Life Table Methodology Applied to NHANES II Database for Analysis of Mortality Associated With Cholesterol/HDL Ratios

David Wesley, MD; Douglas Ingle

This article describes the association between total cholesterol to high-density lipoprotein (HDL)-cholesterol ratios and all-cause mortality in a large cohort of Americans with nearly 17 years of follow-up. Detailed actuarial life table methodology was used. It concludes that the relationship is best described as a J-shaped curve.

The total cholesterol (TC) to high-density lipoprotein (HDL)-cholesterol (HDL-C) ratio is commonly used in life underwriting. Our understanding of this ratio and its implications for mortality come from studies relating TC/HDL-C to cardiovascular disease. There are considerable data available supporting a positive correlation between greater risk for coronary heart disease and TC/HDL-C ratios. However, knowledge of its association to all-cause mortality would be far more useful to those who contract to pay death benefits for death due to any cause. A search of the literature found only 2 community-based studies that related all-cause mortality to TC/HDL-C, and these both targeted elderly persons.

Here we present data and analysis from the second National Health and Nutrition Examination Survey (NHANES II). From this study, both the vital status in 1992 and the results of baseline TC and HDL-C measured 16–17 years earlier are known for 7571 adult Americans. This provided an opportunity to study the ability of TC/HDL-C to predict all-cause mortality.

Actuarial life table methodology was chosen because it is the most familiar to the readers of this journal. Estimation of expected mortality from mean age or other oversimplification is a common weakness of the methodology. We took advantage of access to detailed records from the NHANES database and used seriatim techniques to calculate expected mortality on a person-by-person, year-by-year basis.

NHANES II

Congress enacted legislation under the National Health Survey Act of 1956 to promote analysis of morbidity data and related health information for the general US population. The second in a series of surveys, NHANES II was conducted in 1976–80. The Bureau of Census surveyed approximately 21,000 noninstitutionalized members of the US population. This survey obtained information on health histories and lifestyle practices.
initial contact, individuals between the ages of 25 and 74 were offered a physical examination that included, in most cases, an electrocardiogram, chest x-ray, blood profile, and urinalysis.

Population sampling was not entirely random. The initial contact person was chosen at random, but the field interviewer then surveyed a cluster of individuals around that initial contact (family members, neighbors, etc). The National Center for Health Statistics (NCHS) compiled and analyzed survey data. In 1992, NCHS randomly selected 9252 names from the original survey for a passive mortality follow-up. This data collection was completely passive in that no new contact was made with the individuals. Rather, 2 large databases, the National Death Index and the Social Security Administration Death Master File, were queried, enabling status determinations (alive, dead, or lost to follow-up).

**METHODOLOGY**

Individual identities were blinded in the NHANES II database through the assignment of a unique identification number.

Since each individual record contained the date of entry into the study and the date of death or the end of the observation period, it was possible to use actuarial life table analysis to summarize observed mortality. Furthermore, having age and sex for each individual made it possible to calculate expected mortality using a seriatim method. The seriatim method assigns an expected \( d \) based on \( q' \)s by age and sex for each individual for each year in the study. Summation of the \( d' \)s for each study interval produces an accurate expected mortality. When the data are available, this methodology might be considered superior to aggregating cohorts by age band, as the seriatim format appropriately weights expected mortality according to each individual record.

Entrants were drawn from the general US population and exposure to risk of death occurred between 1976 and 1992. Therefore, both the 1979-81 US Decennial Table \(^5\) and the 1989–91 US Decennial Table \(^7\) were potential sources for expected mortality. Since the overall mortality from the NHANES II sample cohort turned out to be better than either US population-based table, the 1989–91 US Population Life Table was chosen.

**Visual Basic for Applications**

Data management was done in Microsoft Access. This commonly available Microsoft Office application has a programming extension named Visual Basic for Applications (VBA). The seriatim methodology required the following determinations for each of 17 yearly intervals and for each of 7568 records: whether the individual was an interval entrant; the age of the individual at interval entry; whether the individual died or withdrew during the interval; the appropriate expected \( q \) for the individual. Running totals for entrants, withdrawals, deaths, and expected deaths had to be summed for each interval. Routines written in VBA formed an appropriate recordset for each interval and iterated through the recordset to create the summations.

**Life Table Analysis**

Entrants, withdrawals, exposures, deaths, expected deaths, cumulative survivals, mortality ratios, and excess death rates were calculated for each of the 17 one-year intervals using the methodologies that are well-established in insurance medicine. Knowledge of actual age, gender, exposure, and outcome status for each individual permitted any number of ways to stratify and/or aggregate the data. We chose to rank individuals by TC/HDL-C value and stratify them into 5 quintiles. Life table analysis was performed for each quintile.

Stratification was then increased 10-fold and a life table was created for each of 50 strata of the ranked TC/HDL-C values. These strata were defined by first identifying the TC/HDL-C ratios for each of the deaths. The
RESULTS

Demographics

There were 7571 records with TC, HDL-C, and mortality follow-up. Three of these contained obvious errors in the data (eg, died before blood was drawn). These were deleted, leaving 7568 records for analysis. Of these, 2127 were age 65 and up at the time of examination. Demographic features are summarized in Tables 1 and 2.

Observed Mortality

Figure 1 shows the observed mortality by quintile of TC/HDL-C. The middle quintile (number 3) experienced the lowest mortality, while higher and lower quintiles had greater mortality. This result suggests a J-shaped relationship between TC/HDL-C and all-cause mortality. This relationship is made clear in Figure 2, where aggregate mortality for each quintile is shown as a percentage of expected mortality (based on the 1989–91 US population table).

To better define the shape of the relationship, TC/HDL-C values were divided into 50 subranges defined by the occurrence of roughly 30 deaths per subrange (1660 total deaths). These data points are plotted as small circles in Figure 3. The data points were then fitted using linear regression to increasing powers of the TC/HDL-C variable. The third-degree polynomial curve has an R-value of .80. Higher degree polynomials improved the R-value only minimally. The fitted equation is

\[ MR = 183 - 48.2r + 6.32r^2 - 0.20r^3, \]

where \( MR \) is the mortality ratio (times 100) and \( r \) is the TC/HDL-C ratio. The fitted curve and 95% confidence limits are shown in Figure 3.
DISCUSSION

Readers of the Journal of Insurance Medicine are familiar with the use of life table methodology to reverse engineer from published cumulative survival curves. In this study, all the data were available and the challenge was to cope with it while using the familiar actuarial life table methods. Having large numbers of both lives and years of follow-up and the actual raw data to work with would seem to provide ample opportunity to use life-table analysis to provide a detailed picture of the relationship between TC/HDL-C and all-cause mortality. However, we quickly found that 17 intervals, 2 genders, 2 or more age groups, etc, rapidly lowered the number of events (deaths) per cell such that erratic numbers from cell to cell obscured any clear association.

We chose to aggregate data. We grouped TC/HDL-C values by quintiles, lumped both sexes, and other than for the cumulative mortality curves, we aggregated deaths and exposures over the entire follow-up period. But we feel our aggregates were of better quality than most summary data because they were built up from very exact calculations of exposures and expected mortality.

The cumulative mortality curve depicted in Figure 1 was the first result. This curve hinted at an unexpected J-shaped relationship that was clarified by the bar graph in Figure 2. This J-shape is similar to that seen between build and mortality in the 1979 Build Study.
and the association reported for total cholesterol and mortality in the MRFIT study. Increased all-cause mortality on the left side of the J-curve may be due to an association between chronic debilitating disease and lowered TC levels. That this may also be the case for TC/HDL-C as well would not be surprising since total cholesterol is the numerator for the ratio.

By defining numerous small subranges for TC/HDL-C, we were able to create enough data points to fit the curve seen in Figure 3. This curve and/or the corresponding equation can help define underwriting guidelines for TC/HDL-C results. Unfortunately, to allow that many data points, factors such as gender and age could not be separated out. Also, there is a fair amount of scatter in the

Figure 2. Mortality compared with US 1989–91.

Figure 3. Mortality ratio fitted curve.
mortality ratios, resulting in wide confidence limits for the fitted curve.

The study reported by Chyou and Eaker used a Cox proportional hazard model to demonstrate a statistically significant association between TC/HDL-C and all-cause mortality. We are currently developing a Cox proportional hazard model of our NHANES II follow-up data. Further study may show that this model can allow calculation of the risk associated with a given value of TC/HDL-C while adjusting for other factors.

REFERENCES