American Journal of Epidemiology
Vol. 183, No. 3
DOI: 10.1093/aje/kwv262
Advance Access publication:
December 24, 2015

Practice of Epidemiology

A System Dynamics Model of Serum Prostate-Specific Antigen Screening for Prostate Cancer

Anton Palma, David W. Lounsbury, Nicolas F. Schlecht, and Ilir Agalliu*

* Correspondence to Dr. Ilir Agalliu, Department of Epidemiology and Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Belfer Bldg., Room 1315-B, New York, NY 10461 (e-mail: ilir.agalliu@einstein.yu.edu).

Initially submitted November 19, 2014; accepted for publication July 30, 2015.

Since 2012, US guidelines have recommended against prostate-specific antigen (PSA) screening for prostate cancer. However, evidence of screening benefit from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening trial and the European Randomized Study of Screening for Prostate Cancer has been inconsistent, due partly to differences in noncompliance and contamination. Using system dynamics modeling, we replicated the PLCO trial and extrapolated follow-up to 20 years. We then simulated 3 scenarios correcting for contamination in the PLCO control arm using Surveillance, Epidemiology, and End Results (SEER) incidence and survival data collected prior to the PSA screening era (scenario 1), SEER data collected during the PLCO trial period (1993–2001) (scenario 2), and data from the European trial's control arm (1991–2005) (scenario 3). In all scenarios, noncompliance was corrected using incidence and survival rates for men with screen-detected cancer in the PLCO screening arm. Scenarios 1 and 3 showed a benefit of PSA screening, with relative risks of 0.62 (95% confidence interval: 0.53, 0.72) and 0.70 (95% confidence interval: 0.59, 0.83) for cancer-specific mortality after 20 years, respectively. In scenario 2, however, there was no benefit of screening. This simulation showed that after correcting for noncompliance and contamination, there is potential benefit of PSA screening in reducing prostate cancer mortality. It also demonstrates the utility of system dynamics modeling for synthesizing epidemiologic evidence to inform public policy.

cancer-specific mortality; policy evaluation; prostate cancer; prostate-specific antigen screening; system dynamics modeling

Abbreviations: CI, confidence interval; ERSPC, European Randomized Study of Screening for Prostate Cancer; PLCO, Prostate, Lung, Colorectal and Ovarian; PrCa, prostate cancer; PSA, prostate-specific antigen; RR, relative risk; SDM, system dynamics modeling; SEER, Surveillance, Epidemiology, and End Results.

Prostate cancer (PrCa) is the most commonly diagnosed solid tumor and the second-leading cause of cancer death among men in the United States (1, 2). In 2012, the US Preventive Services Task Force recommended against use of serum prostate-specific antigen (PSA) screening for PrCa, concluding that there is moderate-to-high certainty that screening yields little benefit and significant potential harm for most men (3). Although this decision considered many aspects of PSA screening, including overdiagnosis and overtreatment, the recommendation was based on the results of 2 large PrCa screening trials conducted in the United States and Europe. In the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, which was conducted in the United States, Andriole et al. (4, 5) reported a relative risk of 1.09 (95% confidence interval (CI): 0.87, 1.36) for PrCa-specific mortality associated with screening after 13 years of follow-up. In contrast, in the European Randomized Study of Screening for Prostate Cancer (ERSPC), Schröder et al. (6–8) reported a relative risk of 0.71 (95% CI: 0.69, 0.91) for PrCa mortality associated with screening at 13 years. Although both trials had large sample sizes, they differed with respect to study population, participant selection, implementation of screening frequency and protocols across various centers, and the PSA screening threshold for biopsy (3 ng/mL vs. 4 ng/mL), as well

as noncompliance and contamination rates (5, 7–11), which could explain the discrepancy in their results.

Simulation modeling can be used to compare results from screening trials and synthesize other sources of population-level data (12–16). To date, several modeling approaches have been applied to explore the impact of intervention contamination, noncompliance, and overdiagnosis in the PLCO trial to determine the robustness of the trial's results, and these efforts have generated effect estimates as high as 52%, 11%, and 84%, respectively (10–13, 15, 17). However, none of the prior simulation studies attempted to extrapolate findings to longer follow-up periods (e.g., 20 years), where one would expect to observe the largest benefit of screening for slowly progressing tumors like PrCa, which has excellent 5- and 10-year survival rates (18).

System dynamics modeling (SDM) is a novel simulation method that can be used to explore different scenarios that could explain differences in results between the PLCO and ERSPC trials. It is a robust, deterministic mathematical modeling approach that has been used primarily to evaluate the potential influence of public health programs and policies across various health disciplines (19–27). In epidemiology, SDM can be used to explore hypothesized causal mechanisms and dynamic relationships in health systems (28) by using differential equations to simulate the transition of people over time between different states (e.g., healthy to diseased). It also offers an opportunity to conduct virtual experiments through simulation of intervention trials to explore alternative scenarios that would not otherwise be practical or ethical in the real world.

SDM is a particularly useful simulation tool in epidemiologic studies due to its "top-down" or macrosimulation approach, which emphasizes model parameterization at the aggregate level (21, 22, 29). As such, it is well-equipped to synthesize common forms of epidemiologic data as reported in the research literature (i.e., group average risks and rates). In this paper, we describe the use of a system dynamics model to replicate the PLCO trial and test alternative outcomes based on 3 different simulated scenarios, while correcting for weaknesses and inconsistencies in trial implementation.

METHODS

Procedures for system dynamics model-building and validation are organized around the purpose of the model and the quality of available evidence from published data that inform its construction, as well as deliberation about key assumptions regarding model parameterization and calibration (22, 30). Our aim in this system dynamics model was to replicate the PLCO trial with specific corrections for contamination and noncompliance, to assess whether benefits of PSA screening for PrCa-specific mortality would be revealed.

Stock-and-flow diagram

Using SDM, we first replicated the PLCO study design and outcomes through 13-year follow-up using published data. The results of the PLCO and ERSPC trials have been previously described (4–8) and are summarized in Table 1. Figure 1 depicts our model structure, presented as a stock-and-flow diagram.

Table 1. Results From 2 Randomized Clinical Trials of Prostate Cancer Screening

Randomized Clinical Trial	No. of Enrolled Participants	Location	Enrollment Dates	Age Range, years	Median Follow-up, years	Frequency of Screening	Control Arm	PSA Cutoff for Biopsy, ng/mL	Actual PSA Screening Rate, %		Prostate Cancer-Specific Mortality	
									Control Arm	Screening Arm	RR	95% CI
PLCO Cancer Screening Trial	76,685	United States	1993–2001	55–74	11.5	Annual PSA testing for 6 years Annual DRE for 4 years	Usual care	4	85	40–52	1.09	0.87, 1.36
ERSPC	162,243	Europe	1990–1994	55–69	11.3	PSA testing every 4 years (in Sweden, every 2 years)[b]	No screening	3 or 4[a]	82	6–12	0.79	0.69, 0.91

Abbreviations: CI, confidence interval; DRE, digital rectal examination; ERSPC, European Randomized Study of Screening for Prostate Cancer; PLCO, Prostate, Lung, Colorectal and Ovarian; PSA, prostate-specific antigen; RR, relative risk.

[a] Most European centers used a PSA cutoff value of 3.0 ng/mL, while in Finland and Italy the PSA cutoff point was 4 ng/mL.

[b] The PSA screening interval at 6 of the 7 centers in Europe was every 4 years, while in Sweden it was every 2 years.

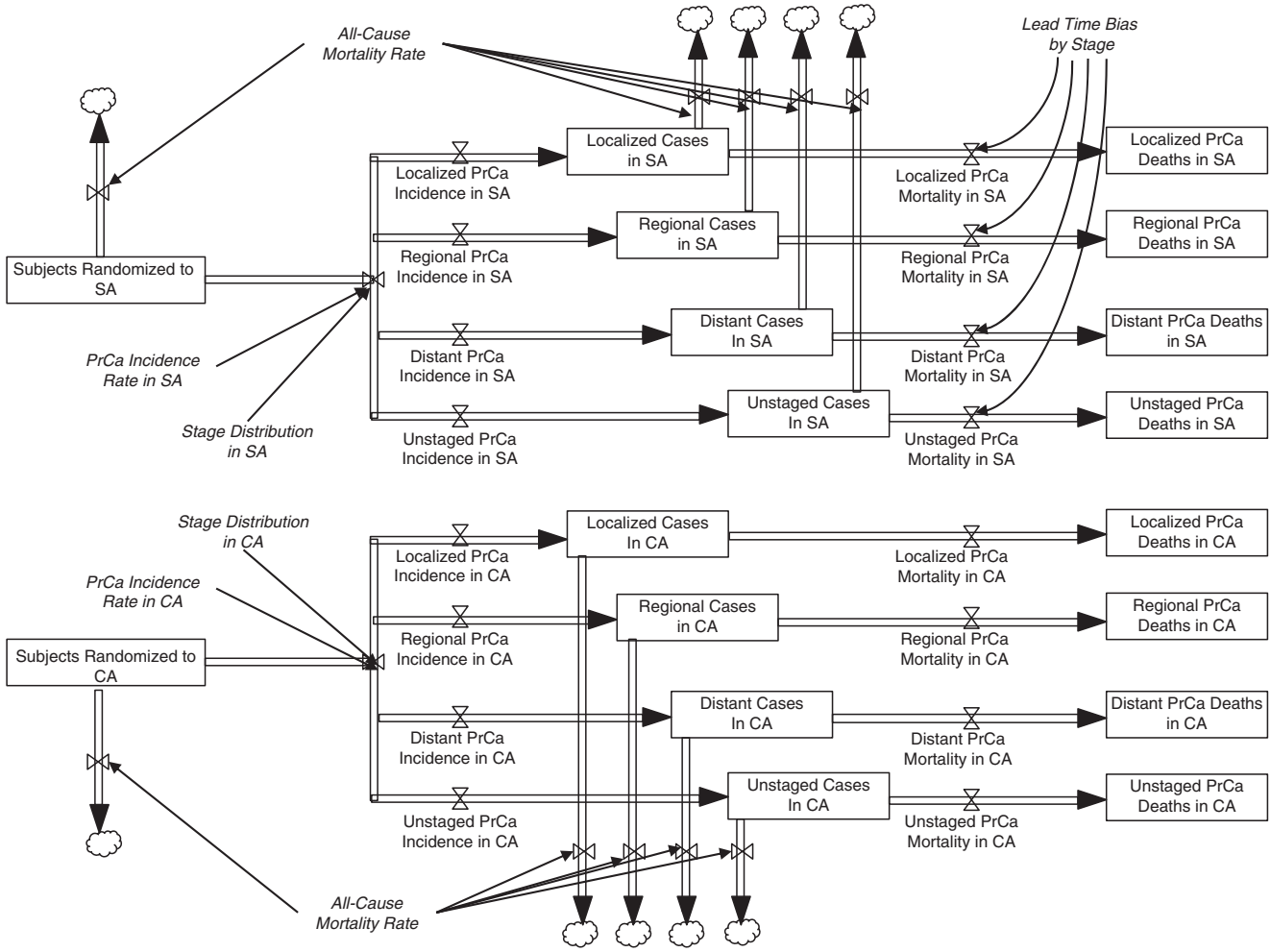

Figure 1. Stock-and-flow diagram for a system dynamics model of prostate cancer (PrCa) screening in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. Stocks, denoted by boxes, represent accumulations of subjects. Flows, denoted by pipes with valves, represent rates of movement between stocks. Auxiliaries, denoted by an *italic* font, represent variables in equations used to calculate rates. For simplicity, arrows linking stocks that are used to calculate rates were omitted from the diagram. Cancer stage was defined using the Surveillance, Epidemiology, and End Results categories as follows: localized (stages I and II), regional (stage III), distant (stage IV), and unstaged/unknown. CA, control arm; SA, screening arm.

Three basic types of SDM structures are shown: stocks, flows, and auxiliary variables. Stocks, represented by boxes, are accumulations of units (e.g., study subjects) in certain states at a given moment. Flows, represented by double-lined arrows with valves, increase or decrease a stock over time. Auxiliary variables, represented by variable names without shapes (denoted by *italic* in Figure 1), are terms used to build equations determining the rates of flows. Their values vary as determined by their relationships with stocks, flows, and other auxiliaries. Relationships between structures are shown as single-lined arrows. A flow structure that ends with a cloud icon represents a sink, where units flow outside the scope of the model. Cloud icons in our model indicate all-cause mortality.

Our stock-and-flow diagram shows 2 stocks initially parameterized to the total number of men randomized into the screening and control (usual-care) arms of the PLCO trial. The model structure represents how subjects were followed from study entry through PrCa incidence to death (either PrCa-specific or other cause of death) or the end of follow-up, with PrCa-specific mortality being the primary outcome of interest. This allowed our simulation results to replicate reported incidence and mortality data by person-years of follow-up in the PLCO trial (4, 5). Other stocks represent accumulation of incident PrCa and PrCa deaths at any given time. We stratified incident cases by tumor stage, using categories from the Surveillance, Epidemiology, and End Results (SEER) program: localized (stages I and II), regional (III), distant (IV), or unstaged (18). The flows represent incidence and mortality rates per person-year of follow-up and were used to compute the numbers of subjects transitioning between stocks at annual intervals.

Data sources and simulated scenarios

Data for simulations of these scenarios were drawn from several sources: 1) reported PLCO and ERSPC trial data, 2) PrCa

incidence and stage-specific survival rates taken from SEER data collected prior to the PSA screening era (1975–1987) and during the PLCO trial period (1993–2001), and 3) US mortality data taken from *Morbidity and Mortality Weekly Reports*. Web Table 1 (available at http://aje.oxfordjournals.org/) describes assumptions and parameters for all scenarios. We calibrated the model using an iterative process of testing several parameter values (taken from published trial data) and compared observed trial data with the simulated behavior of the model (22, 31). We then compared the base case replication with the results of the 3 simulated scenarios correcting for noncompliance and contamination.

Base case scenario (PLCO trial replication). The base case scenario replicated PLCO trial data through 13 years of follow-up and then extrapolated results to 20 years of follow-up. In this scenario, annualized PrCa incidence rates were computed by dividing cumulative incidence by person-years of follow-up, and were weighted by the cancer stage distribution reported in the PLCO trial (4, 5). Mortality was computed in 2 ways. First, we calculated annual mortality rates by dividing cumulative numbers of deaths by person-years of follow-up, stratified by cancer stage in PLCO. Alternatively, we estimated cancer-specific mortality using stage-specific PrCa survival rates obtained from SEER data during the study period. Although the PLCO trial found a healthy volunteer effect (32), both methods yielded comparable results; thus, we report results using directly calculated annualized mortality rates. We compared simulated annual numbers of PrCa incident cases and deaths with the observed data reported by Andriole et al. (5).

Simulated scenarios. In the PLCO trial, Andriole et al. (4) reported 85% compliance with PSA testing in the screening arm and 40%–52% contamination (i.e., receipt of PSA screening) in the control arm. Thus, we simulated 3 scenarios to correct for noncompliance and contamination. In all 3 scenarios, we corrected for noncompliance by using PrCa incidence and stage-specific survival rates of men with screen-detected cancer in the PLCO screening arm (10) (Web Table 1). In scenario 1, we simulated PrCa incidence and stage-specific survival using SEER data collected prior to the PSA screening era (1975–1988) (4, 18). In scenario 2, we used SEER data for calendar period 1993–2001, corresponding to the enrollment period in the PLCO trial (18). In scenario 3, we corrected for contamination in the PLCO control arm using data from the ERSPC control arm (1991–2005), which had a lower reported level of contamination (6%–12% in various countries) (7, 33). To parameterize the model, we estimated annualized PrCa incidence and mortality in the ERSPC control arm from pooled data (7, 8) and adjusted stage distribution, incidence, and mortality such that the corrected PLCO control arm in simulated scenario 3 would reflect what would have been observed had it been similar to that in the ERSPC control arm.

All 3 scenarios were parameterized by estimating PrCa incidence and mortality in men who actually undertook screening compared with those who did not. To account for earlier detection and longer pseudosurvival caused by screening, we accounted for lead-time bias in survival for localized, regional, and unstaged PrCa, whereas for distant PrCa we simulated no lead time, as these cases were most likely diagnosed because of symptoms (4, 5). Lead-time bias in our system dynamics model was derived from Telesca et al. (12), who estimated lead times of 4.59 years and 6.78 years for whites and blacks, respectively. We used the race distribution in the PLCO trial (85% non-Hispanic white, 15% non-Hispanic black or other) (4) and the above data to obtain an overall lead time of 4.92 years. In scenario 1, where survival time was estimated using SEER survival data from the pre-PSA-screening era, we added the lead time to the corrected screening arm; while in scenario 2, where survival time was estimated with SEER data collected during the period when PSA screening was available in the general population, we subtracted lead time from the corrected control arm.

Lastly, we used all-cause mortality rates to account for other causes of death (non-PrCa-specific) observed in the PLCO trial. Data on age-specific all-cause mortality were obtained from *Morbidity and Mortality Weekly Reports* (34), and were weighted by the age distribution of the PLCO cohort at entry and accounted for the aging of the cohort during follow-up. We simulated 8% loss to follow-up at 10 years in SDM for PLCO participants, as reported in the literature (5). The model simulated cumulative PrCa incidence, PrCa-specific deaths, and the relative risk and risk difference (with 95% confidence intervals) for PrCa-specific mortality in the screened arm versus the unscreened (control) arm of the study at 10, 13, and 20 years after enrollment.

External validation. To validate our model, we used published results from the Rotterdam, Netherlands, component (9) of the ERSPC trial to assess the concordance between the observed and simulated data in an independent data set (Web Figure 1). Rotterdam was one of the initial enrollment sites of the ERSPC trial and thus had longer follow-up and more complete data. Using the same procedure as described for the base case scenario, we calculated interval-specific incidence and mortality rates (for every 4 years) using published results from Roobol et al. (9). We estimated incidence in the screening arm for the first 4-year interval using cases occurring in the first interval plus those in the first screening round divided by the number of persons at risk in the beginning of the period. The incidence in the control arm was estimated using 4-year interval incident cases only. Observed cancer-stage distributions were applied to the total incidence rates to obtain stage-stratified incidence within each interval. PrCa-specific mortality rates and other mortality rates for each 4-year interval were calculated in the same way. All model-building and analysis was conducted using Vensim software (Ventana Systems, Inc., Harvard, Massachusetts).

RESULTS

The base case replication model successfully reproduced the overall outcomes of the PLCO trial and the trajectories of PrCa incidence and mortality in both the screening and control arms (Figure 2A and Figure 3A). At 13 years of follow-up, the model simulated a total of 4,447 and 3,915 cumulative incident PrCa cases in the screening and control arms, respectively (Table 2). These numbers represent 5% and 3% increases in PrCa diagnosis in each arm of the trial in comparison with the observed 4,250 and 3,815 incident cases, respectively (5). Our model simulated 134 and 133

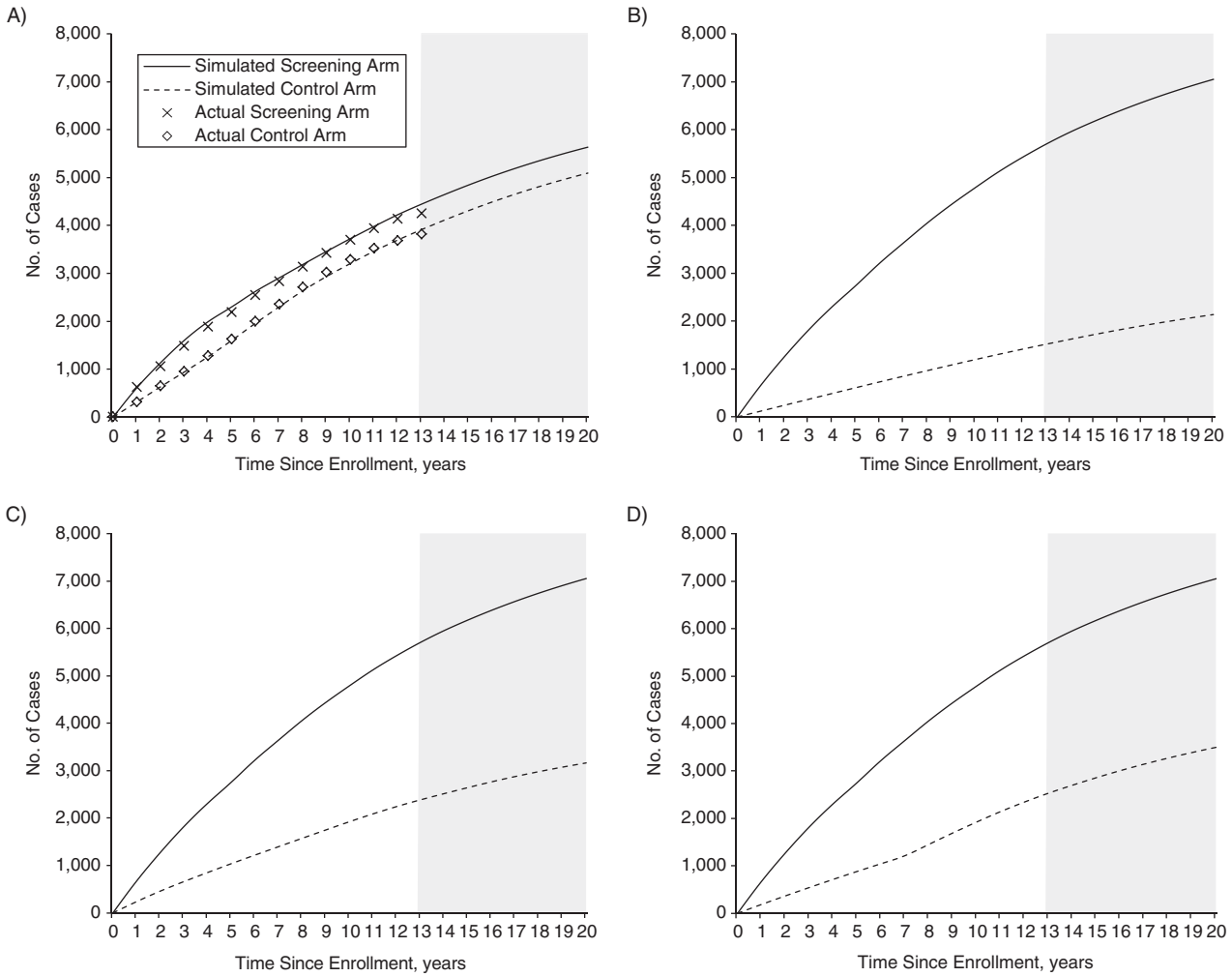

Figure 2. Simulated numbers of incident prostate cancer cases under different scenarios in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. A) Base case scenario: replication of results from the PLCO trial. B) Scenario 1: results corrected to Surveillance, Epidemiology, and End Results (SEER) data from the era before prostate-specific antigen screening (1975–1988). C) Scenario 2: results corrected to SEER data from the period 1993–2001. D) Scenario 3: results corrected to data from the European Randomized Study of Screening for Prostate Cancer control arm (1991–2005). Areas shaded in gray represent data that were extrapolated (from 13 years of follow-up to 20 years of follow-up).

cumulative PrCa-specific deaths in the screening and control arms at 13 years, respectively, which were slightly lower than those reported in the PLCO trial (Figure 3A). This yielded a relative risk of 1.00 (95% CI: 0.79, 1.27) for PrCa-specific mortality associated with PSA screening at 13-year follow-up in our model (Table 2). When the results of the base case scenario were extrapolated to 20 years, there was still no benefit of PSA screening for PrCa-specific mortality (relative risk (RR) = 0.95, 95% CI: 0.79, 1.14).

Scenario 1, which corrected for contamination using SEER PrCa incidence and mortality data collected before the PSA screening era, yielded a total of 5,706 and 1,522 incident PrCa cases and 145 and 224 PrCa-specific deaths in the screening and control arms, respectively, at 13 years of follow-up (Figures 2B and 3B). This corresponded to a statistically significant 36% decrease in PrCa-specific mortality associated with PSA screening (95% CI: 0.52, 0.79; Table 2). At 20 years of follow-up, the model simulated 7,052 and 2,139 incident PrCa cases and 249 and 402 PrCa deaths, respectively, in the screening and control arms (Figures 2B and 3B). This yielded a relative risk of 0.62 (95% CI: 0.53, 0.72) for PrCa-specific mortality at 20 years. The corresponding risk difference for the benefit of PSA screening at 20 years was a reduction of 4.0 deaths per 1,000 men screened (95% CI: −5.30, −2.70; Table 2).

In scenario 2, correcting for contamination using SEER data from the same time period as PLCO enrollment yielded 126 and 125 deaths in the screening and control arms, respectively, at year 13 (Figure 3C). The model yielded no benefit of screening at 13 years (RR = 1.01, 95% CI: 0.79, 1.30); results remained virtually unchanged at 20 years of follow-up (Table 2).

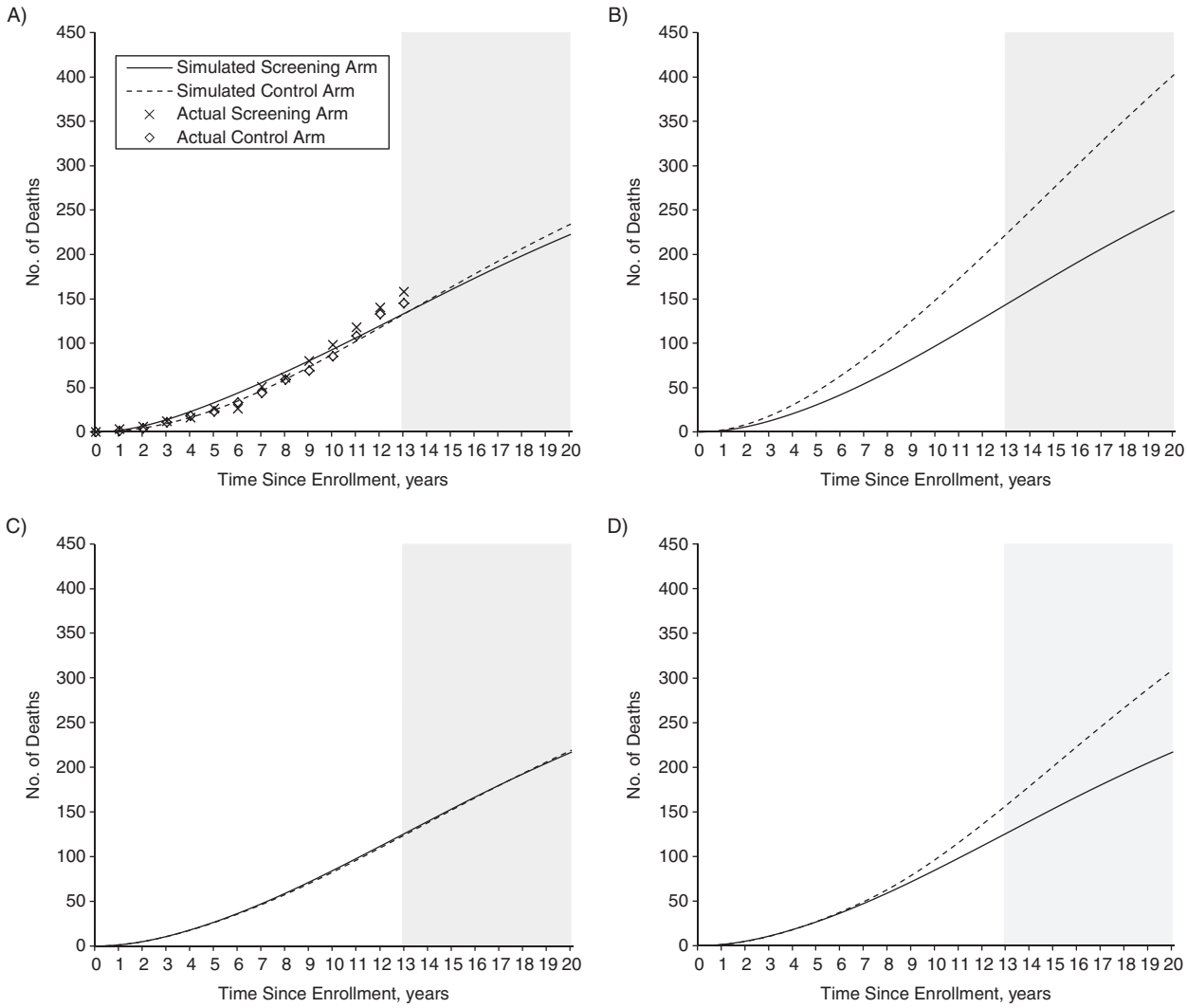

Figure 3. Simulated numbers of prostate cancer deaths under different scenarios in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. A) Base case scenario: replication of results from the PLCO trial. B) Scenario 1: results corrected to Surveillance, Epidemiology, and End Results (SEER) data from the era before prostate-specific antigen screening (1975–1988). C) Scenario 2: results corrected to SEER data from the period 1993–2001. D) Scenario 3: results corrected to data from the European Randomized Study of Screening for Prostate Cancer control arm (1991–2005). Areas shaded in gray represent data that have been extrapolated (from 13 years of follow-up to 20 years of follow-up).

In scenario 3, which corrected for contamination in the PLCO control arm using ERSPC data, simulated results yielded 2,532 incident PrCa cases and 158 PrCa-specific deaths in the PLCO control arm at 13 years of follow-up (Figures 2D and 3D). There was a relative risk of PrCa-specific mortality of 0.80 (95% CI: 0.63, 1.01) associated with PSA screening at 13 years in this scenario (Table 2). The reduction in PrCa-specific mortality was maintained through 20 years of follow-up (RR = 0.70, 95% CI: 0.59, 0.83). The corresponding attributable benefit of PSA screening for this scenario at 20-year follow-up was a reduction of 2.40 deaths per 1,000 men screened (95% CI: −3.57, −1.23).

Finally, we simulated cumulative PrCa incidence and deaths also stratified by tumor stage for all of the above scenarios (Web Table 2). As reported in the PLCO trial (5), the majority of incident prostate tumors were local stage, but with lack of screening (simulated in scenarios 1 and 3) the numbers of regional and distant stage PrCa cases increased. In relation to PrCa deaths, the highest numbers of cumulative deaths were observed among distant stage tumors, and as expected, these numbers increased in scenarios with low or no screening.

Model validation

Consistent with standard SDM methodology, validation involved an iterative process that examined both model structure and behavior (30, 35). We conducted the following tests.

Model structure tests. 1) Structure verification: The model structure (Figure 1) was made parsimonious through

Table 2. Simulated Results From a System Dynamics Model of the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial at 10, 13 and 20 Years of Follow-up

Scenario and Follow-up Time, years	Screening Arm		Control Arm		Relative Risk	95% CI	Risk Difference[a]	95% CI
	Cumulative PrCa Incidence[b]	Cumulative No. of Deaths	Cumulative PrCa Incidence[b]	Cumulative No. of Deaths				
Base case scenario								
10	3,726	93	3,206	88	1.06	0.79, 1.42	0.14	−0.54, 0.83
13	4,447	134	3,915	133	1.00	0.79, 1.27	0.01	−0.82, 0.85
20	5,635	223	5,093	234	0.95	0.79, 1.14	−0.30	−1.39, 0.79
Scenario 1								
10	4,791	98	1,119	150	0.65	0.51, 0.84	−1.36	−2.16, −0.56
13	5,706	145	1,522	224	0.64	0.52, 0.79	−2.08	−3.06, −1.10
20	7,052	249	2,139	402	0.62	0.53, 0.72	−4.00	−5.30, −2.70
Scenario 2								
10	4,791	85	1,923	85	1.02	0.76, 1.38	0.05	−0.61, 0.71
13	5,706	126	2,384	125	1.01	0.79, 1.30	0.04	−0.76, 0.85
20	7,052	217	3,163	219	0.99	0.82, 1.19	−0.06	−1.12, 1.01
Scenario 3								
10	4,791	85	1,924	97	0.88	0.66, 1.17	−0.31	−1.00, 0.38
13	5,706	126	2,532	158	0.80	0.63, 1.01	−0.83	−1.68, 0.03
20	7,052	217	3,497	309	0.70	0.59, 0.83	−2.40	−3.57, −1.23

Abbreviations: CI, confidence interval; PrCa, prostate cancer.
[a] Difference in the number of deaths per 1,000 men screened.
[b] Total number of incident PrCa cases simulated (i.e., accrued) during the specified follow-up period (10, 13, or 20 years).

selection of a minimally sufficient set of parameters to replicate a clinical trial. 2) Dimensional consistency: All variables were labeled by their proper dimensions and verified to be consistent across model equations. 3) Parameter verification: As mentioned in the Methods section, all parameters were drawn from well-documented sources of data (see Web Table 1).

Model behavior tests. 1) Behavior reproduction: These tests evaluated how closely the simulated data fitted the observed trial data (0–13 years). We demonstrated in Figures 2A and 3A that the simulated data fitted the observed PLCO trial data well, for both incidence and mortality. 2) Behavior prediction: Although follow-up in both the PLCO and ERSPC trials has not yet reached 20 years, we used linear extrapolation procedures and 3-year moving averages to extrapolate results. Our simulated data (interval 13–20 years, denoted with shaded gray areas in Figures 2 and 3) are consistent with expected cumulative incidence and mortality in the extended follow-up period, assuming no changes in screening practices or medical care during this time. 3) External validation: Simulated cumulative PrCa incidence and cancer-specific mortality were compared with the observed results from the Rotterdam section of the ERSPC for validation of our system dynamics model with another data set (Web Figure 1). As observed, the simulated results replicated the observed trial data closely but underestimated both incidence and mortality, partly because of a lack of available published data on person-years at risk, which necessitated using average risks to approximate the rates. 4) Parameter sensitivity: We selected specific parameters, which defined the 3 scenarios that corrected for noncompliance and contamination (Figures 2 and 3 and Table 2). Comparison of 3 simulated scenarios with the base case scenario constituted basic sensitivity testing of our system dynamics model. We also carried out other sensitivity tests by varying all-cause mortality, lead-time bias, and the moving average window for both incidence and mortality:

- *All-cause mortality*: We accounted for aging of the PLCO cohort over time, which yielded proportional changes in PrCa incidence and mortality but no qualitative changes in behavior, since all-cause mortality removed men from the undiagnosed and diagnosed pools.
- *Lead time*: We used average lead-time estimates from Telesca et al. (12). However, Draisma et al. (13) estimated a higher lead time of 11.6 years for PSA screening in men aged 55–75 years. We applied both estimates and found that variations in lead time changed the absolute number of deaths, primarily for localized PrCa deaths, but did not qualitatively change results.
- *Moving averages*: In all scenarios, we used a 3-year moving average for data from the first and last 3 years to smooth out random variation during periods in which estimates were subject to greater error due to the small number of events.

We tested a range of moving averages from 2 years to 5 years and observed no significant differences in results for either incidence or mortality curves.

DISCUSSION

SDM belongs to the rapidly evolving domain of system science research (36, 37), which has been used to examine a variety of public health issues and policies, including infectious disease transmission and control (20, 38–46), chronic disease management (26, 29, 47, 48), partner violence (49), and tobacco control (50, 51). As such, SDM could be informative in evaluating the recent US Preventive Services Task Force recommendation guideline against use of PSA screening for PrCa (3, 16).

Our model replicated the PLCO trial, showing no benefit of PSA screening for PrCa-specific mortality at 13 years of follow-up, similar to published PLCO results (4, 5). Extending the follow-up period to 20 years in the PLCO trial did not reveal any further benefits of screening. In contrast, the 3 simulated scenarios showed different results. Scenario 1, which corrected for contamination in the PLCO control arm to simulate a "pristine" unscreened population, yielded the highest benefit of PSA screening. There were statistically significant 35%–38% reductions in PrCa-specific mortality observed from 10 years of follow-up to 20 years of follow-up. In scenario 2, however, correcting for contamination using SEER PrCa incidence and mortality data from the same time period as PLCO trial enrollment did not show any benefit of PSA screening in PrCa-specific mortality. These results were not that surprising, since PSA screening was under way in the general US population between 1993 and 2001, when enrollment in the PLCO trial was ongoing (18, 52).

Finally, correcting for contamination in PLCO using ERSPC data (scenario 3) yielded statistically significant 20% and 30% benefits of PSA screening in PrCa-specific mortality at 13 and 20 years, respectively. Although the ERSPC trial included several study centers with different screening protocols and randomization schemes, the aggregated published data from this trial demonstrated an overall 21% benefit of PSA screening, even with extended follow-up (7, 8), which was also seen in our simulated scenario. It should be noted, however, that the risk differences for both scenario 1 and scenario 3 were relatively low, with risk reductions of 2.4–4.0 PrCa deaths per 1,000 men screened (equivalent to 250–417 men invited to undergo screening in order to prevent 1 death from PrCa).

Results from scenarios 1 and 3 were also similar to those of other published simulation studies (11, 17, 53). For example, Gulati et al. (11) used a natural history of PrCa model from the Cancer Intervention and Surveillance Modeling Network to simulate a virtual PLCO trial, correcting for contamination and noncompliance. They reported that contamination rates in the screening arm of the PLCO trial attenuated the mortality benefit of PSA screening up to 28% at 10 years, which is consistent with our simulated results assuming an unscreened population. They also suggested that the power of the PLCO trial to detect a mortality difference was reduced and that contamination might explain the null findings of the PLCO trial (11).

Nevertheless, the overall estimate of the mortality benefit of PSA screening gleaned from the available clinical trials does not account for the variability in aggressive clinical phenotypes of PrCa. Cooperberg et al. (54) suggested that PSA screening, if used in conjunction with active surveillance, could minimize the harms of overtreating low-risk PrCa patients whose cancer would not have been diagnosed in the absence of screening. We simulated stage-specific incidence and mortality from PrCa in our model, and although there were small variations in simulated outcomes for different scenarios, we did not find major differences in PrCa-specific mortality by tumor stage. We attribute this to the low observed number of deaths for regional and distant stage tumors in the trial.

The benefit of screening for high-risk men (e.g., men with a family history of PrCa or men of African descent) warrants further research, since neither trial reported on the benefit of screening among these high-risk groups. Only 1 study reported a 40% risk reduction in PrCa-specific mortality in men with advanced PrCa, due to PSA screening (53). It is still unclear whether screening may be beneficial for these high-risk patients. Our system dynamics model used published data from the PLCO and ERSPC trials, which did not report on the benefits of screening for men of African origin or those with a family history of PrCa. Should such data be made available, our system dynamics model could be used to simulate results in these high-risk groups. However, in order to better inform individual patient decision-making and health policy decisions, the overall evidence should balance the small benefit of PSA screening observed in only 1 trial (ERSPC) with the overall harms of overdiagnosis and overtreatment at the population level—issues that were considered by the US Preventive Services Task Force (3) and were also recently reviewed in a meta-analysis of PrCa screening trials (55).

In this study, we demonstrated an application of SDM as a new tool in epidemiology to simulate the dynamics of social, biological, and health systems (56). Strengths of our model include the capacity to estimate the benefit of PSA screening up to 20 years of follow-up, accounting for tumor stage distribution, and to correct for contamination using a virtual unscreened population and the experience of the control arm of the ERSPC. Limitations include the use of SEER data to estimate survival in the PLCO trial—a healthy volunteer cohort (32), which might not be representative of the general US population. Our model was limited by lack of access to additional data from the trials on death rates by race/ethnicity, family history of PrCa, and tumor grade (i.e., Gleason score). Another limitation is that we used the aggregate data for the ERSPC, although there were country-specific differences in screening intervals, as well as in contamination and compliance rates. Finally, the model did not incorporate other elaborate dynamic structures that could affect PrCa-specific mortality, such as preclinical state, competing risks, and various PrCa treatments, as well as other potential genetic, behavioral, and social determinants of health.

In summary, our model demonstrates that after correcting for noncompliance and contamination using a truly unscreened population, PSA screening is associated with a reduction in PrCa-specific mortality. This study further demonstrates the utility of system dynamics for synthesizing multiple sources of epidemiologic data to inform public health policy.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York, New York (Anton Palma, David W. Lounsbury, Nicolas F. Schlecht, Ilir Agalliu); Division of Oncology, Department of Medicine, Albert Einstein College of Medicine and Montefiore Medical Center, New York, New York (Nicolas F. Schlecht); Department of Urology, Albert Einstein College of Medicine and Montefiore Medical Center, New York, New York (Ilir Agalliu); and Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York (Anton Palma).

I.A. was supported in part by a research scholar grant (MRSG-11-112-01-CNE) from the American Cancer Society. A.P. was supported in part by a training grant (T32AI114398) from the National Institute of Allergy and Infectious Diseases.

Conflict of interest: none declared.

REFERENCES

1. American Cancer Society. *Cancer Facts and Figures: 2009*. Atlanta, GA: American Cancer Society; 2009.
2. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin*. 2012;62(1):10–29.
3. Moyer VA; U.S. Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2012;157(2):120–134.
4. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med*. 2009;360(13):1310–1319.
5. Andriole GL, Crawford ED, Grubb RL 3rd, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst*. 2012;104(2):125–132.
6. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360(13):1320–1328.
7. Schröder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012;366(11):981–990.
8. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384(9959):2027–2035.
9. Roobol MJ, Kerkhof M, Schröder FH, et al. Prostate cancer mortality reduction by prostate-specific antigen-based screening adjusted for nonattendance and contamination in the European Randomised Study of Screening for Prostate Cancer (ERSPC). *Eur Urol*. 2009;56(4):584–591.
10. Pinsky PF, Blacka A, Kramer BS, et al. Assessing contamination and compliance in the prostate component of the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. *Clin Trials*. 2010;7(4):303–311.
11. Gulati R, Tsodikov A, Wever EM, et al. The impact of PLCO control arm contamination on perceived PSA screening efficacy. *Cancer Causes Control*. 2012;23(6):827–835.
12. Telesca D, Etzioni R, Gulati R. Estimating lead time and overdiagnosis associated with PSA screening from prostate cancer incidence trends. *Biometrics*. 2008;64(1):10–19.
13. Draisma G, Etzioni R, Tsodikov A, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. *J Natl Cancer Inst*. 2009;101(6):374–383.
14. Etzioni R, Gulati R, Tsodikov A, et al. The prostate cancer conundrum revisited: treatment changes and prostate cancer mortality declines. *Cancer*. 2012;118(23):5955–5963.
15. McGregor M, Hanley JA, Boivin JF, et al. Screening for prostate cancer: estimating the magnitude of overdetection. *CMAJ*. 1998;159(11):1368–1372.
16. Habbema JD, Wilt TJ, Etzioni R, et al. Models in the development of clinical practice guidelines. *Ann Intern Med*. 2014;161(11):812–818.
17. Gulati R, Inoue LY, Gore JL, et al. Individualized estimates of overdiagnosis in screen-detected prostate cancer. *J Natl Cancer Inst*. 2014;106(2):djt367.
18. Howlader N, Noone A, Krapcho M, et al. Table 23.1. Cancer of the prostate (invasive). In: *SEER Cancer Statistics Review, 1975–2011*. Bethesda, MD: National Cancer Institute; 1999. http://seer.cancer.gov/archive/csr/1975_2011/results_merged/sect_23_prostate.pdf. Updated March 3, 2013. Accessed December 4, 2014.
19. Homer JB, Hirsch GB. System dynamics modeling for public health: background and opportunities. *Am J Public Health*. 2006;96(3):452–458.
20. Duintjer Tebbens RJ, Pallansch MA, Kew OM, et al. A dynamic model of poliomyelitis outbreaks: learning from the past to help inform the future. *Am J Epidemiol*. 2005;162(4):358–372.
21. Sterman JD. Learning from evidence in a complex world. *Am J Public Health*. 2006;96(3):505–514.
22. Oliva R. Model calibration as a testing strategy for system dynamics models. *Eur J Oper Res*. 2003;151(3):552–568.
23. Wolstenholme E. System dynamics: a system methodology or a system modelling technique. *Dynamica*. 1983;9(2):84–90.
24. Milstein B, Homer J, Hirsch G. Analyzing national health reform strategies with a dynamic simulation model. *Am J Public Health*. 2010;100(5):811–819.
25. Hirsch GB, Edelstein BL, Frosh M, et al. A simulation model for designing effective interventions in early childhood caries. *Prev Chronic Dis*. 2012;9:E66.
26. Homer J, Hirsch G, Minniti M, et al. Models for collaboration: how system dynamics helped a community organize cost-effective care for chronic illness. *Syst Dynam Rev*. 2004;20(3):199–222.
27. Lounsbury DW, Schwartz B, Palma A, et al. Simulating patterns of patient engagement, treatment adherence, and viral suppression: a system dynamics approach to evaluating HIV care management. *AIDS Patient Care STDS*. 2015;29(suppl 1):S55–S63.
28. Richardson G. *Introduction to System Dynamics Modeling with DYNAMO*. 1st ed. Westford, MA: Pegasus Communications; 1981.
29. Hammond RA. Complex systems modeling for obesity research. *Prev Chronic Dis*. 2009;6(3):A97.
30. Martis M. Validation of simulation based models: a theoretical outlook. *Electron J Bus Res Methods*. 2006;4(1):39–46.
31. Wakeland W, Walker RW. Calibration of complex system dynamics models: a practitioner's report. Presented at the 29th International Conference of the System Dynamics Society, Washington, DC, July 24–28, 2011.
32. Pinsky PF, Miller A, Kramer BS, et al. Evidence of a healthy volunteer effect in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial. *Am J Epidemiol*. 2007;165(8):874–881.
33. Kerkhof M, Roobol MJ, Cuzick J, et al. Effect of the correction for noncompliance and contamination on the estimated reduction of metastatic prostate cancer within a randomized screening trial (ERSPC section Rotterdam). *Int J Cancer*. 2010;127(11):2639–2644.

34. Centers for Disease Control and Prevention. MMWR Search. Detailed mortality tables 1999–2013. http://www.cdc.gov/mmwr/distrnds.html/. Published April 2, 2004. Updated October 5, 2012. Accessed October 6, 2012.
35. Barlas Y. Formal aspects of model validity and validation in system dynamics. *Syst Dynam Rev*. 1996;12(3):183–210.
36. Mabry PL, Olster DH, Morgan GD, et al. Interdisciplinarity and systems science to improve population health: a view from the NIH Office of Behavioral and Social Sciences Research. *Am J Prev Med*. 2008;35(2 suppl):S211–S224.
37. Trochim WM, Cabrera DA, Milstein B, et al. Practical challenges of systems thinking and modeling in public health. *Am J Public Health*. 2006;96(3):538–546.
38. Chick SE, Mamani H, Simchi-Levi S. Supply chain coordination and influenza vaccination. *Oper Res*. 2008;56(6):1493–1506.
39. Thompson KM, Tebbens RJD. Using system dynamics to develop policies that matter: global management of poliomyelitis and beyond. *Syst Dynam Rev*. 2008;24(4):433–449.
40. Dangerfield BC, Fang Y, Roberts CA. Model-based scenarios for the epidemiology of HIV/AIDS: the consequences of highly active antiretroviral therapy. *Syst Dynam Rev*. 2001;17(2):119–150.
41. Flessa S. Decision support for AIDS control programmes in eastern Africa. *OR Spectrum*. 2003;25(2):265–291.
42. Goforth RR. A simulation model to assist in managing the HIV epidemic: IMAP2. *Simulation*. 1994;63(2):128–136.
43. Homer J, Ritchie-Dunham J, Rabbino H, et al. Toward a dynamic theory of antibiotic resistance. *Syst Dynam Rev*. 2000;16(4):287–319.
44. Roberts C, Dangerfield BC. Modelling the epidemiological consequences of HIV infection and AIDS: a contribution from operational research. *J Oper Res Soc*. 1990;41(4):273–289.
45. Lebcir R, Atun R, Coker R. System dynamic simulation of treatment policies to address colliding epidemics of tuberculosis, drug resistant tuberculosis and injecting drug users driven HIV in Russia. *J Oper Res Soc*. 2010;61(8):1238–1248.
46. Homer J, St. Clair C. A model of HIV transmission through needle sharing. *Interfaces*. 1991;21(3):26–49.
47. Patterson RE, Colditz GA, Hu FB, et al. The 2011–2016 Transdisciplinary Research on Energetics and Cancer (TREC) initiative: rationale and design. *Cancer Causes Control*. 2013;24(4):695–704.
48. Sabounchi NS, Hovmand PS, Osgood ND, et al. A novel system dynamics model of female obesity and fertility. *Am J Public Health*. 2014;104(7):1240–1246.
49. Hovmand PS, Ford DN. Sequence and timing of three community interventions to domestic violence. *Am J Community Psychol*. 2009;44(3-4):261–272.
50. Tengs TO, Osgood ND, Chen LL. The cost-effectiveness of intensive national school-based anti-tobacco education: results from the tobacco policy model. *Prev Med*. 2001;33(6):558–570.
51. Tobias MI, Cavana RY, Bloomfield A. Application of a system dynamics model to inform investment in smoking cessation services in New Zealand. *Am J Public Health*. 2010;100(7):1274–1281.
52. Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986–2005. *J Natl Cancer Inst*. 2009;101(19):1325–1329.
53. Wu GH, Auvinen A, Yen AM, et al. The impact of interscreening interval and age on prostate cancer screening with prostate-specific antigen. *Eur Urol*. 2012;61(5):1011–1018.
54. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: progress and promise. *J Clin Oncol*. 2011;29(27):3669–3676.
55. Ilic D, Neuberger MM, Djulbegovic M, et al. Screening for prostate cancer. *Cochrane Database Syst Rev*. 2013;1:CD004720.
56. Richardson G. *Feedback Thought in Social Science and Systems Theory*. Waltham, MA: Pegasus Communications, Inc.; 1991.