Appendix III. Statistical Models--Evaluation of Subgroup Differences

Background

Age-adjusted death rates for person of all ages and for premature deaths (under age 75) are used throughout this report to describe mortality trends and patterns. These statistics are summary measures that may disguise important subgroup differences by age, race or year. Health surveillance reports that rely on summary age-adjusted rates often ignore these potential differences. In some cases the simplification offered by age-standardized rates may not be valid due to important age-specific differences. Furthermore, as Fleiss (1981) and others have cautioned, "when stratum-specific comparisons between groups are inconsistent across strata, the relative magnitudes of the overall adjusted measures can largely depend on the weights chosen, and inappropriate adjustment may produce distortion rather than summarization" (Wilcosky T.C., Chambles L.E., 1985). A variety of subgroup differences have been systematically evaluated in this report. This detailed evaluation has allowed us to determine when conclusions based on AAMR and YPLL rates alone may be incomplete. Logistic regression models were used to evaluate these potential differences[i].

The logistic regression models used to evaluate subgroup differences and annual trends offer several advantages over other methods:

- They are widely used in epidemiological studies of dichotomous outcomes (e.g. death-survival) and the techniques are well documented (Hosmer, D. W., Lemeshow S., 2000).
- They can be defined to include age-specific covariates that yield summary estimates (e.g. age-adjusted death rates) that "are almost identical to the direct adjusted measures" (Wilcosky T.C., Chambles L.E., 1985). In addition, the weighted estimates from these stratified models have a similar interpretation to rates that are age-adjusted by the direct method.
- These models can accommodate continuous as well as qualitative predictors, and they make it relatively easy to test for interactions such as the subgroup differences noted.
- In contrast to "standard" analyses that simply use the summary age-adjusted rate as the dependent variable in a regression model, it is important to use an age-specific dependent variable along with age-specific covariates, as was done in this study. As noted by Rosenbaum and Rubin (1984), such standard analyses "are generally inappropriate. The problem arises because outcome variables used in those analyses, such as death rates in various regions, have been age-adjusted, whereas the predictor variables have not been age-adjusted." The models employed in this study all estimate the log-odds of death using a model that includes 18 age-specific covariates to control for age.

Despite the many advantages of model-based summary rates, we have used model-based calculations to augment our analyses rather than as the "starting point." In cases where the model-based tests differ from results based on summary rates, they are identified in the report.

Our emphasis on reporting and discussing the directly adjusted rates is due to a concern that the technical nature of model-based estimates would make the report less accessible to a large fraction of the public health audience we wish to reach. This concern is as important today as it was over 20 years ago when Freeman and Holford (1980) encouraged researchers to balance consideration of the benefits of statistical modeling with knowledge of their intended audience. “In many settings these advantages are overshadowed by the dual requirements for simplicity of presentation and immediacy of interpretation.”

Logistic Regression Model Details. Subgroup differences were evaluated in two specific areas:

1. **Time trends for 1989-1998[ii] were evaluated to determine whether trends vary significantly by year of death, or by age group.**

    a. The estimates from these models provided specific information that has been incorporated into this report. These results include:

        i. Estimates of the average annual percent change in annual death rate

        ii. Identification of individual years between 1989-1998 that depart significantly ($p< 0.005$) from the estimated linear trend. This result was used in the AAMR trend graphs to "flag" specific years identified. (The years identified were the same whether the data were weighted to the 1940 or 2000 standard population.)

        iii. Identification of age groups with trend slopes that differ significantly ($p<0.0014$) from the estimated trend for all age groups.

        iv. The average annual percent change estimated in (i) was reported when no differences were found in stages ii or iii.

    b. Model characteristics:

        i. Logistic regression models were evaluated separately for males and females. Two independent models were used to assess differences in trends by year (Model-1) and by age group (Model-2). First, a global test of the entire interaction category with n-1 df was evaluated, while controlling for age and the annual rate of change. If significant at $p<0.05$, then a detailed assessment of specific interaction terms was made. A stepwise logistic regression procedure was used to test specific interaction terms (in Model-1: C_1-C_{10}; in Model-2: C1-C_{18})

        ii. Observation weights: The data were weighted to conform to the age distribution of the year 2000 standard reference population for Model-1. The age-specific evaluations derived using Model-2 were based on unweighted data.

        iii. Dependent variable (both models) = Log {deaths/survivors}
        Where: Survivors = Population – Deaths (by age/sex /race/year)

        iv. Independent predictors:

        - Model-1 [Tests for single years that do fit the 1989-1998 linear trend model]

            Forced terms: Constant + A(*annual slope*) + B_1(*age1*)+….B_{17}(*age17*) [iii]

            Terms tested for inclusion in forward stepwise evaluation:
            C_1(*year89* **x** *annual slope*) + ….C_{10}(*year98* **x** *annual slope*)

            Where: *age1-age18* = 0/1 dummy variables for ages 0-4 to 85+
            annual slope term = 0,1…9 for 1989,1990…1998
            year89 - year98 = 0/1 dummy variables for 1989 to 1998

        - Model-2 [Tests for Age groups with differing 1989-1998 slopes]

            Forced terms: Constant + A(*annual slope*) + B_1(*age1*)+….B_{17}(*age17*) [3]

            Terms tested for inclusion in forward stepwise evaluation:
            C_1(*age1* **x** *annual slope*) + …C_{18}(*age18* **x** *annual slope)*

            Where: *age1-age18* = 0/1 dummy variables for ages 0-4 to 85+
            annual slope term = 0,1…9 for 1989,1990…1998

2. **Disparities in 1996-1998 AAMRs by race and ethnicity were evaluated to determine whether the degree of disparity varies significantly by age group.** This subgroup analysis was limited to white-black and white-Hispanic differences since the age-specific death counts for other race groups are too small to permit reasonable evaluation.

.

   a. The estimates from these models provided specific information that has been incorporated into this report. These results include:

      i. Estimates the overall disparity as the ratio of the death rate for black or Hispanic population to the rate in the white population.

      ii. Identification of specific age groups where the disparity in 1996-1998 departs significantly ($p < .0014$) from disparity estimated for all ages.

      iii. When an age-specific difference in disparity magnitude is identified (in stage ii) an adjusted disparity level for *all other* age groups is also reported.

   b. Model characteristics:

      i. Logistic regression models were evaluated separately for males and females. First, a global test of the entire interaction category with n-1 df was evaluated, while controlling for age and the annual rate of change. If significant at $p<0.05$, then a detailed assessment of specific interaction terms was made. A stepwise logistic regression procedure was used to test specific interaction terms (C_1-C_{18}).

      ii. The annual figures used were weighted to conform to the year 2000 standard reference population.

      iii. Dependent variable = Log {deaths/survivors}
      Where, Survivors = population – deaths (by age/sex /race/year)

      iv. Independent predictors:
      Forced terms:
      $\text{Constant} + A(\textit{Disparity}) + B_1(\textit{age1})\ldots.B_{17}(\textit{age17})$ [3]

      Terms tested for inclusion in forward stepwise evaluation:
      $C_1(\textit{Disparity}\ \mathbf{x}\ \textit{age1})\ldots C_{18}(\textit{Disparity}\ \mathbf{x}\ \textit{age18})$

      Where:
      Disparity = 0/1 dummy variable for white/black (or white/Hispanic)[iv]
      age1-age18 = 0/1 dummy variables for ages 0-4 to 85+

Multiple comparisons

These subgroup analyses required that many statistical tests be conducted. This necessarily increases the chance of labeling spurious, non-significant results as "significant." To minimize the type-1 error rate we used a Bonferroni adjustment to hold the overall probability level to $p<0.05$ for each set of comparisons. For example, when making 10 independent tests of an hypothesis we have adjusted the threshold for defining a "significant" result by taking the desired p-level ($p=0.05$) and dividing it by the number of planned comparisons (10) resulting in a lower p-level threshold ($.05/10 = .005$). Defining what constitutes an independent set of tests is subject to debate. In this report we have generally treated tests related to each cause of death (males and females) as an independent set of tests. The minor exception to this was the assessment of single-year deviations from the 10-year trend where we adopted a slightly lower threshold, reflecting the importance of identifying these changes. Ultimately other considerations such as the magnitude of the detected differences, and the consistency of the finding with those from other studies are also important to consider. Thus significance levels are just one "screen" through which the findings are passed in the process of interpreting their meaning. The significance levels noted above are based on the number of planned comparisons.

Adjusted Significance Levels based on the number of Planned Comparisons

Type of Comparison	Target P-level (2-tailed)	Number of Planned Comparisons	Adjusted P-level
Single years not on Trend line	0.05	10	0.005
Single Age group (1-18) not on Trend line, Male & Female	0.05	(18 age groups) x (2 sex groups) = 36	0.0014
Single Age group (1-18) not on Trend line, Sex-specific cause	0.05	(18 age groups) x (1 sex group) = 18	0.0028
Black/White Disparity by Age group (1-18) Male & Female	0.05	(18 age groups) x (2 sex groups) = 36	0.0014
Black/White Disparity by Age group (1-18) Sex specific cause	0.05	(18 age groups) x (1 sex group) = 18	0.0028

The logistic regression models, and the graphic illustrations of trends, and age-specific rates were developed using SYSTAT statistical software, version-10 (SPSS Inc., 2000).

References

Fleiss J.L. 1981. *Statistical methods for rates and proportions* (second edition). New York: John Wiley & Sons, Inc.

Freeman D.H., Holford T.R. 1980. Summary rates. *Biometrics* 36:195.

Hosmer, D. W., Lemeshow S. 2000. *Applied logistic regression* (second edition). New York: John Wiley & Sons, Inc.

Inskip, H., Beral V., Fraser P., Haskey, J. 1983. Methods for age-adjustment of rates. *Statistics in Medicine* 2: 455.

Rosenbaum, P.R., Rubin, D.B. 1984. Difficulties with regression analyses of age-adjusted rates. *Biometrics* 4: 437.

SPSS Inc., 2000. SYSTAT version 10 [computer software], Chicago Illinois, SPSS Inc.

Wilcosky T.C., Chambles L.E., 1985. A comparison of direct adjustment and regression of epidemiologic measures. *Journal of Chronic Disease* 38: 849.

---

[i] The logistic regression models actually calculate the ratios of the odds of death (OR), but these values provide a very good approximation of the rate ratio (RR) since deaths are rare events. For example, if we compare male unintentional injury deaths for black and white populations in 1996-1998, the OR and RR values are almost identical. The summary Mantel-Haenszel odds ratio and rate ratio are 1.6842 and 1.6837 respectively. We expect the OR/RR difference to be largest for all causes of death, but even there the figures differ by less than 3% (1.77 and 1.73 respectively).

[ii] Changes in the coding of certain categories of deaths beginning in 1991 may result in an artificial increase in various subcategories of injury deaths when examining trends over time. For this reason, trend analyses for certain subcategories of injury deaths in this report (injury and all other external causes, unintentional injuries, suicide, poisoning, and drug-induced) exclude the years 1989 through 1991 and evaluate trends for the period 1992-1998.

[iii] The base model includes a constant, so it is fully parameterized with respect to age because it includes n-1 or 17 age terms in the model.

[iv] The tables presented in this report provide summary information for persons of white race, without the further delineation of the non-Hispanic portion of this group. While this format is consistent with the mortality tabulations presented by the National Centers for Health Statistics during most of the 1990's and the format of mortality target objectives for 2000, it presents some difficulties when evaluating disparities relative to the Hispanic population. Since the Hispanic and white population groups overlap, comparisons based on the calculated rates for each group will violates the statistical assumption that the comparison groups are independent. In Connecticut, the impact of using a white rather than a white non-Hispanic subgroup to contrast with the Hispanic population is quite small. In 1998 only about 1.5% of Connecticut's white population was estimated to be Hispanic. A separate unpublished series of tabulations and analyses have been performed using the white non-Hispanic deaths and population figures. The disparity assessments based on this series are almost identical to those published in this report. These additional tabulations are available on request.