REVIEW ARTICLE

Calibration Methods Used in Cancer Simulation Models and Suggested Reporting Guidelines

Natasha K. Stout,[1,2] *Amy B. Knudsen,*[3,4] *Chung Yin Kong,*[3,4] *Pamela M. McMahon*[3,4] and *G. Scott Gazelle*[2,3,4]

1   Department of Ambulatory Care and Prevention, Harvard Medical School/Harvard Pilgrim Health Care, Boston, Massachusetts, USA
2   Program in Health Decision Science, Department of Health Policy and Management, Harvard School of Public Health, Boston, Massachusetts, USA
3   Institute for Technology Assessment, Massachusetts General Hospital, Boston, Massachusetts, USA
4   Department of Radiology, Massachusetts General Hospital, Boston, Massachusetts, USA

Contents

Abstract

Increasingly, computer simulation models are used for economic and policy evaluation in cancer prevention and control. A model's predictions of key outcomes, such as screening effectiveness, depend on the values of unobservable natural history parameters. Calibration is the process of determining the values of unobservable parameters by constraining model output to replicate observed data. Because there are many approaches for model calibration and little consensus on best practices, we surveyed the literature to catalogue the use and reporting of these methods in cancer simulation models.

We conducted a MEDLINE search (1980 through 2006) for articles on cancer-screening models and supplemented search results with articles from our personal reference databases. For each article, two authors independently abstracted pre-determined items using a standard form. Data items included cancer site, model type, methods used for determination of unobservable parameter values and description of any calibration protocol. All authors reached consensus on items of disagreement. Reviews and

non-cancer models were excluded. Articles describing analytical models, which estimate parameters with statistical approaches (e.g. maximum likelihood) were catalogued separately. Models that included unobservable parameters were analysed and classified by whether calibration methods were reported and if so, the methods used.

The review process yielded 154 articles that met our inclusion criteria and, of these, we concluded that 131 may have used calibration methods to determine model parameters. Although the term 'calibration' was not always used, descriptions of calibration or 'model fitting' were found in 50% (n = 66) of the articles, with an additional 16% (n = 21) providing a reference to methods. Calibration target data were identified in nearly all of these articles. Other methodological details, such as the goodness-of-fit metric, were discussed in 54% (n = 47 of 87) of the articles reporting calibration methods, while few details were provided on the algorithms used to search the parameter space.

Our review shows that the use of cancer simulation modelling is increasing, although thorough descriptions of calibration procedures are rare in the published literature for these models. Calibration is a key component of model development and is central to the validity and credibility of subsequent analyses and inferences drawn from model predictions. To aid peer-review and facilitate discussion of modelling methods, we propose a standardized Calibration Reporting Checklist for model documentation.

Increasingly, mathematical and computer models are used for economic evaluation of cancer prevention and control policies.[1-5] These models fill an important role in policy making because they are able to synthesize data from multiple sources and estimate the effects of interventions in situations when clinical trials may not be feasible because of time, cost and/or ethical considerations.[6] The US National Cancer Institute's Cancer Intervention and Surveillance Modeling Network (CISNET) recently spurred growth in modelling efforts by funding the development of over 18 models of breast, prostate, colon and lung cancers built to investigate prevention, screening and treatment policy questions in the US. In addition to CISNET, numerous other cancer simulation models have been developed including those for cervical, ovarian and gastric cancers.[7-11]

In general, disease natural history models use a 'systems' approach to simulate the underlying course of disease in individuals and project the overall effect of disease on health in a population. Many of these models are quite complex, describing both unobservable and observable portions of the natural history at an individual level. Capturing the mechanism of disease onset and growth in simulated individuals involves specification and determination of unknown model parameters, many of which cannot be directly informed by data since none may exist. One method for parameter determination is calibration.

Formally, model calibration is the process of determining parameter values so that model output replicates empirical data.[6,12-14] It is performed by comparing model outputs from different input parameter sets with existing data to identify parameter set(s), which produce model outputs that best correspond to existing empirical data. This is often a complex task and there has been little consensus on best practice. Calibration, often termed 'model fitting', may be distinguished from other methods of parameter determination, such as direct 'estimation' of model parameters.[13] For estimation methods, model parameters are estimated in separate processes from the model itself and the overall fit of model output to data is not considered in parameter determination.

Calibration is a key component of model development and together with validation establishes credibility of modelling results. Models are already often criticized for being 'black boxes', with documentation often lacking transparency. If modelling is to gain strength as a tool for informing health policy, it is critical that the assumptions, structure, input data and parameter-estimation methods, including calibration, are well documented and made available to consumers of these models. To understand how calibration methods are currently described, we surveyed the literature to catalogue these methods as they are used in cancer simulation models. We then proposed a framework for reporting the calibration methods used in cancer models and disease simulation models in general.

1. Method of Review

We conducted a focused qualitative review of the literature on calibration methods in cancer simulation models.

1.1 Data Sources

Multiple sources were used to identify relevant published literature. We used MEDLINE to identify English-language articles published in the years 1980 through 2006. The following US National Library of Medicine Medical Subject Headings (MeSH) and keywords were used in the search: 'cancer', 'neoplasm', 'simulation', 'computer simulation', 'natural history' and 'mass screening'. The search results were supplemented by reviewing the reference lists from the articles identified in the search and from articles in our personal reference databases. The full listing of articles retrieved by the search is available from the authors.

1.2 Study Selection

Study-inclusion criteria consisted of articles describing models that explicitly captured the mechanism of the underlying natural history of cancer. Reviews and commentaries were excluded, as were articles describing models of diseases other than cancer. Also excluded were articles unanimously judged by the authors to describe models with no natural history component, including models that begin from the point of cancer detection. We further distinguished purely analytical models that use statistical inference to estimate parameters from micro-simulation models that use calibration to determine underlying parameter values. For analytical models, the process of direct estimation of the unknown natural history parameters to observed data is typically carried out independently of the model itself.[13] We note that some modellers use a 'hybrid' approach with direct estimation of some parameters and subsequent calibration of others.

1.3 Data Abstraction and Analysis

For each article, two investigators independently abstracted pre-determined data elements using a standard form (see the appendix, available in the Supplemental Digital Content 1 http://links.adisonline.com/PCZ/A56). Data elements abstracted included cancer site, type of simulation model, methods used for determination of unobservable parameter values, description of calibration protocol (if any) and whether model validation was mentioned. All investigators reached consensus on items of disagreement. For articles that provided no calibration information but that referenced previous publication(s) with the same model, information from the prior publication(s) was used to supplement the information reported in the primary article. Review data were summarized using descriptive statistics.

Calibration protocols were characterized by the descriptions of five components: calibration targets, goodness-of-fit (GOF) metric, search algorithm, acceptance criteria and the stopping rule. A 'target' refers to observed data that the model attempts to replicate during calibration. The GOF metric is the quantitative measure of the fit or ability of the model to replicate target data for a particular set of parameter values. The search algorithm is the method for selecting alternative model-parameter values to evaluate. The acceptance criteria specify the satisfactory or acceptable levels of fit based on the GOF metric for a particular set of parameter values. Finally, the stopping rule describes the rationale for

Table I. Glossary of calibration components

Term	Definition	Suggested references
Target	Empirical data to be replicated by the model through the process of calibration. By definition, target data are not used as direct model inputs. Targets can be in the form of a single summary statistic (e.g. mean age of cancer diagnosis) or a series of statistics (e.g. age-specific incidence curve). Simultaneously calibrating to multiple targets, along with input parameters taken from multiple sources, permits synthesis of diverse data sources. Choice and number of targets may depend on data availability and quality as well as the number of parameters under calibration and model complexity	Disease dependent
GOF metric	Measure of the model's accuracy at replicating target data as a function of the input parameters. Qualitative metrics include visual inspection of model fit. Quantitative examples include likelihood metrics that quantify the probability of generating the target data with a given set of parameter values and other metrics that measure the distance from model output to the target (e.g. sum of squared errors). An overall GOF metric may be computed as a raw sum or a weighted average across multiple targets	15
Search algorithm	Selection method for parameter sets from all possible feasible parameter values (i.e. the parameter space) to evaluate during the process of calibration. It may be informal such as 'trial-and-error' or formal such as a grid search or directed search. Directed search algorithms use computer algorithms and numerical approximation methods to iteratively select parameter sets to evaluate that are likely to lead to better model fits. Choice of method may depend on the number of parameters under calibration and model complexity. Formal methods may be desired for high-dimensional or complex models. Directed search methods can be useful for improving computational efficiency in searching a high-dimensional parameter space	16-25
Acceptance criteria	Criteria by which the modeller determines if a particular parameter set provides a reasonable fit to the target data. It typically involves assessment of whether the GOF measure meets a threshold (e.g. likelihood ratio less than a predefined criterion). A modeller may ultimately accept multiple parameter sets to characterize uncertainty in parameter values as in a Bayesian calibration approach or instead may proceed with the single best-fitting set identified	12,14,26
Stopping rule	Rule by which the modeller determines if the calibration process is complete or whether additional searching of the parameter space is warranted. The rule may be based on the convergence of a directed search algorithm, the extent of the parameter space searched or other arbitrary criteria	14

GOF = goodness of fit.

ending the search procedure and calibration process as a whole. These components are described in more detail in table I.

2. Search Results

The MEDLINE search and personal databases yielded 169 unique articles. We excluded 18 articles on the basis of the title and/or abstract (five reviews, editorials or commentaries; one meta-analysis; three articles about a disease other than cancer; and nine cancer-related articles that did not report on a model of the natural history of the disease). We reviewed and abstracted data from the remaining 151 articles and excluded an additional 28 articles (five review articles, editorials or commentaries, one article reporting the results of a database analysis; four methodological papers using general models for which calibration was not necessary, eighteen articles on

reporting on models that did not track the underlying natural history of cancer or that did not evaluate cancer-screening programmes). A total of 123 articles met the criteria for inclusion. From the references to prior publications in these articles, we identified 31 additional articles that met the inclusion criteria, yielding a total of 154 articles.

2.1 Model Characteristics

The number of published articles describing cancer simulation models increased substantially over the study period (figure 1). Articles describing models of the natural history of breast or cervical cancers were the most prevalent (table II). Note that multiple articles using the same model were counted separately.

Of these 154 articles, 23 used purely analytical methods to directly estimate parameter values,[27,40,45,50,54,59,62-66,69,73,74,78,95,123,124,149,150,161,174,175] leaving 131 articles for which calibration may
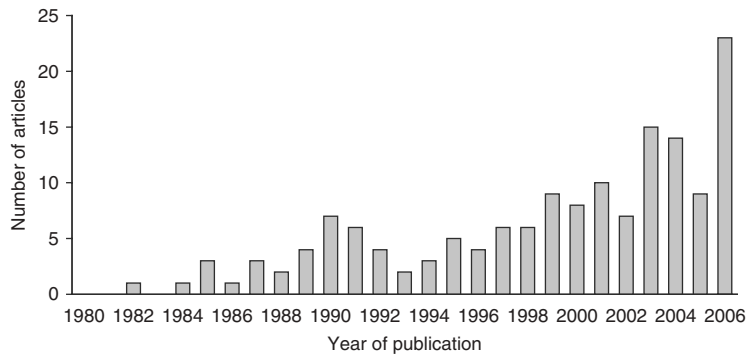
Fig. 1. The distribution by year of publication of the 154 articles describing cancer simulation models included in our review. From 1980 through 2006, the application of cancer simulation models steadily increased. We note that one article from our review was published in 1973 and is not included in this figure, which begins in 1980.

have been used to determine at least some unknown model-parameter values. This subset includes articles that may have used a hybrid or combination approach of both analytical and calibration methods for parameter determination.[49,70,71] Subsequent summary statistics on the proportions of articles reporting calibration details are based on these 131 articles.

2.2 Calibration Details

We found that 66 articles discussed or alluded to calibration in the description of the model in the article itself. In some, calibration was explicitly mentioned in the text,[48,61,120] while in others we inferred calibration was conducted by the authors' use of terms such as 'model fitting' or 'model identification'.[11,93,135] An additional 21 articles did not explicitly state or imply that calibration was used, but did provide references to a prior publication in which calibration methods were mentioned.[29,34,37,38,41,42,47,68,79,82,84,94,98,104,106,111,113,119,121,128,136] Thus, 87 articles (66%) provided some documentation for the calibration protocol used in developing the model.

Of the 87 articles that discussed or provided a reference to calibration methods, 95% (83 of 87) made at least some mention of the data used as calibration targets. Targets included data from cancer registries, observational studies and clinical trials. The vast majority of the models were calibrated to multiple targets, although in most

cases it was unclear if they were calibrated to these data simultaneously or in stages.

GOF metrics were either explicitly or implicitly described in 54% (47 of 87) of the articles that discussed calibration. Visual assessment of fit, a qualitative GOF metric, was used in 20 articles although its use was typically inferred from the article text rather than explicitly stated. For the 27 articles that reported quantitative methods, two used likelihood-based measures[39,77] and 25 used distance measures, such as the absolute or relative differences.[43,48,61,71,81,135] The majority of articles did not describe how GOF metrics for individual targets were combined to yield an overall GOF measure for the model parameters under calibration.

Table II. Published articles describing cancer simulation models by cancer site

Cancer site	Number of articles	References
Breast	56	27-82
Cervical	41	7,9,11,83-120
Colorectal	22	121-142
Oesophageal	1	143
Gastric	4	144-147
Lung	10	148-157
Melanoma	1	158
Oral	1	159
Ovarian	4	10,160-162
Prostate	12	163-174
Multiple cancers	2	175,176
Total	154	

Search algorithms were generally not well described in the articles we reviewed. We inferred that some used informal 'trial-and-error' approaches[101] while others specified 'systematic variation' of model-parameter values and provided no further description.[7,36,81] When further specified, formal search algorithms included grid search[58,61,71,80] and random sampling[31,48] as well as directed or iterative search methods. Directed search methods use computer algorithms and numerical approximation techniques to identify points in the parameter space likely to lead to successively better model fits. Methods used included the Nelder-Mead algorithm,[39,164] as well as a variety of other optimization algorithms from engineering.[39,77,156]

Criteria for identifying parameter sets that provide an acceptable model fit were rarely described in the articles reviewed. Stopping criteria were also not well documented. With few exceptions, modellers ultimately accepted a single best-fitting parameter set to conduct model analyses rather than accept multiple parameter sets to form a posterior distribution that can capture parameter uncertainty.[48]

Model-validation methods or model validity were mentioned in 52% (80 of 154) of the articles, although details as to how validity was assessed were provided in only a few.[10,83]

3. Discussion

Our descriptive review indicates that while an array of techniques is used for model calibration, little attention has been paid to the documentation of these methods in the cancer simulation literature. Calibration is a necessary component of models that simulate an unobserved disease process because it helps to ensure the validity and credibility of inferences drawn from model predictions. We found techniques for calibration used in cancer-screening models vary in level of complexity and rigor. Methods include trial-and-error searching that may employ a more subjective visual assessment of model fit as well as the use of directed search algorithms and more objective quantitative assessments of fit.

Our review also indicates that the depth of documentation of calibration methods in the cancer simulation literature is highly variable. We expected methods would be well documented in all published papers. Descriptions of calibration methods, if included in the articles, ranged in detail and were often only a few sentences in length. In most equivocal cases, details were not reported or the terminology was considered too imprecise to categorize the methods. The very real possibility that misclassification of methods for these models may have occurred emphasizes the need for more clear and consistent reporting. We recognize that journal word count limits may preclude detailed descriptions of calibration protocols and that the technical details of model calibration may be difficult to publish as stand-alone articles because of their specialized nature. However, clear documentation is critical for both modellers and consumers of model results.

To aid the documentation process, we propose a Calibration Reporting Checklist be adopted as standard practice for modellers (figure 2). The checklist can help ensure reporting of details regarding the data, methods and sources for the calibration targets, GOF metric(s), search algorithm(s), acceptance criteria and stopping rules. By encouraging more complete and consistent documentation of all components of calibration, the checklist would aid in the peer-review process. Additionally, with improved transparency, modelling may be less frequently viewed as a black box process.[177]

The use of the checklist will also provide a means of disseminating existing and new calibration methodology from other fields such as engineering and environmental science to the disease-modelling community. Furthermore, it can facilitate comparisons of methods across models, critical for collaborative-modelling projects such as CISNET. However, there are few rules of thumb to guide the choice of methods for any particular model. At present, the process of calibration is often an art, rather than a science. Open research questions remain about the quality and appropriateness of alternative calibration methods. Direct comparisons of methods both within and across models are needed to understand

1) Target data

☐ Description of source data

   ☐ surveillance or registry _____

   ☐ observational study

   ☐ clinical trial _____

   ☐ other _____

   ▪ Do the data span multiple years or ages?

   ☐ Yes   ☐ No

   Details _____

☐ Corresponding model outputs

_____

☐ References

2) Search algorithm

☐ Description or identification of specific algorithm(s)

   ☐ grid _____

   ☐ random _____

   ☐ directed _____

   ☐ trial-and-error _____

   ☐ other _____

☐ If used multiple targets are used, how were they fit?

   ☐ Simultaneously   ☐ Sequentially   ☐ Other

   Details _____

☐ References

3) GOF metric

☐ Description

   ☐ absolute distance _____

   ☐ relative distance _____

   ☐ likelihood-based _____

   ☐ visual _____

   ☐ other _____

☐ Computation of overall "score" if multiple targets are used

   ☐ None   ☐ Raw sum   ☐ Weighted-average   ☐ Other

   Details _____

4) Acceptance criteria

☐ Description

   ☐ statistical significance _____

   ☐ predefined threshold for GOF _____

   ☐ visual assessment _____

   ☐ other _____

☐ Was uncertainty in target data incorporated in acceptance criteria?

   ☐ Yes   ☐ No

   Details _____

☐ Did acceptance criteria allow for multiple parameter sets to be accepted?

   ☐ Yes   ☐ No

   Details _____

5) Stopping rule

☐ Description

_____

6) Validation

☐ Was the validity of the model assessed?

   ☐ Yes   ☐ No

☐ Were methods for assessing validation provided?

   ☐ Yes   ☐ No

   Details _____

☐ Description of validation data

   ☐ surveillance or registry _____

   ☐ observational study _____

   ☐ clinical trial _____

   ☐ other _____

   ▪ Were the data independent from those used in calibration?

   ☐ Yes   ☐ No

☐ References

7) Availability of detailed description of calibration methods

   ☐ published article _____

   ☐ webpage _____

   ☐ technical appendix _____

   ☐ other _____

Fig. 2. The Calibration Reporting Checklist, a standardized framework for describing the model calibration process. The checklist highlights key elements of the process. Inclusion of these key elements in model descriptions can aid in the peer-review process and model comparisons. **GOF** = goodness of fit.

if different methods could lead to different calibration results. These types of comparisons will be important next steps for the advancement of model-calibration methodology as well as disease simulation modelling in general.

Our focus on cancer-screening models for this review does not allow conclusions about the reporting of calibration methods used in models of other diseases. However, similar reporting issues most likely exist in other disease models and the checklist would be applicable for these as well. Because of the imperfect nature of article index-ing, our search may have missed some relevant articles. For example, the term 'natural history' only became a US National Library of Medicine MeSH heading in 1996 and it is a broad term. We attempted to mitigate the potential loss of articles by performing a keyword search and including articles from our personal reprints and those referenced in the retrieved articles. As illustrated by our survey of the literature, this is a growing field with greater numbers of disease-modelling articles published each year. We are aware of several recent articles describing details of calibration

methods;[26,178,179] however, there are likely to be additional examples that have been published after our search period.

4. Conclusions

Disease simulation modelling plays an important role in health policy analysis. Our review indicates the application of such models in cancer screening has become more widespread, particularly within the past decade. The use of disease simulation modelling is likely to expand especially in light of new initiatives in comparative effectiveness research. In addressing policy questions, these models synthesize biological, epidemiological and economic data from diverse sources. With their increasing use for policy making, the models themselves will face additional scrutiny; therefore, careful documentation of methods will be critical for establishing credibility. While there have been efforts to standardize modelling methods in economic evaluation in general,[177,180] less attention has been paid to calibration methods for disease simulation models in particular. Although questions remain about the best calibration practices, the use of the Calibration Reporting Checklist would begin the standardization process, provide a basis for more transparent comparisons across models and facilitate important discussions about methods.

Acknowledgements

References
1. Ramsey SD, McIntosh M, Etzioni RD, et al. Simulation modeling of outcomes and cost-effectiveness. Hematol Oncol Clin North Am 2000 Aug; 14 (4): 925-38

2. Knudsen AB, McMahon PM, Gazelle GS. Use of modeling to evaluate the cost-effectiveness of cancer screening programs. J Clin Oncol 2007; 25: 203-8

3. Feuer EJ, Etzioni RD, Cronin KA, et al. The use of modeling to understand the impact of screening on US mortality: examples from mammography and PSA testing. Stat Methods Med Res 2004 Dec; 13: 421-42

4. Goldie SJ. Chapter 15: public health policy and cost-effectiveness analysis. J Natl Cancer Inst Monogr 2003; 31: 102-10

5. Goldie SJ, Goldhaber-Fiebert JD, Garnett G. Chapter 18: public health policy for cervical cancer prevention. The role of decision science, economic evaluation, and mathematical modeling. Vaccine 2006 Aug 31; 24 (S3): 155-63

6. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics 2006; 24 (11): 1043-53

7. Goldie SJ, Grima D, Kohli M, et al. A comprehensive natural history model of HPV infection and cervical cancer to estimate the clinical impact of a prophylactic HPV-16/18 vaccine. Int J Cancer 2003; 106: 896-904

8. Yeh JM, Kuntz KM, Ezzati M, et al. Development of an empirically calibrated model of gastric cancer in two high-risk countries. Cancer Epidemiol Biomarkers Prev 2008 May 1; 17 (5): 1179-87

9. Mandelblatt JS, Lawrence WF, Womack SM, et al. Benefits and costs of using HPV testing to screen for cervical cancer. JAMA 2002 May 8; 287 (18): 2372-81

10. Urban N, Drescher C, Etzioni R, et al. Use of a stochastic simulation model to identify an efficient protocol for ovarian cancer screening. Control Clin Trials 1997 Jun; 18 (3): 251-70

11. Myers ER, McCrory DC, Nanda K, et al. Mathematical model for the natural history of human papillomavirus infection and cervical carcinogenesis. Am J Epidemiol 2000 June 15, 2000; 151 (12): 1158-71

12. Law AM, Kelton WD. Simulation modeling and analysis. 3rd ed. Boston (MA): McGraw-Hill, 2000

13. Clarke LD, Plevritis SK, Boer R, et al. A comparative review of CISNET breast models used to analyze U.S. breast cancer incidence and mortality trends. J Natl Cancer Inst Monogr 2006; 36: 96-105

14. Banks J, editor. Handbook of simulation: principles, methodology, advances, applications, and practice. New York: John Wiley & Sons. Inc., 1998

15. Bruning JL, Kintz BL. Computational handbook of statistics. 4th ed. Boston (MA): Allyn & Bacon, 1997

16. Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965; 7 (4): 308-13

17. Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science 1983; 220 (4598): 671-80

18. Press WH, Teukolsky SA, Vetterling WT, et al. Numerical recipes in C++. 2nd ed. New York (NY): Cambridge University Press, 2002

19. Wong DF, Leong HW, Liu CL. Simulated annealing for VLSI design. Boston (MA): Kluwer Academic Publishers, 1988

20. Holland JH. Adaptation in natural and artificial systems. Ann Arbor (MI): University of Michigan Press, 1975

21. Goldberg DE. Genetic algorithms in search, optimization and machine learning. Boston (MA): Kluwer Academic Publishers, 1989

22. Manikas TW, Cain JT. Genetic algorithms vs. simulated annealing: a comparison of approaches for solving the circuit partitioning problem. Pittsburgh (PA): University of Pittsburgh, 1996

23. Ingber L, Rosen B. Genetic algorithms and very fast simulated annealing: a comparison. Math Comput Model 1992; 16 (11): 87-100

24. Glover F. Tabu search: part I. ORSA J Computing 1989; 1 (3): 190-206

25. Glover F. Tabu search: part II. ORSA J Computing 1990; 2 (1): 4-32

26. Kong CY, McMahon PM, Gazelle GS. Calibration of disease simulation models using an engineering approach. Value Health 2009; 12 (4): 521-9

27. Baker RD. Use of a mathematical model to evaluate breast cancer screening policy. Health Care Manage Sci 1998 Oct; 1 (2): 103-13

28. Beckett JR, Kotre CJ, Michaelson JS. Analysis of benefit: risk ratio and mortality reduction for the UK Breast Screening Programme. Br J Radiol 2003 May; 76 (905): 309-20

29. Beemsterboer PM, Warmerdam PG, Boer R, et al. Radiation risk of mammography related to benefit in screening programmes: a favourable balance? J Med Screen 1998; 5 (2): 81-7

30. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant treatment on mortality from breast cancer. N Engl J Med 2005 Oct 27; 353 (17): 1784-92

31. Berry DA, Inoue L, Shen Y, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: a Bayesian approach. J Natl Cancer Inst Monogr 2006; 36: 30-6

32. Blanchard K, Colbert JA, Puri D, et al. Mammographic screening: patterns of use and estimated impact on breast carcinoma survival. Cancer 2004 Aug 1; 101 (3): 495-507

33. Boer R, de Koning H, Threlfall A, et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. BMJ 1998 Aug 8; 317 (7155): 376-9

34. Boer R, de Koning H, van Oortmarssen G, et al. Stage distribution at first and repeat examinations in breast cancer screening. J Med Screen 1999; 6 (3): 132-8

35. Carter R, Glasziou P, van Oortmarssen G, et al. Cost-effectiveness of mammographic screening in Australia. Aust J Public Health 1993 Mar; 17 (1): 42-50

36. Carter KJ, Castro F, Kessler E, et al. A computer model for the study of breast cancer. Comput Biol Med 2003 Jul; 33: 345-60

37. Carter KJ, Castro F, Kessler E, et al. Simulation of breast cancer screening: quality assessment of two protocols. J Healthc Qual 2004 Nov-Dec; 26 (6): 31-8

38. Castro F, Carter KJ, Kessler E, et al. The relation of breast cancer staging to screening protocol compliance: a computer simulation study. Comput Biol Med 2005 Feb; 35 (2): 91-101

39. Chia YL, Salzman P, Plevritis SK, et al. Simulation-based parameter estimation for complex models: a breast cancer natural history modelling illustration. Stat Methods Med Res 2004 Dec; 13 (6): 507-24

40. Cong XJ, Shen Y, Miller AB. Estimation of age-specific sensitivity and sojourn time in breast cancer screening studies. Stat Med 2005 Oct 30; 24 (20): 3123-38

41. Connor RJ, Boer R, Prorok PC, et al. Investigation of design and bias issues in case-control studies of cancer screening using microsimulation. Am J Epidemiol 2000 May 15; 151 (10): 991-8

42. de Koning HJ, van Ineveld BM, van Oortmarssen GJ, et al. Breast cancer screening and cost-effectiveness: policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991 Oct 21; 49: 531-7

43. de Koning HJ, Boer R, Warmerdam PG, et al. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. J Natl Cancer Inst 1995 Aug 16; 87 (16): 1217-23

44. Eddy DM. Screening for breast cancer. Ann Intern Med 1989; 111 (5): 389-99

45. Feldstein M, Zelen M. Inferring the natural time history of breast cancer: implications for tumor growth rate and early detection. Breast Cancer Res Treat 1984; 4: 3-10

46. Fett MJ. Computer modelling of the Swedish two county trial of mammographic screening and trade offs between participation and screening interval. J Med Screen 2001; 8 (1): 39-45

47. Fracheboud J, Groenewoud JH, Boer R, et al. Seventy-five years is an appropriate upper age limit for population-based mammography screening. Int J Cancer 2006 Apr 15; 118 (8): 2020-5

48. Fryback DG, Stout NK, Rosenberg MA, et al. The Wisconsin breast cancer epidemiology simulation model. J Natl Cancer Inst Monogr 2006; 36: 37-47

49. Hanin LG, Miller A, Zorin AV, et al. The University of Rochester model of breast cancer detection and survival. J Natl Cancer Inst Monogr 2006; 36: 66-78

50. Hsieh HJ, Chen TH, Chang SH. Assessing chronic disease progression using non-homogeneous exponential regression Markov models: an illustration using a selective breast cancer screening in Taiwan. Stat Med 2002 Nov 30; 21 (22): 3369-82

51. Hunter DJ, Drake SM, Shortt SE, et al. Simulation modeling of change to breast cancer detection age eligibility recommendations in Ontario, 2002-2021. Cancer Detect Prev 2004; 28 (6): 453-60

52. Jacobi CE, Jonker MA, Nagelkerke NJ, et al. Prevalence of family histories of breast cancer in the general population and the incidence of related seeking of health care [letter]. J Med Genet 2003 Jul; 40 (7): e83

53. Jacobi CE, Nagelkerke NJ, van Houwelingen JH, et al. Breast cancer screening, outside the population-screening program, of women from breast cancer families without proven BRCA1/BRCA2 mutations: a simulation study. Cancer Epidemiol Biomarkers Prev 2006 Mar; 15 (3): 429-36

54. Jansen JT, Zoetelief J. MBS: a model for risk benefit analysis of breast cancer screening. Br J Radiol 1995 Feb; 68 (806): 141-9

55. Jansen JT, Zoetelief J. Assessment of lifetime gained as a result of mammographic breast cancer screening using a computer model. Br J Radiol 1997 Jun; 70 (834): 619-28

56. Jansen JT, Zoetelief J. Optimisation of mammographic breast cancer screening using a computer simulation model. Eur J Radiol 1997 Feb; 24 (2): 137-44

57. Knox EG. Evaluation of a proposed breast cancer screening regimen. BMJ 1988 Sep 10; 297 (6649): 650-4

58. Koscielny S, Tubiana M, Valleron AJ. A simulation model of the natural history of human breast cancer. Br J Cancer 1985 Oct; 52 (4): 515-24

59. Lee S, Zelen M. A stochastic model for predicting the mortality of breast cancer. J Natl Cancer Inst Monogr 2006; 36: 79-86

60. Mandelblatt JS, Schechter CB, Yabroff KR, et al. Benefits and costs of interventions to improve breast cancer outcomes in African American women. J Clin Oncol 2004 Jul 1; 22 (13): 2554-66

61. Mandelblatt J, Schechter CB, Lawrence W, et al. The SPECTRUM population model of the impact of screening and treatment on U.S. breast cancer trends from 1975 to 2000: principles and practice of the model methods. J Natl Cancer Inst Monogr 2006; 36: 47-55

62. Manton KG, Stallard E. Demographics (1950–1987) of breast cancer in birth cohorts of older women. J Gerontol 1992 Nov; 47 Spec No: 32-42

63. Michaelson JS, Halpern E, Kopans DB. Breast cancer: computer simulation method for estimating optimal intervals for screening. Radiology 1999 Aug; 212 (2): 551-60

64. Michaelson JS, Satija S, Moore R, et al. Estimates of breast cancer growth rate and sojourn time from screening database information. J Wom Imag 2003 Feb; 5 (1): 11-9

65. Michaelson JS, Satija S, Moore R, et al. Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds. J Wom Imag 2003 Feb; 5 (1): 3-10

66. Myles JP, Nixon RM, Duffy SW, et al. Bayesian evaluation of breast cancer screening using data from two studies. Stat Med 2003 May 30; 22 (10): 1661-74

67. Okubo I, Glick H, Frumkin H, et al. Cost-effectiveness analysis of mass screening for breast cancer in Japan. Cancer 1991 Apr 15; 67 (8): 2021-9

68. Paci E, Boer R, Zappa M, et al. A model-based prediction of the impact on reduction in mortality by a breast cancer screening programme in the city of Florence, Italy. Eur J Cancer 1995; 31A (3): 348-53

69. Plevritis SK. A mathematical algorithm that computes breast cancer sizes and doubling times detected by screening. Math Biosci 2001 Jun; 171 (2): 155-78

70. Plevritis SK, Kurian AW, Sigal BM, et al. Cost-effectiveness of screening BRCA1/2 mutation carriers with breast magnetic resonance imaging. JAMA 2006 May 24; 295 (20): 2374-84

71. Plevritis SK, Sigal BM, Salzman P, et al. A stochastic simulation model of U.S. breast cancer mortality trends from 1975 to 2000. J Natl Cancer Inst Monogr 2006; 36: 86-95

72. Plevritis SK, Salzman P, Sigal BM, et al. A natural history model of stage progression applied to breast cancer. Stat Med 2006; 26 (3): 581-95

73. Shen Y, Huang X. Nonparametric estimation of asymptomatic duration from a randomized prospective cancer screening trial. Biometrics 2005 Dec; 61 (4): 992-9

74. Shen Y, Zelen M. Robust modeling in screening studies: estimation of sensitivity and preclinical sojourn time distribution. Biostatistics 2005 Oct; 6 (4): 604-14

75. Stout NK, Rosenberg MA, Trentham-Dietz A, et al. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst 2006 Jun 7; 98 (11): 774-82

76. Szeto KL, Devlin NJ. The cost-effectiveness of mammography screening: evidence from a microsimulation model for New Zealand. Health Policy 1996 Nov; 38 (2): 101-15

77. Tan SYGL, van Oortmarssen GJ, de Koning HJ, et al. The MISCAN-Fadia continuous tumor growth model for breast cancer. J Natl Cancer Inst Monogr 2006; 36: 56-65

78. Tubiana M, Koscielny S. The natural history of breast cancer: implications for a screening strategy. Int J Radiat Oncol Biol Phys 1990 Nov; 19 (5): 1117-20

79. van der Maas PJ, de Koning HJ, van Ineveld BM, et al. The cost-effectiveness of breast cancer screening. Int J Cancer 1989 Jun 15; 43 (6): 1055-60

80. van Oortmarssen GJ, Habbema JD, Lubbe JT, et al. A model-based analysis of the HIP project for breast cancer screening. Int J Cancer 1990 Aug 15; 46 (2): 207-13

81. van Oortmarssen GJ, Habbema JD, van der Maas PJ, et al. A model for breast cancer screening. Cancer 1990 Oct 1; 66 (7): 1601-12

82. Vervoort MM, Draisma G, Fracheboud J, et al. Trends in the usage of adjuvant systemic therapy for breast cancer in the Netherlands and its effect on mortality. Br J Cancer 2004 Jul 19; 91 (2): 242-7

83. Berkhof J, de Bruijne MC, Zielinski GD, et al. Natural history and screening model for high-risk human papillomavirus infection, neoplasia and cervical cancer in the Netherlands. Int J Cancer 2005 Jun 10; 115 (2): 268-75

84. Berkhof J, de Bruijne MC, Zielinski GD, et al. Evaluation of cervical screening strategies with adjunct high-risk human papillomavirus testing for women with borderline or mild dyskaryosis. Int J Cancer 2006 Apr 1; 118 (7): 1759-68

85. Eddy DM. The frequency of cervical cancer screening: comparison of a mathematical model with empirical data. Cancer 1987 Sep 1; 60 (5): 1117-22

86. Eddy DM. Screening for cervical cancer. Ann Intern Med 1990 Aug 1; 113 (3): 214-26

87. Goldie SJ, Weinstein MC, Kuntz KM, et al. The costs, clinical benefits, and cost-effectiveness of screening for cervical cancer in HIV-infected women. Ann Intern Med 1999 Jan 19; 130 (2): 97-107

88. Goldie SJ, Kuhn L, Denny L, et al. Policy analysis of cervical cancer screening strategies in low-resource settings: clinical benefits and cost-effectiveness. JAMA 2001 Jun 27; 285 (24): 3107-15

89. Goldie SJ, Kim JJ, Wright TCJ. Cost-effectiveness of human papillomavirus DNA testing for cervical cancer screening in women aged 30 years or more. Obstet Gynecol 2004 Apr; 103 (4): 619-31

90. Goldie SJ, Kohli M, Grima D, et al. Projected clinical benefits and cost-effectiveness of a human papillomavirus 16/18 vaccine. J Natl Cancer Inst 2004 Apr 21; 96 (8): 604-15

91. Goldie SJ, Gaffikin L, Goldhaber-Fiebert JD, et al. Cost-effectiveness of cervical-cancer screening in five developing countries. N Engl J Med 2005 Nov 17; 353 (20): 2158-68

92. Gustafsson L, Adami HO. Natural history of cervical neoplasia: consistent results obtained by an identification technique. Br J Cancer 1989 Jul; 60 (1): 132-41

93. Gustafsson L, Adami HO. Cytologic screening for cancer of the uterine cervix in Sweden evaluated by identification and simulation. Br J Cancer 1990 Jun; 61 (6): 903-8

94. Gustafsson L, Adami HO. Optimization of cervical cancer screening. Cancer Causes Control 1992 Mar; 3 (2): 125-36

95. Gyrd-Hansen D, Holund B, Andersen P. A cost-effectiveness analysis of cervical cancer screening: health policy implications. Health Policy 1995 Oct; 34 (1): 35-51

96. Habbema JD, van Oortmarssen GJ, Lubbe JT, et al. Model building on the basis of Dutch cervical cancer screening data. Maturitas 1985 May; 7 (1): 11-20

97. Habbema JDF, Lubbe JTN, van Oortmarssen GJ, et al. A simulation approach to cost-effectiveness and cost-benefit calculations of screening for the early detection of disease. Eur J Oper Res 1987 May; 29 (2): 159-66

98. Helfand M, O'Connor GT, Zimmer-Gembeck M, et al. Effect of the Clinical Laboratory Improvement Amendments of 1988 (CLIA '88) on the incidence of invasive cervical cancer. Med Care 1992 Dec; 30 (12): 1067-82

99. Kim JJ, Wright TCJ, Goldie SJ. Cost-effectiveness of alternative triage strategies for atypical squamous cells of undetermined significance. JAMA 2002 May 8; 287 (18): 2382-90

100. Kim JJ, Leung GM, Woo PP, et al. Cost-effectiveness of organized versus opportunistic cervical cytology screening in Hong Kong. J Public Health (Oxf) 2004 Jun; 26 (2): 130-7

101. Knox EG. A simulation system for screening procedures. In: McLachlan G, editor. Future and present indicatives, problems and progress in medical care. 9th rev. Oxford: Nuffield Provincial Hospitals Trust, Oxford University Press, 1973: 17-55

102. Koong SL, Yen AM, Chen TH. Efficacy and cost-effectiveness of nationwide cervical cancer screening in Taiwan. J Med Screen 2006; 13 Suppl. 1: S44-7

103. Koopmanschap MA, Lubbe KT, van Oortmarssen GJ, et al. Economic aspects of cervical cancer screening. Soc Sci Med 1990; 30 (10): 1081-7

104. Kulasingam SL, Myers ER, Lawson HW, et al. Cost-effectiveness of extending cervical cancer screening intervals among women with prior normal pap tests. Obstet Gynecol 2006 Feb; 107 (2 Pt 1): 321-8

105. Mandelblatt JS, Lawrence WF, Gaffikin L, et al. Costs and benefits of different strategies to screen for cervical cancer in less-developed countries. J Natl Cancer Inst 2002 Oct 2; 94 (19): 1469-83

106. Mandelblatt J, Lawrence W, Yi B, et al. The balance of harms, benefits, and costs of screening for cervical cancer in older women. Arch Intern Med 2004 Feb 9; 164: 245-7

107. Matsunaga G, Tsuji I, Sato S, et al. Cost-effective analysis of mass screening for cervical cancer in Japan. J Epidemiol 1997 Sep; 7 (3): 135-41

108. Office of Technology Assessment. The costs and effectiveness of screening for cervical cancer in elderly women: background paper, OTA-BP-H-65. Washington, DC: U.S. Congress, 1990

109. Parkin DM. A computer simulation model for the practical planning of cervical cancer screening programmes. Br J Cancer 1985 Apr; 51 (4): 551-68

110. Parkin DM, Moss SM. An evaluation of screening policies for cervical cancer in England and Wales using a computer simulation model. J Epidemiol Community Health 1986 Jun; 40 (2): 143-53

111. Radensky PW, Mango LJ. Interactive neural-network-assisted screening: an economic assessment. Acta Cytol 1998 Jan-Feb; 42 (1): 246-52

112. Sato S, Matsunaga G, Tsuji I, et al. Determining the cost-effectiveness of mass screening for cervical cancer using common analytic models. Acta Cytol 1999 Nov-Dec; 43 (6): 1006-14

113. Sawaya GF, McConnell KJ, Kulasingam SL, et al. Risk of cervical cancer associated with extending the interval between cervical-cancer screenings. N Engl J Med 2003 Oct 16; 349 (16): 1501-9

114. Schechter CB. Cost-effectiveness of rescreening conventionally prepared cervical smears by PAPNET testing. Acta Cytol 1996 Nov-Dec; 40 (6): 1272-82

115. Sherlaw-Johnson C, Gallivan S, Jenkins D. Withdrawing low risk women from cervical screening programmes: mathematical modelling study. BMJ 1999 Feb 6; 318 (7180): 356-60

116. Sherlaw-Johnson C, Philips Z. An evaluation of liquid-based cytology and human papillomavirus testing within the UK Cervical Cancer Screening Programme. Br J Cancer 2004 Jul 5; 91 (1): 84-91

117. Shun-Zhang Y, Miller AB, Sherman GJ. Optimising the age, number of tests, and test interval for cervical screening in Canada. J Epidemiol Community Health 1982 Mar; 36 (1): 1-10

118. Sreenivas V, Prabhakar AK, Ravi R, et al. A simulation approach for estimating the loss of woman years due to cervical cancer and probability of developing cervical cancer. Neoplasma 1989; 36 (5): 623-7

119. van Ballegooijen M, van den Akker-van Marle E, Patnick J, et al. Overview of important cervical cancer screening process values in European Union (EU) countries, and tentative predictions of the corresponding effectiveness and cost-effectiveness. Eur J Cancer 2000 Nov; 36 (17): 2177-88

120. van den Akker-van Marle ME, van Ballegooijen M, van Oortmarssen GJ, et al. Cost-effectiveness of cervical cancer screening: comparison of screening policies. J Natl Cancer Inst 2002 Feb 6; 94 (3): 193-204

121. Clemen RT, Lacke CJ. Analysis of colorectal cancer screening regimens. Health Care Manage Sci 2001 Dec; 4 (4): 257-67

122. Frazier AL, Colditz GA, Fuchs CS, et al. Cost-effectiveness of screening for colorectal cancer in the general population. JAMA 2000 Oct 18; 284 (15): 1954-61

123. Gyrd-Hansen D, Søgaard J, Kronborg O. Analysis of screening data: colorectal cancer. Int J Epidemiol 1997 Dec; 26 (6): 1172-81

124. Gyrd-Hansen D, Søgaard J, Kronborg O. Colorectal cancer screening: efficiency and effectiveness. Health Econ 1998 Feb; 7 (1): 9-20

125. Haug U, Brenner H. A simulation model for colorectal cancer screening: potential of stool tests with various performance characteristics compared with screening colonoscopy. Cancer Epidemiol Biomarkers Prev 2005 Feb; 14 (2): 422-8

126. Khandker RK, Dulski JD, Kilpatrick JB, et al. A decision model and cost-effectiveness analysis of colorectal cancer screening and surveillance guidelines for average-risk adults. Int J Technol Assess Health Care 2000 Summer; 16 (3): 799-810

127. Ladabaum U, Chopra CL, Huang G, et al. Aspirin as an adjunct to screening for prevention of sporadic colorectal cancer: a cost-effectiveness analysis. Ann Intern Med 2001 Nov 6; 135 (9): 769-81

128. Ladabaum U, Scheiman JM, Fendrick AM. Potential effect of cyclooxygenase-2-specific inhibitors on the prevention of colorectal cancer: a cost-effectiveness analysis. Am J Med 2003 May; 114 (7): 546-54

129. Lejeune C, Arveux P, Dancourt V, et al. A simulation model for evaluating the medical and economic outcomes of screening strategies for colorectal cancer. Eur J Cancer Prev 2003 Feb; 12 (1): 77-84

130. Lejeune C, Arveux P, Dancourt V, et al. Cost-effectiveness analysis of fecal occult blood screening for colorectal cancer. Int J Technol Assess Health Care 2004 Fall; 20 (4): 434-9

131. Loeve F, Boer R, van Oortmarssen GJ, et al. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. Comput Biomed Res 1999 Feb; 32 (1): 13-33

132. Loeve F, Brown ML, Boer R, et al. Endoscopic colorectal cancer screening: a cost-saving analysis. J Natl Cancer Inst 2000 Apr 5; 92 (7): 557-63

133. Loeve F, Boer R, van Oortmarssen GJ, et al. Impact of systematic false-negative test results on the performance of faecal occult blood screening. Eur J Cancer 2001 May; 37 (7): 912-7

134. Neilson AR, Whynes DK. Cost-effectiveness of screening for colorectal cancer: a simulation model. IMA J Math Appl Med Biol 1995 Sep-Dec; 12 (3-4): 355-67

135. Ness RM, Holmes AM, Klein R, et al. Cost-utility of one-time colonoscopic screening for colorectal cancer at various ages. Am J Gastroenterol 2000 Jul; 95 (7): 1800-11

136. Song K, Fendrick AM, Ladabaum U. Fecal DNA testing compared with conventional colorectal cancer screening methods: a decision analysis. Gastroenterology 2004 May; 126 (5): 1270-9

137. Vijan S, Hwang EW, Hofer TP, et al. Which colon cancer screening test? A comparison of costs, effectiveness, and compliance. Am J Med 2001 Dec 1; 111 (8): 593-601

138. Wagner JL, Herdman RC, Wadhwa S. Cost effectiveness of colorectal cancer screening in the elderly. Ann Intern Med 1991 Nov 15; 115 (10): 807-17

139. Wagner JL, Tunis S, Brown M, et al. Cost-effectiveness of colorectal cancer screening in average-risk adults. In: Young G, Rosen P, Levin B, editors. Prevention and early detection of colorectal cancer. Philadelphia (PA): Saunders, 1996: 321-56

140. Whynes DK, Neilson AR, Walker AR, et al. Faecal occult blood screening for colorectal cancer: is it cost-effective? Health Econ 1998 Feb; 7 (1): 21-9

141. Wong JM, Yen MF, Lai MS, et al. Progression rates of colorectal cancer by Dukes' stage in a high-risk group: analysis of selective colorectal cancer screening. Cancer J 2004 May-Jun; 10 (3): 160-9

142. Yang KC, Liao CS, Chiu YH, et al. Colorectal cancer screening with faecal occult blood test within a multiple disease screening programme: an experience from Keelung, Taiwan. J Med Screen 2006; 13 Suppl. 1: S8-13

143. Garside R, Pitt M, Somerville M, et al. Surveillance of Barrett's oesophagus: exploring the uncertainty through systematic review, expert workshop and economic modelling. Health Technol Assess 2006 Mar; 10 (8): 1-142

144. Davies R, Crabbe D, Roderick P, et al. A simulation to evaluate screening for *Helicobacter pylori* infection in the prevention of peptic ulcers and gastric cancers. Health Care Manage Sci 2002 Nov; 5 (4): 249-58

145. Fendrick AM, Chernew ME, Hirth RA, et al. Clinical and economic effects of population-based *Helicobacter pylori* screening to prevent gastric cancer. Arch Intern Med 1999 Jan 25; 159 (2): 142-8

146. Roderick P, Davies R, Raftery J, et al. The cost-effectiveness of screening for *Helicobacter pylori* to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model. Health Technol Assess 2003; 7 (6): 1-86

147. Roderick P, Davies R, Raftery J, et al. Cost-effectiveness of population screening for *Helicobacter pylori* in preventing gastric cancer and peptic ulcer disease, using simulation. J Med Screen 2003; 10 (3): 148-56

148. Das P, Ng AK, Earle CC, et al. Computed tomography screening for lung cancer in Hodgkin's lymphoma survivors: decision analysis and cost-effectiveness analysis. Ann Oncol 2006 May; 17 (5): 785-93

149. Flehinger BJ, Kimmel M. The natural history of lung cancer in a periodically screened population. Biometrics 1987 Mar; 43 (1): 127-44

150. Flehinger BJ, Kimmel M, Melamed MR. Natural history of adenocarcinoma-large cell carcinoma of the lung: conclusions from screening programs in New York and Baltimore. J Natl Cancer Inst 1988 May 4; 80 (5): 337-44

151. Flehinger BJ, Kimmel M, Polyak T, et al. Screening for lung cancer: the Mayo Lung Project revisited. Cancer 1993 Sep 1; 72 (5): 1573-80

152. Gorlova OY, Kimmel M, Henschke C. Modeling of long-term screening for lung carcinoma. Cancer 2001 Sep 15; 92 (6): 1531-40

153. Mahadevia PJ, Fleisher LA, Frick KD, et al. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost-effectiveness analysis. JAMA 2003 Jan 15; 289 (3): 313-22

154. Marshall D, Simpson KN, Earle CC, et al. Potential cost-effectiveness of one-time screening for lung cancer (LC) in a high risk cohort. Lung Cancer 2001 Jun; 32 (3): 227-36

155. Yamaguchi N, Tamura Y, Sobue T, et al. Evaluation of cancer prevention strategies by computerized simulation

model: an approach to lung cancer. Cancer Causes Control 1991 May; 2 (3): 147-55

156. Yamaguchi N, Mizuno S, Akiba S, et al. A 50-year projection of lung cancer deaths among Japanese males and potential impact evaluation of anti-smoking measures and screening using a computerized simulation model. Jpn J Cancer Res 1992 Mar; 83 (3): 251-7

157. Yamaguchi N, Tamura Y, Sobue T, et al. Evaluation of cancer prevention strategies by computerized simulation model: methodological issues. Environ Health Perspect 1994 Nov; 102 Suppl. 8: 67-71

158. Girgis A, Clarke P, Burton RC, et al. Screening for melanoma by primary health care physicians: a cost-effectiveness analysis. J Med Screen 1996; 3 (1): 47-53

159. Downer MC, Jullien JA, Speight PM. An interim determination of health gain from oral cancer and precancer screening: 2. Developing a model of population screening. Community Dent Health 1997 Dec; 14 (4): 227-32

160. Myers ER, Havrilesky LJ, Kulasingam SL, et al. Genomic tests for ovarian cancer detection and management. Evid Rep Technol Assess (Full Rep) 2006 Oct; (145): 1-100

161. Skates SJ, Singer DE. Quantifying the potential benefit of CA 125 screening for ovarian cancer. J Clin Epidemiol 1991; 44 (4-5): 365-80

162. Skates SJ, Pauler DK, Jacobs IJ. Screening based on the risk of cancer calculation from Bayesian hierarchical change point and mixture models of longitudinal markers. J Am Stat Assoc 2001; 96 (454): 429-39

163. Cowen ME, Chartrand M, Weitzel WF. A Markov model of the natural history of prostate cancer. J Clin Epidemiol 1994 Jan; 47 (1): 3-21

164. Draisma G, Boer R, Otto SJ, et al. Lead times and over-detection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. J Natl Cancer Inst 2003 Jun 18; 95 (12): 868-78

165. Draisma G, De Koning HJ. MISCAN: estimating lead-time and over-detection by simulation. BJU Int 2003 Dec; 92 Suppl. 2: 106-11

166. Draisma G, Postma R, Schroder FH, et al. Gleason score, age and screening: modeling dedifferentiation in prostate cancer. Int J Cancer 2006 Nov 15; 119 (10): 2366-71

167. Etzioni R, Cha R, Cowen ME. Serial prostate specific antigen screening for prostate cancer: a computer model evaluates competing strategies. J Urol 1999 Sep; 162 (3 Pt 1): 741-8

168. Etzioni R, Legler JM, Feuer EJ, et al. Cancer surveillance series: interpreting trends in prostate cancer. Part III: quantifying the link between population prostate-specific antigen testing and recent declines in prostate cancer mortality. J Natl Cancer Inst 1999 Jun 16; 91 (12): 1033-9

169. Etzioni R, Penson DF, Legler JM, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 2002 Jul 3; 94 (13): 981-90

170. Krahn MD, Mahoney JE, Eckman MH, et al. Screening for prostate cancer: a decision analytic view. JAMA 1994 Sep 14; 272 (10): 773-80

171. Parker C, Muston D, Melia J, et al. A model of the natural history of screen-detected prostate cancer, and the effect of radical treatment on overall survival. Br J Cancer 2006 May 22; 94 (10): 1361-8

172. Ross KS, Carter HB, Pearson JD, et al. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA 2000 Sep 20; 284 (11): 1399-405

173. Ross KS, Guess HA, Carter HB. Estimation of treatment benefits when PSA screening for prostate cancer is discontinued at different ages. Urology 2005 Nov; 66 (5): 1038-42

174. Tsodikov A, Szabo A, Wegelin J. A population model of prostate cancer incidence. Stat Med 2006 Aug 30; 25 (16): 2846-66

175. Kimmel M, Flehinger BJ. Nonparametric estimation of the size-metastasis relationship in solid cancers. Biometrics 1991 Sep; 47 (3): 987-1004

176. Wang PE, Wang TT, Chiu YH, et al. Evolution of multiple disease screening in Keelung: a model for community involvement in health interventions? J Med Screen 2006; 13 Suppl. 1: S54-8

177. Garrison LP. The ISPOR good practice modeling principles: a sensible approach. Be transparent, be reasonable. Value Health 2003 Jan-Feb; 6 (1): 6-8

178. Kim JJ, Kuntz KM, Stout NK, et al. Multiparameter calibration of a natural history model of cervical cancer. Am J Epidemiol 2007 July 15, 2007; 166 (2): 137-50

179. Goldhaber-Fiebert JD, Stout NK, Ortendahl J, et al. Modeling human papillomavirus and cervical cancer in the United States for analyses of screening and vaccination. Popul Health Metr 2007; 5 (1): 11 [online]. Available from URL: http://www.pophealthmetrics.com/content/5/1/11 [Accessed 2009 Jul 1]

180. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation. Report of the ISPOR task force on good research practices: modeling studies. Value Health 2003 Jan-Feb; 6 (1): 9-17

Correspondence: Dr *Natasha K. Stout*, Department of Ambulatory Care and Prevention, Harvard Medical School/Harvard Pilgrim Health Care, 133 Brookline Ave, 6th Floor, Boston, MA 02215, USA.
E-mail: natasha_stout@hms.harvard.edu