Maturitas, 7 (1985) 11–20
Elsevier

MAT 00347

Model building on the basis of Dutch cervical cancer screening data

J.D.F. Habbema, G.J. van Oortmarssen, J.T.N. Lubbe and P.J. van der Maas

Department of Public Health and Social Medicine, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands

(Received 26 December 1984; accepted 8 February 1985)

A mass screening programme for cervical cancer is in progress in three pilot regions in The Netherlands. All women living in these regions aged 35–53 are invited to undergo screening at three-year intervals.

The MISCAN simulation model was developed for the analysis and optimization of screening programmes. In this paper the model-based approach to evaluation is first outlined and then illustrated by analysing data from the first two screening rounds in the pilot regions. This analysis resulted in a rather restricted range of data-compatible assumptions for the mean duration of preclinical disease (14–19 yr) and the frequency of spontaneous regression of preinvasive lesions (45–65%), as well as a rather wide sensitivity range for the Pap smear (50–90%).

These preliminary findings are compared with those of a previous MISCAN analysis of cervical cancer screening in British Columbia. On the basis of an assumed 18-yr duration, 50% regression and 70% sensitivity, a number of screening policies relating to the same age ranges but with different intervals are compared. Both the analysis and the policy comparisons are preliminary, but the findings are nevertheless reasonable and consistent with those of previous studies.

A more complete MISCAN-based analysis of the Dutch screening programme and subsequent optimization of screening policies will be possible when further results become available and a cost-effectiveness analysis procedure has been incorporated into the MISCAN programme.

(Key words: Cervical cancer, Natural history, Test sensitivity, Mass screening, Evaluation, Policy optimization, Simulation model, Cost-effectiveness)

Introduction

Following a number of pilot studies [1], a mass screening programme for cervical cancer was started in 1976 in three pilot regions in The Netherlands, comprising the cities of Nijmegen, Rotterdam and Utrecht and their environs. About 25% of the Dutch female population lives in the regions selected. Using information obtained from the population registry, all women aged 35–53 are invited for screening at 3-yr intervals. There is a registration and evaluation office, BERC (Bureau Evaluatie en

Correspondence to: J.D.F. Habbema.

Registratie Cervixcarcinoom) * in each of the three pilot regions while a national committee, EVAC (Evaluatiecommissie inzake de vroege opsporing van cervixcarcinoom) ** evaluates the overall programme. The report on the first and second screening rounds was issued in 1984 [2]. The programme will end after three screening rounds have been completed. The evaluation is complicated by the fact that many preventive Pap smears are taken outside the screening programme, mainly by general practitioners and gynaecologists, and these are incompletely recorded in the BERC offices.

A number of studies have contributed to the evaluation of cervical cancer screening in The Netherlands, one such being the model-based analysis and optimization of cervical cancer screening using the MISCAN-simulation approach [3,4]. The present paper describes this model-based approach and presents a preliminary analysis of some results from the first two rounds of the Dutch screening programme. It goes on to describe a screening-optimization exercise which compares the effectiveness of a number of different screening policies.

Model-based evaluation: the MISCAN simulation approach

The history of mathematical models for cancer screening goes back to the late sixties. A general description and discussion of model-based evaluation clearly lies outside the scope of the present paper but reference is made to a number of recent reviews [5–7]. The basic need for the development of (computerized) mathematical models arose from the complexity of the evaluation problem owing to the many factors involved. Model-based evaluation appears to offer the only effective means of studying the joint impact of all the factors (and their interrelations) on the results, benefits and risks of screening. Some of the more important factors influencing the results of the Dutch screening programme are:

- *duration* of the preclinical lesions (defined in the present paper as histologically-confirmed severe dysplasia, carcinoma in situ, and invasive cancer that is not yet clinically apparent;
- spontaneous *regression* of preinvasive lesions;
- other aspects of the disease process, e.g. *incidence* of preclinical disease and *survival* following diagnosis of clinical cancer;
- population characteristics, e.g. *age-distribution, mortality table, hysterectomies* and *high-risk groups*;
- screening characteristics, e.g. *sensitivity* of the Pap smear, *attendance* at screening sessions, and in The Netherlands at least the pattern of Pap smears taken outside the programme,
- *survival* after early detection.

These factors may be age dependent or may differ between birth cohorts. It is indeed difficult to imagine how the combined influence of these factors could be disentangled without using comprehensive models.

---

* Bureau for evaluation and registration of cervical cancer.

** Evaluation committee on the early detection of cervical cancer.

A number of models for cervical cancer screening have been described elsewhere [e.g. 8–11]. The MISCAN computer-simulation model was developed in order to overcome certain limitations of the earlier models [4]. MISCAN has previously been tested in model-based evaluations of the cervical cancer screening programme in British Columbia [12] and of the HIP (Health Insurance Plan) breast cancer study [13].

A MISCAN-based evaluation proceeds essentially in two stages, viz. *analysis* and *optimization* [3]. At the analysis stage quantitative assumptions regarding the above-mentioned factors are tested for their capability of explaining the observed results of the screening programme. Then, after specifying the model assumptions, MISCAN simulates the screening programme in such a way that corresponding simulated results are obtained for all observed results. The goodness-of-fit between observed and simulated results is assessed by chi-square and Student's *t*-tests. The final output of the analysis stage ideally consists of a demarcation between assumptions which are compatible with the observed results (henceforth referred to as *data-compatible assumptions)* and assumptions which are not.

The data-compatible assumptions are subsequently used during the optimization stage. On the basis of these assumptions, favourable and unfavourable effects of alternative screening policies for cervical cancer screening are simulated. Indices of benefits and costs or a more complete cost-effectiveness or cost-benefit analysis may be used for comparing the merits of the policies. The most important policy questions concern the age range to be screened and the interval between screenings. But other problems can also be addressed, such as the question of selective or more intensive screening of high-risk groups, or the choice between population registry-based screening and general practitioner-based screening.

The model used: structure and assumptions

In any screening model a central role is played by the natural history of cervical cancer and the way it is influenced by screening. The structure of the disease model is shown in Fig. 1. Many of the important factors are indicated in the structure, for example the arrow from 'preinvasive' to 'no cervical cancer' represents spontaneous regression, the sensitivity governs the result of a Pap smear for women in a preclinical stage, and the mean duration of preclinical disease strongly influences the chances of early detection by a screening programme. A common feature shared by these three factors (regression, sensitivity and duration) is that they cannot be quantified in a straightforward way from available data. The analysis stage therefore aims at finding the data-compatible assumptions for these factors by the indirect means of simulations and tests.

To investigate the optimization possibilities, different screening policies are specified in the model and the MISCAN programme then computes their effects. Before an analysis or optimization exercise can start the rest of the factors must be specified in detail. Data are available on some of these factors, e.g. attendance rates can be obtained from the report on the pilot regions programme and survival data

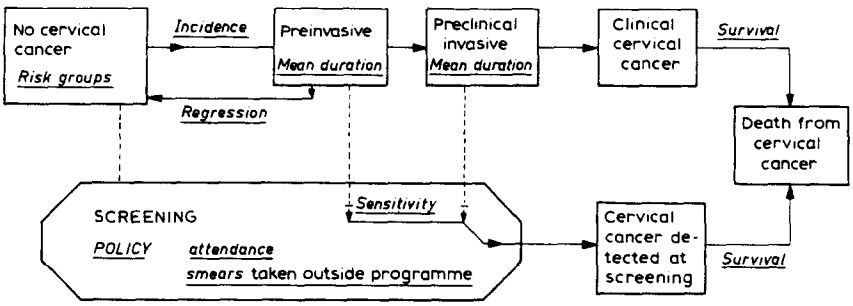


Fig. 1. Structure of the disease process in the MISCAN analysis of cervical cancer screening in the Dutch pilot regions. The stages of the disease process and transitions between the stages are indicated, but two additional stages 'death from other causes' and 'hysterectomy' are omitted in this figure. Some of the factors that greatly influence the effect of screening programmes are indicated in italics.

on cervical cancer cases detected clinically and at screening can be found in the literature. Some other factors, for example the number of smears taken outside the programme, are scarcely documented at all and expert opinion and the limited data available have to be combined to derive assumptions. An account of the factors and the corresponding assumptions made is presented in the Appendix.

Analysis of the results in the pilot regions

Results from the first two rounds of the Dutch screening programme [2,14] were used in the analysis stage of the MISCAN exercise (see Table I). The age structure of the female population in the pilot regions was specified in the model. The screening strategy was also specified: all women aged 35–53 are invited to undergo screening

TABLE I

SCREENING RESULTS FROM THE DUTCH PILOT REGIONS. HISTOLOGICALLY CONFIRMED SEVERE DYSPLASIA OR WORSE, RATE PER 1000 WOMEN SCREENED [2,14]

Stage	All ages	Age at screening						
		35	38	41	44	47	50	53
First round								
Preinvasive	3.4	3.8	3.3	3.8	3.2	3.1	3.4	3.2
Invasive	0.6	0.3	0.4	0.4	0.4	0.7	1.0	1.2
Total detected	4.0							
Stage	All ages	All stages						
Second round								
Preinvasive	1.6	First invitation (age 35)					2.9	
Invasive	0.2	Second invitation (ages 38–53)					1.5	
Total detected	1.8							

at intervals of 3 yr (see Appendix for the other assumptions made, including attendance figures and smears taken outside the screening programme).

The results in Table I were used in deriving data-compatible assumptions for the *mean duration* of the preclinical stage, the percentage spontaneous *regression* of preinvasive lesions and the *sensitivity* of the Pap smear.

Geometrically, all the possible combinations of the assumptions regarding these three factors constitute a three-dimensional space, the data-compatible assumptions being located in a three-dimensional area within that space. A good explanation of the results from the pilot regions is obtained when, for example, a 50% regression rate, a mean duration of 18 yr and 70% sensitivity are assumed.

Reasonable results are also obtained with regression values of between 45 and 65%, a mean duration of between 14 and 19 years, and a sensitivity of between 50 and 90%. However, not all combinations of assumptions from these three ranges are data-compatible. This is illustrated by Fig. 2, which shows a plane for two factors (sensitivity and duration) that emerges when a specific value for the other factor (in this case 50% regression) is chosen. Where 50% regression is assumed, a mean duration of 16 yr is data-compatible only if the sensitivity is as high as 85%. A still lower duration (14 yr) is only data-compatible when combined with a sensitivity of about 80% and a regression rate of 60%.

Fig. 3 shows which data from the screening programme impose a strong restriction on the area of data-compatible assumptions as to sensitivity and mean duration. Clearly, the detection rate of preinvasive cervical cancer at the first screening is most important in this respect. Note that area (a) does not accord with the notion that the detection rate at a first screening should be proportional to the product of sensitivity and mean duration. This is probably due to the fact that a considerable proportion of the women had already had a smear test prior to the first screening round in the pilot regions programme.

The MISCAN approach enables a comparison to be made of the data-compatible assumptions obtained from the analysis of different screening programmes. This is

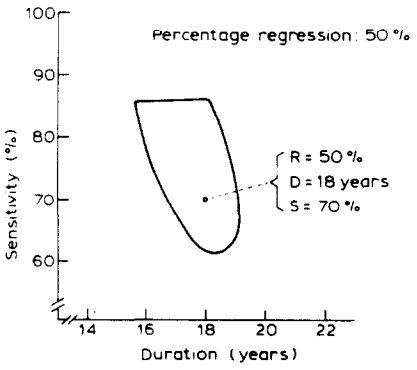
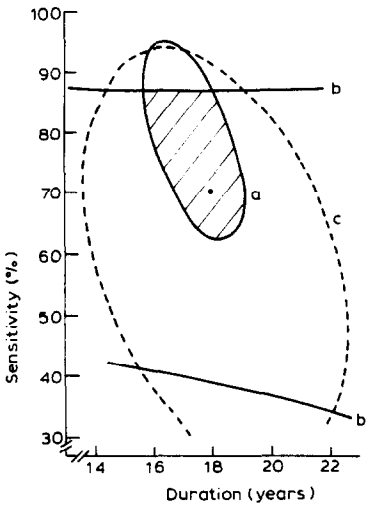


Fig. 2. Assumptions regarding sensitivity of the Pap smear and mean duration of preclinical stages that are consistent with the results of the first and second screening rounds in the Dutch pilot regions. Percentage spontaneous regression is 50%. R = regression; D = duration; S = sensitivity.



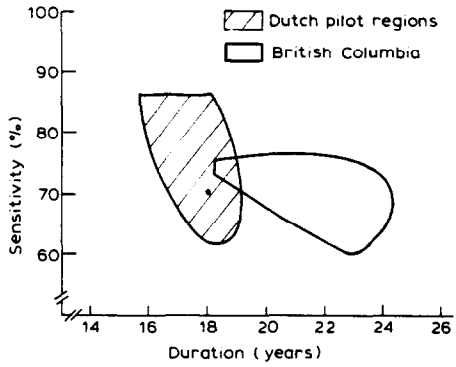


Fig. 3. Consistency of assumptions regarding sensitivity of Pap smear and mean duration of preclinical stages with the results in the Dutch pilot regions: (a) preinvasive stages detected at first screening; (b) invasive stages detected at second screening; (c) all stages detected at second screening, excluding first invitation (age 35).

Fig. 4. Assumptions regarding the sensitivity of the Pap smear and the mean duration of the preclinical stages that are consistent with the results in the Dutch pilot regions and the British Columbia cohort study.

illustrated by Fig. 4, which displays the findings of the Dutch pilot regions analysis and the MISCAN analysis of the British Columbia cohort study [12]. Although such a comparison is stimulating and useful it must be interpreted with great caution, because of important differences between the screening programmes that are not automatically taken into account by the model. For example, the shorter mean duration of the preclinical stages found in the Dutch analysis can be tentatively explained by the broader definition of preclinical disease used in the British Columbia analysis. If moderate dysplasia were to be included in the definition of preclinical disease in the Dutch analysis, the duration would become longer and the two areas of data-compatible assumptions would show more overlap.

Comparison of screening policies

Optimization of screening policies on the basis of the Dutch data is not yet justified, owing to both the incompleteness of the screening data and the preliminary character of the MISCAN analysis. Moreover, a full cost-effectiveness analysis is as yet lacking. We have therefore restricted ourselves to an illustrative example comparing seven screening policies which all address the same age range of 35–55 yr, but differ as regards the interval between screenings (see Table II).

TABLE II

SEVEN SCREENING POLICIES COVERING ROUGHLY THE SAME RANGE (35–55 YEARS) BUT WITH DIFFERENT INTERVALS BETWEEN SCREENINGS

Policy no.	Interval (yr)	No. of screenings	Ages at which screenings are scheduled
1	–	1	35
2	20	2	35 55
3	10	3	35 45 55
4	7	4	35 42 49 56
5	5	5	35 40 45 50 55
6	3	7	35 38 41 44 47 50 53
7	2	11	35 37 39 41 43 45 47 49 51 53 55

The seven policies were simulated for the Dutch female population over the 25-yr period from 1977 to 2002, using the following assumptions: no screening before 1977; no preventative smears taken outside the programme; attendance at screening sessions comparable to the participation rate observed in the Dutch pilot regions; treatment results, mortality from other causes (mortality table) and incidence of hysterectomy taken to be stable throughout the period considered.

The MISCAN simulation of the seven policies was based on the data-compatible assumptions of 50% regression, 18-yr duration and 70% sensitivity (see previous section). The beneficial effects of the seven screening programmes were (crudely) assessed by calculating the number of years of life saved in the Dutch female

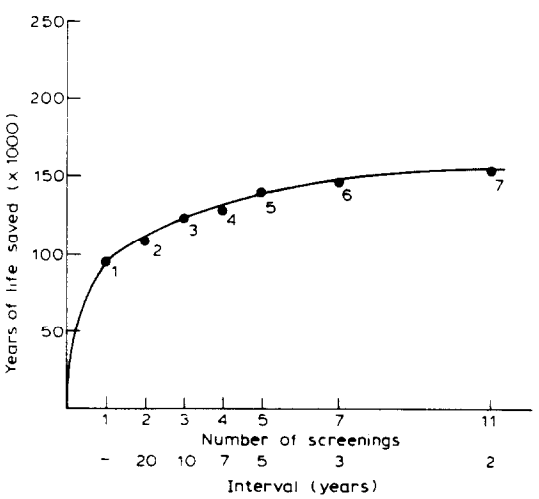


Fig. 5. Total number of years of life saved in the Dutch population with different simulated screening policies. Screening starts in 1977 and ends in 2002. Women in the 35–55 age range are eligible for screening.

population. These years would otherwise have been lost because of cervical cancer mortality. In totalling the years of life no adjustment has as yet been made for time preferences (years in the immediate future are generally preferred to those in the distant future), quality of life (to take account of disability in old age), or straightforward time discounting.

The results of the simulations are shown in Fig. 5. Note the initial sharp rise in the number of years of life saved as the number of screenings increases and the subsequent flattening-off as the intervals become still shorter. This phenomenon has been described earlier [9–12,16,17].

The cost and risks of screening may, as a first approximation, be taken to be proportional to the total number of screenings performed in the population [12]. The ratio of the number of years of life saved to the total number of screenings performed serves as a surrogate cost-effectiveness ratio (see Table III). This ratio declines steadily as the number of screenings increases. The marginal ratio, i.e. the ratio of the additional number of years saved to the additional number of screenings, decreases much more dramatically.

Final remarks

Although the analysis and the policy comparisons that have been made are very preliminary, the results are encouraging. They are in reasonable agreement with those obtained from the analysis of the British Columbia cohort study [12,15], which constitutes the most fully analysed cervical cancer screening programme to date.

A comprehensive, model-based analysis of the Dutch screening programme is scheduled but will require a number of intermediate studies to be completed beforehand. The Dutch screening data will be compared with those from other

TABLE III

YEARS OF LIFE SAVED AND NUMBER OF SCREENINGS PERFORMED IN THE TOTAL DUTCH POPULATION WITH SEVEN DIFFERENT SCREENING POLICIES (SEE TABLE II). (THE RATIO IN COLUMN (6) INDICATES THE NUMBER OF YEARS OF LIFE SAVED PER 1000 SCREENINGS)

Policy			Years of life saved ×1000	No. of screenings made ×1,000,000	Ratio of (4) to (5)
No.	Interval	No. of invitations			
(1)	(2)	(3)	(4)	(5)	(6)
1	–	1	89	2.3	39
2	20	2	116	3.6	32
3	10	3	145	5.2	28
4	7	4	155	5.9	26
5	5	5	179	8.1	22
6	3	7	192	10.7	18
7	2	11	207	14.7	14

countries and any major discrepancies will need to be explained. A cost-effectiveness analysis procedure will also be developed and built into the MISCAN simulation programme. Further results from the screening programme, such as the data from the third round and figures from the cancer registries that are being set up will be taken into account. Optimization of screening will obviously have to be based on the results of this further analysis work.

Acknowledgement

The research reported in this paper was made possible by grant No. 28-935 made available by the Netherlands 'Praeventiefonds' (Prevention Fund).

Appendix

The most important assumptions used in the current model are:
- no risk factors other than age and year of birth are taken into account;
- death from other causes is specified according to the Dutch mortality table for the period 1971–1975;
- hysterectomies for indications other than the treatment of cervical cancer are specified by extrapolation from age-specific hysterectomy rates available for 1965–1982;
- duration of the preclinical stages is governed by a Weibull probability function;
- duration of the preclinical stages is independent of age;
- mean duration of the preinvasive stage is the same for progressive and regressive lesions;
- mean duration of the preclinical invasive stage is 5 yr;
- sensitivity in the preinvasive stage is lower than in the invasive stage (the proportion of cases missed is doubled);
- survival rate in cases of clinically-detected cervical cancer is age-dependent and averages 55%;
- prognosis for cervical cancer detected at screening is a 100% cure rate for preinvasive lesions and a 90% cure rate for invasive lesions;
- attendance figures depend on age and on response to the previous invitation, averaging 70.6% for the first screening and 65% for the second [2,14];
- smears taken outside the programme, by general practitioners, gynaecologists or at special clinics, are assumed to have started in 1967 on a modest scale with a substantial increase after 1975; the number of women assumed to have been screened prior to the first round is in agreement with the proportion (50%) actually reported [2].

N.B.

The incidence of the preinvasive stage is treated as detailed below.

Firstly, a no-regression situation is assumed, the incidence of preinvasive progressive lesions being estimated as to arrive at a good approximation of the cervical

cancer mortality figures. The probability of entering the progressive preinvasive stage before age 75 is taken to be 1.4% on average for all birth cohorts. The relative risks for the birth cohorts in the population are obtained by trend analysis of the cervical cancer mortality rates.

Secondly, when considering different regression probability values, the incidence of the preinvasive stage is corrected in order to maintain a constant rate for progressive lesions. For example, in the case of 50% regression, the (average) cumulated incidence is assumed to be 2.8%. The incidence of regressive preinvasive lesions is in fact concentrated in the younger age groups, mainly those between 20 and 35.

References

1 Collette HJA, Linthorst G, De Waard F. Report of the activities of 'Cyt-U-Universitair' on early diagnosis of cervical cancer, 1970–1973 (in Dutch). Utrecht,1974.

2 EVAC. Report of the evaluation committee on early detection of cervical cancer (EVAC) about the second round of the mass screening program in the pilot regions Nijmegen, Rotterdam and Utrecht (in Dutch). Leidschendam: Ministerie VOMIL, 1984.

3 Habbema JDF, Lubbe JTN, Van der Maas PJ, Van Oortmarssen GJ. A computer simulation approach to the evaluation of mass screening. In: Van Bemmel JH, Ball MJ, Wigertz O, eds. MEDINFO 83, Proceedings of the 4th world conference on medical informatics. Amsterdam: North-Holland Publ. Co, 1983; 1222–1225.

4 Habbema JDF, Van Oortmarssen GJ, Lubbe JTN, Van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comp Prog Biomed (in press).

5 Eddy DM, Shwartz M. Mathematical models in screening. In: Schottenfeld D, Fraumeni JF, eds. Cancer epidemiology and prevention. Philadelphia: WB Saunders, 1982; 1075–1090.

6 Prorok PC. Evaluation of screening programs for the early detection of cancer. In: Cornell RG, ed. Statistical methods for cancer studies. New York: Marcel Dekker, 1984; 267–328.

7 Shwartz M, Plough AL. Models to aid in cancer screening programs. In: Cornell RG, ed. Statistical methods for cancer studies. New York: Marcel Dekker, 1984; 329–416.

8 Coppleson LW, Brown B. Observations on a model of the biology of carcinoma of the cervix: a poor fit between observations and theory. Am J Obstet Gynecol 1975; 122: 127–136.

9 INSERM. Early diagnosis of cervical cancer (in French). Paris: Ministère de la Santé, INSERM, 1975.

10 Eddy DM. Appropriateness of cervical cancer screening. Gynecol Oncol 1981; 12: 168–187.

11 Knox EG. A simulation system for screening procedures. In: McLachlan G, ed. The future and present indicatives. London: Nuffield Provincial Hospitals Trust, 1973; 18–55.

12 Habbema JDF, Lubbe JThN, Van der Maas PJ, Van Oortmarssen GJ. The MISCAN model: a new approach for the evaluation of mass screening. Part I: Mass screening for cervical cancer (in Dutch). Rotterdam: Dept. of Public Health and Social Medicine, Erasmus University, 1982.

13 Habbema JDF, Van Putten DJ, Lubbe JTN, Van der Maas PJ, Van Oortmarssen GJ. The MISCAN model: a new approach for the evaluation of mass screening. Part II: Mass screening for breast cancer (in Dutch). Rotterdam: Dept. of Public Health and Social Medicine, Erasmus University, 1983.

14 EVAC. Mass screening for cervical cancer in the pilot regions Nijmegen, Rotterdam and Utrecht. Report of the first screening round (in Dutch). Leidschendam: Ministerie VOMIL, 1980.

15 Boyes DA, Morrison B, Knox EG, Draper GJ, Miller AB. A cohort study of cervical cancer screening in British Columbia. Clin Invest Med 1982; 5: 1–29.

16 Coppleson LW, Brown B. The prevention of carcinoma of the cervix. Am J Obstet Gynecol 1976; 125: 153–159.

17 Yu Shun-Zhang, Miller AB, Sherman GJ, Optimizing the age, number of tests and test interval for cervical screening in Canada. J Epidemiol Comm Health 1982; 3: 1–10.