COMPARING MORTALITY RISKS WITH CHRONIC CONDITIONS

H. D. Tolley and K. G. Manton, Duke University

This paper presents a method of comparing risks of death after the onset of a chronic disease when onset time is unknown. Using a simple two-step compartment model or primary-secondary decrement model of chronic disease and death, test constraints for equality of mortality rates are formulated. The procedure adjusts for differences in onset and prevalence rates of different chronic conditions and/or different populations. Statistical procedures are based upon large sample results commonly used in categorical data analysis. The methodology is illustrated using U.S. multiple-cause mortality data.

I. INTRODUCTION

In this paper we present a method for testing differences in the risks of death from various acute conditions after the onset of one of several possible chronic diseases when the time of onset of those chronic illnesses is unknown. This method is based upon the simple compartment model representation of the morbidity-mortality process illustrated in Figure 1.

In Figure 1 we view the onset of chronic disease as an irreversible transition of an individual from the well state to one of K chronic illness states. The individual remains in this chronic illness state until he dies of some immediate cause of death. Under most scenarios of the disease process, one would expect that cause specific mortality rates for these immediate causes of death depend on the particular underlying chronic disease condition (i.e., the chronic disease state the individual is in just before death). Therefore, comparisons of observed mortality patterns between cohorts, particularly elderly cohorts where chronic disease is prevalent, must account for any differentiable distribution of chronic conditions. This adjustment is to remove apparent differences in the mortality risks of the immediate or acute causes which are attributed to the differences in underlying chronic illness prevalence and not to true differences in the risks of the immediate causes of death. In a similar fashion, standard mortality ratios are calculated to adjust mortality risks for the effects of identifiable exogenous factors.

This method will be useful for life table studies of the elderly when comparisons of mortality rates of different cohorts is desired when neither the prevalence or incidence rates of any chronic disease in the cohort is known. It is assumed, at death, that the identity of the underlying chronic disease, if there is one, can be determined. Such situations often arise in mortality studies of human populations where the distribution of chronic illness in the population is generally unknown. For example, diabetes may influence the risk of myocardial infarction, and diabetes may be reported as an associated cause of death on the death certificate. However in comparing the risks of myocardial infarction between two small geographic areas, it is necessary, using standard primary-secondary decrement procedures, to know the distribution of diabetes in the population. With the proposed method comparisons can be made using only the information on diabetes reported at the time of death.

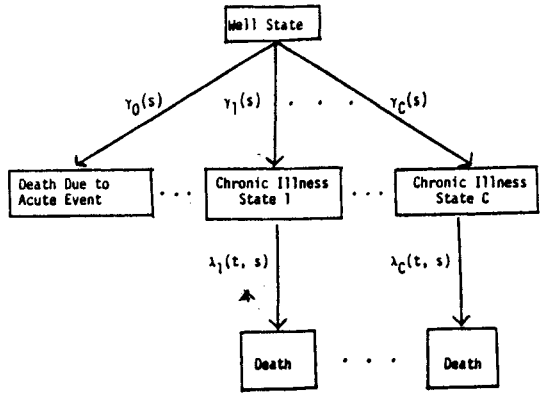

FIGURE 1. Compartment Model Representation of Chronic Disease

Preliminaries

For a theoretical cohort of individuals followed through time for a specific chronic disease, indexed by i, we define the following:

k = index for age of individual or time in the study

N_k = expected number of individuals in the well state at the beginning of Year k (i.e., number aged k)

M_{ik} = expected number of individuals with chronic illness i at beginning of Year k

n_{ik} = expected number of individuals dying with chronic illness i during Year k (i.e., between ages k and k + 1)

$q_k^{\alpha i}$ = probability that an individual in the well state at the beginning of Year k will die with "chronic illness" i before the beginning of Year k + 1

${}_m r_k^i$ = probability that a person in the well state at the beginning of Year k will enter the chronic disease state in the next m years (whether he dies with the chronic disease or lives with disease)

${}_m q_k^i$ = probability that a person in chronic illness state i at the beginning of Year k will die in the next m years.

For convenience we write:

${}_1 r_k^i = r_k^i$;

${}_1 q_k^i = q_k^i$; and

Q_k = probability that a person in the well state at the beginning of Year k will leave the well state due to either death or any chronic disease during the interval k to k + 1.

Model

In this section we consider a single chronic condition. Therefore, for notational simplicity we suppress the argument i designating the particular chronic condition. The expected number of deaths with chronic illness i during Year k, i.e., n_k, consists of all those who were in illness state i at the beginning of the interval and died plus all who were in the well state at time k and died with chronic illness i in one year. In symbols

$$n_k = M_k\, q_k + N_k\, q_k^{\alpha} . \qquad (1)$$

Similarly, the expected number of deaths with chronic illness i for a two-year span is comprised of the following:

$M_k\, {}_2q_k$ = expected number of those chronically ill at time k who die within two years

$N_k\, q_x^{\alpha}$ = expected number of those well at k who die within one year with the chronic illness.

$N_k(r_k - q_k^{\alpha})q_{k+1}$ = expected number well at time k who get ill before k + 1 but die between k + 1 and k + 2

$N_k(1 - Q_k)q_{k+1}^{\alpha}$ = expected number well at time k who stay well until k + 1 but die with chronic illness i before k + 2.

Thus,

$$n_k + M_{k+1} = M_k\, {}_2q_k + N_k \left[q_k^{\alpha} + (r_k - q_k^{\alpha})q_{k+1} + (1 - Q_k)q_{k+1}^{\alpha}\right] . \qquad (2)$$

Similarly,

$$n_k + n_{k+1} + n_{k+2} = M_k\, {}_3q_k + N_k \left[q_k^{\alpha} + (r_k - q_k^{\alpha})\, {}_2q_{k+1}^{\alpha} + (1 - Q_k)q_{k+1}^{\alpha} + (1 - Q_k)\,(1 - Q_{k+1})q_{k+2}^{\alpha} + (1 - Q_k)\,(r_{k+1} - q_{k+1}^{\alpha})q_{k+2}\right] . \qquad (3)$$

Throughout this paper we will be using partial sums of the type given in (1), (2) and (3). We consider k to be a fixed time point for the rest of this paper. In addition, we will need the following assumptions.

Assumption I. For j = 0, 1 or 2, q_{k+j}^{α} is such that

$$q_{k+j}^{\alpha} = 0$$

Assumption II. For $0 \leq t < 3$ the following hold

i) $\quad {}_tr_k = t/3\ {}_3r_k$

ii) $\quad {}_{3-t}q_{k+t} = t/3\ {}_3q_k .$

Assumption I means that persons will not die immediately after the onset of a chronic condition. If an individual dies in the year the chronic condition is initiated, we will assume that this condition is not noted as a chronic condition related to death. It may be an acute analogue to the chronic condition such as an initial stroke or coronary event. This assumption seems appropriate for the intent of examining mortality patterns among the chronically ill.

The first part of Assumption II is usually referred to as the uniform transition rate. This means that the expected number of individuals entering the i^{th} chronic illness state in a fraction of the three-year interval is equal to that corresponding fraction of the total number of individuals entering the chronic disease state during the three-year period.

Part (ii) of Assumption II is commonly referred to in the actuarial literature as the Balducci hypothesis (see e.g., Batten 1978). The Balducci formulation does assume a locally decreasing hazard rate although the long-term hazard can be either increasing or decreasing. For human populations the Balducci hypothesis has been found to provide an adequate approximation of death rates for short periods of time.

Applying Assumption I to Equations (1), (2) and (3), with modifications to the subscripts used, we have

$$n_{k+2} = M_{k+2}\, q_{k+2} \qquad (4)$$

$$n_{k+1} + n_{k+2} = M_{k+2}\, {}_2q_{k+1} + N_{k+1}\, r_{k+1}\, q_{k+2} \qquad (5)$$

$$n_k + n_{k+1} + n_{k+2} = M_k\, {}_3q_k + N_k\, r_k\, {}_2q_{k+1} + N_{k+1}\, r_{k+1}\, q_{k+1} . \qquad (6)$$

To simplify these expressions, note the following hold:

$$M_{k+1} = M_k + N_k\, r_k - n_k$$

$$M_{k+2} = M_k + N_k\, {}_2r_k - n_k - n_{k+1}$$

$$N_{k+1} = N_k\,(1 - Q_k)$$

$$N_{k+2} = N_k\,(1 - Q_k)\,(1 - Q_{k+1}) .$$

Note that by definition of ${}_tr_k$ we have

$$N_{k+1}\, r_{k+1} = N_k\,(1 - Q_k)\, r_{k+1} = N_k\, r_k$$

and similarly

$$N_{k+2}\, r_{k+2} = N_k\, r_k .$$

Using these equations and Assumption II, Equations (4), (5) and (6) may be rewritten as

$$n_{k+2} = (M_k - n_k - n_{k+1})\,\frac{1}{3}\,{}_3q_k + \frac{2}{3}\,B\,{}_3q_k \qquad (7)$$

$$n_{k+1} + n_{k+2} = (M_k - n_k)\,\frac{2}{3}\,{}_3q_k + B\,{}_3q_k \qquad (8)$$

$$n_k + n_{k+1} + n_{k+2} = M_k\, {}_3q_k + B\, {}_3q_k , \qquad (9)$$

where $B = N_k\, r_k$.

Comparison of Risks

The hypothesis that two chronic diseases, indexed i and j, give the same risk of death during the Years k, k + 1, and k + 2 is the same as the hypothesis H_0: ${}_3q_k^{i} = {}_3q_k^{j}$ when

Assumptions I and II hold. One cannot simply compare the number of deaths with each chronic disease to test H_0, however, since the parameters r and M would be expected to be different for the different chronic illnesses. This is due to possible differences in population distributions or chronic disease onset rates. Equations (7), (8) and (9) do provide a method of testing H_0 as follows. Solving Equation (9) for M_{ik} and M_{jk} and Equation (8) for B_i and B_j and substituting into Equation (7) we have that if H_0 is true then

$$(n_{ik} - 2n_{ik+1} + n_{ik+2})(n_{jk} - n_{jk+1}) =$$

$$(n_{jk} - 2n_{jk+1} + n_{jk+2})(n_{ik} - n_{ik+1}). \qquad (10)$$

Multiplying and recombining terms we have that, if H_0 is true, then subject to sampling variability

$$n_{ik}(n_{jk+1} - n_{jk+2}) - n_{ik+1}(n_{jk} - n_{jk+2})$$

$$+ n_{jk+2}(n_{jk} - n_{jk+1}) = 0. \qquad (11)$$

Define the vector

$$\underset{\sim}{n} = (n_{ik}, n_{ik+1}, n_{ik+2}, n_{jk}, n_{jk+1}, n_{jk+2})^T$$

and the matrix

$$X = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Let $\hat{\underset{\sim}{n}}$ denote the estimate of $\underset{\sim}{n}$. Using a first order Taylor expansion of $\tilde{f}(\hat{\underset{\sim}{n}})$ defined as the left hand side of the Equation (11),

$$f(\hat{\underset{\sim}{n}}) - f(\underset{\sim}{n}) = (X\underset{\sim}{n})^T(\hat{\underset{\sim}{n}} - n) \quad . \qquad (12)$$

From (12) we have that the variance of $f(\hat{\underset{\sim}{n}})$ is given by

$$\mathrm{Var}(f(\hat{\underset{\sim}{n}})) = (X\underset{\sim}{n})^T (\mathrm{Var}\ \hat{\underset{\sim}{n}})(X\ \underset{\sim}{n}) \quad . \qquad (13)$$

Rewriting the hypothesis H_0 as H_0': $f(\underset{\sim}{n}) = 0$ we may use the Wald method given in Stroud (1971) to test the hypothesis (also see Tolley and Manton, 1983).

As an example, we compare the cancer mortality rates of white males in Minnesota with those of North Carolina. The data is from death certificates for the years 1975, 1976 and 1977. Table 1 gives the number of deaths for 65 year old cohort followed for three years.

Table 1

NUMBER OF DEATHS WITH CANCER NOTED ON DEATH CERTIFICATES FOR THE 65-YEAR OLD COHORT (WHITE MALES ONLY)

Year	Age	Minnesota	North Carolina
1975	65	97	114
1976	66	118	109
1977	67	124	159

Using formula (14) we have $f(\hat{\underset{\sim}{n}}) = -1080$ where i indicates North Carolina and j indicates Minnesota. However, $f(\hat{\underset{\sim}{n}})$ is highly variable. Using Equation (13), we approximate the standard deviation for this example as 799. The resulting one degree of freedom chi-square test of H_0' is $X_1^2 = 1.83$. The data show no reason to reject equal mortality experience for cancer victims in Minnesota as compared with North Carolina

Bibliography

Batten, R. W., 1978. Mortality Table Construction. Prentice-Hall, Englewood Cliffs, New Jersey.

Tolley, H. D. and K. G. Manton, 1983. "Multiple-Cause Models of Disease Dependency." Scand. Actuarial J. (to appear).

Stroud, T. W. F., 1971. "On Obtaining Large Sample Tests from Asymptotically Normal Estimators." Ann. Math. Statist., 42:1412-1424.

Research support by HCFA Grant Number 18 P-97710 and NIA Grant Number AG01159.