the next generation nonavalent HPV vaccine? Results for four developed countries. *Int J Cancer*. 2016;139(12):2771–80.

7. Sawaya GF, Sanstead E, Alarid-Escudero F, et al. Estimated quality of life and economic outcomes associated with 12 cervical cancer screening strategies: a cost-effectiveness analysis. *JAMA Intern Med*. 2019;179(7):867–78.

8. US Preventive Services Task Force, Curry SJ, Krist AH, et al. Screening for cervical cancer: US Preventive Services Task Force recommendation statement. *JAMA*. 2018;320(7):674–86.

9. Burger EA, de Kok IMCM, Groene E, et al. Estimating the natural history of cervical carcinogenesis using simulation models: a CISNET comparative analysis. *J Natl Cancer Inst*. 2019; djz227 [epub].

10. Miller AB. Quality assurance in screening strategies. *Virus Res*. 2002;89(2):295–9.

11. Habbema D, De Kok IM, Brown ML. Cervical cancer screening in the United States and the Netherlands: a tale of two countries. *Milbank Q*. 2012;90(1):5–37.

12. Yu SZ, Miller AB, Sherman GJ. Optimising the age, number of tests, and test interval for cervical cancer screening in Canada. *J Epidemiol Comm Hlth*. 1982;36:1–10.

13. van Oortmarssen GJ, Habbema JDF. The duration of preclinical cervical cancer and the reduction in incidence of invasive cancer following negative Pap-smears. *Int J Epidemiol*. 1995;24(2):300–7.

14. Insinga RP, Dasbach EJ, Elbasha EH. Epidemiologic natural history and clinical management of human papillomavirus (HPV) disease: a critical and systematic review of the literature in the development of an HPV dynamic transmission model. *BMC Infect Dis*. 2009;9:119.

15. Vink MA, Bootsma MC, Wallinga J. Serial intervals of respiratory infectious diseases: a systematic review and analysis. *Am J Epidemiol*. 2014;180(9):865–75.

16. Burnside ES, Lee SJ, Bennette C, et al. Using collaborative simulation modeling to develop a web-based tool to support policy-level decision making about breast cancer screening initiation age [published online July 8, 2017]. *MDM Policy Pract*.

17. de Koning HJ, Gulati R, Moss SM, et al. The efficacy of prostate-specific antigen screening: Impact of key components in the ERSPC and PLCO trials. *Cancer*. 2018;124(6):1197–206.

18. de Koning HJ, Meza R, Plevritis SK, et al. Benefits and harms of computed tomography lung cancer screening strategies: a comparative modeling study for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2014;160(5):311–20.

19. Knudsen AB, Zauber AG, Rutter CM, et al. Estimation of benefits, burden, and harms of colorectal cancer screening strategies: modeling study for the US Preventive Services Task Force. *JAMA*. 2016;315(23):2595–609.