

An evaluation of screening policies for cervical cancer in England and Wales using a computer simulation model

D M PARKIN¹ AND S M MOSS²

From the International Agency for Research on Cancer,¹ Lyon, France and the Institute for Cancer Research,² Sutton, Surrey, England

SUMMARY Several screening policies have been recommended for implementation in England and Wales in the last 20 years, although no evidence as to their relative effectiveness or efficiency has been provided. Using a computer simulation model, the outcomes expected from those policies had they been implemented over a 30 year period (1961–90) have been examined. The original policies based on five-yearly testing of women aged over 35 appear to be the most cost-effective, and extension of screening to younger age groups leads to loss of efficiency. Attempts to use non-screening health care contacts in order to take cervical smears (eg, during pregnancy, family planning, at gynaecology clinics) produce few advantages and considerably complicate the establishment of regular testing for the individual. The achievement of higher attendance rates is as important to the outcome of screening as concentration on more intensive or complex policies.

Cervical cytology screening is undoubtedly effective in reducing the incidence of invasive carcinoma and in preventing subsequent death and disability. However, there is disagreement on the most satisfactory screening policy which should be employed, even within a single population such as that of England and Wales. There are many considerations to bear in mind, the most important of which is the natural history of the disease in terms of the incidence rate and sojourn-time of the pre-invasive stages of disease, and their variation with age, which can be used in mathematical models to estimate the apparently optimal spacing of a given number of screening tests.^{1–3} In addition, however, several other factors influence the precise policy adopted. These include epidemiological knowledge of groups of the population at potentially higher risk of abnormalities of the cervix (eg, women attending family planning or gynaecological clinics), practical considerations such as the opportunity to examine women who might not normally attend for screening (eg, during or soon after pregnancy), and often intuitive judgements about the desirability of different outcomes. An example of the latter is the evident concern engendered by relatively few cases of cervical cancer in young women^{4,5} compared with the ineffectiveness of the current screening programmes in preventing disease in the elderly.⁶

In this paper we compare the predicted results of implementing several different screening policies which have been recommended for England and Wales since 1965. This was done using a computer simulation model which reproduces the demographic structure of the female population of England and Wales over a 30-year period (1961–90) and simulates the occurrence of cervical cancer in this population, using different assumptions concerning natural history. This allows, in effect, the comparison of input and outcomes of different interventions in similar populations, in an analogy with a series of (very large) clinical trials.

It is intuitively obvious that the outcome of screening, however measured, increases with the volume of input—usually expressed as the number of tests performed.⁷ It is therefore essential to compare policies in terms of their cost-effectiveness: the results obtained in proportion to the resources deployed.⁸

Methods

COMPUTER SIMULATION MODEL

This has been described in detail elsewhere.⁹ It employs a microsimulation approach, where the life histories of individuals in a random sample of size 100 000 of the female population of England and Wales are modelled during the 30 years 1961–90.

The natural history of cervix cancer is reproduced by specifying sets of transition rates between eight different states (normal, dysplasia, carcinoma in situ, occult/microinvasive carcinoma, clinical cancer, hysterectomy, dead of cancer of cervix, and dead of other causes). The transfer probabilities can be made conditional on age, duration in state, parity, and marital status. For the present study, two sets of transition rates were used, corresponding to the natural histories H1 and H3 previously described.⁹ These natural histories are quite different, yet both reproduce the observed prevalence of preclinical disease and age specific incidence rates observed in England and Wales in the early 1960s, before screening was introduced. In both, dysplasia is a transient condition—75–80% of cases regress—and the median duration is only two years. However, the natural history of carcinoma in situ is very different, as shown in figure 1.

The incidence of carcinoma of the cervix (and its precursors) is changing quite rapidly at present, with

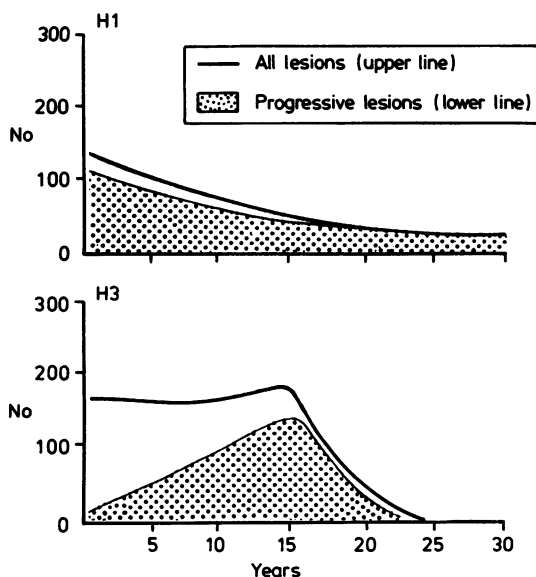


Fig 1 Sojourn time distributions of carcinoma in situ. The curves represent the distribution of the durations of carcinomas in situ for two different sets of assumptions concerning natural history (H1 and H3). The upper curve is that for all lesions, the lower for those which progress to become invasive. The ratio of the areas below the curves (open and shaded) represents the ratio of regressive to progressive lesions.

In H1, the distribution of sojourn times is exponential in form, their median duration is 10 years, and only 15% of lesions regress.

In H3, 50% of cases regress; progressive lesions have a skew (approximately log-normal) distribution of sojourn times; their median duration is 12 years.

marked differences in risk of disease in different birth cohorts;¹⁰ in an unscreened population these changes would give rise to increases in incidence, particularly among younger women belonging to cohorts born since 1931.⁶ In order to simulate this increase, two different approaches were used:

(1) The incidence of dysplasia was weighted according to the relative risk of preclinical disease by marital status and parity.¹¹ The changing pattern of marriage and divorce and, to a much lesser extent, childbearing during the 30-year period 1961–90 results in increasing prevalence of preclinical disease and a progressive increase in incidence.⁹ However, the observed changes in the shape of the age-specific incidence curve are not well reproduced by this method.

(2) For each five-year birth cohort, the transfer rates from normal to dysplasia were multiplied by a cohort factor, *c*. These were calculated from estimated rates of incidence of clinical cancer, by age group and calendar period, in the absence of screening,⁹ it being assumed that the relative risk of dysplasia between cohorts would be the same as that for clinical cancer. The cohort factors were calculated by fitting an age, period, and cohort model¹² using the GLIM package; addition of period effects was not found to improve the fit of the model, so that such effects were ignored. Since the simulation deals with women born between 1861 and 1991, some projection of cohort factors was required. Those before 1881 (the base year, *c*=1.0) were set to 1.0. Projection beyond the latest data point (1956) can be done in various ways;¹³ we chose to assume that the rate of increase in *c* observed since 1946 continued until 1981 and then remained constant. The values of *c* applied are shown in figure 2.

Natural histories H1 and H3 were used with both the marital/parity weighting (H1M, H3M) and with the cohort factors (H1C, H3C).

SCREENING POLICIES

Seven screening policies which have been recommended for implementation in England and Wales were simulated.

(1) (MOH) The original policy promulgated by the Ministry of Health in 1966¹⁴ was for the screening of women aged 35 and over at five-yearly intervals. No upper age limit for testing was suggested; we have assumed that screening will cease after age 65 (a maximum of seven tests per lifetime).

(2) (MOH+) The provision for beginning screening after a third pregnancy was added to this policy in 1973.

(3) (BSCC) The British Society for Clinical Cytology made a new set of proposals in 1977:¹⁵

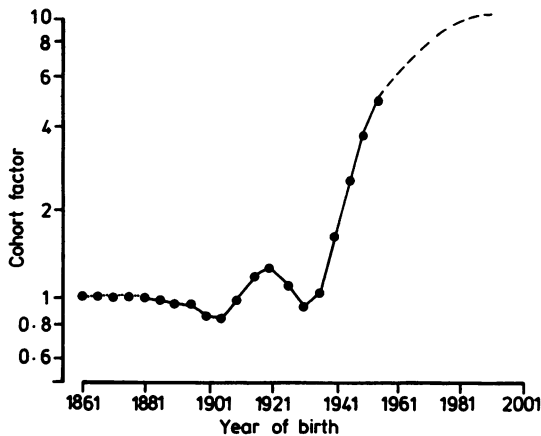


Fig 2 Cohort factors (c) for onset of dysplasia. The increase in risk of dysplasia for women of different birth cohorts relative to that of 1881. The solid line is derived from modelling of England and Wales incidence rates for cervix cancer; the dotted lines are projections from this (see text).

Start screening around age 25 at any consultation for contraception, pregnancy or venereal disease; any sexually active woman who has not been tested should have a first smear at age 30. Tests should be performed at five-yearly intervals until age 70 (no age limit for a first test).

Since the model simulates the occurrence of births, testing during pregnancy is easily incorporated. In the absence of information concerning individual probability of consulting for contraception or venereal disease, we have assumed an annual probability of doing so of 0.13 (so that 50% of women will be contacted between the ages of 25 and 29), and that the prevalence of preclinical abnormalities in women who attend such consultations is twice that in those who do not.

(4) (BSCC+) The same group suggested that screening should be performed at "... three yearly intervals in those aged over 35 if resources permit".

(5) (CGC) The policy recommended in 1982 by the Committee on Gynaecological Cytology¹⁸ is as follows:

Smears should be taken at the ages 35, 40, 45, 50, 55, 60, 65, and also

- early in the course of each pregnancy;
- at age 22, or the next visit thereafter, for women attending for family planning advice, and who have not previously been screened;
- at age 30, for women attending for family planning advice, and who have not had a smear during the previous five years.

Any other woman aged between 22 and 35 who is, or has been, sexually active should be screened on one occasion in this age interval if she requests a test.

The complexity and ambiguity of this policy make precise specifications for the model rather difficult. The probability of attending for family planning advice between ages 22 and 24 was set at 0.5, and (for women unscreened in the preceding five years) at 0.3 at age 30. For attenders at such screening, prevalence of preclinical abnormality of the cervix is assumed to be double that of non-attenders.

(6) (CGC+) The committee making the above recommendations noted: "The practice of taking smears at STD and gynaecological clinics is a useful one since women attending these clinics are likely to be in high risk groups". Screening at clinic attendance was thus added to the policy defined in (5). Hospital admission rates, by age, for gynaecology are available from HIPE, but for outpatients only total numbers of attendances (new and old) are recorded.¹⁷ The annual number of cytology tests from hospital clinics between 1965 and 1980¹⁸ corresponds very closely with the number of gynaecology admissions. Clinic screening is thus modelled using age-specific gynaecology admission rates for 1966–80 with extrapolations beyond these years based on the observed annual increase of 2%. For individuals with preclinical disease, attendance rates are multiplied by 3 to correspond with the observed relative prevalence of abnormalities in clinic attenders compared to other sources of tests.¹⁹

(7) (CGC II) The Committee for Gynaecological Cytology has recently revised its recommendations.^{20,21} Emphasis is placed on screening women over 35 and those who have had three or more pregnancies, at five-yearly intervals. Screening should start "for any woman who is or has been sexually active on her first presentation for contraceptive advice, or whenever she first requests screening. Screening should be repeated ... at the ages of 20, 25, 30, and 35 and not on any other occasion, except that every woman should be screened early in ... each pregnancy".

For simulation of screening related to contraceptive advice (or spontaneous demand) the same parameters as Policy 3 were used, that is, an annual probability of 0.13 (starting at age 16) with prevalence of abnormality in attenders double that in non-attenders. Screening is stopped at age 65 for women with two consecutive negative smears.

For comparison with the seven "recommended" policies, the results from two very simple schedules of tests were studied:

(8) Five-yearly tests from age 25 to 65 (maximum of nine per lifetime);

(9) Three-yearly tests from 25 to 64 (maximum of 14 per lifetime).

All of the policies tested comprise mixtures of "incidental" tests (taken during attendance for some

other purpose) and "routine" tests at predetermined intervals. We examined the above policies under two assumptions of attendance for routine testing—that 80% of those invited would attend and that 50% would do so. We also assumed, in accordance with observation,²² that a proportion of individuals with screen-detected disease escape adequate follow-up or treatment; the proportions were 0.08 for dysplasias and 0.04 for carcinoma in situ, for three years after detection (one-fifth and one-tenth of the totals, respectively). The sensitivity of the screening test was set at 70% (for all grades of preclinical disease) and the specificity at 99.5%, values similar to those estimated from the results of screening programmes.^{23, 24}

OUTCOME MEASURES

The results of screening policies applied over the 30 years of simulation are conveniently measured in terms of the reduction in the adverse effects of cervix cancer. The simplest measure is the reduction in mortality. This can be expressed as:

the reduction in the number of deaths in the 30-year observation period;

the reduction in person-years of life lost (PYLL) from the normal expectation of life (which gives more weight to deaths avoided at young ages, eg, 32 years at age 35, 12 years at age 70);

the reduction in discounted life-years lost (DYLL). Life years are discounted at 5%; in cost-benefit studies it is usual to value savings in the distant future less highly than immediate savings. This provides a weighting intermediate between the first two.

INPUT MEASURES

The input measures to be considered comprise the costs involved in performing screening examinations and in following up any abnormalities detected; in addition, account must be taken of the costs involved in treatment and care of cases of clinical cancer. Arbitrary units may be used; we have chosen to compare costs of different procedures with that of a routine screening test. Thorn *et al*²⁵ provide the most comprehensive estimates of monetary costs based on those for Aberdeen in 1971.

Screening: taking and interpreting routine smears	£1.15
Inpatient treatment: Clinical cancer	£487
Microinvasive, carcinoma in situ,	
severe atypia	£93
Moderate atypia or less	£63

The assumptions in the simulation model are that 10% of in-situ lesions and 20% of dysplasias will escape adequate follow-up. With the two natural histories used (H1+H3), 40% of dysplasias will regress to normality within two years; since management of milder grades of abnormality found on screening usually entails a series of follow-up

smears, we have assumed that only the remaining 60% would receive some form of intervention. We have used the following costs: routine tests, 1; clinical cancer cases, 400; positive tests (cis, microinvasive cancer), 75; positive tests (dysplasia), 30. We have assumed that the cost of an incidental smear will be only half that of one derived from a routine screening attendance, and that women with false-positive tests will be subjected to two follow-up smears (cost=2).

An alternative set of costings, to allow for advances in treatment of preclinical disease,²⁶ and the diagnosis of clinical cases at less advanced stages was also tested, using the ratios 300 (clinical cancer) : 37.5 (positive, in situ) : 15 (positive, dysplasia); these are referred to as "lower" unit costs as compared to the "higher" unit costs above.

Results

SIMULATION OF NATURAL HISTORY

There are some small differences between the four different natural histories in the results obtained in the absence of screening. When the entire simulation period (30 years) is considered, H1 gives rise to 4% more cases and 4–7% more deaths than H3, and weighting by parity/marital status produces 9% more cases and 8–11% more deaths than weighting by cohort effect. Time trends in incidence and mortality are shown in fig 3; all of the natural histories used show increases, although these are progressive throughout the period in the marital/parity weighted histories, whereas the cohort-weighted histories show a rise in incidence rates only in the second half of the simulation period. In reality, incidence and mortality in England and Wales have fallen progressively since 1961, although this is in large measure due to screening, without which a rise would have occurred.⁸ Study of the change in age-specific incidence rates during the period of simulation shows that the cohort-weighted natural histories reproduce some of the features expected in the population if no screening had occurred, especially the striking rise in incidence in young women (fig 4). The marital/parity weighted histories do not produce this change in shape of the age-specific incidence curve, which remains similar in configuration to that of 1966.

RESULTS OF SCREENING

Table 1 shows the number of tests performed with the nine different policies at two levels of attendance for routine testing, and the proportion that are incidental tests (assumed to have only half the cost).

Since the specificity of the screening tests was set at 99.5%, the number of false-positive examinations resulting from screening is 0.5% of the total number of tests.

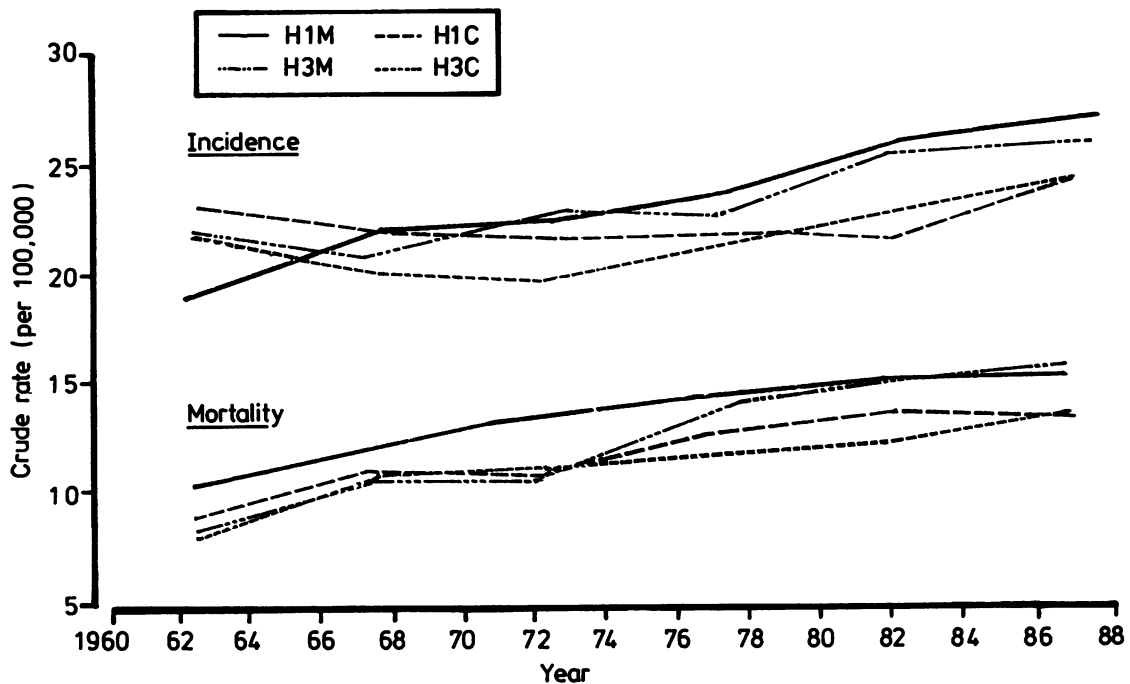


Fig 3 Time trends in incidence and mortality rates assuming no screening. Incidence and mortality rates over the 30 years of simulation. Four different natural histories:

— H1M;
 - - - H1C;
 - · - · - H3M;
 ····· H3C.

Table 1 Number of tests performed in 30 years screening

Policy	Number of tests: total ($\times 10^6$) and % incidental	
	80% attendance	50% attendance
1 MOH	193 (0)	121 (0)
2 MOH+	204 (0)	128 (0)
3 BSCC	262 (12)	177 (18)
4 BSCC+	372 (9)	246 (13)
5 CGC	300 (36)	229 (48)
6 CGC+	364 (48)	293 (59)
7 CGC II	390 (31)	290 (42)
8 5 yearly	257 (0)	161 (0)
9 3 yearly	397 (0)	249 (0)

Percentages in parentheses

Comparison of results obtained with the four different natural histories show that the savings in cases, lives, and life-years are rather greater with natural histories H3 than with H1. Table 2 shows the results assuming regular five-yearly testing (policy 8). This result might be anticipated from the distribution of sojourn times of preclinical lesions in H1 and H3 (fig 1); with H3 there are relatively few in-situ lesions with very short sojourn times that are able to progress to invasive disease in the inter-screening intervals.

Table 2 Savings achieved assuming different natural histories. Figures are percentage reductions from baseline of no screening. Policy: 5 yearly tests age 25-65

Natural history	Percentage reduction					
	80% attendance			50% attendance		
	Cases	Deaths	PYLL	Cases	Deaths	PYLL
H1M	59.8	52.6	58.0	41.5	37.9	37.9
H1C	59.0	55.0	60.9	42.0	37.3	40.1
H3M	69.7	61.9	66.5	51.6	44.9	48.6
H3C	67.7	58.9	63.3	51.6	42.8	49.8

The corollary of this greater reduction in clinical disease achieved by assuming natural history H3 is the detection of larger numbers of preclinical lesions with natural history H3 than with H1 (table 3). This table also shows that when natural histories are weighted by the cohort factors, the increasing prevalence of preclinical disease at young ages leads to more preclinical lesions detected with policies (such as 5 and 9) where screening begins at young ages.

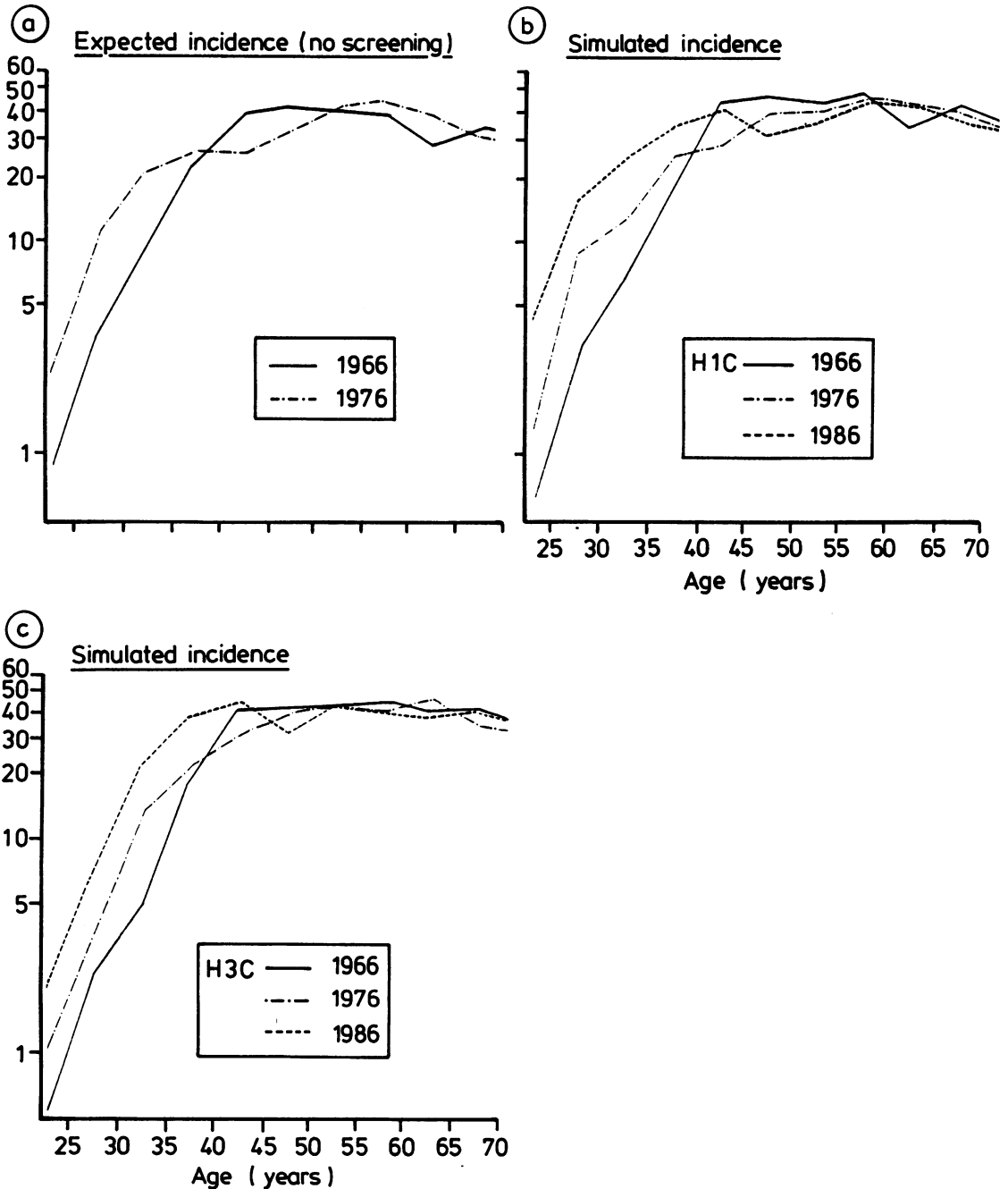


Fig 4 Expected and simulated age-specific incidence of cervix cancer in England and Wales in the absence of screening.
 A Incidence expected in the absence of screening (from reference 6) for 1966 and 1976.
 B Results of simulation, using natural history H1C. The curves show the averages for the decades centred upon 1966, 1976, and 1986.
 C As for B, but using natural history H3C.

Table 3 Preclinical lesions detected (true positives) with different natural histories

Policy (50% attend)	H1		H3	
	M	C	M	C
1 MOH	1515	1326	1721	1454
5 CGC I	2665	3550	2918	4579
9 3 yearly	2964	3143	3370	3589

In order to illustrate the effects of the nine screening policies under opposing assumptions concerning natural history, results obtained with natural histories H1M and H3C are presented in tables 4 and 5. The cost-effectiveness ratios have been calculated assuming higher unit costs (see methods). It is evident that, in each of these tables, the savings in cases, deaths, and person-years of life

Table 4 Results of screening for 30 years; natural history H1M

Screening	True-positive tests (×100)						Cost-effectiveness (Savings per 1000 units input)*		
	Dysplasia	Cis/miv	Cases	Deaths	PYLL	DYLL	Deaths	PYLL	DYLL
None	0	0	753	409	8111	4776			
80% attendance									
1 MOH	14.4	8.1	323	199	3878	2289	1.66	33.5	19.6
2 MOH+	16.1	8.4	301	190	3568	2191	1.61	33.5	19.1
3 BSCC	23.2	9.4	242	163	3033	1839	0.97	20.1	11.6
4 BSCC+	27.7	10.2	211	145	2406	1480	0.88	19.0	11.0
5 CGC I	23.3	9.2	280	185	3361	2073	1.12	23.9	13.6
6 CGC+	32.9	9.5	224	151	2710	1658	1.08	22.7	13.6
7 CGC II	29.6	8.8	262	167	2924	1742	0.83	17.8	10.4
8 5 yearly	22.2	9.1	303	194	3410	2111	1.00	21.8	12.4
9 3 yearly	31.2	9.9	217	149	2632	1508	0.73	15.5	9.3
50% attendance									
1 MOH	9.2	6.0	459	260	5113	3019	1.94	39.0	22.8
2 MOH+	10.4	6.5	426	249	5068	2966	2.07	39.3	23.4
3 BSCC	16.0	6.9	379	220	4099	2516	1.68	35.7	20.1
4 BSCC+	19.8	8.4	311	174	3272	1985	1.32	27.2	15.7
5 CGC I	19.3	7.4	369	212	3970	2475	1.44	30.3	16.9
6 CGC+	29.9	8.8	312	205	3445	2172	1.09	24.8	13.9
7 CGC II	23.3	7.4	385	230	4198	2577	0.85	18.6	10.5
8 5 yearly	15.2	6.6	440	254	4826	2968	1.19	23.5	13.9
9 3 yearly	21.4	8.2	332	197	3469	2101	1.01	22.2	12.8

*Input calculated using "higher" unit costs.
Cis=carcinoma in situ; miv=microinvasive.

Table 5 Results of screening for 30 years; natural history H3C

Screening	True-positive tests (×100)						Cost-effectiveness (Savings per 1000 units input)*		
	Dysplasia	Cis/miv	Cases	Deaths	PYLL	DYLL	Deaths	PYLL	DYLL
None	0	0	668	355	7038	4171			
80% attendance									
1 MOH	13.2	8.4	268	158	3264	1953	1.43	27.5	16.2
2 MOH+	14.5	8.6	258	152	3007	1787	1.36	26.9	15.9
3 BSCC	28.3	11.1	217	148	2967	1735	0.88	17.2	10.3
4 BSCC+	31.9	11.6	147	110	2165	1266	0.73	14.6	8.7
5 CGC I	35.1	10.8	203	133	2433	1476	0.89	18.5	10.8
6 CGC+	52.4	11.7	169	121	2111	1104	0.72	15.1	9.4
7 CGC II	49.0	10.9	207	143	2377	1523	0.57	12.4	7.1
8 5 yearly	29.3	10.9	216	146	2584	1583	0.84	18.0	10.4
9 3 yearly	40.5	11.5	155	108	1758	923	0.61	13.1	8.1
50% attendance									
1 MOH	8.7	5.8	381	227	4528	2741	1.66	32.5	18.5
2 MOH+	9.4	6.5	368	211	4068	2427	1.67	34.6	20.3
3 BSCC	21.9	8.9	300	195	3556	2206	1.08	23.5	13.3
4 BSCC+	25.2	10.3	243	141	2861	1663	1.00	19.4	11.7
5 CGC I	37.1	8.7	287	180	3102	1935	0.87	19.6	11.1
6 CGC+	48.1	10.0	228	170	2924	1616	0.73	16.3	10.1
7 CGC II	39.8	9.0	289	176	3114	1939	0.67	14.6	8.4
8 5 yearly	18.3	8.0	323	203	3535	2229	1.09	25.2	13.9
9 3 yearly	26.2	9.7	242	164	3045	1756	0.82	17.2	10.4

*Input calculated using "higher" unit costs.
Cis=carcinoma in situ; miv=microinvasive.

vary considerably between the different policies. When attendance rates for routine screening are high (80%), the rankings of policies in terms of lives saved, or life-years saved, are practically the same, no matter which natural history is assumed. Clearly, the greatest reductions are achieved with the more intensive policies. However, the greater savings that can be achieved by changing to more complex and intensive policies are in general no more than are possible by increasing attendance rates for the same policy from 50% to 80%. When comparing policies by the measures of efficiency (savings per 1000 units of input), the rankings are very largely reversed, although the correspondence is not exact. The most efficient are the two low intensity policies (1 and 2).

The CGC II policy (7) includes a large number of incidental tests taken during pregnancy—89 000 over the 30-year period. These tests account for relatively small savings; pregnancy testing *alone* over the 30 years gives rise to a saving of 33 lives and 999 life-years with natural history H1M. The cost-effectiveness of this policy is correspondingly poor—the lowest in rank of all those examined (except under the assumption of high attendance rates and natural history H1M).

The original CGC policy (5) performs moderately well, its efficiency at high attendance rates is third in rank (after the low intensity policies), and fourth when attendance for routine screening is 50%. Under the assumptions shown in tables 4 and 5 the addition of screening of symptomatic women (CGC+, policy 6) leads to a fall in the cost-effectiveness ratio. This is partly because of the increase in preclinical cases (especially dysplasia) which are detected and require treatment, the costs of which are usually sufficient to offset the moderate reductions in clinical cancers and deaths.

When attendance rates for routine screening are low, the simple BSCC policy (3) scores well in efficiency. It is interesting to compare this policy, which, in effect, adds three screenings to the simple MOH policy (incidental tests around 25–29, rescreen at 30, and a test at age 70) with policy 8, which adds two routine screens (at ages 25 and 30) to the simple MOH policy. The marginal gains from the additional routine tests in young women are small (2–4% increase in lives saved, 9–11% in life-years, depending on attendance rates). The BSCC schedule has the advantage of testing young women at higher than average risk (family planning attenders) and adding a test in older women, whose risk of disease is still high. The additional costs incurred by the discovery of more preclinical lesions is offset by the fall in the number of cases of clinical invasive cancer that require treatment.

The differences between screening with a three-yearly interval between tests compared to five-yearly can be seen by comparing policies 8 and 9. Firstly, it is clear that screening five-yearly with an 80% attendance rate results in fewer cases, deaths and person-years of life lost than screening three-yearly with 50% attendance. The differences in cost-effectiveness are, however, small. With constant attendance rates, increasing frequency of testing leads to a 55% increase in the number of tests performed, but this is accompanied by a much smaller increase in the number of lives saved during the 30 year period (10–14%, depending on natural history and attendance rate). The cost of implementing the more intensive policy is also increased by the need to investigate and treat the additional positives generated by more frequent testing (although at low attendance rates this extra cost is offset by the reduction in cost of treatment of clinical cancer cases). The cost-effectiveness ratios are thus 15–30% lower (depending on the index chosen). The differences between the two BSCC policies (3 and 4) are rather small, since the more frequent schedule of testing in the BSCC+ policy begins only after the age of 35.

The calculation of cost-effectiveness ratios using the lower unit costs can be done using the data presented in tables 1, 4, and 5. All of the values for savings per 1000 units of input are increased, the size of this increase depending on the numbers of preclinical lesions and invasive cancers that require treatment. Thus the rankings of the different policies, and the relative differences between them, remain almost unchanged.

The comparisons between the different policies above have been made in terms of their costs and savings relative to those incurred over a 30-year period during which absolutely no screening activity is undertaken. It could be argued that this is an unrealistic baseline, since even in the absence of any screening programme, cytological examinations are often carried out on women attending medical services with gynaecological symptoms. Although such examinations can be regarded as a form of selective screening (searching for unrecognised disease in a known high risk group), in practice they are usually considered to be diagnostic tests. We have therefore considered the results of the different screening policies against a background of "diagnostic testing", making the assumption that the rate of such testing is one half that used for gynaecology testing in policy 6 (see methods), and that the rate of abnormality in women receiving such tests is three times that in the general population.

Over the 30-year simulation period, this diagnostic testing gave rise to the 33 500 cytological tests. Using

Table 6 Cost-effectiveness of screening policies in the presence of diagnostic testing; natural history: H3C; "higher" unit costs

	Savings per 1000 units input		
	Deaths	PYLL	DYLL
80% attendance			
1 MOH	0.86	15.8	10.1
2 MOH+	0.78	14.2	9.0
3 BSCC	0.67	12.3	7.7
4 BSCC+	0.52	9.2	5.9
5 CGC I	0.65	12.6	8.3
6 —	—	—	—
7 CGC II	0.44	8.7	5.3
8 5 yearly	0.71	13.8	8.4
9 3 yearly	0.45	8.9	5.4
50% attendance			
1 MOH	1.29	22.2	14.4
2 MOH+	1.28	23.7	14.8
3 BSCC	0.85	15.8	9.9
4 BSCC+	0.75	14.1	8.7
5 CGC I	0.72	14.0	8.6
6 —	—	—	—
7 CGC II	0.50	10.2	6.1
8 5 yearly	0.84	15.9	9.9
9 3 yearly	0.61	11.6	7.2

natural history H3C, these tests generate 146 false positives and 1886 true positives (1384 dysplasia, 502 cis/microinvasive). In the absence of any additional screening, there are 510 cases of clinical cancer and 295 deaths (life-years lost=5442, "discounted" life-years=3362). These results can be compared with those obtained in the total absence of screening (table 5)—there is a 23.7% reduction in cases of invasive cancer and a 16.9% reduction in deaths. Table 6 shows the cost-effectiveness of the different screening policies in the presence of diagnostic testing, using the marginal costs and savings incurred. For each policy, the net costs are increased, and the net savings decreased, compared to those calculated against a background of no screening (table 5). The ratios of savings:costs are hence all reduced. However, the relative efficiency of the different policies is almost the same (results for the CGC+ policy, which include testing of clinic attenders, have not been calculated). The CGC II policy (7) scores lowest, and the low intensity policies (1 and 2) the highest. Of the remainder, the attendance rate influences their relative rankings—at 80% attendance the five-yearly policy (8) and CGC policy (5) perform well, while the BSCC policy (3) seems to be more cost-effective at the lower rates of attendance.

Discussion

The recommendations of different national bodies concerning ages and frequencies for cervical cytology tests (Pap smears) vary considerably, ranging from examinations at one- to three-yearly intervals in all

women under age 60 who have had sexual intercourse—as suggested in the USA,²⁷ to schedules involving five-yearly examinations of women in older age groups, as in Finland²⁸ or, originally, in England and Wales.¹⁴ The reason for such wide disparity between countries is not clear; it is not related to possible differences in the natural history of disease, and presumably reflects availability of resources and the priority given to the control of cervix cancer compared to other health care programmes.

Simple deterministic simulation models can be used to demonstrate what is intuitively obvious—that a progressive increase in the frequency of screening leads to a progressive reduction in the marginal benefit per unit input (usually expressed as number of screening tests).^{3 7 29 30} The choice of age range for screening is determined in part by the natural history of cervix cancer and partly by subjective judgements as to the value of different outcomes (eg, the relative utility of cancers avoided or lives saved in women of different ages). A natural history consistent with observed data on age-specific incidence and prevalence of preclinical and invasive disease cannot be reproduced with many models; in particular, it is clear that there are marked changes in risk of disease by birth cohort in some populations.^{6 31} The microsimulation model used in this study is able to simulate such changes. In addition, there is the ability to simulate rather complex screening policies, where attendance for screening is dependent on some past event (for example, three or more pregnancies, non-attendance at any other clinic in the preceding five years). The screening policies proposed for England and Wales since the original 1966 proposals¹⁴ have all been of this type; they recognise the very complex nature of current screening activity, which takes place in many different locations,^{18 32} and seek to build on this background. These include extensive screening during or soon after pregnancy, and in relation to attendance for family planning advice and at hospital clinics.

We used four different sets of data to reproduce natural history which, although relatively extreme, are plausible and reproduce observed frequencies of preclinical and clinical disease and, in those incorporating cohort factors, the rising incidence of disease in recent generations of women. The dichotomisation of preclinical disease into dysplasia (CIN I and II) and carcinoma in situ (CIN III) was adopted for convenience of simulation. However, in view of the results of recent virological studies,^{33 34} this may also prove to have biological validity. Infection by human papilloma virus (HPV) types 6 and 11 gives rise to a histological appearance characterised by koilocytosis which has generally been called mild or moderate dysplasia; many such

lesions are quite transient. Infection by HPV types 16 and 18, by contrast, is characterised by nuclear atypia, regular and atypical mitotic figures, and aneuploidy. In advanced lesions the pattern corresponds to severe dysplasia/carcinoma in situ; regression is less common than with HPV types 6 and 11.

The outcome measures chosen relate solely to mortality from cervix cancer. These are convenient and easily recorded and have been the most frequently used in other assessments of costs and benefits of screening for cervix cancer.^{29 30 35} The choice of index has little effect on the evaluation of the different policies. These are not the only outcomes of a screening programme that are important, however. The main objective of cervix cancer screening is the prevention of invasive cancer, and, of the cases so far prevented, only a proportion would have been fatal and counted as deaths or life-years lost. Screening will also lead to a reduction in non-fatal cancers, which might be counted as one of the beneficial outcomes of the programme (savings in disability, pain, etc) as well as contributing to a reduction in cost.

Only the outcomes observed during the 30 years of screening have been included in the evaluation, along with the costs incurred during this period. Although the benefits of a screening programme will continue to accrue after the actual screening has ceased, we have not included these in the evaluation. It is most unlikely that cervix cancer screening will ever be stopped, so that post-programme evaluation is a rather theoretical idea. Furthermore, there is no reason to suppose that, in relation to the policies examined here, addition of post-screening benefits would make any difference to their relative efficiency. Since the items contributing to the costs (tests, follow-up investigations, treatment) are spread over the entire period, the total cost has been reckoned at constant prices, with no attempt at discounting to a base year. Two sets of costs were examined, with almost no effect on the relative cost-effectiveness of the different policies. The costs include only the direct financial costs of the services provided, and we have not attempted to estimate so-called indirect or intangible costs to the patients (or their relatives)—for example, due to loss of work during investigation and treatment, or the anxiety engendered by a false-positive test.³⁶

From a cost-effectiveness point of view, the original policies advocated for England and Wales, involving seven to eight tests per lifetime, perform rather well. Increasing the frequency of testing and lowering the age of onset of testing both improve outcomes, but with a lowering of the cost-effectiveness ratio. The most recent policy

recommended by the Committee for Gynaecological Cytology^{30 21} seems to be unsatisfactory, based on the criteria that we have adopted here. The concentration of screening at young ages, even though such tests may be less costly than usual (eg, in pregnancy) and involve higher risk groups (contraceptive users), leads to high costs for relatively small gains in outcome. It may be that this relative inefficiency will be less marked in the longer term if the incidence of disease in young women continues to rise at rates even more marked than those incorporated in this simulation (fig 4). Nevertheless the wisdom of concentrating screening tests at young ages, as in Canada (the latest recommendations suggest 18 tests up to age of 35 and five thereafter³⁷), must be seriously questioned. The rationale is presumably to ensure that some women who may not volunteer to start a screening programme will receive at least a first test during a visit for another purpose and may thereafter be more readily induced to re-attend. It might reasonably be questioned whether the complexity of these programmes makes this possible advantage worthwhile. The absolute benefits from a programme of five-yearly testing after age 25 are only slightly inferior to those achieved by the much more complex policies recommended by the Committee for Gynaecological Cytology, and, in comparison with the latest recommendations, it is more efficient also. A system to ensure that all women receive invitations for screening (and that those who fail to attend receive reminders) should be easy to install if any sort of age-sex register for general practice is in existence. The large volume of rather haphazard incidental testing in relation to pregnancy and family planning might then be drastically curtailed.

Requests for reprints to Dr D M Parkin, International Agency for Research on Cancer, Lyon, France.

References

- ¹ Knox EG. Ages and frequencies for cervical cancer screening. *Br J Cancer* 1976; **34**: 444–52.
- ² Albert A, Gertman PM, Louis TA, Liu S. Screening for the early detection of cancer II. The impact of screening on the natural history of the disease. *Math Biosci* 1975; **40**: 61–109.
- ³ Yu Shan-Zhang, Miller AB, Sherman GJ. Optimising the age, number of tests and test interval for cervical cancer screening in Canada. *J Epidemiol Community Health* 1982; **36**: 1–10.
- ⁴ Andrews FJ, Linehan JJ, Melcher DH. Cervical cancer in younger women. *Lancet* 1978; **ii**: 776–8.
- ⁵ Cotton R. Why DHSS screening policy puts young women unnecessarily at risk. *World Medicine* 1980; **15** (no 13): 21–6.

- ⁶Parkin DM, Nguyen-Dinh X, Day NE. The impact of screening on the incidence of cervical cancer in England and Wales. *Br J Obstet Gynaecol* 1985; **92**: 150-7.
- ⁷Eddy DM. American Cancer Society: guidelines for the case-related checkup: recommendations and rationale. *CA* 1980; **30**: 194-240.
- ⁸Simpson P. Economic aspects of screening. In: Prorok PC, Miller AB, eds. Screening for Cancer. UICC Technical Report Series, Vol 78. Geneva: UICC, 1984.
- ⁹Parkin DM. A computer simulation model for the practical planning of cervical cancer screening programmes. *Br J Cancer* 1985; **51**: 551-68.
- ¹⁰Cook GA, Draper GJ. Trends in cervical cancer and carcinoma in situ in Great Britain. *Br J Cancer* 1984; **50**: 367-75.
- ¹¹Parkin DM, Hodgson P, Clayden AD. Incidence and prevalence of preclinical carcinoma of cervix in a British population. *Br J Obstet Gynaecol* 1982; **89**: 564-70.
- ¹²Osmond C, Gardner MJ. Age, period and cohort models applied to cancer mortality rates. *Statistics in Medicine* 1982; **1**: 245-59.
- ¹³Osmond C. Using age, period and cohort models to estimate future mortality rates. *Int J Epidemiol* 1985; **14**: 124-9.
- ¹⁴Ministry of Health. Population screening for cancer of the cervix. NHS circular HM (66)76; 1966.
- ¹⁵Spriggs AI, Husain OAN. Cervical smears. *Br Med J* 1977; **i**: 1516-8.
- ¹⁶Draper GJ. Screening for cervical cancer: revised policy. The recommendations of the DHSS Committee on Gynaecological Cytology. *Health Trends* 1982; **14**: 37-40.
- ¹⁷Department of Health and Social Security. Health and Personal Social Services Statistics. London: HMSO, 1983.
- ¹⁸Roberts A. Cervical cytology in England and Wales, 1965-80. *Health Trends* 1982; **14**: 41-3.
- ¹⁹Parkin DM, Leach K, Cobb P, Clayden AD. Cervical cytology screening in two Yorkshire areas: results of testing. *Public Health (London)* 1982; **95**: 3-14.
- ²⁰Lancet. New advice on cervical screening. *Lancet* 1984; **i**: 1196.
- ²¹Department of Health and Social Security. Health Services Development; screening for cervical cancer. Circular 1984; HC (84) 17.
- ²²Elwood JM, Cotton RE, Johnson GM, Curnow J, Beaver MW. Are patients with abnormal cervical smears adequately managed? *Br Med J* 1984; **289**: 891-4.
- ²³Husain OAN. Quality control in cytological screening for cervical cancer. *Tumori* 1976; **62**: 303-14.
- ²⁴Boyes DA, Morrison B, Knox EG, Draper GJ, Miller AB. A cohort study of cervical screening in British Columbia. *Clin Invest Med* 1982; **5**: 1-29.
- ²⁵Thorn JB, MacGregor JE, Russell EM, Swanson K. Costs of detecting and treating cancer of the uterine cervix in north-east Scotland in 1971. *Lancet* 1975; **i**: 674-6.
- ²⁶Singer A, Walker P. Commentary: what is the optimum treatment of cervical intraepithelial neoplasia? *Br J Obstet Gynaecol* 1982; **89**: 335-7.
- ²⁷National Institutes of Health, Consensus statement. Cervical cytology screening: the Pap smear. *Br Med J* 1980; **281**: 1264-6.
- ²⁸Hakama M, Räsänen-Virtanen U. Effect of a mass screening program on the risk of cervical cancer. *Am J Epidemiol* 1976; **103**: 512-7.
- ²⁹Knox EG. A simulation system for screening procedures. In: McLachlan G, ed. The future and present indicatives. (*Problems and Progress in Medical Care* Vol IX). Oxford: Nuffield Provincial Hospitals Trust, 1973.
- ³⁰Eddy DM. *Screening for cancer: theory, analysis and design*. Englewood Cliffs, New Jersey: Prentice Hall, Inc, 1980.
- ³¹Hakama M. Trends in the incidence of cervical cancer in the Nordic countries. In: Magnus K, ed. *Trends in cancer incidence*. Washington: Hemisphere Publishing Corp, 1982.
- ³²Parkin DM, Collins W, Clayden AD. Cervical cytology screening in two Yorkshire areas: pattern of service. *Public Health (London)* 1981; **95**: 311-21.
- ³³Gissmann L. Papillomaviruses and their association with cancer in animals and in man. *Cancer Sur* 1984; **3**: 161-81.
- ³⁴Schneider A, Kraus H, Schuhmann R, Gissmann L. Papillomavirus infection of the lower genital tract: detection of viral DNA in gynaecological swabs. *Int J Cancer* 1985; **35**: 443-8.
- ³⁵Dickinson L. Evaluation of the effectiveness of cytological screening for cervical cancer. III Cost-benefit analysis. *Mayo Clin Proc* 1972; **47**: 550-5.
- ³⁶Klarman HE. Application of cost-benefit analysis to health-systems technology. *J Occup Med* 1974; **16**: 172-86.
- ³⁷Department of National Health and Welfare. Cervical cancer screening programs: summary of the 1982 Canadian task force report. *Can Med Assoc J* 1982; **127**: 581-9.