Chapter 3
Experiences from Forecasting Mortality in Finland

Juha M. Alho

3.1 Modeen and Törnqvist

The first official cohort-component projection of the population of Finland was prepared by Gunnar Modeen (1934a), an actuary with the Central Statistical Office of Finland at the time. Modeen's work had elements of genuine forecasting in that he commented on past trends in fertility, mortality, and migration, and discussed their possible long-term implications (Modeen 1934b). On the other hand, the work was rather schematic in nature. In particular, age-specific mortality was assumed not to change during the projection period, although Modeen was aware of its declining trend since the late nineteenth century (Modeen 1934a, 38). Unable to pinpoint the future rate of decline exactly, Modeen rejected any alternative assumption as speculative (cf., Modeen and Fougstedt 1938).

Analyses of mortality trends in presumably more advanced countries were used as *leading indicators* in the United States by Whelpton et al. (1947), for example. In Finland, Leo Törnqvist (1949) proposed similar methods. In particular, he used Swedish mortality as a target towards which he assumed Finnish mortality to converge. Both series were first transformed via a logistic type transformation. Then, the curves were aligned, and the Finnish curve was prolonged in accordance with the Swedish development.

A problem with Modeen's projection was that it soon became outdated. Fertility started to increase at the time the projection was published, and mortality continued to decline. Modeen's calculations suggested that the Finnish population would never exceed four million, but this mark had already been crossed by 1950. Törnqvist's collected works (Viren et al. 1981) do not mention the error of past forecasts as a motivation for his own early work. Nevertheless, the future statistics professor, a

J. M. Alho (✉)
University of Joensuu, Joensuu, Finland
e-mail: juha.alho@helsinki.fi

specialist in time-series analysis (among other fields, cf. Nordberg 1999), was well aware that forecasts cannot be made without error. He appears to have been the first to formulate the problem of uncertainty in population forecasting in probabilistic terms in Törnqvist (1949). Later, Törnqvist also conducted what must be one of the earliest assessments of the empirical accuracy of Finnish forecasts (including his own!) in Törnqvist (1967).

In this note, we will outline current developments in Finnish mortality forecasting. In Sect. 3.2, we describe the methods used by official forecasters. These derive mainly from the tradition of early cohort-component forecasters (cf. DeGans 1999). In Sect. 3.3, we discuss how uncertainty can be taken into account using probabilistic models and present-day computing facilities. We conclude in Sect. 3.4 by commenting on some applications for which mortality forecasts are particularly relevant.

3.2   Official Forecasts[1]

The arithmetic underlying cohort-component forecasts was understood a hundred years ago (DeGans 1999). Since the method relies on detailed assumptions concerning future age-specific rates, the real key to forecast accuracy lies with those assumptions. One would think that major improvements would have occurred during the past century, judging from the way the assumptions are formulated. Yet, the methods were essentially perfected by Whelpton back in the 1940s.

The two producers of official population forecasts in Finland are Statistics Finland and the Social Insurance Institution of Finland (or KELA, an abbreviation of the Finnish name). Since the forecasters of the two institutions cooperate on an informal basis, the forecasts have many similarities.

Both institutions produce forecasts approximately every 3 years. More frequent updates are made if unexpected developments occur. Both disaggregate the population by sex and single years of age (0, 1, 2, ..., 99, 100+). Currently both organizations forecast until 2050.

KELA produces a national forecast only, whereas Statistics Finland forecasts the population of every one of the 448 municipalities of Finland. In the case of mortality, the country is divided into three relatively homogeneous areas: Northern and Eastern Finland, which have a high level of mortality (due in particular to cardio-vascular diseases among males); the Swedish-speaking coastal areas, with low mortality; the rest of the country, with intermediate mortality. The reason for the low mortality among the Swedish speakers has not been established, but both socio-economic and lifestyle factors apparently play a part (Koskinen and Martelin 1995).

---

[1]The author would like to thank Mr. Matti Saari, Statistics Finland, and Mr. Markku Ryynänen, KELA, for information on the practice of forecasting. Any misunderstandings are the sole responsibility of the author.

Neither organization uses cause-specific mortality data in the preparation of their assumptions. This is contrast with the U.S. Office of the Actuary, for example (*e.g.*, Wade 1987). However, we have argued elsewhere that cause-specific information cannot be expected to increase forecast accuracy unless one of two conditions are met: either leading indicators can be identified in the preparation of forecasts, or structural changes can be anticipated based on other available information (as in the case of AIDS, for example) (Alho 1991).

Both organizations use trend extrapolation as a basis for their mortality forecast. Starting from a target value for life expectancy at birth, e0, Statistics Finland adjusts future age-specific mortality rates so that the implied increase in life expectancy gradually slows down until the target of e0 is reached. Age-dependent proportional adjustment is used to modify the jump-off rates. In KELA the starting point is a classification of individual ages into aggregates with similar mortality levels. Regression analysis is used in the log-scale to estimate rates of decline that gradually decelerate. The assumption, made by both organizations, that the rate of decline eventually falls off, is far from self-evident. In fact, we have used U.S. data to show that such an assumption has historically made the U.S. mortality forecasts *worse* than simpler trend extrapolations (Alho 1990).

Neither organization formulates their targets on a cohort basis although both occasionally examine cohort trends to see whether there are any irregularities. A current example of such an irregularity was reported by KELA: the female cohorts born in the 1950s appear to have higher mortality than cohorts born earlier, during WWII.

The methods of trend extrapolation used by the organizations blend judgment and empirical analysis. Neither organization has experimented with the method proposed by Lee and Carter (1992). Its performance in regard to ages 65+ in Finland, was investigated in a University of Joensuu *pro gradu* thesis by Eklund (1995), who found that a one-dimensional singular-value decomposition produced a good fit to the data. Because of random variation, however, the resulting forecast was not always an increasing function of age.

In addition to the trend forecast, KELA produces another mortality variant in which it is assumed (as Modeen did) that mortality will remain at the jump-off level. Statistics Finland limits itself to a single variant even though high and low variants have previously been used in national forecasts.

3.3 Predictive Distribution of Mortality

A major contribution by Törnqvist (1949) was that he was apparently the first to maintain that since the future values of a vital rate cannot be totally known, they must be treated as random variables. The actual future values are then "samples" from their distributions. In modern terminology, the uncertainty of the future value is

expressed in terms of a predictive distribution that represents both our best guess and its uncertainty. The distribution is conditioned on all information available at the jump-off time of the forecast (*e.g.*, Gelman et al. 1995, 9).

Törnqvist's contribution may have been ahead of its time. In particular, correct formal treatment of the predictive distribution would have been difficult before the availability of high-speed computing. In recent years, the potential usefulness of a probabilistic approach to uncertainty has been noted on several occasions.[2] At the University of Joensuu, we have written a computer program, PEP (Program for Error Propagation), which is capable of simulation samples from a wide range of predictive distributions.

The main concept of PEP is that it allows us to describe the uncertainty connected with a forecast at the time it is being made. All sources of uncertainty – age-specific fertility and age and sex-specific mortality and migration – are taken into account and propagated throughout to derive the predictive distribution of the population. In this sense, PEP is merely a stochastic version of the cohort-component bookkeeping system. The usefulness of the results depends on the assumptions underlying the calculations. The user of PEP must specify a point forecast for each of the vital rates for all future years, just as in ordinary cohort-component forecasting. An additional step is required in the form of specifying the uncertainty surrounding the forecast.

Suppose $R(j,t)$ is the mortality rate for age $j = 0, 1, \ldots, \omega$ in a future year $t = 1, 2, \ldots, T$. PEP assumes that

$$R(j, t) = \exp(\hat{r}(j, t) + X(j, t)),$$

where $\hat{r}(j, t)$ is the point forecast of the log-rate, and $X(j,t)$ is a random error with a mean of $E[X(j,t)] = 0$. The random error can always be written in the form

$$X(j, t) = \varepsilon(1, t) + \cdots + \varepsilon(j, t).$$

In PEP, the *error increments* $\varepsilon(j,t)$ are *assumed* to be of the form

$$\varepsilon(j, t) = S(j, t)\left(\eta_j + \delta(j, t)\right),$$

where the $S(j,t)$'s are known scale factors that can be chosen to match any sequence of error variances $\text{Var}(X(j,t))$ that increases with t. Fixing j, we can think of the terms η_j as representing errors in forecasted trends. In the case of mortality, the trend corresponds to the rate of decline, for example. Since the terms $\delta(j,t)$ are independent for any fixed j, they represent unpredictable random variation. The relative roles of the two types of uncertainties derive from the assumption $\eta_j \sim N(0, \kappa_j)$, and $\delta(j, t) \sim N$

---

[2]Review of Land (1986); "Special Section on Statistical Analysis of Errors in Population Forecasting and Its Implications on Policy," *Journal of Official Statistics*, September 1997; "Frontiers of Population Forecasting," *Population and Development Review*, 1998 Supplement; review of U.N. forecasts by the National Research Council (2000).

$(0, 1 - \kappa_j)$, where $0 \leq \kappa_j \leq 1$. The terms η_j are assumed to be independent of the terms $\delta(j,t)$. Finally, the terms η_j can either have a constant correlation across j, or an AR(1) type correlation. The same is true for the $\delta(j,t)$'s, when t is fixed. This *scaled model* for error was introduced in Alho and Spencer (1997).

In Alho (1998) we provide details of the application of PEP to the population of Finland for 1999–2050. The point forecasts for each vital rate were as specified by Statistics Finland. We now present some details on the treatment of uncertainty in the mortality forecast.

Age-specific mortality data in 5-year age-groups 0–4, 5–9, ..., 75–79, and 80+ were available for the years 1900–1994. After a preliminary analysis, the data were aggregated into the broader age groups 0–4, 5–34, 35–59, 60–79, and 80+ by adding the age-specific rates together. This increased the stability of the trends. The analysis was carried out in terms of the logarithm of the sum (cf. Alho 1998, Figures 5a–e, pp. 19–21). The unusual values produced by the civil war in 1918 and WWII in 1939–1944 were smoothed using values from the previous year. For each of the five broad age groups, we produced baseline forecasts as follows:

- Starting from year y = 1915, we used the data for the previous 15 years (y, y – 1, ..., y – 15) to calculate a trend forecast for all future years until 1994.
- A linear trend was estimated from the first and the last observation of the 15-year data period.
- In case the linear trend was positive, it was replaced by a constant value (*i.e.*, slope = 0).

For each y = 1915, 1916, we calculated the empirical forecast error for lead times t = 1, 2, ..., 50. For each lead time t, we could then estimate the standard deviation of the error around zero (*i.e.*, assuming that the forecasts are unbiased). This would give us estimates of Var(X(j,t)) directly, from which the scales S(j,t) could be deduced. However, it turned out that especially for younger ages the estimates were somewhat erratic because of the large random (Poisson-type) variation in the counts. Therefore, final estimates were produced by averaging the estimates from the six time series corresponding to the three broad age groups of 35–59, 60–79, and 80 + for males and females. The resulting estimate of the standard deviation of the relative error starts from approximately 0.06 at t = 1 and increases in a linear fashion to about 0.6 at t = 50. Otherwise expressed, the relative error one might expect for a single age group increases from 6% to roughly 60% in 50 years (cf., Alho 1998, Fig. 6, p. 22). These estimates were used for all ages.

The results were checked by fitting an ARIMA(1,1,0) model to the data series, and similar results were obtained (Alho 1998, Fig. 6, p. 22).

The parameter κ was estimated by the least-squares method. The single value $\kappa = 0.149$ was applied for all ages.

An AR(1) process was used to model the autocorrelation of the error terms η_j and $\delta(j,t)$ across age j. Otherwise expressed, the correlation was assumed to be $\rho^{|i - j|}$ for any two single years of age i and j, where the empirical estimate $\rho = .945$ was used for η_j's and $\rho = .977$ was used for $\delta(j,t)$'s. Finally, a parameter for contemporaneous
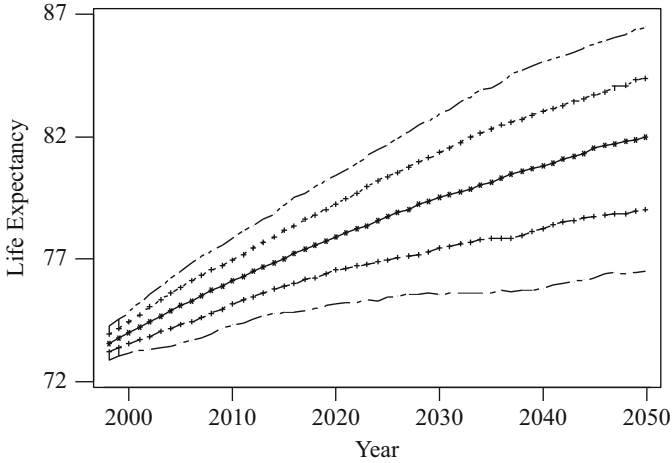
Fig. 3.1 Predictive distribution of male life expectancy in Finland in 1998–2050
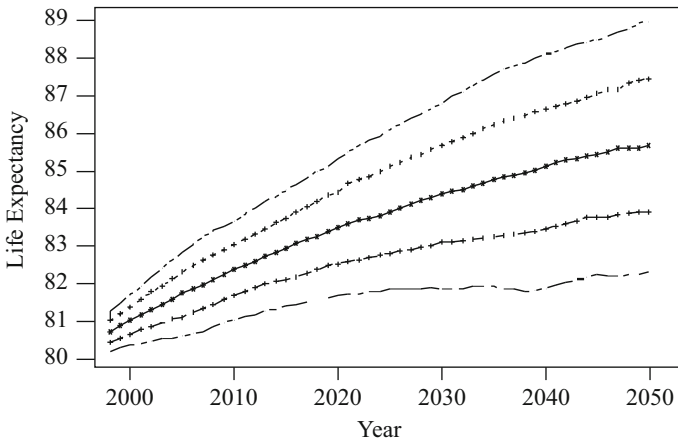

Fig. 3.2 Predictive distribution of female life expectancy in Finland in 1998–2050

crosscorrelation between the error of male mortality and the female mortality was estimated as .795.

The details are fairly complex. One way to assess the reasonableness of the procedures is to consider their implications for life expectancy. Figure 3.1 has a predictive distribution for male life expectancy at birth, and Fig. 3.2 has a plot for female life expectancy. The median of the predictive distribution is 82.0 for males and 85.6 for females in 2050. A 50% prediction interval (or interval between the first and third quartile) is [79.0, 84.4] for males and [83.9, 87.5] for females. An 80% prediction interval for males is [76.5, 86.5] and [82.3, 89.0] for females. The narrower spread for females is probably due to their lower level of mortality.

Two concerns can be raised concerning the intervals. First, the long-term point forecasts are based on an eventual slowdown of the decline in mortality; this may make the Finnish forecast too conservative, as it did in the U.S. earlier. However, we may note that the life expectancy implied by the current Swedish forecasts for males is 82.6 and for females 86.5 years in 2050. In the intermediate variant of the Norwegian forecast, the corresponding ages assumed are 80.0 and 84.5 years. We see that despite the assumption of a slow-down in the mortality decline in Finland, the Finnish forecast is the most optimistic of the three in terms of improvement, since the current life expectancy in Finland is the lowest. Even though the Finnish point forecast may be too low at the end of the forecast period, from this perspective the Finnish forecast appears less conservative.

Second, in view of the vast potential for new medical advances, one could argue that the range of uncertainty expressed by the widths of the intervals might be overly narrow. Two arguments seem relevant here. For the U.S. (both sexes combined), Lee and Carter (1992, p. 660, Fig. 3.1) calculated model-based 95% intervals for life expectancy 50 years ahead. The width of these intervals was approximately 8.4 years. In a normal model, the corresponding width for an 80% interval would be approximately 5.5 years. Thus, our intervals are clearly wider. In a discussion of the paper by Lee and Carter, we noted that by including all sources of variation, the Lee and Carter intervals would have been approximately one half wider (Alho 1992). This would have resulted in estimates close to ours. (One could also argue that in a large country with heterogeneous sub-populations there might be some offsetting variation, resulting in a national average more stable than in a small homogeneous country. While conceivable, this possibility does not seem to be an adequate reason for inflating the Finnish intervals, since it does not show up in the Finnish time series).

A related criticism suggests that future advances in medical knowledge may be so unprecedented that intervals based on past outcomes are too narrow. However, in 1915, at the start of our observation period, both life expectancy for both males and females was substantially lower – 43.2 and 49.2 years (Kannisto and Nieminen 1996); the improvements by the end of the twentieth century were 27.5 and 29.5 years, respectively. Our estimates reflect the variation in this turbulent period of major improvement. The future can be even more volatile, but the advantage of our intervals is that they correspond to actual past variation, rather than to a subjective assessment. Of course, a subjective assessment may be used as a basis for other calculations that would complement our own.

3.4   Applications

Now that we have the analytical capability to produce predictive distributions for future vital rates and future population, it is of some interest to consider how they might be applied. There are two aspects to this.

First, it is critical that we understand how the predictive distribution can be understood. As noted, *e.g.*, in Alho (1998), predictive distributions can be based on (1) formal statistical models, (2) errors of past forecasts can be used to estimate the error of future forecasts, (3) errors of baseline forecasts can be used to estimate future error, and (4) error specification can be purely subjective. Of course, any mixture of the four is also a possibility. The results we have shown rely primarily on (3), although they have elements of (1) and (4), as well. The aim was to provide an empirical assessment of the difficulty of forecasting (or "forecastability") of the vital processes for different times. As such, the *results for mortality correspond to the uncertainty in mortality forecasting during the twentieth century*. One can reasonably question whether forecasting is now easier, or more difficult, than in the past, but at least we now have a quantitative empirical assessment of how things were before.

Second, the predictive distribution can be used to address numerous social-policy issues that depend on future population and its age structure. For example, in Alho et al. (2001) we review an example in which output from PEP was used in combination with the Finnish overlapping-generations model to devise alternative pension-funding rules, and another example in which output from PEP was used to assess the stability of the current rules for state aid to municipalities. In a University of Joensuu *pro gradu* thesis, Polvinen (2001) used PEP to form a predictive distribution of the so-called generational accounts. All these questions are of fundamental concern for the long-term planning of the social-support systems in Finland. In no case has the effect of uncertain population age-distribution previously been recognized. Other applications have been presented by Lee and Tuljapurkar (1998) for the Social Security system of the United States, for example. Further research opportunities are discussed in Auerbach and Lee (2001).

References

Alho, J. (1990). Stochastic methods in population forecasting. *International Journal of Forecasting, 6*, 521–530.

Alho, J. (1991). Effect of aggregation on the estimation of trend in mortality. *Mathematical Population Studies, 3*, 53–67.

Alho, J. (1992). Comment. *Journal of the American Statistical Association, 87*, 673–675.

Alho, J. (1998). *A stochastic forecast of the population of Finland* (Reviews 1998/4). Helsinki: Statistics Finland.

Alho, J., & Spencer, B. D. (1997). The practical specification of expected error of population forecasts. *Journal of Official Statistics, 13*, 203–225.

Alho, J., Lassila, J., Salo, M., Spencer, B. D., & Valkonen, T. (2001). *Predictive distributions for future population with applications to aging*. Manuscript.

Auerbach, A. J., & Lee, R. D. (Eds.). (2001). *Demographic change and fiscal policy*. Cambridge: Cambridge University Press.

DeGans, H. A. (1999). *Population forecasting 1895–1945*. Dordrecht: Kluwer.

Eklund, K. (1995). *Kuolevuuden mallittaminen ja ennustaminen Suomessa vanhusväestön keskuudessa*. Pro Gradu thesis, University of Joensuu.

Gelman, A., Carlin, J. B., Rubin, D. B., & Stern, H. S. (1995). *Bayesian data analysis*. London: Chapman-Hall.

Kannisto, V., & Nieminen, M. (1996). *Revised life tables for Finland. Population 1996:2*. Helsinki: Statistics Finland.

Koskinen, S., & Martelin, T. (1995). *Why is mortality low among the Swedish speaking minority in Finland?* Paper presented at the 11. Nordic Demographic Symposium, June 1995, in Helsinki, Finland.

Land, K. C. (1986). Methods for national population forecasts: A review. *Journal of the American Statistical Association, 81*, 888–901.

Lee, R. D., & Carter, L. (1992). Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association, 87*, 659–671.

Lee, R. D., & Tuljapurkar, S. (1998). Uncertain economic futures and social security finances. *American Economic Review*, May, 237–241.

Modeen, G. (1934a). *Finlands framtida befolkningsutveckling* (Tilastokatsauksia No: 10). Helsinki: Tilastokeskus.

Modeen, G. (1934b). Suomen väkiluvun tuleva kehitys ja sen taloudelliset seuraukset. *Kansantaloudellinen aikakauskirja, VI*, 351–378.

Modeen, G., & Fougstedt, G. (1938). *Laskelmia Suomen vastaisesta väestönkehityksestä* (Sarja VI, Väestötilastoa 89). Helsinki: Tilastokeskus.

National Research Council. (2000). *Beyond six billion: Forecasting the world's population*. Washington, DC: National Academies Press.

Nordberg, L. (1999). Leo Törnqvist – The "Grandfather" of Finnish Statistics. In J. Alho (Ed.), *Statistics, registries and science. Experiences from Finland*. Helsinki: Statistics Finland.

Polvinen, A. (2001). *Stokastisen ennusteen käyttö julkisen talouden kehityksen arvioinnissa* (Eläketurvakeskuksen raportteja 2001:23). Helsinki: Eläketurvakeskus.

Törnqvist, L. (1949). Näkökohdat, jotka ovat määränneet primääristen prognoosiolettamusten valinnan. In J. Hyppölä, A. Tunkelo, & L. Törnqvist (Eds.), *Suomen väestöä, sen uusiutumista ja tulevaa kehitystä koskevia laskelmia* (Tilastollisia tiedonantoja 38). Helsinki: Tilastokeskus.

Törnqvist, L. (1967). The post-war development of Finland compared with predictions made after the war. In *World population conference, 1965, Vol. III, Projections, measurement of population trends*. New York: United Nations.

Viren, M., Alho, K., Teräsvirta, T., Vartia, P., & Vartia, Y. (1981). *Collected scientific papers of Leo Törnqvist* (Series A7). Helsinki: The Research Insititute of the Finnish Economy.

Wade, A. (1987). *Social security area population projections 1987*. Baltimore: U.S. Office of the Actuary.

Whelpton, P. K., Eldridge, H. T., & Siegel, J. S. (1947). *Forecasts of the population of the United States 1945–1975*. Washington, DC: U.S. Bureau of the Census.