Journal of Public Health Medicine
DOI: 10.1093/pubmed/fdg075

Vol. 25, No. 4, pp. 325–335
Printed in Great Britain

Systematic review of the use and value of computer simulation modelling in population health and health care delivery

David Fone, Sandra Hollinghurst, Mark Temple, Alison Round, Nathan Lester, Alison Weightman, Katherine Roberts, Edward Coyle, Gwyn Bevan and Stephen Palmer

Abstract

Background The objective of the review was to evaluate the extent, quality and value of computer simulation modelling in population health and health care delivery.

Methods A narrative systematic review was carried out of world literature from 1980 to 1999, searching Medline, INSPEC, Embase, HealthSTAR, Science Citation Index, CINAHL, MathSci, INFORMS Online and SIGLE databases, and researchers in the field were contacted. Papers were included if they contained a computer simulation model of individuals in a stochastic system and the topic or setting related to population health or health service delivery.

Results A total of 182 papers met the inclusion criteria. Simulation modelling has been undertaken in a wide range of health care topic areas, including hospital scheduling and organization, communicable disease, screening, costs of illness and economic evaluation. However, the quality of published papers was variable and few reported on the outcomes of implementation of models, so that the value of modelling could not be assessed.

Conclusion Simulation modelling is a powerful method for modelling both small and large populations to inform policy makers in the provision of health care. It has been applied to a wide variety of health care problems. Although the number of modelling papers has grown substantially over recent years, further research is required to assess the value of modelling.

Keywords: systematic review, simulation modelling, health care

Introduction

Computer models are used extensively in many areas of systems management. They provide an insight into the working of a system and can be used to predict the outcome of a change in strategy. This is particularly useful when the system is very complex and/or when experimentation is not possible. With the increased use and capacity of computer technology, modelling techniques have developed rapidly and there are now a large number of approaches available, including decision analysis, Markov processes, mathematical modelling, systems dynamics and simulation modelling.

The intrinsic uncertainty of health care needs, demands and outcomes requires that if health care policy and management are to be based on evidence, then they must be designed to cope with a wide range of complex systems. It seems self evident that computer modelling should be valuable in providing evidence of how to cope with these stochastic problems, perhaps as an alternative to learning by doing or empirical research.[1]

One approach is discrete event simulation modelling (DESM) and closely related methods, which Taylor and Lane[2] suggest are especially appropriate when there are 'multiple variables which potentially can produce an enormous number of possible connections and effects'. The health care sector abounds with such detail complexity where a wide range of possible outcomes may

[1]Gwent Health Authority, Mamhilad House, Mamhilad Park Estate, Pontypool NP4 0YP.
[2]University of Wales College of Medicine, Heath Park, Cardiff CF14 4XN.
[3]Bro Taf Health Authority, Temple of Peace and Health, Cathays Park, Cardiff CF1 3NW.
[4]N&E Devon Health Authority, Dean Clarke House, Southernhay East, Exeter EX1 1PQ.
[5]London School of Economics and Political Science, Finsbury Tower, 103–105 Bunhill Row, London EC1 8TG.

David Fone,[1,2] Consultant in Public Health Medicine[1], Hon. Senior Lecturer in Department of Epidemiology and Public Health[2]

Sandra Hollinghurst,[1] Health Economist (Public Health)

Mark Temple,[3] Consultant in Public Health Medicine

Alison Round,[2,4] Senior Lecturer in Epidemiology and Public Health[2], Consultant in Public Health Medicine[4]

Nathan Lester,[1] Research Officer

Alison Weightman,[2] Information Resources Manager, Library Services

Katherine Roberts,[2] Information Officer, Library Services

Edward Coyle,[1] Director of Public Health

Gwyn Bevan,[5] Reader in Policy Analysis

Stephen Palmer,[2] Professor of Epidemiology and Public Health

Address correspondence to David Fone, at National Public Health Service, Mamhilad Park Estate, Pontypool NP4 0YP.
E-mail: david.fone@nphs.wales.nhs.uk.

arise from any policy change. Simulation modelling is able to deal with detail complexity by simulating the life histories of individuals and then estimating the population effect from the sum of the individual effects. Each member of the population (entity) included in a simulation model is tracked through a network of options. At each decision point a variety of choices are available, and the outcome will depend on, for example, the characteristics of the entity and resources, previous movement through the model, and the choices other entities have made. In a stochastic system these choices are made by random sampling from defined probability distributions. The majority of models that fulfil these requirements use DESM, although others include, for example, Monte Carlo simulation. By simulating individuals through a system, these models are more understandable and more closely resemble reality, than say, Markov models, in which transition probabilities between health states are applied equally to all individual members of a pre-defined cohort. The individualistic nature of DESM makes simulation modelling especially appropriate for modelling small populations.[3]

With modern developments of software and computing, we might expect simulation modelling to be an essential and routine element in modern management and evaluation of health care. But although modelling has been undertaken in some areas of health care, the use and value of these models remains unclear. The quality of models is not routinely evaluated, and no formal investigation has been carried out to establish the extent to which findings have been translated into policy. No comprehensive systematic review of the use of modelling in health care planning has been published. Previous reviews have been confined to either a particular type of modelling or to a particular application,[4–6] with no attempt to critically appraise the quality of modelling studies or assess the influence of these studies on policy. We therefore conducted a systematic review to evaluate the extent, quality and value of computer simulation modelling in population health and health care delivery.

Methods

Search strategy and selection of papers

We searched Medline, INSPEC (which covers literature on physics, electronic engineering & electronics, computers & information technology), Embase, Science Citation Index, CINAHL (Cumulative Index to Nursing and Allied Health Literature), MathSci, INFORMS Online (an online information service covering Operations Research and Management Sciences, at http://www.informs.org), and SIGLE (System for Information on Grey Literature in Europe) for papers published between 1980 and 1999 inclusive.

Combinations of relevant text words and database subject headings describing simulation modelling in health care settings were searched. The initial search strategy was validated using 14 'key' papers known to the authors, and further refinements were made to improve sensitivity, including cross referencing of citation lists. Table 1 shows the final search strategy and an example of the results of a Medline search from 1997 to 1999.

We compiled a list of researchers currently active in the field from a search of the National Research Register[7] and contacted known researchers in the United Kingdom and the Netherlands to supplement the electronic search and to identify unpublished and current studies. We selected abstracts on the basis of title and keywords. If in doubt we included an abstract for review. Three members of the review team assessed each selected abstract to identify papers suitable for full text review. Each of these reviewers assessed all abstracts independently and results were compared. Where there was discrepancy, the conservative route was taken and the full text of the paper was requested.

Inclusion criteria

We included papers published in peer-reviewed journals or published as a full paper in conference proceedings that contained (1) a computer simulation model of individuals in a stochastic system and (2) the topic or setting related to population health or health service delivery. Papers for which two reviewers were independently in agreement were included; if there was doubt a third reviewer arbitrated.

Critical appraisal

Two reviewers independently critically appraised all papers fulfilling the inclusion criteria. The appraisal was carried out systematically, using a standard proforma (Table 2), developed

Table 1 Search strategy example: MEDLINE <1997 to 1999>

1	exp Computer simulation/	(7011)
2	exp stochastic processes/	(1112)
3	(simulation adj3 (computer: or model: or stochastic:)).mp. [mp=title, abstract, cas registry/ec number word, mesh subject heading]	(1356)
4	soft systems methodology.mp.	(6)
5	(computer adj3 management adj3 model:).mp.	(2)
6	(Markov adj3 model:).mp.	(369)
7	operation: research.mp.	(45)
8	microsimulation.mp.	(18)
9	discrete event.mp.	(28)
10	(stochastic and (process: or model:)).mp.	(697)
11	or/1–10	(8930)
12	exp Cost-benefit analysis/	(6873)
13	exp Patient admission/	(2031)
14	exp Health care costs/	(5696)
15	exp Patient care planning/	(5831)
16	exp Health planning/	(21 456)
17	exp Patient care planning/	(5831)
18	exp 'Appointments and schedules'/	(1200)
19	or/12–18	(38 773)
20	11 and 19	(414)
21	(health or hospital or patient or NHS or screening or out-patient: or outpatient: or clinic:).mp.	(487 075)
22	19 or 21	(501 385)
23	11 and 22	(1983)

Table 2 Critical appraisal sheet

Reference manager ID number: Main author:

Systematic review of the use and value of stochastic computer simulation modelling in population health and health care delivery: critical appraisal questions

Derived from Weightman AL, Barker JM and Lancaster J (2000). *Health Evidence Bulletins Wales Project Methodology 3*. Cardiff: Department of Information Services, UWCM, 2000.

Paper details: Authors:

Title:

Source:

(A) Screening questions

	Yes	Cannot tell	No
General:	Continue:		Discard:
Is the paper relevant to the objectives of the systematic review of simulation modelling?			
Is the choice of a modelling approach appropriate?			
Does it describe a simulation model?			
1–3. Does the paper address a clearly focused issue?			
in terms of:			
• the health care setting			
• the population studied			
• the interventions/scenarios modelled			
• the outcomes considered			
Are the aims and objectives clearly stated and focused?			
Is there a critical review of relevant literature to justify undertaking the study?			
• Has a sufficiently complete search of the relevant literature been carried out?			
• Have the authors reflected the current state of knowledge according to an unbiased review of the literature?			

(B) Assessment of validity

	Yes	Cannot tell	No
4. Is there a description of the model?			
• Flow chart			
• Text description			
5. Are the parameters of the model specified and appropriate?			
6. Identification and quality of data sources?			
Check to see whether:			
• parameter data are observed or modelled and the data sources are stated and described			
• the statistical distributions of the parameter data are described and justified			
• if using local data did the investigators use an appropriate sample and method to obtain the data?			
7. Are the assumptions in the model explicit and reasonable?			
• Are the assumptions described and justified?			
• Is there evidence to support the assumptions?			
8. Is there an assessment of the validity of the model?			
• Is the duration of the simulation explicit and reasonable?			
• Have sensitivity analyses been performed?			
• Has the model output been compared with observed data?			
• Has the specification of the model and parameters been validated?			

(C) What were the results and conclusions?

	Yes	Cannot tell	No
9. What are the primary findings?			
Consider whether the results:			
• address the research question			
• relate to the aims, objectives and chosen outcome measures			
• present estimates of precision			
10. Interpretation of the results?			
• Check sufficient runs have been performed for adequate precision			
• Are the explanations for the results plausible and coherent?			
• Were the results interpreted and discussed in context?			
• Are the results of the study compared with those from other studies?			
11. Can the results only be applied to the study situation?			
Consider differences between the study and other populations (e.g. cultural, geographical, ethical) which could affect the relevance of the study.			
12. Does the paper state that the delivery of health care or health care policy altered as a result?			

from a previously published format.[8] This was designed to identify the strengths and weaknesses of each paper. Part A of the appraisal proforma distinguished between population health and health care delivery as the main topic of the paper and requires that the research question is focused in terms of the setting and/or population under study, the interventions under assessment and the outcomes to be considered. A review of relevant literature is required to justify undertaking the study and to demonstrate that the study will add to the body of knowledge and resolve uncertainty where uncertainty exists. If this is achieved then the study should have clearly stated and focused aims and objectives.

Part B assessed the validity of the model. The reviewer must be satisfied that the modelling approach used was the correct study design to answer the research question and that the main parameters are specified, together with justification for the statistical distributions of the parameter data. The validity and reliability of the data sources should be explicitly discussed, and the model assumptions are discussed, reasonable and supported, where possible, by evidence. Sufficient runs of the model to estimate the range of model outputs with appropriate precision must be demonstrated.

Part C assesses the overall results and study conclusions. Generalizability of the findings and assessment of whether all potentially important outcomes were considered is required. Finally, the impact of the study on the delivery of health care or policy development was considered, if stated in the paper.

To achieve a 'qualitative overview' of studies[9] in which the highest quality papers were given the greatest weight, each paper was judged against 10 criteria (Table 3) and awarded 0, 1 or 2 (poor to good) for its quality in each criterion. The total of the scores assigned each paper to one of four categories, A, B, C or D.

Data extraction

We extracted the following data: author, title, ID, date, topic and setting, research question, type of model or software package, main findings, and comments on quality, country, journal type, grade awarded, validation, generalizability, usefulness and evidence of implementation.

Results

Overview

The flow chart of the review is shown in the Figure. We assessed 2226 abstracts from which 990 full text papers were reviewed. A total of 182 papers met the inclusion criteria and were appraised, including 17 papers that were identified from personal contact. Four databases (Medline, INSPEC, Embase, HealthSTAR) held 90 per cent of the papers appraised. We identified five broad topic areas: hospital scheduling and organization, infection and communicable disease, costs of illness and economic evaluation, screening and miscellaneous. The number of published papers increased over time, with two-thirds published after 1990. Sixty per cent of papers were set in the United States, but the number and proportion of UK or European papers also increased after 1990. The quality of papers was variable but overall nearly two-thirds of papers were rated in the two higher quality categories. There was some suggestion of improvement in quality over time, moving from around 60 per cent higher quality papers published before 1990 to 70 per cent in more recent papers. Overall UK papers scored higher than US papers, many of which were published in the pre-1980s. We found an increasing trend for publication in clinical or health services research and a decreasing trend in non-medical operational research journals and conference proceedings. We did not find any difference in quality of papers between clinical or health services research journals and non-medical journals, but the quality standards in full text conference proceedings was lower. Table 4 summarizes data on quality assessment, country of origin, year of publication and type of journal for the 182 included papers.

Hospital scheduling and organization

This section included 94 (52 per cent) of the 182 papers in the review. Seventy-three (78 per cent) of these studies originated in the United States, with a higher proportion of older papers in this topic area compared with the other topic areas (Table 4). Patient scheduling and admissions policies were popular topics for modelling, covering outpatient clinics,[10] a walk-in clinic[11] and operating room scheduling.[12,13] These models investigated the trade-off between patient waiting times and staff utilization to assess whether improvements can be made, either in waiting times or staff utilization with differing booking systems.

A number of systems within health care delivery have been modelled, including outcome measures such as overall productivity, patient waiting times, staff utilization and costs. Models include medical resident staffing schedules,[14] out-patient pharmacies,[15,16] a university infirmary,[17] and utilization of surgical staff and facilities during an urban terrorist bomb incident.[18]

Modelling of bed requirements and the key efficiency trade-offs between bed availability, utilization and waiting lists allows many alternative configurations to be assessed before expensive schemes are implemented. Our review included papers on emergency bed planning,[19] in-patient bed planning[20] and planning for a coronary care unit.[21]

The location of ambulance services depends on a number of interrelated factors such as population density, accident density, road transport systems, and the distribution of hospital and trauma facilities. These data are available from existing ambulance records and lives saved can be estimated under various possible scenarios.[22]

In general, many papers in this section reported potential improvements to the organization and delivery of care, and identified trade-offs that are critical for decision-making. Despite this, very few papers reported that models had been implemented.

Table 3 Quality criteria summary sheet

Quality criterion	Score (0, poor to 2, good)
1 Clarity of aims and objectives	
2 Intervention(s)/scenario(s) modelled adequately defined	
3 Outcome measures defined and appropriate	
4 Model adequately described	
5 Parameters specified	
6 Quality of data sources	
7 Explicitness and appropriateness of assumptions	
8 Evaluation/validation criteria defined	
9 Presentation of appropriate results with estimation of precision	
10 Results interpreted and discussed in context	
Overall (category D <11, C 11–13, B 14–16, A>16)	

Quality criteria summary sheet: scoring criteria

Quality criterion	Score (0, poor to 2, good)
1 Clarity of aims and objectives	Stated 1 Stated and focused 2
2 Intervention(s)/changes under test adequately defined	Vague 1 Concise and supported by literature review 2
3 Outcome measures defined and appropriate	Vague 1 Concise 2
4 Model adequately described	Flow chart 1 Good general text description 1
5 Parameters specified	Incomplete 1 Complete, reproducible 2
6 Quality of data sources	Yes to some of Q6 in appraisal 1 Yes to all of Q6 in appraisal 2
7 Explicitness and appropriateness of assumptions	Explicit or appropriate 1 Explicit and appropriate 2
8 Validation of model	States criteria for validation 1 Evidence of validation undertaken 1
9 Presentation of appropriate results with estimation of precision	Results flow from objectives 1 Precision of estimates stated 1
10 Results interpreted and discussed in context	Vague 1 Concise 2

Infection and communicable disease

Seven papers modelled infection and communicable disease, covering the prevention and control of HIV/AIDS and sexually transmitted disease, where individual human characteristics and behaviour are important determinants of spread.[23,24] Simulation models have addressed the complex issues of ensuring the effective use of immunization, either by improving the understanding of the biases inherent in field trials[25] or by exploring different strategies to deliver vaccine to the population,[26] and one aspect of preventing nosocomial infections.[27]

Costs and economic evaluation

Excluding screening papers (see below), costs were included in 17 papers. The majority of these were published since 1995, indicating the growing concern over cost containment in health services. Many papers took a straightforward approach to the costing and used only direct costs. In some cases, clinical or operational outcomes were modelled and costs were applied to the results; in others the costs were integrated into the model. Very few of these papers are generalizable to other settings. Alternative clinical interventions were analysed in five papers[28–32] and McHugh[33] estimated the costs of alternative staffing arrangements for nurses. These studies used well-specified models that included costs alongside outcomes, which allowed the modelling of sub-populations to identify optimal strategies.

Screening

We appraised 44 papers that modelled screening. This topic area has also increased in popularity over the years, with 37 of the 44 papers published since 1990. Papers in this section were generally of high quality.

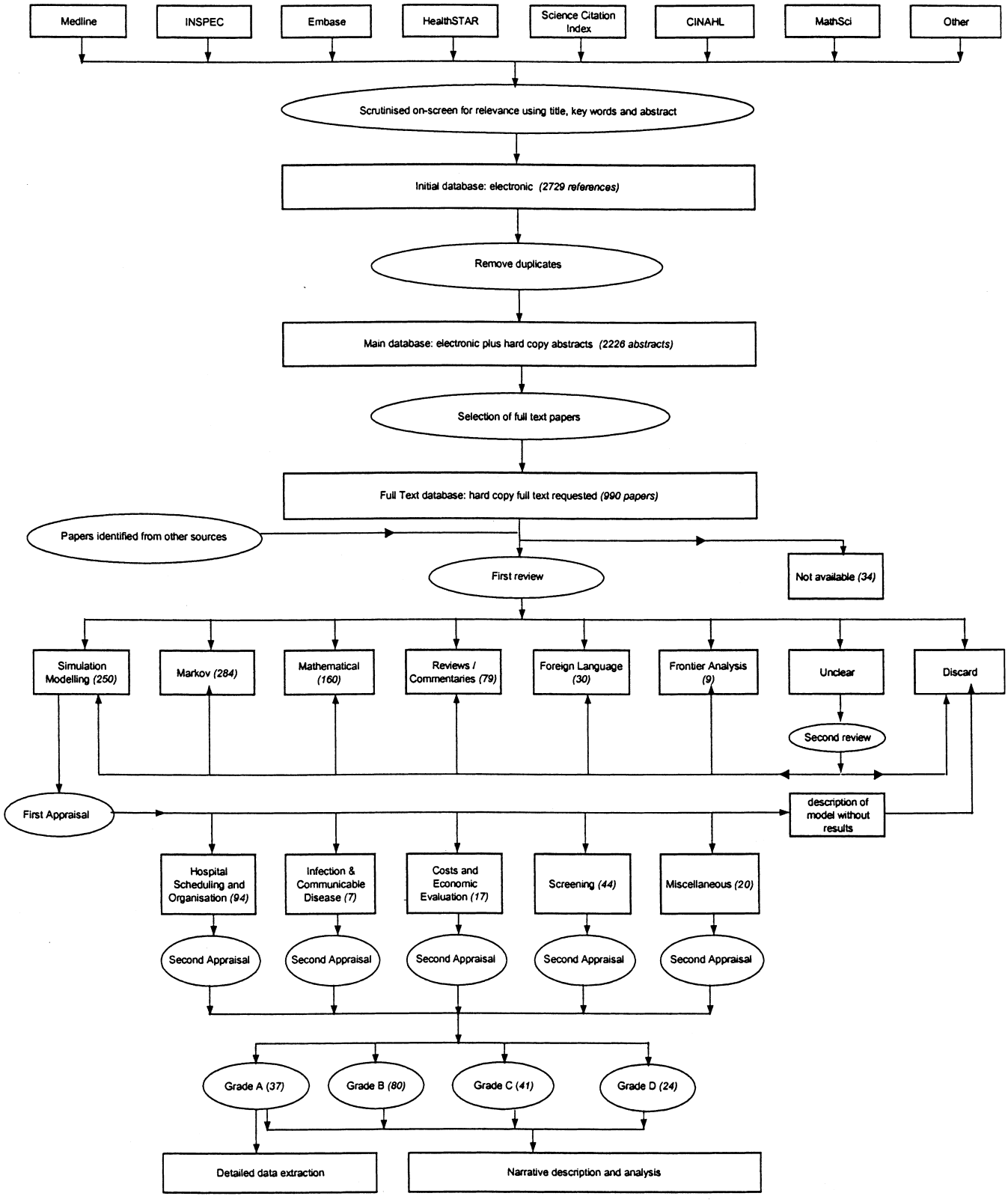

Figure Flow chart of review.

Table 4 Summary data on papers appraised

		Grade				Country			Year of publication					Journal type		
	Total	A	B	C	D	UK	USA	Other	pre-1980	1980–1984	1985–1989	1990–1994	1995–1999	Medical	Non-medical	Conference proceedings
Hospital scheduling and organization																
Number	94	16	36	26	16	8	73	13	26	11	12	19	26	34	30	30
Percentage		17.0	38.3	27.7	17.0	8.5	77.7	13.8	27.7	11.7	12.8	20.2	27.7	36.2	31.9	31.9
Infection and communicable disease																
Number	7	2	2	2	1	1	2	4	0	1	0	2	4	4	1	2
Percentage		28.6	28.6	28.6	14.3	14.3	28.6	57.1		14.3		28.6	57.1	57.1	14.3	28.6
Costs and economic evaluation																
Number	17	5	7	2	3	2	9	6	0	1	0	1	15	16	0	1
Percentage		29.4	41.2	11.8	17.6	11.8	52.9	35.3		5.9		5.9	88.2	94.1		5.9
Screening																
Number	44	11	26	7	0	4	14	26	0	2	5	17	20	39	5	0
Percentage		25.0	59.1	15.9		9.1	31.8	59.1		4.5	11.4	38.6	45.5	88.6	11.4	
Miscellaneous																
Number	20	3	9	5	3	6	11	3	1	0	3	4	12	15	4	1
Percentage		15.0	45.0	25.0	15.0	30.0	55.0	15.0	5.0		15.0	20.0	60.0	75.0	20.0	5.0
All papers																
Number	182	37	80	42	23	21	109	52	27	15	20	43	77	109	39	34
Percentage		20.3	44.0	23.1	12.6	11.5	59.9	28.6	14.8	8.2	11.0	23.6	42.3	59.9	21.4	18.7

Screening lends itself well to simulation modelling. Issues such as frequency, the number of tests performed, and the threshold for investigation can be investigated in a setting that would not be possible under experimentation. In most screening programmes, more testing is intuitively associated with increased effectiveness. However, this is likely to be at a greater cost, so for policy development, an economic evaluation needs to be included. Of the screening papers appraised, 27 included a cost element.

Simulation modelling has been most widely applied to cancer screening. The MISCAN (MIcrosimulation SCreening ANalysis) model, which uses Monte Carlo microsimulation of a large number of life histories according to the epidemiology of the disease in question, has been widely used over many years in modelling cervical and breast cancer screening. Applications in cervical screening include cost-effectiveness studies of screening policies,[34–36] and comparison with other screening tests, such as Human Papilloma Virus screening.[37] Breast cancer screening models have compared alternative screening programmes[38,39] and quality of life studies[40,41] in the Netherlands, predicted the age range of screening that remains cost-effective[42] and the screen interval in older age groups.[43] MISCAN has also been used extensively to model various breast screening issues in a number of other countries, including the relative cost-effectiveness of breast screening in the United Kingdom, Spain and France in comparison with the Netherlands,[44] the cost-effectiveness of breast screening in Australia,[45] Germany,[46,47] Italy,[48] Spain[49,50] and using a similar model 'MICROLIFE', partly populated with MISCAN parameter data, in New Zealand.[51] MISCAN has recently been applied to the UK population[52] and has shown how the model can be used to predict the reduction in mortality expected from screening and the proportion of total breast cancer mortality that is attributable to screening. A more recent application of MISCAN has been to compare the effectiveness of two colorectal screening strategies (faecal occult blood versus flexible sigmoidoscopy), measured by years of life saved.[53] Other cancer screening models reviewed included two papers modelling the effect of screening for ovarian cancer[54,55] and an evaluation of cervical screening policies in England and Wales.[56]

Modelling has also been used extensively in screening for diabetic retinopathy. Javitt[57] developed the PROPHET (PROspective Population Health Event Tabulation) model to simulate the effect and estimate the cost of five screening and treatment strategies for diabetic retinopathy in type I diabetes. This work has been followed up in further published papers[58–63] that refine the model and extend its application to type II diabetes and retinopathy of prematurity.

Miscellaneous

The papers in this section covered a range of disparate topics. For example, Bronnum-Hansen[64] used discrete event simulation to validate the epidemiological model PREVENT. Warner *et al.*[65] modelled the effect of a smoking cessation policy, and Zenios *et al.*[66] used a Monte Carlo simulation to investigate the equity and efficiency implications of alternative strategies that allocate kidneys for transplantation.

Discussion

In this systematic review we aimed to assess the extent and quality of simulation modelling in health care and population health, rather than report on the content and results of the models appraised. We found that simulation modelling was a popular tool applied to the problem of hospital scheduling (individual units and whole systems) before 1980, particularly in the United States. Following refinement of modelling techniques in the 1990s, the number of papers published that describe modelling in health care has grown substantially, extending to a wider range of settings and problems. Hospital scheduling, which remains the most popular, and screening are the principal areas of interest, because here the method is theoretically well suited to answer the research questions. In addition, there is clearly interest in simulation modelling of communicable disease problems and incorporating costs into models to provide useful comparison of alternative strategies.

Assessment of the quality of simulation modelling papers was a fundamental part of the review. Critical appraisal of epidemiological studies to assess the validity and generalizability of published papers is widely taught, at both undergraduate and post-graduate level, and assessment pro formas to systematize this process with agreed criteria are widely used and available.[8,67] However, no criteria are available to aid the critical appraisal and judgement of quality of simulation modelling papers. In view of the increasing numbers of these papers now published and the implications for policy development of implementing model results that may not be robust, we suggest that criteria for a system of critical appraisal and quality assessment are essential.

The method of critical appraisal we used was adapted from a previously developed peer-reviewed pro forma.[8] This was assessed and refined during the appraisal process and we found that it provided a robust, analytical approach to evaluating the studies. Of note was the wide range in the quality of published papers. Any published modelling paper should as a minimum clearly state the aims and objectives and provide some information on model specification, parameter data, assumptions, validation and results. Generally, the difference between 'A' and 'B' grade papers was that the 'A' grade papers provided more depth of information and detail on model validation, and presented the precision of the results and a balanced interpretation and discussion of the main findings. Although one-third of papers were in the lower quality categories there does seem to be a general trend of increasing quality over time, which may reflect the decreasing numbers of conference proceedings and the increasing number of papers published in peer-reviewed journals in the 1990s at the time when evidence-based practice

and critical appraisal of the quality of papers was starting to be accepted in health care. Generally the papers reporting screening model results were of higher quality than hospital scheduling papers, reflecting the more recent and focused interest in cancer screening models. We were unable to draw conclusions about the relative quality of papers in the other topic areas because of the small numbers in the communicable disease, economic evaluation and miscellaneous categories.

It should be noted that we could only judge the information given in the published papers, not the actual model or conduct of the study. Because of pressure on journal space and the complexities of many modelling studies it is likely that some papers did not do justice to the modelling study undertaken. The use of criteria for assessing the quality of a paper should also guide researchers to the minimum content of a journal paper for it to be assessed as a quality paper.

The papers included in the review were published in a wide range of journals; less than a third of the studies were found in 18 journals. This suggests there is no 'natural home' for papers reporting on modelling in health care and may partly account for the lack of uptake of the findings of health care modelling by clinicians, health service managers and policy developers. Our review concentrated on practical applications of modelling (rather than, for example, methodological descriptions) and we found that place of publication is governed more by the subject area than the method employed. The journal of publication may also reflect the interests of the authors, as many of the older studies were conducted by operational researchers and published in non-medical journals or presented at conferences. The more recent interest in simulation modelling includes clinicians with an interest in modelling and consequently the proportion of modelling papers published in clinical or health services research journals has grown steadily over the years.

Devising the search strategy for the review was not straightforward. Our experience of electronic searching leads us to conclude that the large number of potentially relevant MeSH headings in the current Medline and HealthSTAR thesauri are confusing and may also be to the indexers. Our search strategy was a trade-off of higher specificity for lower sensitivity. The only search term that would have identified the 17 papers identified from personal contact was 'models, theoretical', which produced 68 000 references. Therefore few modelling search strategies are likely to approach 100 per cent sensitivity (unless specificity is very low) and future systematic reviews in the field of modelling should include reference list follow-up and contact with researchers in the field.

Despite the increasing numbers of quality papers published in medical or health services research journals we were unable to reach any conclusions on the value of modelling in health care because the evidence of implementation was so scant. Evidence is needed on the outcomes of model implementation, including, for example, assessment of the change in parameters of the problem being modelled, an assessment of the magnitude of meaningful change, and the further insights into the system offered by an evaluated implementation. Our inability to assess the value of modelling in this review causes some concern. One likely explanation for this omission in published papers is the timescale required to publish papers. It is likely that many modelling studies are published before validation is complete and before implementation has been carried out (and assessed). We could, perhaps, conclude that the volume of funded work in the field of health care modelling bears testament to its value. Nevertheless, further research to assess model implementation is required to assess the value of modelling.

Conclusions

Simulation modelling is a powerful tool that has been applied to a wide range of topic areas and research questions in population health and health care. The number of papers published on modelling in health care has grown substantially over recent years. The quality of papers is variable, but improving over time. The potential of simulation modelling to inform evidence-based policy development for the provision of health care is clear, but information on the outcomes of model implementation and hence the value of modelling requires further research. We will report on the content and results of the models appraised in future publications.

Acknowledgements

We should like to thank the following researchers, who provided useful comments and drafts of unpublished material: Adrian Bagust, Pelham Barton, Timothy Coats, Elisabeth Fenwick, Laura Ginnelly, Paul Harper, Alyson Kettles, Peter Millard, Gerrit van Oortmarssen, Jackie Riley, Mark Sculpher, Chris Sherlaw-Johnson and David Spiegelhalter. This study was funded by a grant from the Wales Office of Research & Development.

References

1 Royston G. Commentary: trials versus models in appraising screening programmes. *Br Med J* 1999; **318:** 360–361.

2 Taylor K, Lane D. Simulation applied to health services: opportunities for applying the system dynamics approach. *J Hlth Serv Res Policy* 1998; **3:** 226–232.

3 Temple JMF, Fone DL. Implementing national guidance on prevention of coronary heart disease: clinical governance and computer simulation modelling. *Br J Clin Gov* 2002; **7**(1): 27–33.

4 Davies R. An assessment of models of a health system. *J Op Res Soc* 1985; **36**(8): 679–686.

5 Lehaney B, Hlupic V. Simulation modelling for resource allocation and planning in the health sector. *J R Soc Hlth* 1995; **115**(6): 382–385.

6 Jun JB, Jacobson SH, Swisher JR. Applications of discrete-event simulation in health care clinics: a survey. *J Op Res Soc* 1999; **50**(2): 109–123.

7 National Research Register 1999, Issue 4. [On-line.] Available at http://www.doh.gov.uk/research/nrr.htm (6 October 2003, date last accessed).

8 Weightman AL, Barker J, Lancaster J. Health Evidence Bulletins – Wales: a methodical approach to identifying the evidence. Project methodology 4. Cardiff: University of Wales College of Medicine, Department of Information Services, 2001. [On-line.] Available at http://hebw.uwcm.ac.uk/projectmethod/title.htm (6 October 2003, date last accessed).

9 Mulrow CD, Oxman AD, eds. *The Cochrane Collaboration Handbook Version 4.0* (Updated 1 March 1997). In: The Cochrane Library. Oxford: The Cochrane Collaboration, 1996.

10 Stafford EF Jr, Aggarwal SC. Managerial analysis and decision-making in outpatient health clinics. *J Op Res Soc* 1979; **30:** 905–915.

11 Reilly TA, Marathe VP, Fries BE. A delay-scheduling model for patients using a walk-in clinic. *J Med Sys* 1978; **2:** 303–313.

12 Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA. An operating room scheduling strategy to maximize the use of operating room block time: computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. *Anaes Analg* 1999; **89:** 7–20.

13 Fitzpatrick KE, Baker JR, Dave DS. An application of computer simulation to improve scheduling of hospital operating room facilities in the United States. *Int J Comp Apps Tech* 1993; **6:** 205–224.

14 Dittus RS, Klein RW, DeBrota DJ, Dame MA, Fitzgerald JF. Medical resident work schedules: design and evaluation by simulation modelling. *Man Sci* 1996; **42:** 891–906.

15 Hendershott TH. Evaluating process change proposals in an outpatient pharmacy using simulation. *Proc Hlth Inform Manage Syst Soc* 1996; 165–174.

16 Vemuri S. Simulated analysis of patient waiting time in an outpatient pharmacy. *Am J Hosp Pharm* 1984; **41:** 1127–1130.

17 Taylor BW, Keown AJ. A network analysis of an inpatient/outpatient department. *J Op Res Soc* 1980; **31:** 169–179.

18 Hirshberg A, Stein M, Walden R. Surgical resource utilization in urban terrorist bombing: a computer simulation. *J Trauma-Injury Infect Crit Care* 1999; **47:** 545–550.

19 Altinel IK, Ulas E. Simulation modeling for emergency bed requirement planning. *Ann Op Res* 1996; **67:** 183–210.

20 Vassilacopoulos G. A simulation model for bed allocation to hospital inpatient departments. *Simulation* 1985; **45:** 233–241.

21 Cohen MA, Hershey JC, Weiss EN. Analysis of capacity decisions for progressive patient care hospital facilities. *Hlth Serv Res* 1980; **15:** 145–160.

22 Goldberg J, Dietrich R, Chen JM, *et al*. A simulation model for evaluating a set of emergency vehicle base locations: development, validation, and usage. *Soc Econ Plan Sci* 1990; **24:** 125–141.

23 Auvert B, Moore M, Bertrand WE, *et al*. Dynamics of HIV infection and AIDS in central African cities. *Int J Epidemiol* 1990; **19:** 417–428.

24 Kretzschmar M, van Duynhoven YT, Severijnen AJ. Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *Am J Epidemiol* 1996; **144:** 306–317.

25 Nourjah P, Frerichs RR. Minimum attack rate for measuring measles vaccine efficacy. *Int J Epidemiol* 1995; **24:** 834–841.

26 De Boo ThM, Van Druten JAM, Plantinga AD. Predicting the dynamic effects of rubella vaccination programmes. *Statist Med* 1987; **6:** 843–851.

27 Sebille V, Valleron AJ. A computer simulation model for the spread of nosocomial infections caused by multidrug-resistant pathogens. *Comp Biomed Res* 1997; **30:** 307–322.

28 Eastman RC, Javitt JC, Herman WH, *et al*. Model of complications of NIDDM. I. Model construction and assumptions. *Diab Care* 1997; **20:** 725–734.

29 Eastman RC, Javitt JC, Herman WH, *et al*. Model of complications of NIDDM. II. Analysis of the health benefits and cost-effectiveness of treating NIDDM with the goal of normoglycemia. *Diab Care* 1997; **20:** 735–744.

30 Nathan DM. Lifetime benefits and costs of intensive therapy as practiced in the diabetes control and complications trial. *JAMA* 1996; **276:** 1409–1415.

31 Ramsdell JW, Grossman JA, Stephens JM, Botteman MF, Arocho R. A short-term cost-of-treatment model for type 2 diabetes: comparison of glipizide gastrointestinal therapeutic system, metformin, and acarbose. *Am J Man Care* 1999; **5:** 1007–1024.

32 Roberts SD, Maxwell DR, Gross TL. Cost-effective care of end-stage renal disease: a billion dollar question. *Ann Int Med* 1980; **92:** 243–248.

33 McHugh ML. Cost-effectiveness of clustered unit vs. unclustered nurse floating. *Nurse Econ* 1997; **15:** 294–300.

34 van Ballegooijen M, Koopmanschap MA, van Oortmarssen GJ, *et al*. Diagnostic and treatment procedures induced by cervical cancer screening. *Eur J Cancer* 1990; **26:** 941–945.

35 Koopmanschap MA, van Oortmarssen GJ, van Agt HM, *et al*. Cervical-cancer screening: attendance and cost-effectiveness. *Int J Cancer* 1990; **45:** 410–415.

36 van Ballegooijen M, Habbema JD, van Oortmarssen GJ, *et al*. Preventive pap-smears: balancing costs, risks and benefits. *Br J Cancer* 1992; **65:** 930–933.

37 van Ballegooijen M, van den Akker-van Marle ME, Warmerdam PG, *et al*. Present evidence on the value of HPV testing for cervical cancer screening: a model-based exploration of the cost-effectiveness. *Br J Cancer* 1997; **76:** 651–657.

38 van der Maas PJ, de Koning HJ, van Ineveld BM, *et al*. The cost-effectiveness of breast cancer screening. *Int J Canc* 1989; **43:** 1055–1060.

39 de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials. *J Nat Cancer Inst* 1995; **87:** 1217–1223.

40 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; **49:** 531–537.

41 de Haes JC, de Koning HJ, van Oortmarssen GJ, *et al*. The impact of a breast cancer screening programme on quality-adjusted-life-years. *Int J Cancer* 1991; **49:** 538–544.

42 Boer R, de Koning HJ, van Oortmarssen GJ, van der Maas PJ. In search of the best upper age limit for breast cancer screening. *Eur J Cancer* 1995; **31A:** 2040–2043.

43 Boer R, de Koning HJ, van der Maas PJ. A longer breast carcinoma screening interval for women aged older than 65 years. *Cancer* 1999; **86:** 1506–1510.

44 van Ineveld BM, van Oortmarssen GJ, de Koning HJ, Boer R, van der Maas PJ. How cost-effective is breast cancer screening in different EC countries? *Eur J Cancer* 1993; **29A:** 1663–1668.

45 Carter DM, Glasziou P, van Oortmarssen GJ, *et al*. Cost-effectiveness of mammographic screening in Australia. *Aust NZ J Publ Hlth* 1993; **17:** 42–50.

46 Beemsterboer PM, de Koning HJ, Warmerdam PG, *et al*. Prediction of the effects and costs of breast cancer screening in Germany. *Int J Cancer* 1994; **58:** 623–628.

47 Warmerdam PG, de Koning HJ, Boer R, *et al*. Quantitative estimates of the impact of sensitivity and specificity in mammographic screening in Germany. *J Epidemiol Commun Hlth* 1997; **51:** 180–186.

48 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:** 348–353.

49 van den Akker-van Marle ME, Reep-van den Bergh CM, Boer R, *et al*. Breast cancer screening in Navarra: interpretation of a high detection rate at the first screening round and a low rate at the second round. *Int J Cancer* 1997; **73:** 464–469.

50 Beemsterboer PMM, Warmerdam PG, Boer R, *et al*. Screening for breast cancer in Catalonia: which policy is to be preferred? *Eur J Publ Hlth* 1998; **8:** 241–246.

51 Szeto KL, Devlin NJ. The cost-effectiveness of mammography screening: evidence from a microsimulation model for New Zealand. *Hlth Pol* 1996; **38:** 101–115.

52 van den Akker-van Marle ME, de Koning HJ, Boer R, van der Maas PJ. Reduction in breast cancer mortality due to the introduction of mass screening in the Netherlands; comparison with the United Kingdom. *J Med Screen* 1999; **6:** 30–34.

53 Loeve F, Boer R, van Oortmarssen GJ, van Ballegooijen M, Habbema JD. The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening. *Comp Biomed Res* 1999; **32:** 13–33.

54 Skates SJ, Singer DE. Quantifying the potential benefit of CA 125 screening for ovarian cancer. *J Clin Epidemiol* 1991; **44:** 365–380.

55 Urban N, Drescher C, Etzioni R, Colby C. Use of a stochastic simulation model to identify an efficient protocol for ovarian cancer screening. *Cont Clin Trials* 1997; **18:** 251–270.

56 Parkin DM, Moss SM. An evaluation of screening policies for cervical cancer in England and Wales using a computer simulation model. *J Epidemiol Commun Hlth* 1986; **40:** 143–153.

57 Javitt JC, Canner JK, Sommer A. Cost effectiveness of current approaches to the control of retinopathy in type I diabetics. *Ophthalm* 1989; **96:** 255–264.

58 Fendrick AM, McCort JT, Chernew ME, *et al*. Immediate eradication of *Helicobacter pylori* in patients with previously documented peptic ulcer disease: clinical and economic effects. *Am J Gastroent* 1997; **92:** 2017–2024.

59 Javitt JC, Canner JK, Frank RG, Steinwachs DM, Sommer A. Detecting and treating retinopathy in patients with Type I diabetes mellitus. *Ophthalmologica* 1990; **97:** 483–495.

60 Javitt JC, Aiello LP, Bassi LJ, Chiang YP, Canner JK. Detecting and treating retinopathy in patients with type I diabetes mellitus. Savings associated with improved implementation of current guidelines. American Academy of Ophthalmology. *Ophthalmologica* 1991; **98:** 1565–1573.

61 Javitt JC, Dei CR, Chiang YP. Cost-effectiveness of screening and cryotherapy for threshold retinopathy of prematurity. *Pediatrics* 1993; **91:** 859–866.

62 Javitt JC, Aiello LP, Chiang Y, *et al*. Preventive eye care in people with diabetes is cost-saving to the federal government. Implications for health-care reform. *Diab Care* 1994; **17:** 909–917.

63 Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Ann Int Med* 1996; **124:** 164–169.

64 Bronnum-Hansen H. How good is the Prevent model for estimating the health benefits of prevention? *J Epidemiol Commun Hlth* 1999; **53:** 300–305.

65 Warner KE, Smith RJ, Smith DG, Fries BE. Health and economic implications of a work-site smoking-cessation program: a simulation analysis. *J Occup Environ Med* 1996; **38:** 981–992.

66 Zenios SA, Wein LM, Chertow GM. Evidence-based organ allocation. *Am J Med* 1999; **107:** 52–61.

67 Public Health Resource Unit, Oxford Critical Appraisal Skills Programme. Learning resources appraisal tools [On-line.] Available at http://www.phru.org.uk/~casp/appraisal.htm (6 October 2003, date last accessed).

Accepted on 3 June 2003