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

Anton Palma[1,2], David W. Lounsbury[1], Nicolas F. Schlecht[1], Ilir Agalliu[1]

[1] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, New York

[2] Department of Epidemiology, Columbia University, Mailman School of Public Health, New York, New York

RUNNING HEAD: System Dynamics Modeling in Epidemiologic Research

Correspondence to: Ilir Agalliu, M.D., Sc.D.

Department of Epidemiology and Population Health

Albert Einstein College of Medicine

1300 Morris Park Avenue, Belfer Bldg. Room 1315-B

Bronx, NY 10461

Tel: 1-718-430-2929

E-mail: ilir.agalliu@einstein.yu.edu

Abstract

Since 2012, the guidelines recommended against routine prostate-specific antigen (PSA) screening for prostate cancer. However, evidence for screening benefit from Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial and European Randomized Study of Screening for Prostate Cancer (ERSPC) was inconsistent, partly due to differences in noncompliance and contamination. Using system dynamics (SD) modeling, we replicated PLCO trial and extrapolated follow-up to 20 years. We then simulated three scenarios correcting for contamination in PLCO control arm using SEER incidence and survival data prior to PSA-screening era (scenario 1), during PLCO trial-period (scenario 2), and using ERSPC control arm data (scenario 3). In all scenarios noncompliance was corrected using incidence and survival rates of screen-detected men in PLCO screening arm. Both scenarios 1 and 3 showed PSA screening benefit with relative risks of 0.62 (95% CI 0.53, 0.72) and 0.70 (0.59, 0.83) for cancer-specific mortality at 20-year follow-up, respectively. In scenario 2, however, there was no benefit of screening, similar to PLCO published results. 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 SD for synthesizing epidemiologic evidence to inform public policy.

Keywords: cancer-specific mortality, policy evaluation, prostate cancer, PSA screening, system dynamics modeling

Manuscript

Prostate cancer (PrCa) is the most commonly diagnosed solid tumor and the second leading cause of cancer deaths among men in the US (1, 2). In 2012, the US Preventive Services Task Force (USPSTF) recommended against use of serum prostate specific antigen (PSA) screening for PrCa, concluding that there is moderate to high certainty that screening yields very small 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 results of two PrCa screening trials conducted in the US and Europe. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial conducted in the US reported a relative risk (RR) of 1.09 (95% CI 0.87, 1.36) of PrCa-specific mortality associated with screening after 13 years of follow-up (4, 5). In contrast, the European Randomized Study of Screening for Prostate Cancer (ERSPC), reported a RR=0.71 (95% CI 0.69, 0.91) of PrCa-mortality associated with screening at 13 years (6-8). Although both trials included large sample sizes, they differed with respect to study population, selection into the trial, implementation of screening frequency and protocols across various centers, PSA screening threshold for biopsy (3 vs. 4 ng/ml), as well as non-compliance and contamination rates (5, 7-11), which could explain the discrepancy in their results.

Simulation modeling can serve to compare results from these screening trials and synthesize other sources of population-level data (12-16). To date several modeling approaches have been used to explore the effects of contamination, noncompliance and overdiagnosis in the PLCO trial to determine sensitivity of the trial's results; and obtained 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 expects to observe the largest benefit of screening for slow-progressing tumors like PrCa, which has excellent 5- and 10-year survival rates (18).

System dynamics (SD) modeling is a novel simulation method that can be used to explore different scenarios that could explain differences in results between PLCO and ERSPC trials. SD is a robust, deterministic mathematical modeling approach that has been used primarily to evaluate the potential impact of public health programs and policies across various health disciplines (19-27). In epidemiology, SD can 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). SD modeling 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.

SD modeling is particularly a useful simulation tool in epidemiological studies due to its "top-down" or macro-simulation approach, which emphasizes model parameterization at the aggregate level (21, 22, 29). As such, SD is well-equipped to synthesize common forms of epidemiologic data as reported in the research literature (i.e., group average risks and rates). This paper describes a SD model to replicate the PLCO trial and test alternative outcomes based on three different simulated scenarios, while correcting for weaknesses and inconsistencies in trial implementation.

METHODS

Procedures for SD model-building and validation are organized around the purpose of the model, the quality of available evidence and published data to inform its purpose, as well as

deliberation about key assumptions regarding model parameterization and calibration (21, 30). The aim of this SD model was to replicate the PLCO screening trial with specific corrections for contamination and non-compliance, to assess whether benefits of PSA screening on PrCa-specific mortality would be revealed.

Stock-and-flow diagram

Using SD, we first replicated PLCO study design and outcomes through 13-year follow-up using published data. The results of PLCO and ERSPC trials have been previously described (4-8), and are summarized in Table 1. Figure 1 depicts our SD model structure presented as a stock-and-flow diagram. Three basic types of 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 in *italics* 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 two stocks initially parameterized to the total number of men randomized into the screening and control (usual care) arm of the PLCO trial. The model structure represents how subjects were followed from study entry through PrCa incidence until death (either PrCa-specific or other causes of death) or 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 PLCO (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 tumors (18). The flows represent incidence and mortality rates per person-year of follow-up and were used to compute the number of subjects transitioning between stocks at annual intervals.

Data sources and simulated scenarios

Table 2 describes assumptions and parameters for all 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 the SEER data prior to the PSA screening era (1975-1987) and during the PLCO trial period (1993-2001), and 3) the US mortality data taken from the Morbidity and Mortality Weekly Reports. We calibrated the model using an iterative process of testing several parameter values (taken from published trial data) and then compared observed trial data to simulated behavior of the model (21, 31). We then compared the base case replication with the results of three simulated scenarios correcting for non-compliance and contamination:

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

Simulated scenarios. The PLCO trial reported 85% compliance with PSA testing in the screening arm, and 40-52% contamination (i.e., receipt of PSA screening) in the control arm (4). Thus, we simulated three scenarios to correct for noncompliance and contamination. In all three scenarios, we corrected for noncompliance by using PrCa incidence and stage-specific survival rates of screen-detected men in PLCO screening arm (10) (Table 2). In scenario 1, we simulated PrCa incidence and stage-specific survival using SEER data 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 into PLCO trial (18). In scenario 3, we corrected for contamination in PLCO control arm using data from the ERSPC trial control arm, which reported lower contamination (6-12% in various countries) (7, 33). To parameterize SD model, we estimated annualized PrCa incidence and mortality in 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 ERSPC control arm.

All three scenarios were parameterized by estimating PrCa incidence and mortality in men who actually undertook screening compared to those who did not. To account for earlier detection and longer pseudo-survival 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 due to symptoms (4, 5). Lead time bias in our SD model was derived from Telesca and colleagues (12), who estimated lead times of 4.59 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 above data to yield an overall lead time of 4.92 years. In scenario 1, where survival time was estimated using pre-PSA screening era SEER survival data, we added the lead time to the corrected screening arm; while in scenario 2, where survival time was estimated with SEER data 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 the US Morbidity and Mortality Weekly Reports (34), and were weighted by age distribution of the PLCO cohort at entry and accounted for aging of cohort during follow-up. We simulated 8% loss to follow-up at 10 years in SD model for PLCO participants, as reported in literature (5). The model simulated cumulative PrCa incidence, PrCa-specific deaths as well as RR and risk difference (RD) with 95% confidence intervals (CI) for PrCa mortality in screened vs. unscreened (control) arms of the study cohort at 10, 13 and 20 years after enrollment.

External validation of SD model: To validate our SD model, we used published results from the Rotterdam component (9) of the ERSPC trial to assess the concordance between the observed and simulated data in an independent dataset. The Rotterdam was one of the initial enrollment sites of 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. Prostate cancer-specific 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, Harvard, MA).

RESULTS

The base case replication model successfully reproduced overall outcomes of the PLCO trial and the trajectories of PrCa incidence and mortality in both screening and control arms (Figure 2A and 3A). At 13 years, the SD model simulated a total of 4,447 and 3,915 cumulative incident PrCa in the screening and controls arms, respectively (Table 3). These numbers represent a 5% and 3% increase in PrCa diagnosis in each arm of the trial in comparison to the observed 4,250 and 3,815 incident cases, respectively (5). Our model simulated 134 and 133 cumulative PrCa-specific deaths in the screening and control arm at 13 years, respectively, which were slightly lower than those reported in PLCO trial (Table 3, Figure 3A). This yielded a RR 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 3). 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 (RR=0.95; 95%CI 0.79, 1.14).

Scenario 1, which corrected for contamination using SEER PrCa incidence and mortality data before PSA screening, yielded a total of 5,706 and 1,522 PrCa incident cases, and 145 and 224 PrCa-specific deaths in the screening and control arms, respectively, at 13 years follow-up (Figures 2B and 3B). This corresponded to a significant 36% decrease in PrCa-specific mortality associated with PSA screening (95%CI 0.52, 0.79; Table 3). At 20 years of follow-up, the SD model simulated 7,052 and 2,139 incident PrCa, and 249 and 402 PrCa-deaths, respectively, in screening and control arms (Figures 2B and 3B). This yielded a RR=0.62 (95%CI 0.53, 0.72) for PrCa-specific mortality at 20 years. The corresponding risk difference (RD) for the benefit of PSA screening at 20 years was a reduction of 4.0 deaths per 1000 men screened (95%CI -5.30, -2.70).

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

In scenario 3, which corrected for contamination in the PLCO control arm using ERSPC trial data, simulated results yielded 2,532 incident PrCa and 158 PrCa-specific deaths in the PLCO control arm at 13 years follow-up (Figure 2D and 3D). There was a RR=0.80 (95%CI 0.63, 1.01) of PrCa-specific mortality associated with PSA screening at 13 years in this scenario (Table 3). The reduction in PrCa-specific mortality was maintained through 20 years of follow-up (RR=0.70, 95% CI 0.59, 0.83). However, the corresponding attributable benefit of PSA screening for this scenario at 20-year follow-up was -2.40 men per 1000 screened (95%CI -3.57, -1.23).

Finally, we simulated cumulative PrCa incidence and deaths stratified by tumor stage for all the above scenarios (see Supplementary Table S1). As reported by PLCO trial, the majority of incident PrCa were local, but with lack of screening (simulated in scenarios 1 and 3) the number of regional and distant stage PrCa 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 SD methodology, validation involved an iterative process that examined both model structure and behavior (30, 35). We conducted the following tests:

Model structure tests. (a) *Structure-verification*: The model structure (Figure 1) was parsimonious through selection of a minimally sufficient set of parameters to effectively replicate a clinical trial; (b) *Dimensional-consistency*: All variables were labeled by their proper dimensions and verified to be consistent across model equations; (c) *Parameter-verification*: As mentioned in methods, all parameters were drawn from well-documented sources of data (presented in Table 2).

Model behavior tests. (a) *Behavior-reproduction*: These tests evaluated how closely the simulated data fit observed data (0-13 years). We demonstrated in Figures 2A and 3A that simulated data fit the observed PLCO trial data well, both for incidence and mortality; *(b) Behavior-prediction*: Although follow-up on both PLCO and ERSPC trials has not reached 20 years yet, we used linear extrapolation procedures and 3-year moving averages to extrapolate results. Our simulated data (13-20 years) is 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; *(c) External validation:* Simulated PrCa cumulative incidence and cancer-specific mortality were compared to the observed results from the Rotterdam section of the ERSPC trial for validation of SD model in another dataset (presented in Supplementary Figure 1). As observed, the simulated results replicated the observed trial data closely, but underestimated both incidence and mortality partly due to lack of available published data on person-years at risk, which necessitated using average risks to approximate the rates;
(d) *Parameter sensitivity*: We selected specific parameters, which defined the three scenarios that corrected for non-compliance and contamination (Table 3, Figures 2 and 3). Comparison of three simulated scenarios to base-case scenario constituted basic sensitivity testing of our SD model. We also carried out other sensitivity tests by varying all-cause mortality, lead-time bias, and moving average window for both incidence and mortality:

(i) *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.

(ii) *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.

(iii) *Moving averages*: In all scenarios, we used a 3-year moving average for the first and last three-year data to smooth out random variation where estimates are subject to greater error due to small number of events. We tested a range of moving averages from two to five years, and observed no significant impact on results for both incidence and mortality curves.

DISCUSSION

System dynamics modeling 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 infection disease transmission and control (20, 38-46), chronic disease management (29, 47-49), partner violence (50) and tobacco control (51, 52). As such, SD modeling could be informative in evaluating the recent USPSTF recommendation guideline against use of PSA screening for PrCa (3, 16).

Our SD model replicated the PLCO trial, showing no benefit of PSA screening on PrCa-specific mortality at 13 years of follow-up, similar to published PLCO results (4, 5). Extending the follow-up to 20 years in the PLCO trial did not reveal any further screening benefits. In contrast, the three 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 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 of PLCO trial enrollment did not show any benefit of PSA screening on PrCa-specific mortality. These results were not that surprising as between 1993 and 2001, where enrollment in the PLCO trial was ongoing, PSA screening was underway in the general US population (18, 53).

Finally, correcting for contamination in PLCO using ERSPC data (scenario 3), provided statistically significant 20% and 30% benefits of PSA screening on PrCa-specific mortality at 13 and 20 years, respectively. Although the ERSPC trial included several centers with different screening protocols and randomization issues, the aggregated published data from this trial

reported 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 scenarios 1 and 3 were relatively low, with risk reductions of 2.4 to 4.0 PrCa deaths per 1,000 men (equivalent to 250 to 417 men invited to screen to prevent one death from PrCa).

Results of our scenarios 1 and 3 were also similar to other published simulation studies (11, 17, 54). 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 non-compliance. 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. Their paper (11) 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 PLCO trial.

Nevertheless, the overall estimate of 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. (55) suggested that PSA screening, if used effectively 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 of PrCa in our model, and although there were small variations in simulated outcomes for different scenarios, we did not find major differences of PrCa-specific mortality by tumor stage. We attribute this to low observed number of deaths for regional and distant stage tumors in the trial.

The benefits of screening for high-risk men (i.e., men with a family history of PrCa or men of African descent) warrants further research, since both trials did not report on the benefit

of screening among these high-risk groups. We note that only one study reported a 40% risk reduction in PrCa specific mortality in men with advanced PrCa, due to PSA screening (54). It is still unclear if screening may be beneficial for these high-risk patients. Our SD model used currently published data in PLCO and ERSPC trials that did not report on the benefits of screening for black men or those with a family history of PrCa. Should such data be made available, our SD model could be used to simulate effects by these high-risk subgroups. 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 on PrCa-mortality observed in only one trial (ERSPC) with overall harms of overdiagnosis and overtreatment at the population level; issues that were considered by the USPSTF (3), and were also recently reviewed in a meta-analysis of screening trials (56).

In this study we demonstrate an application of SD modeling as a new tool in epidemiology to simulate the dynamics of social, biological and health systems (57). Strengths of our SD 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 European trial. Limitations include the use of the 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 SD 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 trial, although there were country-specific differences in both screening interval, as well as 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 effects of various PrCa treatments, as well as other potential genetic, behavioral, and social determinants of health.

In summary, our SD model demonstrates that, after correcting for non-compliance 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 SD for synthesizing multiple sources of epidemiological data to inform public health policy.

ABBREVIATIONS

CI – Confidence interval

CISNET – Cancer Intervention and Surveillance Modeling Network

ERSPC – European Randomised Study of Screening of Prostate Cancer

PLCO – Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial

PrCa – Prostate cancer

PSA – Prostate-specific antigen

RR – Relative risk

SD – System dynamics

SEER – Surveillance, Epidemiology and End Results

Table 1. Comparison of Two Randomized Clinical Trials of Prostate Cancer Screening

Screening Trial	**PLCO (US)**	**ERSPC (Europe)**
Enrolled participants	76,685	162,243
Enrollment period	1993 – 2001	1990 – 1994
Age range (years)	55 –74	55 – 69
Median follow-up (years)	11.5	11.3
Frequency of screening	Annual PSA for 6 yrs DRE for 4 yrs	PSA once every 4 yrs (Sweden 2 years)†
Control arm	Usual care	No Screening
PSA cut-off point for biopsy	4 ng/ml	3 or 4 ng/ml‡
Actual PSA screening rate in screening arm	85%	82%
Actual PSA screening rate in control arm	40% - 52%	6% - 12%
RR (95% CI) of prostate cancer-specific mortality	1.09 (0.87, 1.36)	0.79 (0.69, 0.91)

Abbreviations: PSA – Prostate-specific Antigen; PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; ERSPC – European Randomized Study of Screening for Prostate Cancer; RR – Relative Risk

† The PSA screening interval at six of the seven centers in Europe was every 4 years; while in Sweden it was every 2 years.

‡Most centers used a PSA cutoff value of 3.0 ng/ml, while in Finland and Italy the PSA cutoff was 4 ng/ml.

Table 2. Assumptions and Parameters of System Dynamics Model for Simulated Scenarios of the PLCO Trial

Scenario		**Base case scenario PLCO replication**	**Scenario 1 Corrected to SEER (1975-1988)**	**Scenario 2 Corrected to SEER (1993-2001)**	**Scenario 3 Corrected to ERSPC**	
Assumptions						
Correction for Non-Compliance		None	Screen-detected men	Screen-detected men	Screen-detected men	
Correction for Contamination		None	SEER PrCa incidence, stage distribution and survival data prior to PSA screening era (1975-1988)	SEER PrCa incidence, stage distribution and survival data for PLCO trial period (1993-2001)	ERSPC control arm data	
Parameters	Subgroup	Values				Units
PrCa incidence[a], mean (range)	*Screening arm*	1096 (893-1656)	1402 (1094-1803)	1402 (1094-1803)	1402 (1094-1803)	cases per 100,000 person-years
	Control arm	974 (830-1117)	391 (323-458)	577 (546-631)	634 (490-808)	
Stage distribution	*Screening arm*					% of cases by stage within study arm
	Localized	96.0%	97.3%	97.3%[b]	97.3%	
	Regional	1.4%	1.5%	1.5%	1.5%	
	Distant	2.1%	1.0%	1.0%	1.0%	
	Unstaged	0.4%	0.2%	0.2%	0.2%	
	Control arm					
	Localized	94.3%	60%	88.2%	76.1%	
	Regional	1.9%	15%	3.4%	16.7%	
	Distant	2.7%	15%	4.4%	4.0%	
	Unstaged	1.1%	10%	4.0%	3.2%	
PrCa mortality, mean (range)	*Screening arm*	49.4 (7.8-103.2)	*N/A*	*N/A*	*N/A*	deaths per 100,000 person-years
	Control arm	50.6 (2.6-112.8)				
5-year survival rate[b]	*Localized*	*N/A*	90.0%	100%	100%	% of diagnosed cases surviving at 5 years
	Regional		79.8%	88.7%	88.7%	
	Distant		32.9%	36.5%	36.5%	
	Unstaged		78.4%	87.1%	87.1%	
All-cause mortality[c]		37.1 (16.5 – 67.0)	37.1 (16.5 – 67.0)	37.1 (16.5 – 67.0)	37.1 (16.5 – 67.0)	deaths per 1,000 person-years
Lead time bias[d] (screening arm only)	*Localized*	*N/A*	4.92	4.92	4.92	years
	Regional		4.92	4.92	4.92	
	Distant		0.00	0.00	0.00	
	Unstaged		4.92	4.92	4.92	

Abbreviations: PrCa – Prostate cancer; PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; SEER – Surveillance, Epidemiology and End Results; ERSPC – European Randomised Study of Screening for Prostate Cancer

[a] The average annual PLCO cases during 20 years of follow up (3-year moving average)

[b] SEER 5 year survival by stage (1995-2001)

[c] 5-year age-specific rates for all-cause mortality

[d] Estimated lead time weighted by PLCO race distribution

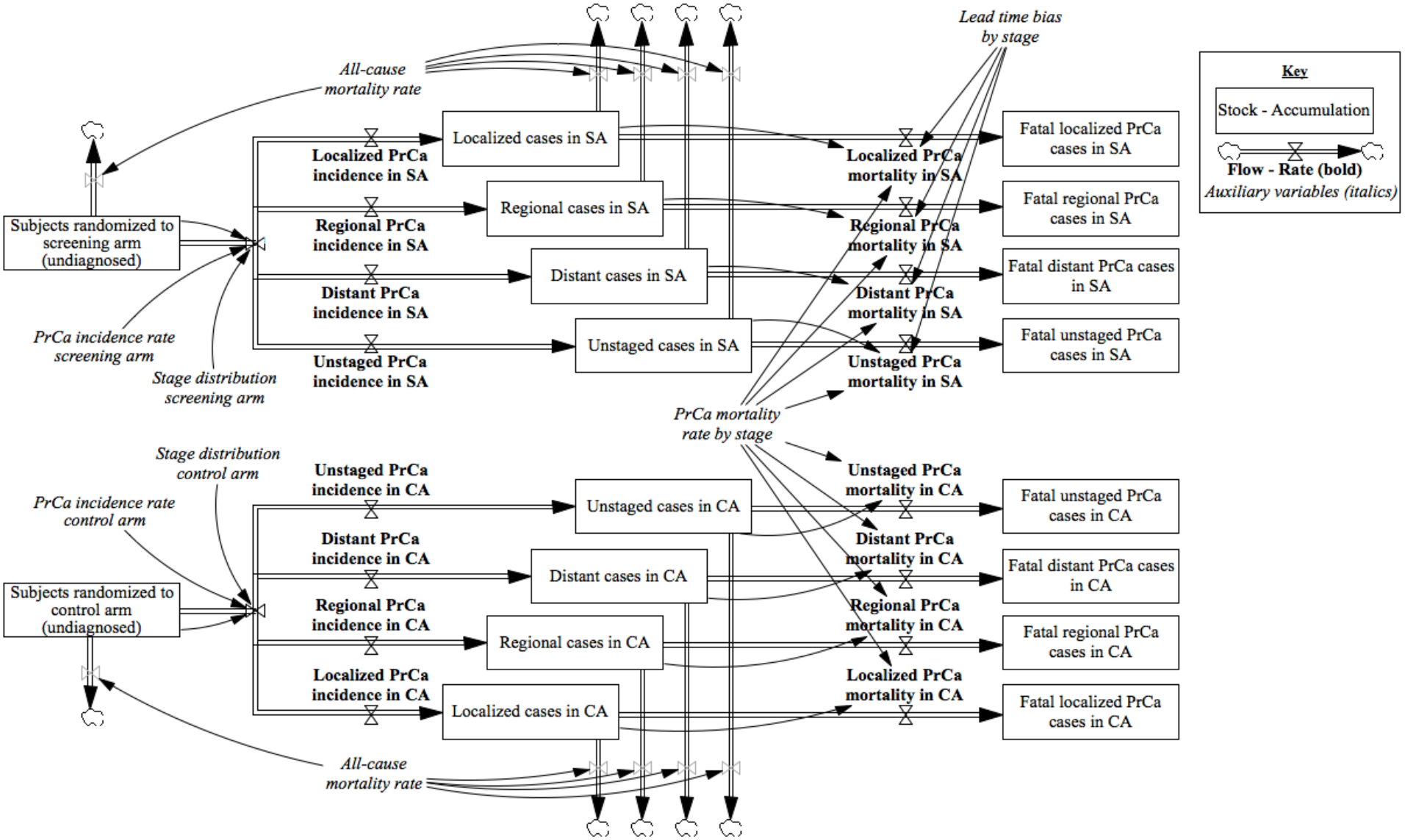
Table 3. Results for Simulated Scenarios of the PLCO Trial at 10, 13 and 20 Years

Follow-up	10 years	13 years	20 years
Base case scenario - PLCO replication			
Cumulative incidence (screening arm)	3,726	4,447	5,635
Cumulative incidence (control arm)	3,206	3,915	5,093
Cumulative deaths (screening arm)	93	134	223
Cumulative deaths (control arm)	88	133	234
RR (95% CI)	*1.06 (0.79, 1.42)*	*1.00 (0.79, 1.27)*	*0.95 (0.79, 1.14)*
RD (95% CI) [a]	*0.14 (-0.54, 0.83)*	*0.01 (-0.82, 0.85)*	*-0.30 (-1.39, 0.79)*
Scenario 1 – Corrected to SEER (1975-1988)			
Cumulative incidence (screening arm)	4,791	5,706	7,052
Cumulative incidence (control arm)	1,199	1,522	2,139
Cumulative deaths (screening arm)	98	145	249
Cumulative deaths (control arm)	150	224	402
RR (95% CI)	*0.65 (0.51, 0.84)*	*0.64 (0.52, 0.79)*	*0.62 (0.53, 0.72)*
RD (95% CI) [a]	*-1.36 (-2.16, -0.56)*	*-2.08 (-3.06, -1.10)*	*-4.00 (-5.30, -2.70)*
Scenario 2 – Corrected to SEER (1993-2001)			
Cumulative incidence (screening arm)	4,791	5,706	7,052
Cumulative incidence (control arm)	1,923	2,384	3,163
Cumulative deaths (screening arm)	85	126	217
Cumulative deaths (control arm)	83	125	219
RR (95% CI)	*1.02 (0.76, 1.38)*	*1.01 (0.79, 1.30)*	*0.99 (0.82, 1.19)*
RD (95% CI) [a]	*0.05 (-0.61, 0.71)*	*0.04 (-0.76, 0.85)*	*-0.06 (-1.12, 1.01)*
Scenario 3 – Corrected to ERSPC			
Cumulative incidence (screening arm)	4,791	5,706	7,052
Cumulative incidence (control arm)	1,924	2,532	3,497
Cumulative deaths (screening arm)	85	126	217
Cumulative deaths (control arm)	97	158	309
RR (95% CI)	*0.88 (0.66, 1.17)*	*0.80 (0.63, 1.01)*	*0.70 (0.59, 0.83)*
RD (95% CI)[a]	*-0.31 (-1.00, 0.38)*	*-0.83 (-1.68, 0.03)*	*-2.40 (-3.57, -1.23)*

Abbreviations: PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; SEER – Surveillance, Epidemiology and End Results; ERSPC – European Randomised Study of Screening for Prostate Cancer; RR – Relative Risk; RD – Risk Difference; CI – Confident Interval

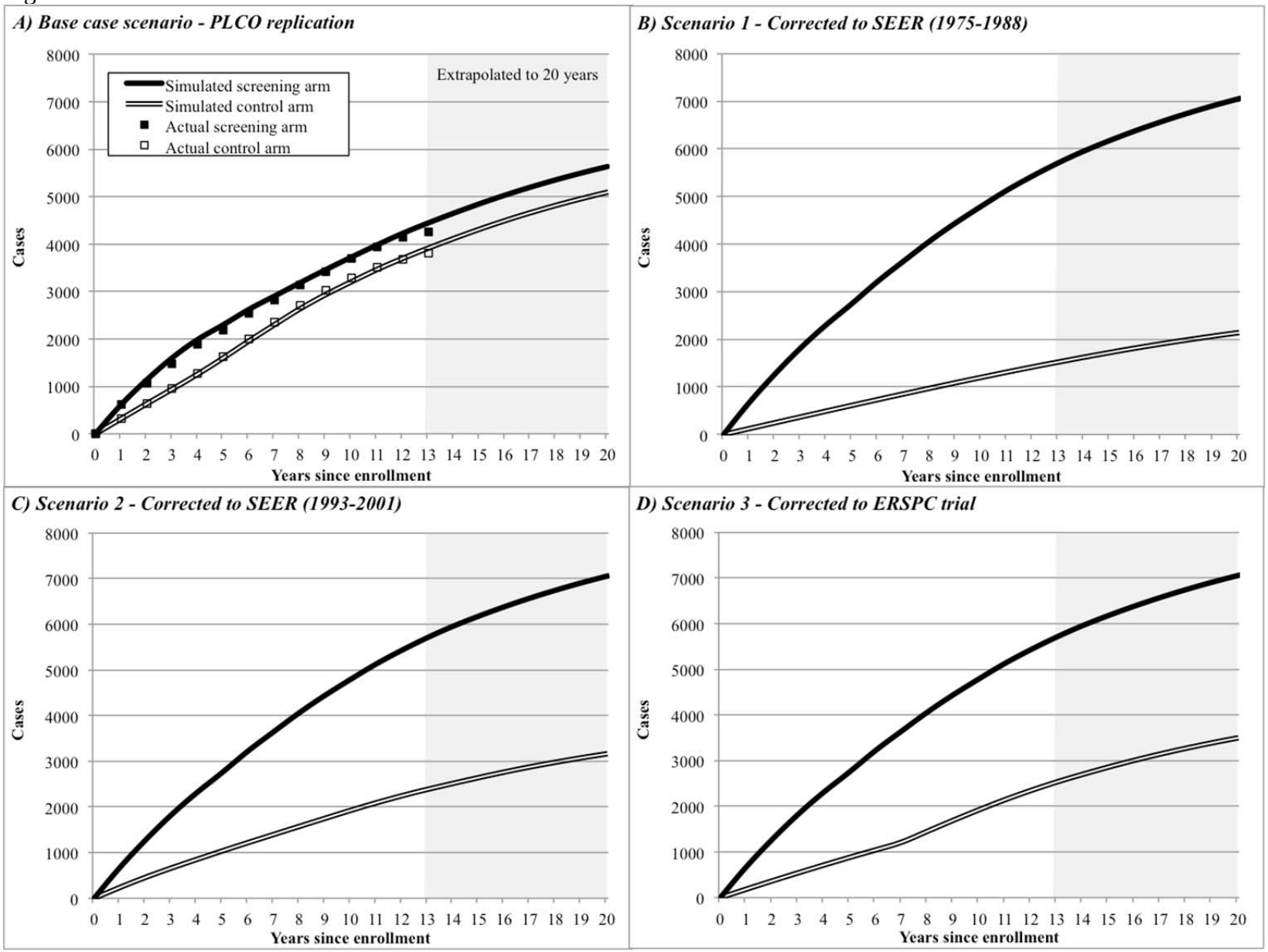
[a]Risk differences are reported as number of deaths per 1,000 men screened.

Figure 1. Stock-and-Flow Diagram for System Dynamics Model of PLCO Trial



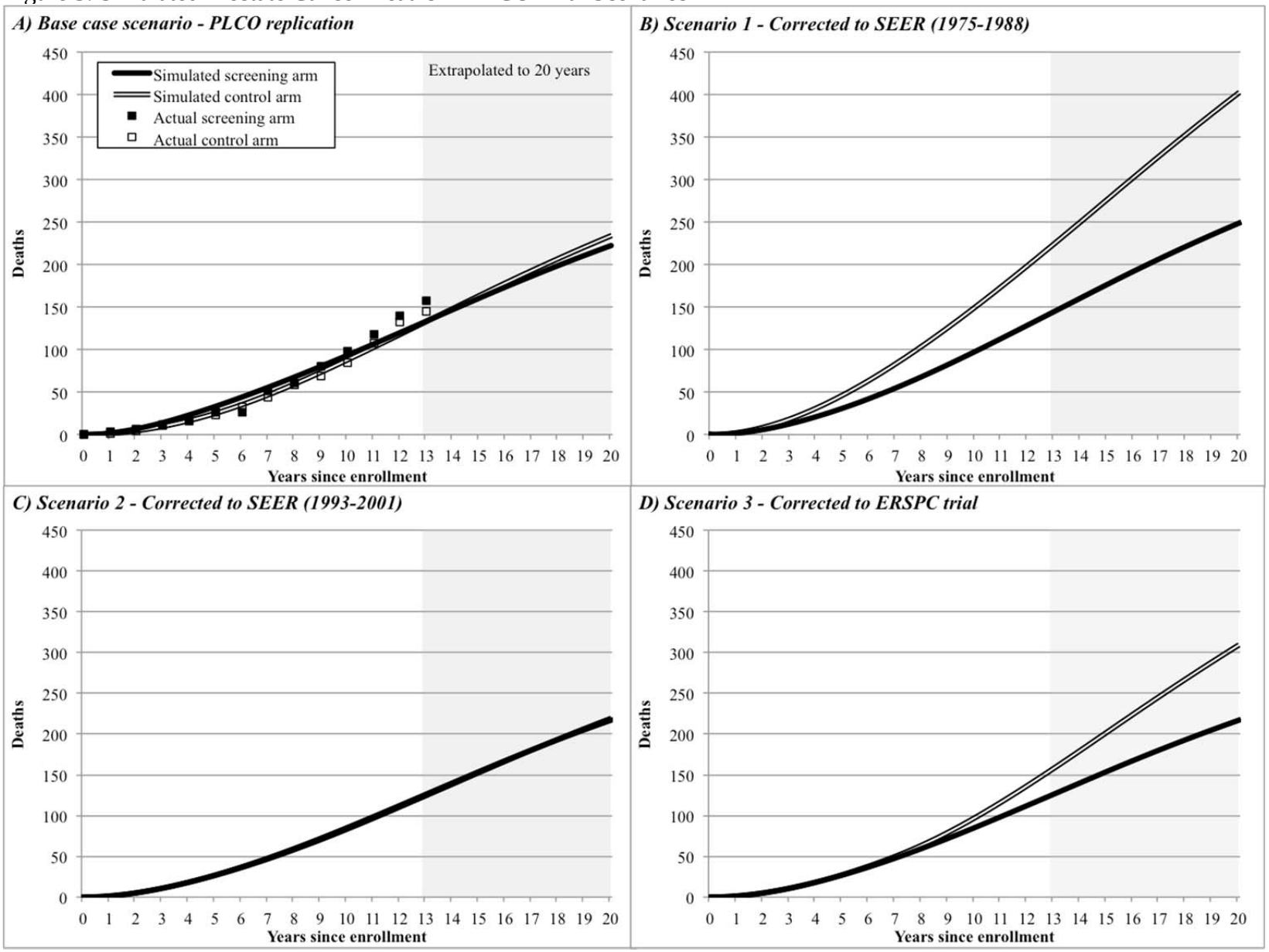
Abbreviations: PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; PrCa – Prostate Cancer; SA – Screening Arm; CA – Control Arm

Figure 2. Simulated Prostate Cancer Incident Cases of PLCO Trial Scenarios



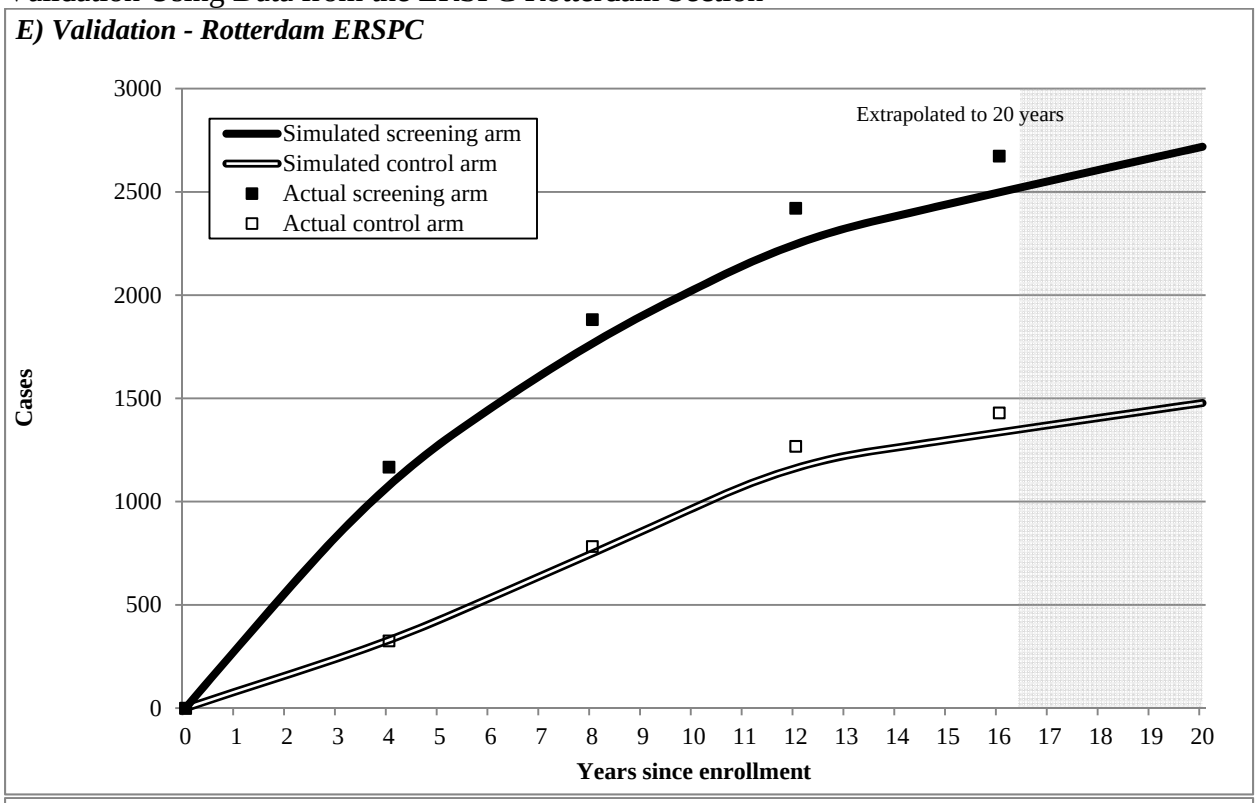
Abbreviations: PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; SEER – Surveillance, Epidemiology and End Results; ERSPC – European Randomised Study of Screening for Prostate Cancer

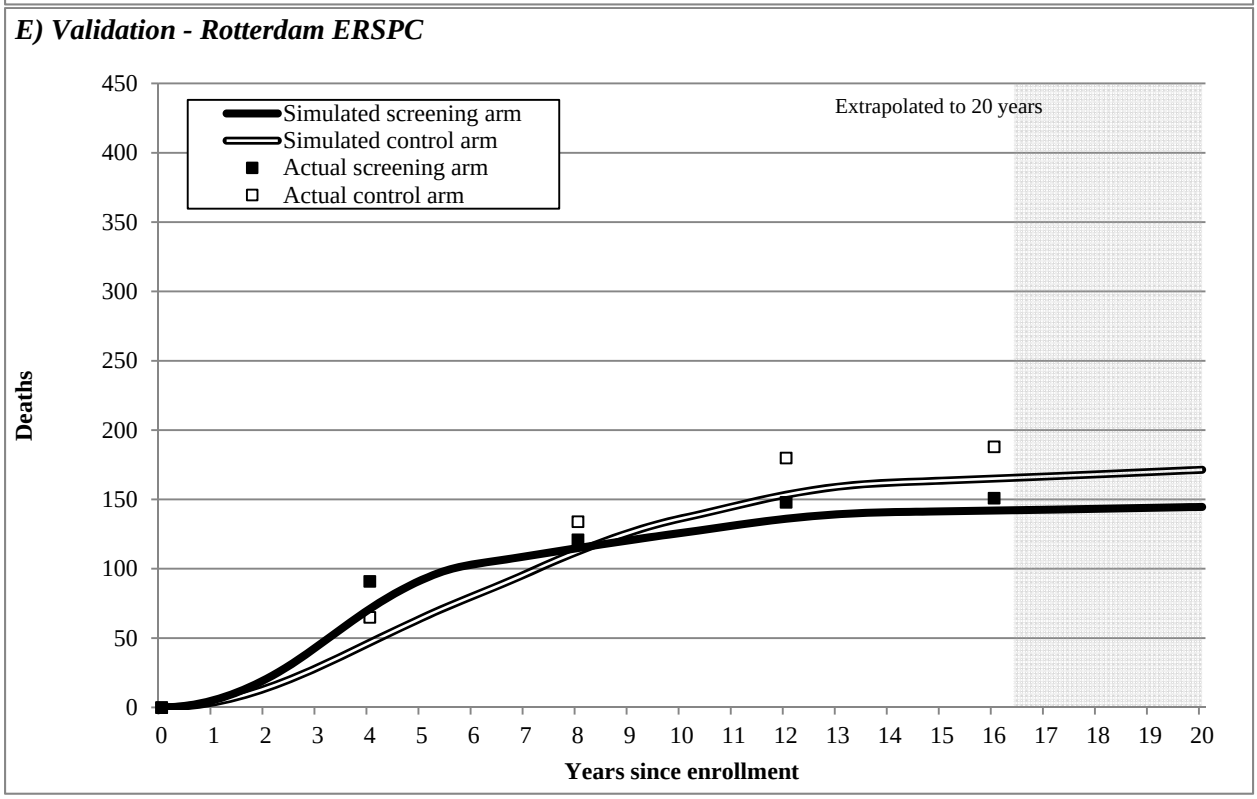
Figure 3: Simulated Prostate Cancer Deaths in PLCO Trial Scenarios



Abbreviations: PLCO – Prostate, Lung, Colorectal Ovarian Cancer Screening Trial; SEER – Surveillance, Epidemiology and End Results; ERSPC – European Randomised Study of Screening for Prostate Cancer

Supplemental Figure S1: Simulated Prostate Cancer Incident Cases and Deaths in the SD Model Validation Using Data from the ERSPC Rotterdam Section





Abbreviations: ERSPC – European Randomised Study of Screening for Prostate Cancer

ACKNOWLEDGEMENTS

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

Dr. Ilir Agalliu was supported in part by a research scholar grant (MRSG-11-112-01-CNE) from the American Cancer Society. Anton Palma was supported in part by the training grant T32AI114398 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Conflicts of interest: none declared.

REFERENCES

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA: a cancer journal for clinicians. 2012;62(1):10-29.
2. American Cancer Society (ACS). Cancer Facts & Figures 2014.
3. Moyer et al. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;157(2):120-34.
4. Andriole G, et al.,. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med. 2009;360(13):1310-9.
5. Andriole G, 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;104:125-32.
6. Schroder F, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med. 2009;2009(360):13.
7. Schroder F, et al. Prostate-cancer mortality at 11 years of follow-up. N Engl J Med. 2012;366:981-90.
8. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Zappa M, Nelen V, 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):2027-35.
9. Roobol M, 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). European Urology. 2009;56:584-91.
10. Pinsky P, Black, A, Kramer, BS, Miller, A, Prorok, PC, Berg, C. Assessing contamination and compliance in the prostate component of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. Clin Trial. 2010;7:303-11.
11. Gulati R, et al. The impact of PLCO control arm contamination on perceived PSA screening efficacy. Cancer Causes Control. 2012;23:827-25.
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):10-9.

13. Draisma G, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst. 2009;101(6):374-83.
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):5955-63.
15. McGregor M, Hanley, JA, Boivin, JF, Mclean, RG. Screening for prostate cancer: estimating the magnitude of overdetection. CMAJ. 1998;159(11):1368-72.
16. Habbema JD, Wilt TJ, Etzioni R, Nelson HD, Schechter CB, Lawrence WF, et al. Models in the development of clinical practice guidelines. Ann Intern Med. 2014;161(11):812-8.
17. Gulati R, Inoue LY, Gore JL, Katcher J, Etzioni R. Individualized estimates of overdiagnosis in screen-detected prostate cancer. J Natl Cancer Inst. 2014;106(2):djt367.
18. Howlader N, Noone A, Krapcho M, Garshell J, Miller D, Altekruse S, et al. SEER Cancer Statistics Review, 1975-2011, National Cancer Institute. Bethesda, MD, http://seer.cancer.gov/csr/1975_2011/, based on November 2013 SEER data submission, posted to the SEER web site, April 2014. National Cancer Institute. Bethesda, MD, 2014.
19. Homer J, Hirsch, GB. System dynamics modeling for public health: background and opportunities. Am J Public Health. 2006;96(3):452-8.
20. Duintjer Tebbens RJ, Pallansch MA, Kew OM, Caceres VM, Sutter RW, Thompson KM. A dynamic model of poliomyelitis outbreaks: learning from the past to help inform the future. Am J Epidemiol. 2005;162(4):358-72.
21. Oliva R. Model calibration as a testing strategy for system dynamics models. Eur J Oper Res. 2003;151(3):552-68.
22. Sterman J. Learning from evidence in a complex world. Am J Public Health. 2006;96(3):505-14.
23. Wolstenholme EF. System dynamics: A system methodology or a system modelling technique. Dynamica. 1983;9(II):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):811-9.
25. Hirsch G, Edelstein, BL, Frosh, M, Anselmo, T. A simulation model for designing effective intervention in early childhood caries. Prev Chronic Dis. 2012;9:E66.

26. Homer J, Hirsch, G, Minniti, M, Pierson, M. Models for collaboration: how system dynamics helped a community orgnaize cost-effective care for chronic illness. System Dynamics Review. 2004;20:199-222.
27. Lounsbury D, Schwartz, B, Palma, A, Blank, A. Simulating patterns of patient engagement, treatment adherence, and viral suppression: a system dynamics approach to evaluating HIV care management. AIDS Patient Care STDS. 2015;Suppl 1:S55-63.
28. Richardson G. Introduction to system dynamics modeling with DYNAMO. 1st ed. Westford, MA: Pegasus Communications; 1981.
29. Hammond R. Complex systems modeling for obesity research. Prev Chronic Dis. 2009;6(3):A97.
30. Martis M. Validation of simulation based models: A theoretical outlook. Electronic Journal of Business Research Methods. 2006;4(1):39-46.
31. Walker R, Wakeland W. Calibration of complex system dynamics models: A practitioner's report. Portland, Oregon: 2011.
32. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial. Am J Epidemiol. 2007;165(8):874-81.
33. Kerkhof M, Roobol MJ, Cuzick J, Sasieni P, Roemeling S, Schroder FH, 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). International journal of cancer Journal international du cancer. 2010;127(11):2639-44.
34. Centers for Disease Control and Prevention. Mortality tables - MMWR Morbidity and Mortality Weekly Reports. 2012:Retrieved from http://www.cdc.gov/mmwr/distrnds.html.
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, Abrams DB. 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):S211-24.
37. Trochim WM, Cabrera DA, Milstein B, Gallagher RS, Leischow SJ. Practical challenges of systems thinking and modeling in public health. Am J Public Health. 2006;96(3):538-46.

38. Chick SE, Mamani H, Simchi-Levi S. Supply chain coordination and Influenza vaccination. Operation Research. 2008;56(6):1493-506.
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):433-49.
40. Dangerfield BC, Fang YX, Roberts CA. Model-based scenarios for the epidemiology of HIV/AIDS: the consequences of highly active antiretroviral therapy. Systems Dynamics Review. 2001;17(2):119-50.
41. Flessa S. Decision support for AIDS control programmes in eastern Africa. Or Spectrum. 2003;25(2):265-91.
42. Goforth RR. A simulation-based model to assist in managining the HIV epidemic. Simulation. 1994;63(2):128-36.
43. Homer J, Ritchie-Dunham J, Rabbino H, Puente LM, Jorgensen J, Hendricks K. Toward a dynamic theory of antibiotic resistance. Systems Dynamics Review. 2000;16(4):287-319.
44. Lebcir RM, Atun RA, Coker RJ. System dynamic simulation of treatment policies to address colliding epidemics of tuberculosis, drug resistant tuberculosis and injecting drug users driven HIV in Russia. Journal Operational Research Society. 2010;61(8):1238-48.
45. Roberts C, Dangerfield BC. Modeling the epidemiologic consequences of HIV-infection and AIDS: A contribution from operational research. Journal Operational Research Society. 1990;41(4):273-89.
46. Homer J, St. Clair, CL. A model of HIV transmission through needle sharing. Interfaces. 1991;21:26-49.
47. Homer J, Hirsch G, Minniti M, Pierson M. Models for collaboration: how system dynamics helped a community organize cost-effective care for chronic illness. System Dynamics Review. 2004;20(3):199-222.
48. Patterson RE, Colditz GA, Hu FB, Schmitz KH, Ahima RS, Brownson RC, et al. The 2011-2016 Transdisciplinary Research on Energetics and Cancer (TREC) initiative: rationale and design. Cancer causes & control : CCC. 2013;24(4):695-704.
49. Sabounchi NS, Hovmand PS, Osgood ND, Dyck RF, Jungheim ES. A novel system dynamics model of female obesity and fertility. American journal of public health. 2014;104(7):1240-6.

50. Hovmand PS, Ford DN. Sequence and timing of three community interventions to domestic violence. American journal of community psychology. 2009;44(3-4):261-72.
51. Tengo T, Osgood, ND, Chen, LL. The cost-effectiveness of intensive national school-based antitobacco education: results from the tobacco policy model. Prev Med. 2001;33:558-70.
52. 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):1274-81.
53. 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):1325-9.
54. Wu GH, Auvinen A, Yen AM, Hakama M, Tammela TL, Stenman UH, et al. The impact of interscreening interval and age on prostate cancer screening with prostate-specific antigen. European urology. 2012;61(5):1011-8.
55. Cooperberg MR, Carroll PR, Klotz L. Active surveillance for prostate cancer: progress and promise. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2011;29(27):3669-76.
56. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev. 2013;1:CD004720.
57. Richardson GP. Feedback Thought in Social Science and Systems Theory. Waltham, MA: Pegasus Communications, Inc.; 1991.