STATISTICS IN MEDICINE, VOL. 6, 469–481 (1987)

MODELS FOR TEMPORAL VARIATION IN CANCER RATES. II: AGE–PERIOD–COHORT MODELS

D. CLAYTON

Department of Community Health, Clinical Sciences Building, Leicester Royal Infirmary, PO Box 65, Leicester LE2 7LX, U.K.

AND

E. SCHIFFLERS

Facultés Universitaires N.-D. de la Paix, Rue de Bruxelles, 61, B5000, Namur, Belgium

SUMMARY

Our first paper reviewed methods for modelling variation in cancer incidence and mortality rates in terms of either period effects or cohort effects in the general multiplicative risk model. There we drew attention to the difficulty of attributing regular trends to either period or cohort influences. In this paper we turn to the more realistic problem in which neither period nor cohort effects alone lead to an adequate description of the data. We describe the age–period–cohort model and show how its ambiguities surrounding regular trends 'intensify'. We recommend methods for presenting the results of analyses based upon this model which minimize the serious risk of misleading implications and critically review previous suggestions. The discussion is illustrated by an analysis of breast cancer mortality in Japan with special reference to the phenomenon of 'Clemmesen's hook'.

KEY WORDS Cohort analysis Cancer trends Age–period–cohort models Standardized rates

INTRODUCTION

In our first paper[1] we described an approach to the analysis of data on the variation of cancer incidence or mortality with time. This approach is in the main stream of modern methodology in chronic disease epidemiology, being based upon the proportional hazards model. This is an empirically based general model which, in its simplest form, holds that the ratios of age-specific rates between two groups of individuals with different exposures to carcinogenic influences are constant for all age groups. This general model underpins traditional methods of age standardization[2] as well as relative risk analyses in age-matched case–control studies.[3]

We described two different models for variation over time, the age–period model and the age–cohort model. These models predict constant ratios of age-specific rates

(i) between different periods, that is the calendar periods during which the incidence (or mortality) rates were observed, or
(ii) between different cohorts, that is longitudinally observed groups of people born within specific periods.

Cross-sectional tables of rates by age and calendar period allow us to fit the former (age–period) model precisely and the latter (age–cohort) model to a close approximation, at least when the data are grouped with almost equal time intervals on both age and calendar period axes (say 5 years).

0277–6715/87/040469–13$06.50

Received 6 March 1986
Revised 5 September 1986

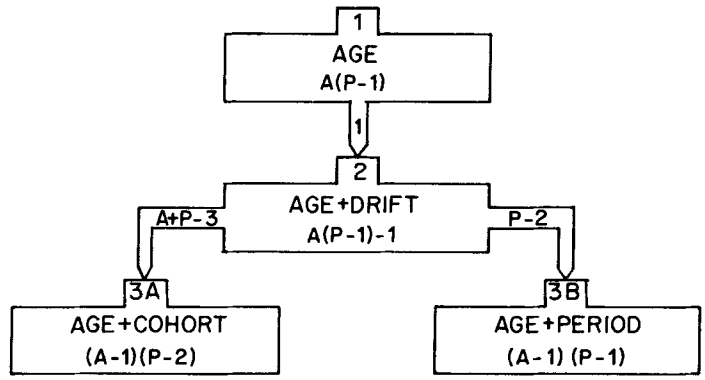

Figure 1. Logical order in which to consider models with at most one type of temporal variation. The formula on the second line of a box gives the number of degrees of freedom for a table with A age classes and P time periods

Finally, we considered the problem posed by a specific type of regular trend in which the ratio of age-specific rates between two adjacent time periods is not only constant across age groups, but is constant for any pair of adjacent time periods. We showed this model to be indistinguishable from the equivalent regular age–cohort trend model in which the relative risk between adjacent birth cohorts is constant. The models are indistinguishable, in the sense that either may generate identical predictions for the data using the same trend parameter – the relative risk between adjacent periods or cohorts. However, in so doing, the models must adopt different age curves so that if the relationship between age and incidence or mortality rates were known then the models would be distinguishable. In the absence of such knowledge, however, the models are equivalent and we introduced the term drift to describe this regular trend, unattributable to specifically period or cohort influences.

We also introduced the terms cross-sectional and longitudinal age curves for the age relationships estimated by the age–period and age–cohort versions, respectively, of the regular drift model. It follows directly from our remarks of the previous paragraph that, in the absence of extra information allowing us to determine which model is true, we are likewise unable to decide which of these represents the true age curve. This fundamental difficulty is central to understanding the problems of interpretation with the more complex models described in this paper.

Figure 1 illustrates the logical order in which to consider the models we have encountered so far. The first model is the null hypothesis of no temporal variation, while the second model is the model of regular drift unattributable to period or cohort influences. Only if this model does not adequately describe the data are we justified in considering either specifically age–period or age–cohort models.

In our first paper[1] we also discussed the assessment of the goodness-of-fit of models and showed how the change in a global measurement of goodness-of-fit (or badness-of-fit, deviance) may be used to construct statistical significance tests. Thus, comparison of model (2) with model (1) provides a one-degree of freedom test for trend (drift). Comparison of model (3A) with model (2) provides a $(P-2)$ degree-of-freedom test for irregular trends incidence (death) attributable to period while comparison of models (3B) and (2) provides a $(C-2)$ degree-of-freedom test for irregular cohort effects. Note that models (3A) and (3B) may not be compared directly in this way and it is not possible to construct a formal test of whether the age–cohort model is significantly better than the age–period model.

Table I. Age-specific mortality rates (per 100,000 person-years observation) of breast cancer in Japan, during the period 1955–1979. Numbers of cases on which rates are based are in parentheses (source: WHO mortality data base).

Age/period	1955–1959		1960–1964		1965–1969		1970–1974		1975–1979	
25–29	0·44	(88)	0·38	(78)	0·46	(101)	0·55	(127)	0·68	(179)
30–34	1·69	(299)	1·69	(330)	1·75	(363)	2·31	(509)	2·52	(588)
35–39	4·01	(596)	3·90	(680)	4·11	(798)	4·44	(923)	4·80	1056)
40–44	6·59	(874)	6·57	(962)	6·81	(1171)	7·79	(1497)	8·27	(1716)
45–49	8·51	(1022)	9·61	(1247)	9·96	(1429)	11·68	(1987)	12·51	(2398)
50–54	10·49	(1035)	10·80	(1258)	12·36	(1560)	14·59	(2079)	16·56	(2794)
55–59	11·36	(970)	11·51	(1087)	12·98	(1446)	14·97	(1828)	17·79	(2465)
60–64	12·03	(820)	10·67	(861)	12·67	(1126)	14·46	(1549)	16·42	(1962)
65–69	12·55	(678)	12·03	(738)	12·10	(878)	13·81	(1140)	16·46	(1683)
70–74	15·81	(640)	13·87	(628)	12·65	(656)	14·00	(900)	15·60	(1162)
75–79	17·97	(497)	15·62	(463)	15·83	(536)	15·71	(644)	16·52	(865)

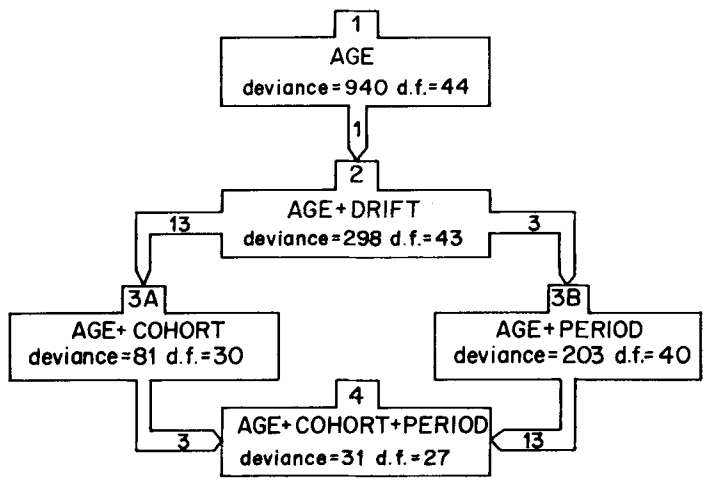

Figure 2. Deviances obtained from fitting various multiplicative models to the data of Table I. Numbers in the connecting arrows give the loss of degrees of freedom (d.f.)

In this paper we consider the case where neither the age–period nor the age–cohort model provides an adequate fit to the observed table of rates.

THE AGE–PERIOD–COHORT MODEL

Table I shows mortality from breast cancer in Japan during the period 1955–1979, and the upper section of Figure 2 shows the results of fitting the four models of Figure 1.

In this case the fit of the models is measured using a deviance or likelihood-ratio criterion which assesses the deviations between observed and fitted rates relative to Poisson variability expected on the basis of the numerators of the observed rates (see paper I, third section).[1] If any model gives a

true description of the underlying rates, the corresponding deviance should be distributed as chi-squared with the appropriate degrees of freedom.

It can be seen that although the age–cohort model is clearly the best of those considered it does not fit the data adequately; a value of chi-squared of 81 on 30 degrees of freedom is highly significant.

This suggests that the model to consider next is one in which both cohort and period effects are included; this is the age–period–cohort model which has received much attention in recent literature. As for the age–period and age–cohort models, it may be written either as a model for the log-rates in which the effects of age, period and cohort combine additively, or as a model for the rates themselves in which the effects combine multiplicatively. That is, writing Y_{ap} for the logarithms of the rates, the age–period–cohort model is

$$Y_{ap} = \alpha_a + \beta_p + \gamma_c, \tag{1}$$

$$c = A - a + p.$$

Here, as in our first paper,[1] c indexes diagonals of age × period table, which approximate to birth cohorts.

Writing, ρ_{ap} for the untransformed rates and α'_p, β'_p and γ'_c for the antilogs of the corresponding parameters, the multiplicative form of the model is

$$\rho_{ap} = \alpha'_a \beta'_p \gamma'_c \tag{2}$$

so that age-specific rates, α'_a, are multiplied by factors β'_p corresponding to the calendar period of incidence or death, and by factors γ'_c corresponding to the birth cohort of individuals affected.

This model may be fitted to the data either by weighted least-squares or by Poisson maximum likelihood. In these papers the latter method has been used throughout. When fitted to the data of Table I, model (1) gives a deviance of 30·7 on 27 degrees of freedom, which is consistent with chance (Poisson) fluctuations ($P \simeq 0{\cdot}28$). Figure 2 shows the sequence of models leading to this final, acceptable model. By comparing models 3B and 4 we conclude that, after adjusting for period effects, cohort effects are statistically significant (deviance = 172, on 13 degrees of freedom). Likewise, comparing models 3A and 4 we see that, after adjusting for cohort effects, the effect of calendar period is significant (deviance = 50, on 3 degrees of freedom). Note that the degrees of freedom for each of these tests mirrors the value for the test corresponding to the opposite corners connecting models 2, 3A, 3B and 4. These tests are the corresponding crude tests. For example, comparison of 3A and 2 tests for cohort effects but does not adjust for period effects.

Exactly as the adjusted tests, (3A − 2) and (3B − 2), test for attributable cohort and period effects, respectively (that is effects over and above regular drift), so do the corresponding tests based upon the full age–period–cohort model. The impossibility of ascribing drift to either specifically period or specifically cohort influences must and does persist. This in turn presents us with serious problems in displaying and interpreting the estimates of the model parameters. We shall discuss this issue in the next section.

MEASURING AGE, PERIOD AND COHORT EFFECTS

The fundamental problem in interpreting parameter estimates from the age–period–cohort model, (1) or (2), is that there is no single unique solution; indeed there are infinitely many.[4–7] We have already encountered this in our first paper,[1] in which we demonstrated that the period drift and the cohort drift models lead to identical fitted rates. Thus there are clearly limitations to the interpretation of such data. We must examine what the infinitely many possible solutions have in

Table II. Three sets of age, period and cohort effects that give identical best fitting expected rates for Table I

Set no.	(1)	(2)	(3)	Set no.	(1)	(2)	(3)
Age $\alpha'_a \times 100\,000$							
25–29	0·55	0·38	0·27	Cohort $\gamma'_c \times 100$			
30–34	2·14	1·58	1·17	1880	190·3	149·7	117·8
35–39	4·30	3·38	2·66	1885	162·0	135·3	113·0
40–44	6·84	5·71	4·77	1890	133·9	118·7	105·3
45–49	9·30	8·25	7·32	1895	113·6	107·0	100·8
50–54	11·00	10·37	9.76	1900	100·0	100·0	100·0
50–59	11·23	11·23	11·23	1905	96·3	102·3	108·6
60–64	10·38	11·02	11·70	1910	94·2	106·2	119·7
65–69	9·85	11·11	12·52	1915	92·7	111·0	132·9
70–74	9·46	11·33	13·56	1920	90·3	114·8	145·9
75–79	9·43	11·99	15·24	1925	84·9	114·6	154·7
Period $\beta'_p \times 100$				1930	77·4	111·0	159·1
1955–1959	100·0	100·0	100·0	1935	69·3	105·5	160·6
1960–1964	106·2	100·0	94·2	1940	67·1	108·4	175·2
1965–1969	121·3	107·5	95·4	1945	67·8	116·3	199·6
1970–1974	147·2	123·0	102·7	1950	71·4	130·1	237·1
1975–1978	173·9	136·8	107·6				

common with one another, for it is this we may interpret. Although this is obvious, it seems to have largely escaped attention in the (futile) search for a mathematical 'solution' to the 'problem' of identifiability. Such attempts can only invite conclusions unsupported by the data. The model as specified has more parameters than may be estimated from the data. We might attempt to proceed, as previously, to find a parameterization which has α'_a representing fitted age-specific rates by choosing one reference period and setting the corresponding β'_p to 1, and a reference cohort so that one γ'_c is also taken as 1. This would leave as unknown A age parameters, $(P-1)$ period parameters and $(C-1)$ cohort parameters. Unfortunately, however, this does not work; there is no unique solution. Table II displays three possible sets of parameter estimates and it may be verified easily that each set gives an identical prediction for the observed table. In this table the reference period and cohort are $p = 1$ and $c = 5$, respectively.

These solutions are chosen to illustrate how the unwary could be led to unjustified conclusions. Incidentally, they are not too different from age relationships observed for breast cancer cross-sectionally in different countries. In the first solution the period effects show a strong increase from the first period onwards, while the cohort effects show a reverse gradient. In this solution the age curve is unusual, with rates increasing until 55–59 and thereafter decreasing. In the second solution, there is an upward trend with calendar period but a U-shaped cohort curve with a minimal risk for the 1900 cohort. The age curve shows the phenomenon of Clemmesen's hook; rates increase to a maximum at 50–54 then fall back slightly before continuing their upward trend from the age of 65 onwards. Finally, the last solution yields U-shaped period and cohort curves but, compared to the previous parameterization, a more pronounced increase over successive cohorts from 1900 onwards. This solution, has no Clemmensen's hook, the age curve being uniformly increasing.

The reason for these seemingly paradoxical results lies with the problem of drift, which, as we showed earlier, is not specifically attributable either to period or cohort effects and is described by a single parameter in addition to the age parameter, α_a. Adoption of the age–period model adds $(P-2)$ extra parameters expressing irregular period effects (see Figure 2). Likewise, adoption of the age–cohort model adds $(C-2)$ parameters to the regular age–drift model. Finally, the

age–period–cohort model includes: (i) drift, (ii) non-drift period effects and (iii) non-drift cohort effects, that is $1 + (P-2) + (C-2)$ parameters in addition to the age-curve parameters. Table II, however, purports to estimate $(P-1)$ period effects and $(C-1)$ cohort effects which include two parameters for period drift and cohort drift. We already know that the data are not capable of distinguishing between these two effects and it is not surprising that we get into difficulty when trying to estimate two indistinguishable parameters!

To clarify the position, let us look more closely at such an attempt when there are no non-drift period or cohort effects. As in our first paper[1] we shall wirte δ_P and δ_C for the parameters of period–drift and cohort–drift, respectively. Thus, without non-drift effects, the age–period–cohort model for the logarithms of the rates is

$$Y_{ap} = \alpha_a + \delta_P(p - p_0) + \delta_C(c - c_0) \tag{3}$$

where p_0, c_0 are reference period and cohort, respectively. Thus, the antilogs of δ_P and δ_C, δ'_P and δ'_C are the relative risks between adjacent periods and adjacent cohorts, respectively (constant across age).

As previously, however, the cohort passing through age group a at period p is totally determined by a and p according to the relationship $c = A - a + p$, or equivalently by $p = c + a - A$. We can substitute either of these expressions for c or for p in (3) and obtain, respectively:

$$Y_{ap} = \alpha_a - \delta_C(a - a_0) + (\delta_P + \delta_C)(p - p_0) \tag{4}$$

which is the age + period–drift model, with drift parameter $\delta = (\delta_P + \delta_C)$, and:

$$Y_{ap} = \alpha_a + \delta_P(a - a_0) + (\delta_P + \delta_C)(c - c_0) \tag{5}$$

which is the age + cohort–drift model, again with drift parameter δ.

This corresponds to the problem we first discussed in the fifth section of our first paper.[1] Not only are the (age + period–drift) and (age + cohort–drift) models indistinguishable from each other, they are also indistinguishable from any (age + period–drift + cohort–drift) model in which the net drift, $\delta = (\delta_P + \delta_C)$, is held constant. It is only this net drift which can be estimated using only the data in the age × period table of rates.

Are, then, all such models identical? Again unfortunately not – they differ in the age curves which must be assumed to represent the observed data, as may be seen by comparing (4) and (5). To identify the true age curve we therefore need to partition the net drift between age and cohort influences, and this we cannot do, at least without further information or assumptions.

In our first paper[1] we suggested the term cross-sectional age curve for the age effects estimated when fitting the age–period model and longitudinal age curve for the age effects estimated when fitting the age–cohort model. Inspection of (4) shows that the cross-sectional age curve differs in gradient from the true age curve by (minus) the cohort drift, δ_C. Likewise, the longitudinal age curve differs in gradient from the true age curve by the period drift, δ_P.

We now return to the results of Table II. These represent various parameterizations of the full age–period–cohort model. This differs from the model discussed above only in that it also includes non-drift period effects and non-drift cohort effects. The difficulties concerning drift remain; the three solutions displayed all show the same net drift, but differ according to how it is partitioned between period and cohort components. Thus, the three sets of parameters all predict identical fitted rates, but suggest different age relationships. In the first parameterization there is a strong positive period drift, a strong negative cohort drift and an inverted U age curve. The other two solutions have milder period drift compensated for by an equal increase in the cohort drift. This transfer is matched by an increase in the age gradient. Note again that the transfer of drift onto the age curve causes the shape of the curve to change, in particular local extrema may be induced or erased.

How then can we present the parameters of the age–period–cohort model? It is our belief that any parameterization in a form such as we used in Table II runs the risk of over-interpretation and should be avoided, unless extra information has effectively resolved the partition of drift between period and cohort influences. However carefully one might deal with the problem in the text of a paper, the selection of one arbitrary parameterization for a table or a graphical display can be grossly misleading. It would seem wiser to report the net drift as some overall summary of the relative risks betwen adjacent intervals, and to report only non-drift period or cohort effects. This has been suggested recently by Holford.[7]

Three methods have been proposed for presentation of period or cohort effect. We shall describe these methods in relation to period effects, but the same considerations apply for cohort effects. We start from one arbitrary parameterization, say any one of the columns in Table II. If these parameters are β_p, then we may de-trend them by adding in a log-linear drift term to give new parameters

$$\beta_p^* = \beta_p + \delta(p - p_0)$$

where we choose δ so that the resultant β_p^* are free of drift. However, this raises a question as to how we define β_p^* as being free of drift. Holford[7] suggests to interpret this such that the linear regression line of the parameters β_p^* against the periods, p, has zero slope. This has the advantage that the resulting β_p^* are identifiable, that is do not depend on the repartition of drift. An even simpler alternative is based upon drift being defined as the average of the successive first differences, $(\beta_2 - \beta_1)$, $(\beta_3 - \beta_2)$. This leads to a choice of δ such that $\beta_1^* = \beta_P^*$; the period curve is restrained to return to the same level as it commenced. Since period 1 is usually taken as reference, so that $\beta_1^* = 0$, this leads also to taking $\beta_p^* = 0$ which is in computational terms very straightforward. However, it must be kept in mind that the proper interpretation of such β_p^* is not straightforward. For instance, β_p^* obtained by the latter method should be interpreted as: $\beta_p^* = (\beta_p - \beta_1) - (p-1)(\beta_P - \beta_1)/(P-1)$. The third method derives from a consideration of what it is that defines non-drift period effects. Non-drift effects operate in such a way that the relative risks between adjacent periods are not identical. Non-drift effects are, therefore, expressible as contrasts between such relative risks. Perhaps the simplest such contrast would be the ratio of two adjacent relative risks (see Figure 3):

$$\frac{\beta'_3/\beta'_2}{\beta'_2/\beta'_1}, \quad \frac{\beta'_4/\beta'_3}{\beta'_3/\beta'_2} \cdots$$

Note that in Table II these contrasts are identical in all the parameterizations. For example,

$$\frac{\beta'_3/\beta'_2}{\beta'_2/\beta'_1} = \frac{1{\cdot}213/1{\cdot}062}{1{\cdot}062/1{\cdot}00} = \frac{1{\cdot}075/1{\cdot}00}{1{\cdot}00/1{\cdot}00} = \frac{0{\cdot}954/0{\cdot}942}{0{\cdot}942/1{\cdot}00} = 1{\cdot}075$$

so that the relative risk of period 3 versus period 2 is 8 per cent higher than that of period 2 versus period 1. This may be thought of as a measure of acceleration of period trend during the time around period 2. On the logarithmic scale, these contrasts are the 'second differences': $(\beta_3 - \beta_2) - (\beta_2 - \beta_1) = \beta_3 - 2\beta_2 + \beta_1, \beta_4 - 2\beta_3 + \beta_2$, which are well-known measures of curvature. Zero value indicates that the log-risk versus calendar time curve is locally a straight line, while positive or negative values indicate convex or concave relationships, respectively.

Figure 3 shows graphically the identifiable second differences parameter estimates for our breast cancer example. While such contrasts are unfamiliar in epidemiology, they have the important property of representing characteristics specifically attributable to age, period or cohort without any arbitrary repartition of drift components. Undoubtedly there are other possibilities for presentation of the identifiable information which might be helpful. In our example the second differences show two irregularities in the birth cohort effects, indicating sudden changes in the

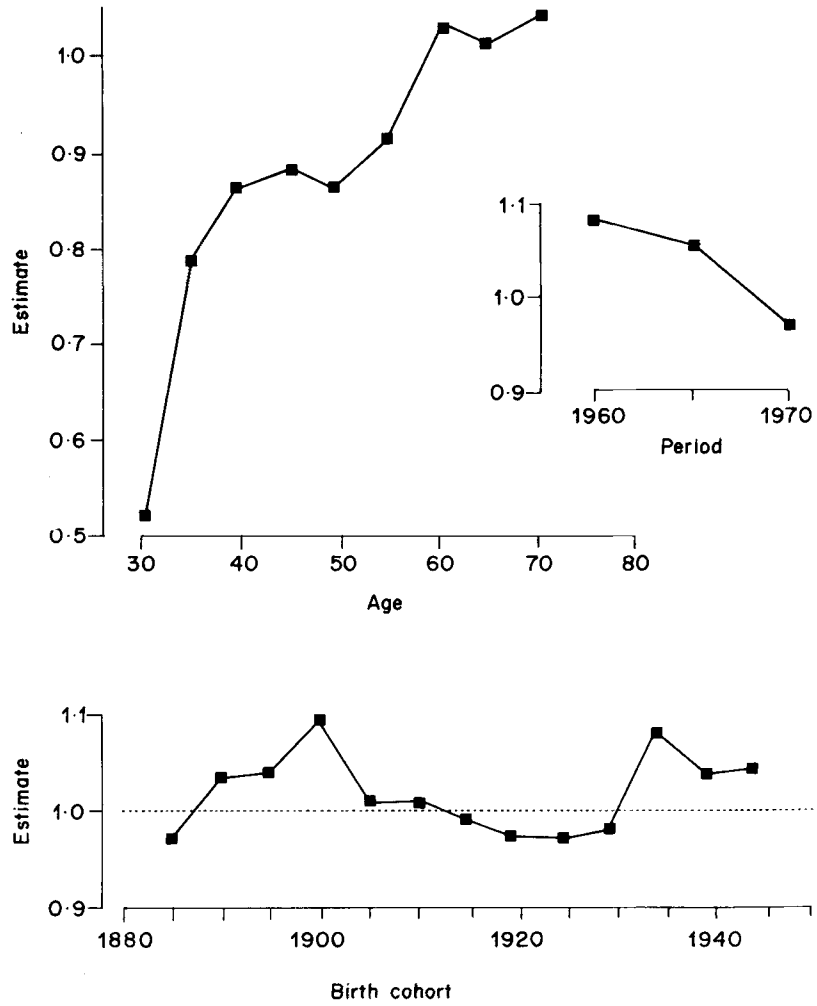

Figure 3. Mortality from breast cancer in Japan: estimates of the identifiable non-drift effects

cohort trend around 1900 and around 1935, and also the distinctive dip in the age curve around menopause, which, depending upon the partition of drift, may manifest itself as Clemmesen's hook.

These second differences have the important practical advantage that the value taken is affected by only neighbouring data. For example, the second difference around period 2 is not perceptibly influenced by trends occurring after period 3; this is not the case for the other methods we have outlined.

A NOTE ON THE EFFECT OF GROUPING

In these papers[1] we have accepted the identification of birth cohorts and diagonals of regular tables of vital rates. As stated earlier this is only approximate. Our first paper[1] draws attention to a consequence of the grouping, namely the spurious cohort effects which can result from a sudden change in birth rate. We illustrate this now.

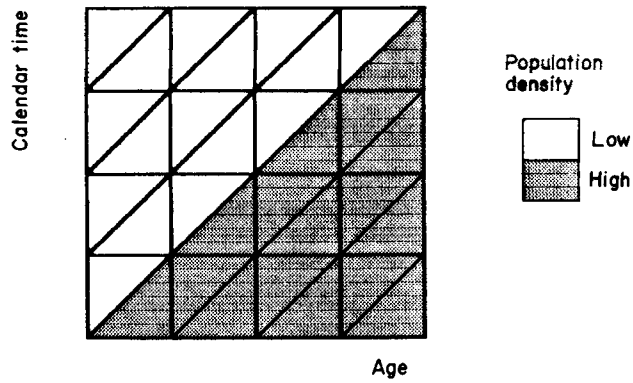

Figure 4. Lexis diagram

Figure 4 is a Lexis diagram plotting calendar time against age. The horizontal and vertical lines represent the usual age and period grouping, and birth cohorts correspond to diagonals. Suppose, for the sake of our argument, that all birth cohorts after a specified date are less numerous than before as shown by the shadings. Clearly the mean age in a square cell affected by the change in birth rates will be higher that the mid-point of the age group. Such effect will not occur elsewhere. This shift in the mean age will result in shifted (usually increased) rates. These excessive rates, aligned on a diagonal, mascarade a cohort effect.

Failure to take account of distortions due to grouping have led Boyle *et al.*[8] and more recently Boyle and Robertson[9,10] to a fallacious elimination of non-identifiability. Their argument requires access to individual records (or at least to more detailed tabulation). All arguments we presented earlier about identifiability depend in no way upon the degree of grouping, and it follows that finer grouping itself cannot resolve the non-identifiability of the model. The fallacy can be illustrated using again Figure 4. The diagonal lines delineate birth cohorts with grouping interval equal to that used for age and calendar time. Thus each age × period cell contains two triangular regions refering to adjacent birth cohorts. Consider cell (a, p), containing, say, cohorts c and $c+1$. Boyle *et al.*[8] and Boyle and Robertson[9,10] suggested modelling the log-rates in these two regions by:

$$Y_{ap}^{(1)} = \alpha_a + \beta_p + \gamma_c \tag{6}$$

$$Y_{ap}^{(2)} = \alpha_a + \beta_p + \gamma_{c+1}. \tag{7}$$

By calculating rates for the $2 \times A \times P$ triangular regions it can be shown that this model is fully identifiable. The authors claim that the fine grouping has solved the problem, but the true source of the solution is the assumption that the age and period effects are identical between the opposed triangles of each cell. Indeed the model (6) and (7) implies that the age incidence curve is a step function (similarly the period trend). If this functional form may be taken as known fact, then the model is indeed identifiable. However what if this assumption is false (as it must be)?

We have shown that the model in which period and cohort effects consist of equal and opposite drift (the EOD model, say) leads to rates which vary only with age. However, if there are no trends (the NT model) the rates also vary only with age at least when there is no constraint on the age parameters. Equations (6) and (7) imply we would be able to differentiate between these results since the EOD model predicts discrepancies between $Y_{ap}^{(1)}$ and $Y_{ap}^{(2)}$, while the NT model predicts equality. However if the age curve is not constant within cells this will also lead to discrepancies

between $Y_{ap}^{(1)}$ and $Y_{ap}^{(2)}$. Thus the truth of the conclusion to which one is led by equations (6) and (7) depends entirely upon the validity of constant incidence rates within age groups.

DISCUSSION

In this final section we address the question as to whether previous claims to have 'solved' the identifiability problem have been well founded and, if not, whether these difficulties negate the usefulness of this approach as we already stated. It is mistaken to regard the non-identifiability of drift component as a problem which needs only an advance in methodology for its solution. We have tried to demonstrate in these papers that this is a scientific problem in which our data do not discriminate between different models or explanations. A good statistical analysis will not only summarize the data in a succint and meaningful way, but will also make clear its limitations. By considering the age–period–cohort model in the context of the sequence of models set out in Figure 2, we would hope that this is achieved. Nevertheless, the view that non-identifiability is a methodological problem is prevalent as illustrated by the papers discussed in the previous paragraph as well as by the four different approaches we shall discuss next.

Recently, Osmond and Gardner[5,6] introduced a mathematical constraint in the model. Essentially, they choose one of the infinitely many possible solutions on the ground that it has certain mathematical properties. Such a strategy can only be justified if the property which identifies the unique solution has any biological basis and no such justification has been offered. Their solution, is therefore, totally arbitrary. The mathematical constraint they proposed is difficult to explain in non-technical language, but its effect is to partition the drift between period and cohort curves in a ratio which depends upon the relative magnitude of non-drift effects. Thus, if the age–cohort model is better than the age–period model then in the age–period–cohort model drift will be concentrated into the cohort effects. It is interesting to speculate on likely results of the Osmond and Gardner method when applied to data such as those for lung cancer mortality in Belgian females (see Table VIII of first paper[1]) which is well described by the age–drift model. The solution obtained would, of necessity, be determined by statistically insignificant fluctuations. There seems no scientific reason to prefer such a solution to any other.

Day and Charnay[11] considered the extra information available when analysing data from several cancer registries. They pointed out that, if the age effects may be assumed equal for different registries, then the identifiability problem is partially resolved. This assertion is undoubtedly accurate, but it is very doubtful whether one would be prepared to make such an assumption. There are different levels of carcinogenic exposure in different registry areas and it is quite conceivable that these will result in age curves of a different shape; for example, the age relationship for lung cancer differs markedly between persons with different smoking histories. A method which assumes the form of the relationship between disease rates and age to be an immutable biological constant unaffected by environmental exposure is unlikely to command widespread support.

Similarly, the identifiability problem theoretically disappears if we are prepared to assume a precise mathematical form for the age curve (an approach similar to that suggested by Boyle *et al.*,[8] and Boyle and Robertson[9,10]), provided that form does not contain a log-linear component. Now, this mathematical function must be chosen on the ground of compelling biological evidence, otherwise the whole process, even if confirmed by a good fit, amounts to a complicated but still arbitrary repartition of drift. One such curve is the Weibull law in which incidence rates are proportional to the power of age so that log-rates are linearly related to the logarithm of age rather than to age itself. This relationship is suggested by the multi-stage model for carcinogenesis[12,13] and by empirical evidence from animal carcinogenesis experiments. With the Weibull model, the

log-linear components of both cohort and period effects are identifiable. However, as already stressed, their identifiability depends upon the difference between the Weibull law and the log-linear (Gompertz) law. This difference is small and the resultant solution is unstable, the estimates of the newly identifiable linear components having very large variances and covariances. There is a further difficulty in that the empirical evidence for the Weibull law requires measurement of age not from birth but from some predefined starting point. This point may be thought of as the end of a guarantee period during which the disease may not be detected. Unfortunately the statistical information for estimation of the guarantee period is very limited since it is drawn almost entirely from the incidence rates observed at the youngest ages which are by their very nature estimated from the sparsest data. The estimates obtained for the age–period–cohort model must be expected to be heavily dependent upon the choice of guarantee period. We must therefore conclude that this approach also cannot provide a satisfactory resolution of the problem.

The last approach we consider is an extended form of age–period–cohort model, originally proposed by Moolgavkar, Stevens and Lee,[14] and discussed in detail by James and Segal.[4] In this extended model, the age effects are allowed to vary over calendar periods in such a manner that the age curve during one period, expressed as additive effects upon logarithms of rates, is a fixed multiple of the corresponding curve during another period. These multiples are an extra set of parameters over and above those required by the age–period–cohort model. Rather strangely, this extended form of the model does not suffer the same identifiability problems of the basic model. However, the model is difficult to interpret and, of course, depends upon a lack of fit of the age–period–cohort model. If the age–period–cohort model fits adequately then the extended model will degenerate to the basic form and the identifiability problem reappears. This is, therefore, not likely to prove a widely useful approach.

It is clear from the above discussion that there has been no satisfactory resolution of the problem of identifiability of log-linear trend components in age–period–cohort models. This led Kupper *et al.*[15] to conclude that, at present, such models offer little or no advantage over simple graphical methods. The same authors have recommended that future research efforts should be directed to develop and evaluate methods which bypass the identifiability problem. We would disagree with both statements. The simple facts of the information available and of the relationship between the three variables ensure that any research efforts directed at the search for this philosopher's stone of modern epidemiology is both futile and pointless. It is the purpose of statistical analysis to extract from research data the maximum information in as parsimonious and comprehensive manner as possible. No sophistication of method can create information where that information is lacking and there can surely be no other conclusion but that the observation of incidence and mortality rates in populations over time does not provide sufficient information to ascribe smooth trends to period or cohort influences with any reliability, but this is not to deny all uses for such models. In replying to the remarks of Kupper *et al.*,[15] Holford[16] pointed out that there are other aspects of such data that can be identified, and models can still provide a more parsimonious representation of the data than simply graphing the full data. While this parsimony might be considered of dubious value given the overhead of understanding necessary for interpretation of an analysis of a single table, the same cannot be said of more complex analyses over numerous registries and for many sites. It is for that purpose that we believe that the age–cohort model will continue to be of some value. Further work is necessary but should be directed at finding the most comprehensible parameterization of the model and for presenting the statistical reliability of identifiable estimates. This latter problem has been largely ignored. The problem of the analysis of tables in which the width of observation periods and of age groups is unequal (Schifflers *et al.*[17]) also requires further work, particularly as the interpretation of graphical displays is more difficult in this case.

Finally we should address a few words to the problem of forecasting future cancer rates. At first sight it might appear, since cohort risks are estimated from past observations, such forecasting is achievable without undue extrapolation. In recent years, there have been several attempts to use an age–period–cohort model fitted to past data to forecast rates. It should come as no surprise to a reader of these papers that we would in most cases doubt the wisdom of this course! Holford[16] states that, for the purpose of forecasting, the partition of drift between period and cohort components is irrelevant. Unfortunately this is only true if we are prepared to assume that the log-linear period drift which has occurred in the past will continue unchanged into the future. This is a strong assumption which will rarely be justified in practice. It is not possible to use the model to forecast under the assumption of no future period effect. It follows that forecasting is not possible without sufficient knowledge of the epidemiology of a given cancer and of the concomittant trends in population exposure to the major etiological factors to be able to resolve the underlying ambiguities. In certain situations this will be the case, for example for mesothelioma, but, when such detailed understanding is missing, we believe the place of the age–period–cohort model is in descriptive epidemiology. In this setting it has its place, provided the researcher is aware of the limits to inference from the data it is used to summarize.

ACKNOWLEDGEMENTS

This collaborative work was supported by I.A.R.C. (Lyon). We should like to thank Dr. C. S. Muir in particular for his support and encouragements. We also thank Mr. A. De Coninck for computing assistance and Miss F. De Carli for typing the manuscript.

REFERENCES

1. Clayton, D. and Schifflers, E. 'Models for temporal variation in cancer rates. I: Age–period and age–cohort models', *Statistics in Medicine*, **6**, 449–467 (1987).
2. Breslow, N. E. and Day, N. E. 'Indirect standardization and multiplicative models for rates, with reference to the age adjustment for cancer incidence and relative frequency data', *Journal of Chronic Diseases*, **28**, 289–303 (1975).
3. Breslow, N. E. and Day, N. E. *Statistical Methods in Cancer Research, Vol. I—The Analysis of Case–Control Studies*, International Agency for Research on Cancer, Lyon, (IARC Scientific Publications No. 32), 1980.
4. James, I. R. and Segal, M. R. 'On a method of mortality analysis incorporating age–year interaction, with application to prostate cancer mortality', *Biometrics*, **38**, 433–443 (1982).
5. Osmond, C. and Gardner, M. J. 'Age, period and cohort models applied to cancer mortality rates', *Statistics in Medicine*, **1**, 245–259 (1982).
6. Gardner, M. J. and Osmond, C. 'Interpretation of time trends in disease rates in the presence of generation effects', *Statistics in Medicine*, **3**, 113–130 (1984).
7. Holford, T. R. 'The estimation of age, period and cohort effects for vital rates', *Biometrics*, **39**, 311–324 (1983).
8. Boyle, P., Day, N. E. and Magnus, K. 'Mathematical modelling of malignant melanoma trends in Norway 1953–1978', *American Journal of Epidemiology*, **118**, 887–896 (1982).
9. Boyle, P. and Robertson, C. 'Statistical modelling of lung cancer and laryngeal cancer incidence in Scotland, 1960–1979', *American Journal of Epidemiology* (in press).
10. Robertson, C. and Boyle, P. 'Age, period and cohort models: the use of individual records', *Statistics in Medicine*, **5**, 527–538 (1986).
11. Day, N. E. and Charnay, B. 'Time trends, cohort effects and aging as influence on cancer incidence', in Magnus, K. (ed.) *Trends in Cancer Incidence: Causes and Practical Implications*, Hemisphere Publishing Corporation, 1982, pp. 51–65.
12. Armitage, P. and Doll, R. 'Stochastic models for carcinogenesis', *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, **4**, 19–39 (1961).
13. Day, N. E. and Brown, C. C. 'Multistage models and primary prevention of cancer', *Journal of the National Cancer Institute*, **64**, (4), 977–989 (1980).

14. Moolgavkar, S. H., Stevens, R. G. and Lee, J. A. H. 'Effect of age on incidence of breast cancer in females', *Journal of the National Cancer Institute*, **62**, (3), 493–501 (1979).
15. Kupper, L. L., Janis, J. M., Karmous, A. and Greenberg, B. G. 'Age–period–cohort analysis: a review and critique', *Journal of Chronic Diseases*, **38**, 10, 811–830 (1984).
16. Holford, T. R. 'An alternative approach to statistical age–period–cohort analysis', *Journal of Chronic diseases*, **38**, 10, 831–840 (1985).
17. Schifflers, E., Smans, M. and Muir, C. S. 'Birth cohort analysis using irregular cross-sectional data: a technical note', *Statistics in Medicine*, **4**, 63–75 (1985).