Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2011 July ; 20(7): 1263–1268. doi:10.1158/1055-9965.EPI-11-0421.

Age-Period-Cohort Models in Cancer Surveillance Research: Ready for Prime Time?

Philip S. Rosenberg, PhD¹ and William F. Anderson, MD, MPH^{2,*}

¹Division of Cancer, Epidemiology & Genetics, National Cancer Institute

²Division of Cancer Epidemiology and Genetics, National Cancer Institute

Abstract

Standard descriptive methods for the analysis of cancer surveillance data include canonical plots based on the lexis diagram, directly age-standardized rates (ASR), estimated annual percentage change (EAPC), and joinpoint regression. The age-period-cohort (APC) model has been used less often. Here, we argue that it merits much broader use. Firstly, we describe close connections between estimable functions of the model parameters and standard quantities such as the ASR, EAPC, and joinpoints. Estimable functions have the added value of being fully adjusted for period and cohort effects, and generally more precise. Secondly, the APC model provides the descriptive epidemiologist with powerful new tools, including rigorous statistical methods for comparative analyses and the ability to project the future burden of cancer. We illustrate these principles using invasive female breast cancer incidence in the United States, but these concepts apply equally well to other cancer sites for incidence or mortality.

Keywords

Cancer Surveillance Research; age standardized rates (ASR); estimated annual percentage change (EAPC); joinpoint regression; Age-period-cohort (APC) model; APC estimable parameters; APC linear trends; APC deviations; APC drifts; APC fitted age at onset curve

Introduction

Cancer incidence and mortality rates are closely monitored to track the burden of cancer and its evolution in populations (1-4), provide etiological clues (5-11), reveal disparity (12-14), and gauge the dissemination of screening modalities (15-17) and therapeutic innovations (18, 19). A standard “toolbox” of graphical and quantitative methods has evolved to handle the needs of cancer surveillance researchers. Perhaps the most widely used methods include classical descriptive plots based on the lexis diagram (20-22), directly age-standardized rates (ASR) (23), estimated annual percentage change (EAPC) (24), and the joinpoint regression method (25). The underlying philosophy is agnostic and empirical; hence standard tools are particularly well suited to descriptive, exploratory, and hypothesis-generating studies.

At the same time, the age-period-cohort (APC) model has been developed in the statistics literature as a mathematical counter-point to purely descriptive approaches (20, 26-33). The APC model is based on fundamental generalized linear model theory (34); in principle, it allows the descriptive epidemiologist to both *generate* and *test* hypotheses. However,

*Corresponding Author: William F Anderson, Division of Cancer Epidemiology and Genetics, National Cancer Institute, EPS 8036, 6120 Executive Blvd, Rockville, MD, 20852-7244, United States wanderso@mail.nih.gov.

DISCLAIMER: None of the co-authors has a financial conflict of interest that would have affected this research.

although the APC model is generally accepted, our sense is it remains more of a niche methodology than an integral part of mainstream practice.

We believe two misunderstandings have slowed the uptake of the APC approach. Firstly, there are concerns about the “identifiability problem” of the APC model (27, 28). Secondly, close connections between the classical toolbox and the APC model have not been clearly spelled out in the literature. In this commentary, we will attempt to clarify both misunderstandings and thereby make the case that the APC model merits much wider use.

Data, Methods, and Results

Example: Breast cancer incidence data

We will develop this commentary using as a concrete example the incidence of invasive female breast cancers in the United States. For this purpose, we obtained age-specific case and population data from the National Cancer Institute’s Surveillance, Epidemiology, and End Results 9 Registries Database (SEER9) for the 36-year time period from 1973 through 2008 (November 2010 submission) (35).

In general, for any given cancer and population group, the matrix $\mathbf{Y} = [Y_{pa}, p = 1, \dots, P, a = 1, \dots, A]$ contains the number of cancer diagnoses in calendar period p and age group a , and the matrix $\mathbf{O} = [O_{pa}, p = 1, \dots, P, a = 1, \dots, A]$ contains the corresponding person-years. The observed incidence rates per 100,000 person-years are $\lambda_{pa} = 10^5 Y_{pa}/O_{pa}$, and the expected log rates are $\rho_{pa} = \log(E(Y_{pa})/O_{pa})$.

It is instructive to think of the rate matrix in terms of its corresponding Lexis diagram (Figure 1), which makes visually clear how the diagonals of matrices $\mathbf{Y}$ and $\mathbf{O}$, from upper right to lower left, represent successive birth cohorts indexed by $c = p - a + A$, from the oldest ($c = 1$) to the youngest ($c = C \equiv P + A - 1$). From this perspective, it becomes clear that a new cohort enters prospective follow-up with each consecutive calendar period. For this reason, one can think of a registry as a “cohort of cohorts.” Because cancer registries are operated in perpetuity, over time, a substantial number of birth cohorts are followed. Our example includes $C = 24$ nominal 8-year cohorts born from 1892 through 1984 (referred to by mid-year of birth).

The APC model: formulation

APC analysis is based on a log-linear model for the expected rates with additive effects for age, period, and cohort:

$$\rho_{pa} = \alpha_a + \pi_p + \gamma_c \quad (1)$$

The generic additive effects in equation (1) can be partitioned into linear and non-linear components (28). There are number of equivalent ways to make this partition while incorporating the fundamental constraint that $c \equiv p - a$. Two of the most useful (36) are the age-period form

$$\rho_{pa} = \mu + (\alpha_L - \gamma_L)(a - \bar{a}) + \tilde{\alpha}_a + (\pi_L + \gamma_L)(p - \bar{p}) + \tilde{\pi}_p + \tilde{\gamma}_{p-a+A} \quad (2)$$

and the age-cohort form

$$\rho_{ca} = \mu + (\alpha_L + \pi_L)(a - \bar{a}) + \tilde{\alpha}_a + (\pi_L + \gamma_L)(c - \bar{c}) + \tilde{\pi}_{c+a-A} + \tilde{\gamma}_c \quad (3)$$

Notation and parameters are summarized in Table 1. Importantly, all the parameters in equations (2) and (3) can be estimated from the data without imposing additional constraints, and fitted rates from both forms are identical.

There is a close correspondence between APC parameters and estimable functions in Table 1 and fundamental aspects of the data investigated using the standard descriptive toolbox. Before highlighting some of these connections below, we hopefully can shed further light on the much discussed identifiability problem.

Identifiability: “problem” or uncertainty principle?

The aspect of identifiability in question concerns whether log-linear trends in rates can uniquely be attributed to the influences of age, period, or cohort, quantified by parameters α_L , π_L , and γ_L . Mathematically, it has been shown by Holford (28) that one cannot do this without imposing additional unverifiable assumptions, because the three time scales are co-linear (cohort equals period minus age, $c = p - a$). This issue has often implicitly been held out as a unique and unfortunate limitation of the APC model. In fact, the same issue affects time-to-event analysis of any cohort study.

To see this, consider the following thought experiment. Suppose one enrolls a cohort of exchangeable persons of identical age (e.g., the 1956 birth cohort in Figure 1) and follows them longitudinally over a decade for cancer. At the end of the study, one observes that the log incidence rate increases linearly with age. It is natural to attribute this trend entirely to the effects of ageing, and equate the age-associated slope to the value of a parameter α_L .

However, suppose one had also assembled an identical cohort of persons of the same age, but this study had been conducted ten years earlier. It is possible that the age-associated slopes of the two studies would be very different, if disease-causing exposures out of experimental control had been increasing or decreasing in prevalence over time. Hence, the observed age-associated slope actually estimates parameter ($\alpha_L + \pi_L$) or longitudinal age trend (LAT in Figure 1) (32), where α_L is the component of the trend that is attributable to aging and π_L is the component of the trend due to the net impact of unknown and uncontrollable exposures over successive calendar-periods.

A similar issue affects any cross-sectional analysis. To “control” for the effects of ageing, suppose one studied in succession over time an event rate in persons of the same age (e.g., age group 65-69 years in Figure 1), to estimate the slope of the time-trend π_L . By definition, each successive group in this cross-sectional study was born a year later. Hence, both unknown factors and factors out of experimental control associated with birth cohort could also play a role. Therefore, the observed slope over time actually estimates a parameter ($\pi_L + \gamma_L$) or net drift in Figure 1 (29, 30), where π_L is the component of the trend that is attributable to calendar time and γ_L is the component of the trend attributable to the successive cohorts enrolled in the study.

These simple thought experiments, Figure 1, and Table 1 illustrate an important ‘uncertainty principle’ regarding the measurement of absolute rates in cohorts. Interestingly, this principle is seldom considered in the context of most epidemiological cohort and case-control studies, perhaps because these studies have a fairly narrow accrual window and often focus on relative rates rather than absolute rates. In contrast, this issue is often central in the analysis of registry data, because the follow-up has sufficient breadth and depth to reveal long-term secular trends in the population associated with age, period, and cohort. Indeed, a unique role of registry studies is to identify and quantify such trends, thereby providing direction and guidance regarding the needs for targeted analytical studies.

Estimable functions: separating signal from noise

The APC model provides a unique set of best-fitting log incidence rates, $\hat{\rho}_{pa}$ or equivalently $\hat{\rho}_{ca}$, obtained by plugging in maximum likelihood estimators into equations (2) or (3), respectively. The corresponding variances are readily calculated. In our experience the fitted rates have an appealing amount of smoothing, and we use them routinely in our studies (36-45), especially for rare cancer outcomes. Experience suggests that for “moderate” sized rate matrices (in terms of A and P), the APC model smoothes the data conservatively, about as much as a 3-point moving average, yielding around a 40-60% reduction in the width of the confidence intervals. Of course, the precise amount of noise reduction depends on a number of technical details including whether over-dispersion is present or accounted for.

This application of the APC model is illustrated in Figure 2 for the breast cancer data. The age-standardized rates (ASRs) over time calculated using the observed rates are nearly identical to the ASRs calculated using the APC fitted rates. However, the point-wise confidence intervals for the fitted rates are substantially narrower, by around 40% averaged over the 10-year time period.

Estimable functions: connections to the classical approaches

The APC parameter called the net drift (Table 1 and equations (2) and (3)) estimates the same quantity as the EAPC of the ASR, i.e. the overall long-term secular trend. The point estimates for these quantities are almost identical for the breast cancer data in Figure 2; net drift = 0.83% per year (95% CI: 0.78 to 0.85%/yr) and EAPC = 0.78% percent per year (0.18 to 1.39%/yr). However, for this example, the estimated confidence bands are much narrower for the net drift.

We introduced a novel estimable function called the fitted age-at-onset curve to summarize the longitudinal (i.e. cohort-specific) age-associated natural history (Table 1 and figure 3) (46). By construction, the fitted curve extrapolates from observed age-specific rates over the full range of birth cohorts to estimate past, current, and future rates for the referent cohort, e.g., the 1932 cohort in this example. The fitted age-at-onset curve provides a longitudinal age-specific rate curve that is adjusted for both calendar-period and birth-cohort effects. We view it as an improved version of the cross-sectional age-specific rate curve, improved because the cross-sectional curve is not adjusted for period and cohort effects (47). The fitted curve has proven very useful in practice (38-40, 42-44, 46, 48).

Finally, period deviations in the APC model (Table 1) identify changes over time; such change points are often analyzed non-parametrically using joinpoint regression methods (25). Similarly, cohort deviations can provide an explanation for joinpoint patterns in age-specific rates over time.

APC analysis: beyond the basics

There are many useful extensions to the basic APC model. Estimable functions are amenable to formal hypothesis tests (29, 30). Parameters associated with age, period, and cohort can be smoothed (49). Parametric assumptions about the shape of the age incidence curve derived from mathematical models of carcinogenesis can be incorporated (50). Other extensions have included parametric (33) and nonparametric (51, 52) assessments of changes in period and cohort deviations, and simultaneous modeling of a moderate or large number of strata, such as geographic areas, using Bayes and Empirical Bayes methods (53).

Recently, we developed novel methods to compare age-related natural histories and time-trends between distinct event rates assuming that separate APC models hold for each (36). Using this approach one can formally contrast the incidence of a given tumor such as breast

cancer in two populations, say Black versus White women (46), or the incidence of two tumor subtypes in the same population, say, ER positive versus ER negative breast cancers ((46), supplemental Figure). We demonstrated that two event rates are proportional over age, period, or cohort if and only if certain sets of APC parameters are all equal across the respective event-specific models (36). We also developed corresponding tests of proportionality and estimators of rate ratios.

A number of authors have forecast future cancer rates using the APC model (54-58). Projections quantify the future implications of current trends, for example, the impact of a net drift of 1% versus 2% over time, or the future impact of recent changes in birth cohort patterns.

Discussion

Successful technological evolution builds on effective design. This is just as true for statistical methods as for computers and cellular phones. We have argued here that the APC model provides a useful evolutionary extension to the standard armamentarium of methods available to the descriptive epidemiologist. The APC model is not a replacement for existing methods, which are popular and successful. Rather, it provides a refined means of estimating the same quantities, while also adding useful new capabilities, such as formal methods for comparing two sets of rates or projecting the future cancer burden.

Using the APC model, cancer registry data can be analyzed in the same spirit as any other epidemiological cohort using the same concepts, such as proportional hazards, confounding, and effect modification/interaction. Importantly, because cancer registries follows a cohort of cohorts, analysis of registry data can reveal fundamental changes in population rates that are not usually discernable in standard cohort or case-control studies.

Currently, software for APC analysis is available only through fairly specialized packages (SAS, R, Matlab). Development of good stand-alone software, in addition to education and training, are needed if the full potential of the APC model is to be exploited by descriptive epidemiologists.

Acknowledgments

This research was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute. All of the authors had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Parkin DM. The evolution of the population-based cancer registry. *Nat Rev Cancer* Aug. 2006; 6(8): 603–612.
2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. Mar-Apr; 2011 61(2):69–90. [PubMed: 21296855]
3. Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics, 2010. *CA Cancer J Clin*. Jul 7.2010
4. Kohler BA, Ward E, McCarthy BJ, et al. Annual Report to the Nation on the Status of Cancer, 1975-2007, Featuring Tumors of the Brain and Other Nervous System. *J Natl Cancer Inst*. Mar 31.2011
5. Bergstrom R, Adami HO, Mohnen M, et al. Increase in testicular cancer incidence in six European countries: a birth cohort phenomenon. *J Natl Cancer Inst*. Jun 5; 1996 88(11):727–733. [PubMed: 8637026]

6. Verhoeven R, Houterman S, Kiemeny B, Koldewijn E, Coebergh JW. Testicular cancer: marked birth cohort effects on incidence and a decline in mortality in southern Netherlands since 1970. *Int J Cancer*. Feb 1; 2008 122(3):639–642. [PubMed: 17764111]
7. Liu S, Semenciw R, Waters C, Wen SW, Mery LS, Mao Y. Clues to the aetiological heterogeneity of testicular seminomas and non-seminomas: time trends and age-period-cohort effects. *Int J Epidemiol*. Oct; 2000 29(5):826–831. [PubMed: 11034964]
8. Bray F, Richiardi L, Ekblom A, et al. Do testicular seminoma and nonseminoma share the same etiology? Evidence from an age-period-cohort analysis of incidence trends in eight European countries. *Cancer Epidemiol Biomarkers Prev*. Apr; 2006 15(4):652–658. [PubMed: 16614105]
9. Spix C, Eletr D, Blettner M, Kaatsch P. Temporal trends in the incidence rate of childhood cancer in Germany 1987–2004. *Int J Cancer*. Apr 15; 2008 122(8):1859–1867. [PubMed: 18076067]
10. McNally RJ, Cairns DP, Eden OB, Kelsey AM, Taylor GM, Birch JM. Examination of temporal trends in the incidence of childhood leukaemias and lymphomas provides aetiological clues. *Leukemia*. Oct; 2001 15(10):1612–1618. [PubMed: 11587220]
11. Svensson E, Moller B, Tretli S, et al. Early life events and later risk of colorectal cancer: age-period-cohort modelling in the Nordic countries and Estonia. *Cancer Causes Control*. Apr; 2005 16(3):215–223. [PubMed: 15947873]
12. Chu KC, Tarone RE, Brawley OW. Breast cancer trends of black women compared with white women. *Arch Fam Med*. 1999; 8(6):521–528. [PubMed: 10575392]
13. Sim X, Ali RA, Wedren S, et al. Ethnic differences in the time trend of female breast cancer incidence: Singapore, 1968–2002. *BMC cancer*. 2006; 6:261. [PubMed: 17078893]
14. Chie WC, Chen CF, Lee WC, Chen CJ, Lin RS. Age-period-cohort analysis of breast cancer mortality. *Anticancer Res*. 1995; 15(2):511–515. [PubMed: 7763032]
15. Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: interpreting trends in prostate cancer—part II: Cause of death misclassification and the recent rise and fall in prostate cancer mortality. *J Natl Cancer Inst*. Jun 16; 1999 91(12):1025–1032. [PubMed: 10379965]
16. Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Chapter 4: changing patterns in breast cancer incidence trends. *J Natl Cancer Inst Monogr*. 2006; (36):19–25. [PubMed: 17032890]
17. Feuer EJ, Etzioni R, Cronin KA, Mariotto A. The use of modeling to understand the impact of screening on U.S. mortality: examples from mammography and PSA testing. *Stat Methods Med Res*. Dec; 2004 13(6):421–442. [PubMed: 15587432]
18. Jatoti I, Chen BE, Anderson WF, Rosenberg PS. Breast cancer mortality trends in the United States according to estrogen receptor status and age at diagnosis. *J Clin Oncol*. 2007; 25(13):1683–1690. [PubMed: 17404367]
19. Pinder MC, Duan Z, Goodwin JS, Hortobagyi GN, Giordano SH. Congestive heart failure in older women treated with adjuvant anthracycline chemotherapy for breast cancer. *J Clin Oncol*. Sep 1; 2007 25(25):3808–3815. [PubMed: 17664460]
20. Carstensen B. Age-period-cohort models for Lexis diagram. *Statistics in Medicine*. 2007; 26:3018–3045.
21. Keiding N. Statistical inference in the Lexis Diagram. *Phil Trans R Soc Lond A*. 1990; 332:487–509.
22. Vandeschrick C. The Lexis diagram, a misnomer. *Demographic Research*. 2001; 4:97–124.
23. Last, JM. *A Dictionary of Epidemiology*. Third ed. Oxford: Oxford University Press; 1995.
24. Fay MP, Tiwari RC, Feuer EJ, Zou Z. Estimating average annual percent change for disease rates without assuming constant change. *Biometrics*. Sep; 2006 62(3):847–854. [PubMed: 16984328]
25. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. *Stat Med*. Feb 15; 2000 19(3):335–351. [PubMed: 10649300]
26. Holford TR. The estimation of age, period and cohort effects for vital rates. *Biometrics*. Jun; 1983 39(2):311–324. [PubMed: 6626659]
27. Holford TR. Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annu Rev Public Health*. 1991; 12:425–457. [PubMed: 2049144]
28. Holford, TR. Age-Period-Cohort Analysis. In: Armitage, P.; Colton, T., editors. *Encyclopedia of Biostatistics*. Second Edition. Vol. 1. West Sussex: John Wiley & Sons Ltd; 2005. p. 105–123.

29. Clayton D, Schifflers E. Models for temporal variation in cancer rates. I: Age-period and age-cohort models. *Stat Med.* Jun; 1987 6(4):449–467. [PubMed: 3629047]
30. Clayton D, Schifflers E. Models for temporal variation in cancer rates. II: Age-period-cohort models. *Stat Med.* Jun; 1987 6(4):469–481. [PubMed: 3629048]
31. Robertson C, Boyle P. Age-period-cohort models of chronic disease rates. II: Graphical approaches. *Stat Med.* Jun 30; 1998 17(12):1325–1339. [PubMed: 9682323]
32. Robertson C, Boyle P. Age-period-cohort analysis of chronic disease rates. I: Modelling approach. *Stat Med.* Jun 30; 1998 17(12):1305–1323. [PubMed: 9682322]
33. Tarone RE, Chu KC. Evaluation of birth cohort patterns in population disease rates. *Am J Epidemiol.* Jan 1; 1996 143(1):85–91. [PubMed: 8533751]
34. McCullagh, P.; Nelder, JA. *Generalized Linear Models.* New York: Chapman and Hall; 1989.
35. SEER-9. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence-SEER 9 Regs Research Data, Nov 2009 Sub (1973-2008) Katrina/Rita Population Adjustment - Linked To County Attributes - Total U.S., 1969-2007 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2011, based on November 2010 submission. 2011
36. Rosenberg PS, Anderson WF. Proportional hazard models and age-period-cohort analysis of cancer rates. *Stat Med.* 2010; 29:1228–1238. [PubMed: 20209480]
37. Bradford PT, Anderson WF, Purdue MP, Goldstein AM, Tucker MA. Rising melanoma incidence rates of the trunk among younger women in the United States. *Cancer Epidemiol Biomarkers Prev.* 2010; 19:2401–2406. [PubMed: 20826837]
38. Mbulaiteye SM, Anderson WF, Bhatia K, Rosenberg PS, Linet MS, Devesa SS. Trimodal age-specific incidence pattern for Burkitt Lymphoma in the United States, 1973-2005. *Int J Cancer.* 2010; 126:1732–1739. [PubMed: 19810101]
39. Anderson WF, Jatoi I, TSE J, Rosenberg PS. Male breast cancer: a population-based comparison with female breast cancer. *J Clin Oncol.* 2010; 28(2):232–239. [PubMed: 19996029]
40. Anderson WF, Pfeiffer RM, Tucker MA, Rosenberg PS. Divergent cancer pathways for early-onset and late-onset cutaneous malignant melanoma. *Cancer.* 2009; 115:4176–4185. [PubMed: 19536874]
41. Menashe I, Anderson WF, Jatoi I, Rosenberg PS. Underlying causes of the Black-White racial disparity in breast cancer mortality: a population-based analysis. *JNCI.* 2009; 101:993–1000. [PubMed: 19584327]
42. Grimley PM, Matsuno RK, Rosenberg PS, Henson DE, Schwartz AM, Anderson WF. Qualitative age interactions between low and high grade serous ovarian carcinomas. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(8):2256–2261. [PubMed: 19622723]
43. Reimers LL, Anderson WF, Rosenberg PS, Henson DE, Castle PE. Etiological heterogeneity for cervical carcinoma by histopathological type, using age-period-cohort (APC) models. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(3):792–800. [PubMed: 19258470]
44. Kilfoy BA, Devesa SS, Ward MH, et al. Gender is an age-specific effect modifier for papillary cancers of the thyroid gland. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(4):1092–1100. [PubMed: 19293311]
45. Anderson WF. Cancer Surveillance Research (CSR). *Cancer Epidemiol Biomarkers Prev.* 2009; 18(6):1669–1671. [PubMed: 19505896]
46. Anderson WF, Rosenberg PS, Menashe I, Mitani A, Pfeiffer RM. Age-related crossover in breast cancer incidence rates between Black and White Ethnic Groups. *J Natl Cancer Inst.* 2008; 100(24):1804–1814. [PubMed: 19066264]
47. Parkin DM, Bray FI, Devesa SS. Cancer burden in the year 2000. The global picture. *Eur J Cancer.* Oct; 2001 37(Suppl 8):S4–66. [PubMed: 11602373]
48. Anderson WF, Chen BE, Brinton LA, Devesa SS. Qualitative age interactions (or effect modification) suggest different cancer pathways for early-onset and late-onset breast cancers. *Cancer Causes and Control.* 2007; 18(10):1187–1198. [PubMed: 17823850]
49. Heuer C. Modeling of time trends and interactions in vital rates using restricted regression splines. *Biometrics.* Mar; 1997 53(1):161–177. [PubMed: 9147590]

50. Holford TR, Zhang Z, McKay LA. Estimating age, period and cohort effects using the multistage model for cancer. *Stat Med*. Jan 15; 1994 13(1):23–41. [PubMed: 9061838]
51. Tarone RE, Chu KC. Implications of birth cohort patterns in interpreting trends in breast cancer rates. *J Natl Cancer Inst*. Sep 16; 1992 84(18):1402–1410. [PubMed: 1512791]
52. Tarone RE, Chu KC. Nonparametric evaluation of birth cohort trends in disease rates. *Journal of epidemiology and biostatistics*. 2000; 5(3):177–191. [PubMed: 11051114]
53. Robertson C, Ecob R. Simultaneous modelling of time trends and regional variation in mortality rates. *Int J Epidemiol*. Oct; 1999 28(5):955–963. [PubMed: 10597997]
54. Bray F, Moller B. Predicting the future burden of cancer. *Nat Rev Cancer*. Jan; 2006 6(1):63–74. [PubMed: 16372017]
55. Peto J, Decarli A, La Vecchia C, Levi F, Negri E. The European mesothelioma epidemic. *Br J Cancer*. Feb; 1999 79(3-4):666–672. [PubMed: 10027347]
56. Woo PP, Thach TQ, Choy ST, McGhee SM, Leung GM. Modelling the impact of population-based cytologic screening on cervical cancer incidence and mortality in Hong Kong: an age--period--cohort approach. *Br J Cancer*. Oct 31; 2005 93(9):1077–1083. [PubMed: 16205695]
57. Cleries R, Martinez JM, Escriba JM, et al. Monitoring the decreasing trend of testicular cancer mortality in Spain during 2005-2019 through a Bayesian approach. *Cancer Epidemiol*. Jun; 2010 34(3):244–256. [PubMed: 20381445]
58. Cleries R, Ribes J, Esteban L, Martinez JM, Borrás JM. Time trends of breast cancer mortality in Spain during the period 1977-2001 and Bayesian approach for projections during 2002-2016. *Ann Oncol*. Dec; 2006 17(12):1783–1791. [PubMed: 16971661]

Figure 1.

Rate matrix or Lexis diagram (20-22) for invasive female breast cancer. Data from the National Cancer Institute's Surveillance, Epidemiology, and End Results 9 Registries Database (SEER 9) for cases diagnosed between 1973 through 2008 (35). Sixteen 4-year age groups (21-24, 25-29, ..., 81-84 years) and nine 4-year time periods (1973-1976, 1977-1980, ..., 2005-2008) span 24 partially overlapping 8-yr birth cohorts, with age groups in the columns, time-periods in the rows, and birth-cohorts along the diagonals. Conditioned upon time period (e.g., 1973-1977), the cross sectional age trend (CAT) cuts across increasing age groups and decreasing birth cohorts, $CAT = \alpha_L - \gamma_L$. Conditioned upon birth cohort (e.g., 1956), the longitudinal age trend (LAT) cuts across increasing age groups and time periods, $LAT = \alpha_L + \pi_L$. Conditioned upon age group (e.g., 65-69 years), the net drift cuts across increasing time periods and birth cohorts, Net drift $t = (\pi_L + \gamma_L)$.

Figure 2. Age standardized rates (ASRs, 2000 standard US population) for invasive female breast cancer. Data from the National Cancer Institute's SEER 9 Database. ASRs calculated using observed rates (grey) and age-period-cohort fitted rates (red). Point estimates (solid circles) and 95% confidence limits (shaded areas) are shown.

Figure 3.

Cohort-specific age-specific incidence rates for invasive female breast cancer. Data from the National Cancer Institute's SEER 9 Database, stratified by 8-year birth cohorts. The age-period-cohort (APC) fitted age at onset curve (red line and shaded 95% confidence intervals) 'stitches together' the experience of 11 cohorts observed over staggered age intervals. By construction, the fitted (or longitudinal) curve is centered on the referent-cohort; e.g., the 1932 birth-cohort. Because these breast cancer incidence rates are increasing over time with a net drift of 0.83% per year (figure 2 and text), the fitted rates for the referent-cohort lie below the observed rates for the younger cohorts and above the rates for the older cohorts. In other words, with rising incidence rates from older to younger generations (figure 2 and text), the referent-cohort rates on average are lower than rates for the younger cohorts and higher than rates for the older cohorts, respectively. See the text and Table 1 for additional details.

Table 1Some estimable parameters and functions in the APC model^{*}

Quantity [*]	Nomenclature
μ	Grand mean
$\tilde{\alpha}_a$, $\tilde{\pi}_p$, and $\tilde{\gamma}_c$	Age, period, and cohort deviations
$(\alpha_L + \pi_L)$	Longitudinal age trend
$(\alpha_L - \gamma_L)$	Cross-sectional age trend
$(\pi_L + \gamma_L)$	Net drift $\approx$ EAPC of the ASR ^{**}
$\mu + (\alpha_L + \pi_L)(a - \tilde{a}) + \tilde{\alpha}_a$	Fitted longitudinal age-at-event curve
$\mu + (\alpha_L - \gamma_L)(a - \tilde{a}) + \tilde{\alpha}_a$	Fitted cross-sectional age-at-event curve
$\mu + (\pi_L + \gamma_L)(p - \tilde{p}) + \tilde{\pi}_p$	Fitted temporal trends

^{*} The APC model is defined over a $P \times A$ event matrix $\mathbf{Y}$ and corresponding matrix of person-years $\mathbf{O}$. The referent age, period, and cohort are $\tilde{a} = [(A + 1)/2]$, $\tilde{p} = [(P + 1)/2]$, and $\tilde{c} = \tilde{p} - \tilde{a} + A$, respectively, where P and A are the total numbers of period and age groups and $[\cdot]$ is the greatest integer function.

^{**} EAPC, estimated annual percentage change; ASR, direct age standardized incidence rate (ASR) (23).