Age period cohort analysis of cancer mortality data: methods and application to Italian male mortality data for gastric cancer and cancers of the oral cavity and pharynx

Matteo Malvezzi[1,2,3], Paola Bertuccio[1,2], Valeria Edefonti[1]

[1] Dipartimento di Medicina del Lavoro "Clinica del Lavoro Luigi Devoto", Sezione di Statistica Medica e Biometria "Giulio A. Maccacaro", Università degli Studi di Milano, Via Vanzetti 5, 20133 Milan, Italy.
[2] Istituto di Ricerche Farmacologiche "Mario Negri", Via La Masa 19, 20156 Milan, Italy.
[3] Struttura Complessa di Statistica Medica, Biometria e Bioinformatica, Fondazione IRCCS Istituto Nazionale Tumori di Milano, Via Venezian 1, 20133 Milan, Italy.

Corresponding Author:
Dr. Matteo Malvezzi,
Dipartimento di Medicina del Lavoro "Clinica del Lavoro Luigi Devoto", Sezione di Statistica Medica e Biometria "Giulio A. Maccacaro", Università degli Studi di Milano, Via Vanzetti 5, 20133 Milan, Italy.
Tel.: +39-0250320873; Fax:+39-0250320866
E-mail: matteo.malvezzi@unimi.it

Summary

Objectives. The paper's objective is to provide an in depth age period cohort analysis of gastric and oral cancers in Italian men.

Methods. Mortality data from gastric cancer and oral cavity and pharynx cancers in Italian men was obtained from the WHO mortality database for the period 1950-2003. An in depth age period cohort analysis was performed on the data.

Results. Analysis of gastric cancer mortality showed descending trends for both period and cohort effects. For oral cancer, period effects rose steeply until the late 1980's to then descend to the last period; cohorts had a downward trend up to 1910 to then rise up the 1960's, displaying a descent for the effects of the last two birth cohorts.

Conclusions. The descending effects of period and cohort from gastric cancer in men reflect the effect of the improving quality of life in Italy and predict further falls in mortality.

Period effects in oral cancer mortality reflect the influence of smoking cessation in Italian men, with a downward trend for the most recent years. The cohort effects are probably influenced by alcohol consumption, which has longer lasting consequences, and only show a favourable pattern for the most recent cohorts.

KEY WORDS: *Age period cohort analysis, mortality, gastric cancer, oral cancer, Italy.*

Introduction

Until the mid-1990's, gastric cancer was the most common cause of cancer death worldwide. Rates have been declining substantially for several decades, and gastric cancer has become a relatively rare cancer in Western European countries (1). In Italian males, mortality from this cancer remains common, with nearly 6,400 certified deaths in 2003, and a standardised death rate of about 10/100,000, accounting for nearly 7% of male cancer mortality (2).

Oral and pharyngeal cancer is the seventh most common cancer and the ninth site of cancer death in Europe, with an estimated number of 67,000 incident cases and 26,000 deaths in 2004 (3). In Italy, male mortality from oral and pharyngeal cancer peaked in the late 1980's and decreased from around 6/100,000 to about 4/100,000 in 2003 (2, 4). Italy had one of the hi-

ghest levels of alcohol consumption in Europe up to the early 1980's, but this has substantially declined over the last few decades (5), resulting in favourable trends in oral cancer mortality over the last 20 years (6).

Standardised rates are a useful descriptive tool when comparing different populations, but offer no analytical insight. On the other hand, joinpoint analysis offers a form of analytical insight on trends and their changes in time, but completely ignores information on the cohort structure of the analysed data (7). Age-period-cohort (APC) analysis performs a simultaneous study of the effects of age, period and cohort (8). The age effects correspond to the variability explained by physiological differences that characterise the different age groups. Period effects usually portray the consequences of the introduction of new therapies or screening interventions; they are also affected by exposures that manifest their consequences on the studied event over short periods of time. Cohort effects are usually determined by gestational exposures, or by those exposures that influence event frequencies over long time periods.

However useful, the APC model suffers from an intrinsic structural issue: the age, period and cohort variables have an exact linear dependence (i.e. age=period–cohort (A=P-C)), which causes the identifiability problem that makes this model complicated to treat (9). In recent decades, many efforts have been spent in an attempt to overcome this issue in both parametric and non parametric contexts (10). A few examples are the use of estimable functions, the addition of extra constraints to the model or the use of a two-stage cascade approach, performing sequential fittings of a two-factor model and then regressing over the residuals with the remaining factor. More recently, techniques involving smoothers and Bayesian methods have also been introduced (10-13). All these solutions have their merits and setbacks: estimable functions are epidemiologically hard to interpret; adding constraints requires biological and statistical *a priori* knowledge and is not always justifiable; cascade regression estimates do not always agree with each other and also require biological justifiability. In this paper, we propose a step by step approach to APC analysis that includes the use of a likelihood penalising function (14, 15), which offers a good balance between trade-offs and functionality, while maintaining the characteristics of a factorial model. We illustrate the use of this method applying it to Italian male mortality data for gastric and oral cancers.

Materials and methods

Italian male certified deaths from gastric cancer and cancers of the oral cavity and pharynx were derived from the WHO mortality database (WHOSIS) for the 1950-2003 period stratified by 5-year age groups (16). Estimates of the resident population, based on official censuses, were obtained from the same WHO database. During the period considered, four different revisions of the International Classification of Diseases (ICD7, ICD8, ICD9 and ICD10) were used. Classification of cancer deaths was recoded for all calendar periods according to the 9th revision of the ICD (17).

From these matrices, we calculated age-specific mortality rates per 100,000 inhabitants for 5-year age groups (from 30-34 to 75-79 years) and 5-year periods (from 1950-54 to 2000-03, data for 2004 was not available). Cohorts were defined according to their central year of birth. Hence, the earliest possible cohort is the one centred in 1875 and includes subjects that could have been born in any of the 10 years between 1870 and 1879, which are individuals aged 75 to 79 who died in the quinquennium 1950-54.

To illustrate APC model issues, in their methodological review Clayton and Schifflers tackle the problem starting from performing analysis of variance (ANOVA) of successively more complex Poisson models (18, 19). They start with a simple age model, moving on to a linear drift model that adds either a cohort or a period drift (which are equivalent for the purposes of ANOVA) to the age model, to then apply the age-period (AP) and age-cohort (AC) models and only in a final instance, if justified by the ANOVA, apply one of the possible APC solutions (18, 19).

The age model analyses the rate trends using age as a factor and can be expressed with the following expression:

$$\log(E[r_i]) = \alpha_i$$

where α_i are the log-estimates of the age-specific rates r_i, which can be expressed as $EXP(\alpha_i)$*100,000 inhabitants.

The AP model assumes that the age-specific rates maintain the same structure during all the studied periods, but vary according to the time period, and can be written as:

$$\log(E[r_{ij}]) = \alpha_i + \beta_j$$

where α_i is the log-estimate of the i-th age-specific rate

in the reference period β_0, β_j is the additive effect of the j-th period on the logarithm of the age-specific rates, with r_{ij} being the estimate of the rate of the i-th age group in the j-th period.

This model is over-parameterised as it has one parameter for each age group and one for each period, but this is easily solved by constraining one of the period effects to be equal to 0 and setting it as the reference β_0.

The AC model is structurally similar to the AP model, but replaces the period parameters with the cohort ones:

$$\log(E[r_{ik}]) = \alpha_i + \gamma_k$$

where α_i is the log-estimate of the i-th age-specific rate in the reference cohort γ_0, γ_k is the additive effect of the k-th cohort on the logarithm of the age-specific rates, and r_{ik} is the estimate of the rate of the i-th age group in the k-th cohort. Similarly to the previous model, the over parameterisation problem is solved taking a reference cohort γ_0, referred to which γ_k is the log-relative risk for the k-th cohort.

Observing the previous two models, the question arises as to what would happen if the period and cohort effects estimates were substituted with a single log-linear trend. This is called the linear drift model, where the expected rates, on a logarithmic scale, can be expressed either as a function of period or as a function of cohort:

$$\log(E[r_{ij}]) = \alpha_i + \beta(j - j_0)$$

$$\log(E[r_{ik}]) = \alpha_i + \gamma(k - k_0).$$

The resulting effects are age-specific rates calculated in the reference period or cohort (j_0 or k_0 respectively) and a linear trend component (on a logarithmic scale) representing the linear drift.

The main feature of these two models is that they are equivalent as far as the expected value estimates are concerned. The models provide an age-specific rate structure that is influenced either by a constant linear cohort effect or by a constant linear period effect.

In the case where age is the most important scale (as is the case in most cancer mortality studies) and cohort or period can be prioritised according to *a priori* biological or statistical information, a full APC model can be approximated by regressing an AP or AC model first, to then obtain an estimate of the effect of the remaining factor by regressing it over the residuals of the first model.

In the case of an AC model followed by a period model, what is being obtained is the residual period effect conditional the age and cohort estimates from the preceding model:

$$\log(E[r_{ij}]) = \hat{\alpha}_i + \hat{\gamma}_k + \beta_j$$

which for the expected cases becomes:

$$\log(E[O_{ij}]) = \hat{\alpha}_i + \hat{\gamma}_k + \log(N_{ij}) + \beta_j$$

where O_{ij} and N_{ij} are the observed cases and populations for the i-th age group and j-th period respectively. This is the general expression of a Poisson model, where the offset is now given by the log of the fitted values from the AC model.

This procedure estimates marginal age and cohort effects and conditional period effects. Accordingly, the corresponding confidence intervals result from marginal standard error estimates for the age and cohort effects and conditional standard error estimates for the period effects. This procedure can obviously be used starting with an AP model and then regress the cohort effects. This method can also be used by fitting one of the possible age-drift models and then regressing over its residuals to obtain estimates for the remaining parameters . In the case of an age-drift-period model, period estimates conditional to the age and cohort drift ones are provided. This translates to:

$$\log(E[r_{ij}]) = \hat{\alpha}_i + \hat{\gamma}(k - k_0) + \beta_j$$

and for the expected cases:

$$\log(E[O_{ij}]) = \hat{\alpha}_i + \hat{\gamma}(k - k_0) + \log(N_{ij}) + \beta_j .$$

This gives us a set of period effect estimates unaffected by linear drift, as this has already been absorbed by the first model. The same can be done to obtain an age-drift-cohort model where the cohort estimates are drift free. A full APC model is written as:

$$\log(E[r_{ij}]) = \alpha_i + \beta_j + \gamma_k .$$

To fit this model the identifiability problem must be resolved or worked around. The solution that we propose here is to use a likelihood penalising function to average the three possible two-factor models (14, 15). A brief explanation of the method follows.

An APC model can be considered as a parameter estimate problem in a log-linear Poisson model, such as those illustrated up till now:

$$\log(E[O_{ij}]) = \log N_{ij} + \log a_i + \log p_j + \log c_k$$

where a_i (i=1, ..., I), p_j (j=1, ..., J) and c_k (k=1, ..., K) are the multiplicative parameters for age, period and co-

hort respectively. These can be estimated by minimising the following expression using the weighted least squares method:

$$f(a,p,c) = \sum O_{ij}(\log O_{ij} - \log N_{ij} - \log a_i - \log p_j - \log c_k)^2 .$$

Due to the linear relation between A, P and C, the solution set X (a, p, c) is infinite. To work around this issue the solution set can be re-parameterised in λ as X(λ)=(a, p, c, λ) given by:

$$\log \hat{a}'_i = \log \hat{a}_i + \lambda(I - i)$$

$$\log \hat{p}'_j = \log \hat{p}_j + \lambda j$$

$$\log \hat{c}'_k = \log \hat{c}_k - \lambda k .$$

This parameterisation makes it possible to calculate a goodness of fit statistic (G^2) that is independent from λ.

The identifiability problem is then solved by taking the solutions estimated by the three two-factor models:

$$X_c = (\hat{a}_c, \hat{p}_c, c_0); X_p = (\hat{a}_p, p_0, \hat{c}_p); X_a = (a_0, \hat{p}_a, \hat{c}_a)$$

where c_0 and p_0 are unit vectors of length K and J respectively, and a_0 is a vector whose elements take the form:

$$a_{0i} = \exp\left[\frac{\sum_j O_{ij}(\log O_{ij} - \log N_{ij})}{\sum_j O_{ij}}\right],$$

and placing the natural logarithms of the estimates of these solutions in the real space R^{I+J+K}. Their Euclidean distances from the λ- parameterised saturated model solutions X(λ) are defined as:

$$d_c(\lambda) = \|X_c - X(\lambda)\|$$

$$d_p(\lambda) = \|X_p - X(\lambda)\|$$

$$d_a(\lambda) = \|X_a - X(\lambda)\| .$$

The sum of these distances weighted with a function of goodness of fit and degrees of freedom:

$$g(X) = \frac{d_c(\lambda)}{\frac{G_c^2}{(I-1)(J-1)}} + \frac{d_p(\lambda)}{\frac{G_p^2}{(I-1)(J-2)}} + \frac{d_a(\lambda)}{\frac{G_a^2}{(I-2)(J-1)}}$$

can be minimised in λ. This action gives a solution X'(λ) that minimises the distance of the saturated model from the three two-factor models, constructing a geometrical weighted average. Consequently, the drift is distributed according to the goodness of fit statistics.

This modelling technique does not allow for the calculation of confidence intervals in a conventional manner, hence, an *ad-hoc* parametric bootstrap simulation technique was implemented. Data for each 5-year age-specific number of deaths in all time periods was obtained by random extractial from a Poisson distribution characterised by the observed number of deaths for that period and age-group (20). The resulting datasets are fed through the model and, for every parameter, values for the 2.5th and 97.5th percentiles from the resulting estimates are taken as an approximation of a 95% confidence interval. For this paper 1,000 simulations per model were used.

Results

Table 1 shows ANOVA results from the successively more complex Poisson models as applied to mortality data from the two cancer sites in study. As would be expected from cancer mortality data, the age model gives a very strong fit for both cancers. The same cannot be said for the linear drift models: gastric cancer shows a valuable improvement in fit, while for oral cancer there is very little change in fit compared to the simpler age model. Both datasets respond well to the addition of period factors to the model, while the addition of cohort factors to the age models only seems to be useful with gastric cancer mortality data. The full APC model has an edge over the preceding simpler models for both cancers, but works better on oral cancer data, where the AC and linear drift models show a weaker fit.

Figures 1 and 2 show, respectively for male gastric and oral cancer mortality, log-scale plots of: age-specific and cohort-specific mortality rates plotted against period of death, the AP and AC two-factor models, the two-stage cascade fit age-drift period and age-drift cohort models and the full APC model fitted using the penalised likelihood method.

The first plot shows that gastric cancer age-specific and cohort-specific death rates have strong parallelism and a definite log-linear behaviour, justifying the good results obtained from the age drift model fit in ANOVA. A more complex structure emerges from the same plot with oral cavity and pharynx cancer death rates. These behaviours are also reflected in the two-factor AP and AC models, where, for gastric cancer, cohort and

Table 1. Analysis of variance results for regression models applied to gastric cancer and oral cavity and pharynx cancers.

	Gastric Cancer		Oral Cavity and Pharynx Cancers	
Models	**Degrees of Freedom**	**Deviance**	**Degrees of Freedom**	**Deviance**
Null Model	110	1479717519	110	1484057559
Age-Model	100	70183	100	3746
Age-Drift Model	99	4714	99	3725
Age-Period Model	90	1713	90	2407
Age-Period-Cohort Model	72	306	72	187
Null Model	110	1479717519	110	1484057559
Age-Model	100	70183	100	3746
Age-Drift Model	99	4714	99	3725
Age-Cohort Model	81	927	81	3215
Age-Period-Cohort Model	72	306	72	187

period factors display little variation along a linear (in a logarithmic scale) descending trend in effects, while in the case of oral cancer they reflect the afore mentioned complexity. The period effect estimates rise up to the late 1980's then showing a definite linear downward trend; on the other hand, estimates for cohort effects seem to have an overall descending trend, albeit displaying an oscillation centred around the reference birth cohort with central year in 1920. The confidence intervals for these estimates widen exponentially towards the latest cohorts that are composed of fewer cells and have smaller numbers of recorded deaths, due to the younger ages involved.

The age-drift models followed by period and cohort regressions over residuals display what the period and cohort effect estimates would look like devoid of the linear drift term. Gastric cancer data, which gave a very strong response to the drift model (see in table 1), shows cohort and period conditional estimates that differ from the simple two-factor models in that, instead of displaying descending monotone linear trends, they have a rising slope in the earliest birth cohorts and periods of death, a brief plateau, and then take to falling faster and further than the corresponding two-factor model effects. In oral cancer, period effects maintain the shape of the two-factor model, the only difference being that the initial rise and successive fall are both sharper and longer. The cohort conditional estimates change similarly to gastric cancer: the approximately linear downward trend changes into an initial rise until the late nineteenth century birth cohort, where it reaches a plateau that lasts until the 1930's to then fall steeply up to the most recent cohorts.

The full APC model fitted using the penalised likelihood method gives the full picture with a single set of estimates. As expected from the ANOVA results for gastric cancer, the period and cohort effect estimates follow the ones from the AP and AC models closely, displaying a descending tendency throughout the periods and cohorts studied. The same analysis performed on the oral cavity and pharynx cancer mortality data gives a more complex picture. The period effects rise up to the late 1980's to then descend up to the early 2000's similarly to the age-drift period model ones. On the other hand, the cohort effect estimates are transformed by the full APC model, maintaining the basic structure of those from the AC model, but, where this model could be considered approximately linear, this new set of estimates has a strong fall in effects up to the cohort centred in 1910, to then rise sharply up to those born around 1960, and finally shows a fall in effects for the last cohort. This change is explained by the poor ANOVA results for the AC model, which let the APC model give greater weight to the other effects, as reflected by the very wide intervals on the last cohorts.

Discussion

Due to the identifiability problem, APC analysis is an intricate process. The steps illustrated in the methods section and the results of their application to Italian male mortality data for gastric and oral cancers are meant to demonstrate this complexity, as well as its usefulness. Particularly for APC analysis, the investigation process should always start from the examination of the raw data

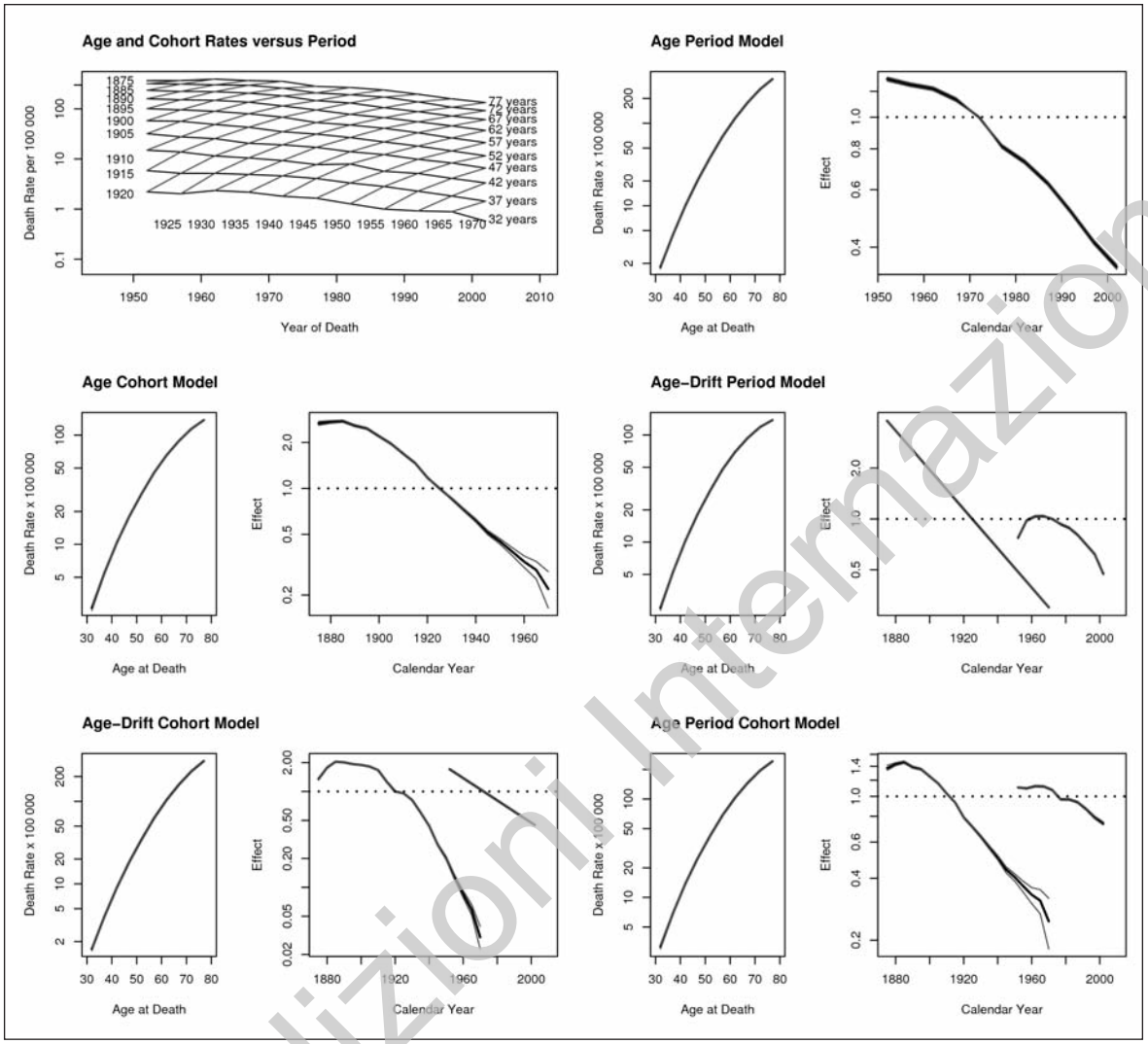

Figure 1. Italian male gastric cancer mortality log-plots of: age-specific (black lines) and cohort-specific (grey lines) mortality rates plotted against period of death, the AP and AC two-factor models, age-drift period model, age-drift cohort model and the full APC model (black lines: effect estimates, grey lines: 95% effect confidence intervals).

in its tabular form, followed by simple graphing techniques, followed by fitting successively more complex models and performing ANOVA. Only then should a full APC model be applied, whether it is the penalised likelihood technique described here, or one of the recently proposed techniques involving smoothers or Bayesian methods (10-13, 15).

Before proceeding to the interpretation of results, a few general considerations on the chosen model are due. The interpretation of age and cohort estimates should be carried out with due caution, whereas variance issues are minor when related to period of death values, as these are built on relatively similar numbers over subsequent calendar periods. For age values, the issue tends to manifest with the younger age groups, as the absolute numbers of deaths for these pathologies are low. In cohort effects, these problems are potentially greater at both ends of the curve: as reflected by the simulated intervals, the earliest and latest cohorts are only based on a single observation, while the number of observations go up moving towards the central cohorts. In addition to this, the more recent cohorts are based on smaller numbers of deaths because they represent the youngest age groups. It follows that changes in trends in these recent cohorts should be interpreted with caution, even though they provide important information towards future trends.

Another limit of this APC model is that it has difficulties

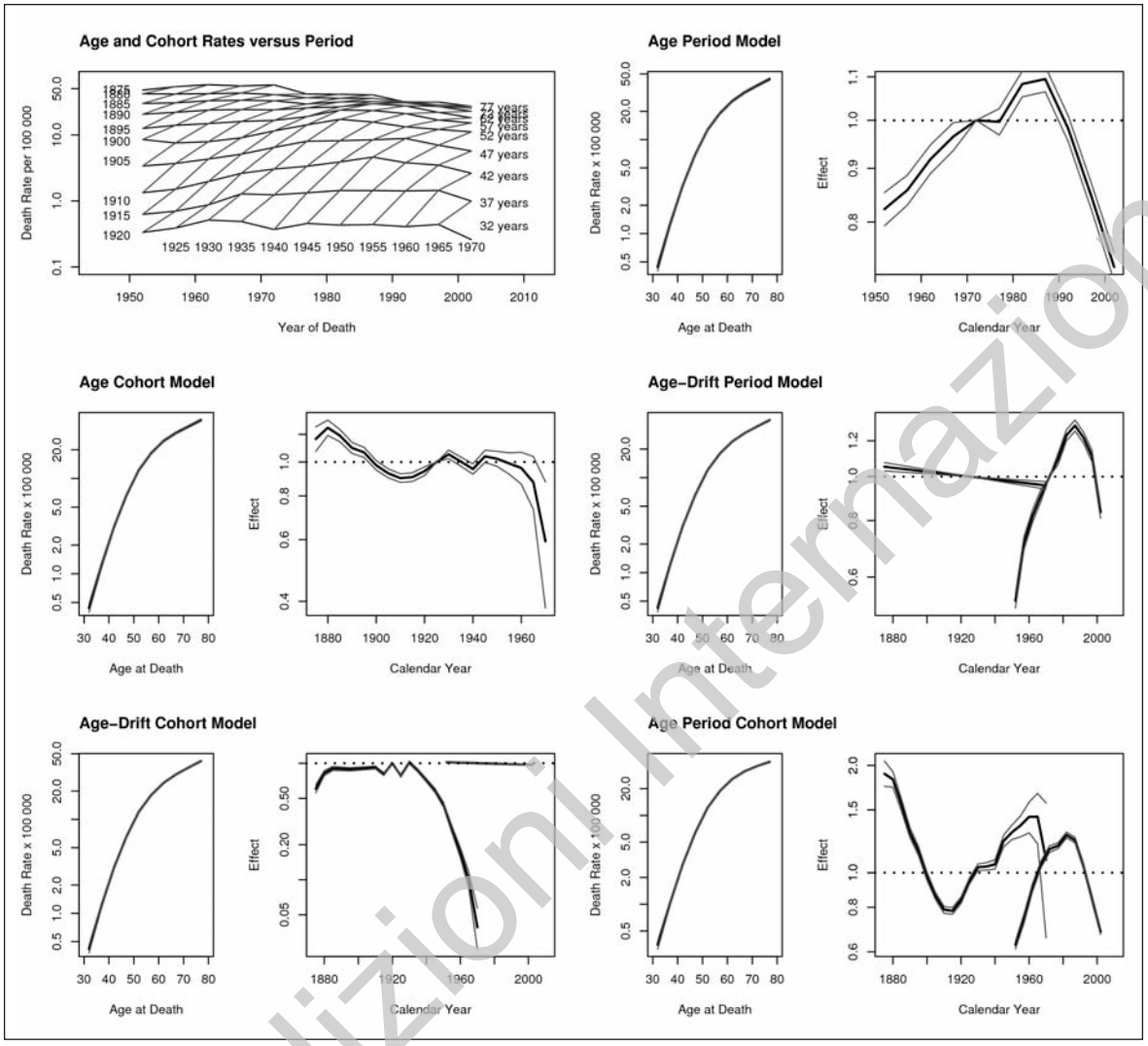

Figure 2. Italian male oral cavity and pharyngeal cancer mortality log-plots of: age-specific (black lines) and cohort-specific (grey lines) mortality rates plotted against period of death, the AP and AC two-factor models, age-drift period model, age-drift cohort model and the full APC model (black lines: effect estimates, grey lines: 95% effect confidence intervals).

discerning whether the major underlying trend is a cohort or a period one when both their estimated effects share the same direction (21), as in the case of the downward trends in effects recorded from the analysis of gastric cancer mortality. This model also has a systematic tendency to favour cohort effects as they have greater weight due to their larger number of categories.

The causes behind the favourable trends in cohort of birth and period of death effects in gastric cancer mortality are not clearly understood. It is highly probable that they reflect the effects of: a diet that is more affluent and richer in fresh fruit and vegetables, better food conservation (including refrigeration) as well as better hygiene, with a lower level of *Helicobacter pylori* infection (22, 23). An association with a reduced consumption of salt and salted foods has also been observed consistently (24-26). Other methods of food conservation, such as smoking and curing, have also been connected to stomach cancer, but there is less consistent evidence. Tobacco smoking also has an important role in stomach cancer mortality and has been estimated to be responsible for about 10% of all recorded cancer cases (27).

The favourable period effect trend may also be explained by the effects of improved and newly developed methods of diagnosis and treatment, but this effect remains hard to quantify.

The favourable trends, seen in both cohort and period

effects, are essentially the result of a more developed socio-economic environment and healthier lifestyle, and don't seem to slow down even in the most recent cohorts. Therefore, if the positive lifestyle and socio-economical conditions are maintained and further nurtured, there should be room for additional improvements in mortality from gastric cancer in Italian men.

The main risk factors for oral and pharyngeal cancer mortality are tobacco smoking and alcohol consumption (27, 28). Prevalence of tobacco smoking in Italian men has been falling consistently in the last decades and the effect of smoking cessation is evident within very few years (29-31), as portrayed by the period effects of the APC model. Conversely, risks associated with alcohol drinking are persistent for several years. Therefore, even though alcohol consumption has been falling over the last decades, the effect may only be apparent for the most recent cohorts (32, 33). The role of other factors is likely smaller and remains largely unquantified. A dietary aspect that might have influenced oral and pharyngeal cancer mortality is low consumption of fruit and vegetables, which may account for a non negligible fraction of these neoplasms in Italy (34, 35). Low consumption of β-carotene, considered an indicator of fruit and vegetable intake, accounted for 24% of oral and pharyngeal cancers in a study conducted in Italy (36).

It is unlikely that trends in oral and pharyngeal cancer have been appreciably influenced by changes in diagnosis and improved certification of the disease, since oral and pharyngeal cancer is relatively easy to diagnose, and no major change in certification has been introduced for this neoplasm across the subsequent ICD revisions considered in the present analysis.

Considering the afore mentioned risk factors and the results of the APC analysis it is reasonable to expect further decreases in mortality from oral cancer in Italian men as long as the measures for tobacco and alcohol consumption control are maintained and reinforced.

APC analysis is an extremely useful tool in the study of mortality data, particularly in the study of cohort effects, but it must be used with caution in order to avoid erroneous conclusions. From the analysis of Italian male mortality data from gastric cancer and cancers of the oral cavity and pharynx, the importance of efforts towards the successful control of smoking and drinking habits emerges, as well as the role of a healthy diet that is rich in fruit and vegetables, and that of socio-economic progress on hygienic conditions.

Acknowlegements

This work was conducted with the contribution of the Italian Association for Cancer Research (AIRC). Paola Bertuccio and Matteo Malvezzi were supported by a fellowship from the Italian Foundation for Cancer Research (FIRC). The authors thank Professor Rocco Micciolo for providing the R macro for APC analysis adapted from the original in GLIM.

References

1. Bertuccio P, Chatenoud L, Levi F, Praud D, Ferlay J, Negri E, Malvezzi M, La Vecchia C. Recent patterns in gastric cancer: a global overview. Int J Cancer 2009;125:666-73.
2. Malvezzi M, Bertuccio P, Chatenoud L, Negri E, La Vecchia C, Decarli A. Cancer mortality in Italy, 2003. Tumori 2009;95:655-64.
3. Boyle P, Ferlay J. Cancer incidence and mortality in Europe, 2004. Ann Oncol 2005;16:481-88.
4. La Vecchia C, Lucchini F, Negri E, Levi F. Trends in oral cancer mortality in Europe. Oral Oncol 2004;40:433-39.
5. Bosetti C, Levi F, Lucchini F, Zatonski WA, Negri E, La Vecchia C. Worldwide mortality from cirrhosis: An update to 2002. J Hepatol 2007;46:827-39.
6. Malvezzi M, Bosetti C, Negri E, La Vecchia C. Cancer mortality in Italy, 1970-2002. Tumori 2008;94:640-57.
7. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335-51.
8. Holford TR. The estimation of age, period and cohort effects for vital rates. Biometrics 1983;39:311-24.
9. Frost WH. The age selection of mortality from tuberculosis in successive decades. 1939. Am J Epidemiol 1995;141:4-9; discussion 3.
10. Carstensen B. Age-period-cohort models for the Lexis diagram. Stat Med 2007;26:3018-45.
11. Yang Y, Schulhofer-Wohl S, Fu W, Land K. The intrinsic estimator for age-period-cohort analysis: What it is and how to use it. AJS 2008;113:1697-736.
12. Held L, Schmid V. Bayesian age-period-cohort modeling and prediction - BAMP. J Stat Softw 2007;21.
13. Smith HL. Advances in age-period-cohort analysis. Sociological Methods Research 2008;36:287-96.
14. Osmond C, Gardner MJ. Age, period and cohort models applied to cancer mortality rates. Stat Med 1982;1:245-59.
15. Decarli A, La Vecchia C. Age, period and cohort models: a review of knowledge and application in GLIM. Rivista di Statistica Applicata 1987;20:397-410.
16. World Health Organization Statistical Information System. WHO mortality database. Available at: http://www3.who.int/whosis/menu.cfm, 2007.
17. World Health Organization. International Classification of Disease: 9th revision. Geneva, 1977.

18. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: Age-period and age-cohort models. Stat Med 1987;6:449-67.
19. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. Stat Med 1987;6:469-81.
20. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. Boca Raton: Chapman & Hall/CRC, 1993.
21. Robertson C, Gandini S, Boyle P. Age-period-cohort models: a comparative study of available methodologies. J Clin Epidemiol 1999;52:569-83.
22. Parkin DM. The global health burden of infection-associated cancers in the year 2002. Int J Cancer 2006;118:3030-44.
23. IARC. IARC Handbooks of Cancer Prevention. Fruit and vegetables, Vol. 8 Lyon: International Agency for Research on Cancer, 2003.
24. Shikata K, Kiyohara Y, Kubo M, Yonemoto K, Ninomiya T, Shirota T, Tanizaki Y, Doi Y, Tanaka K, Oishi Y, Matsumoto T, Iida M. A prospective study of dietary salt intake and gastric cancer incidence in a defined Japanese population: the Hisayama study. Int J Cancer 2006;119:196-201.
25. Negri E, La Vecchia C, D'Avanzo B, Gentile A, Boyle P, Franceschi S. Salt preference and the risk of gastrointestinal cancers. Nutr Cancer 1990;14:227-32.
26. La Vecchia C, Negri E, Franceschi S, Decarli A. Case-control study on influence of methionine, nitrite, and salt on gastric carcinogenesis in northern Italy. Nutr Cancer 1997;27:65-8.
27. IARC. IARC Monographs on the evaluation of carcinogenic risks to humans. Tobacco smoke and involuntary smoking. Vol. 8 Lyon: International Agency for Research on Cancer, 2004.
28. IARC. IARC Monographs on the evaluation of carcinogenic risks to humans. Alcohol drinking. Vol. 44. Lyon: International Agency for Research on Cancer, 1988.
29. Gallus S, Zuccaro P, Colombo P, Apolone G, Pacifici R, Garattini S, Bosetti C, La Vecchia C. Smoking in Italy 2005-2006: effects of a comprehensive national tobacco regulation. Prev Med 2007;45:198-201.
30. La Vecchia C, Franceschi S, Bosetti C, Levi F, Talamini R, Negri E. Time since stopping smoking and the risk of oral and pharyngeal cancers. J Natl Cancer Inst 1999;91:726-28.
31. Bosetti C, Gallus S, Peto R, Negri E, Talamini R, Tavani A, Franceschi S, La Vecchia C. Tobacco smoking, smoking cessation, and cumulative risk of upper aerodigestive tract cancers. Am J Epidemiol 2008;167:468-73.
32. World Health Organization Statistical Information System. Health topics. Alcohol drinking. Available at: http://www.who.int/topics/alcohol_drinking/en/, 2006.
33. Franceschi S, Levi F, Dal Maso L, Talamini R, Conti E, Negri E, La Vecchia C. Cessation of alcohol drinking and risk of cancer of the oral cavity and pharynx. Int J Cancer 2000;85:787-90.
34. Lucenteforte E, Garavello W, Bosetti C, La Vecchia C. Dietary factors and oral and pharyngeal cancer risk. Oral Oncol 2009;45:461-67.
35. Kreimer AR, Randi G, Herrero R, Castellsague X, La Vecchia C, Franceschi S. Diet and body mass, and oral and oropharyngeal squamous cell carcinomas: analysis from the IARC multinational case-control study. Int J Cancer 2006;118:2293-297.
36. Negri E, La Vecchia C, Franceschi S, Tavani A. Attributable risk for oral cancer in northern Italy. Cancer Epidemiol Biomarkers Prev 1993;2:189-93.