Practice of Epidemiology

Comparison of Bayesian Random-Effects and Traditional Life Expectancy Estimations in Small-Area Applications

Marcel F. Jonker*, **Frank J. van Lenthe, Peter D. Congdon, Bas Donkers, Alex Burdorf, and Johan P. Mackenbach**

* Correspondence to Marcel F. Jonker, Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands (e-mail: m.jonker@erasmusmc.nl).

There are several measures that summarize the mortality experience of a population. Of these measures, life expectancies are generally preferred based on their simpler interpretation and direct age standardization, which makes them directly comparable between different populations. However, traditional life expectancy estimations are highly inaccurate for smaller populations and consequently are seldom used in small-area applications. In this paper, the authors compare the relative performance of traditional life expectancy estimation with a Bayesian random-effects approach that uses correlations (i.e., borrows strength) between different age groups, geographic areas, and sexes to improve the small-area life expectancy estimations. In the presented Monte Carlo simulations, the Bayesian random-effects approach outperforms the traditional approach in terms of bias, root mean square error, and coverage of the 95% confidence intervals. Moreover, the Bayesian random-effects approach is found to be usable for populations as small as 2,000 person-years at risk, which is considerably smaller than the minimum of 5,000 person-years at risk recommended for the traditional approach. As such, the proposed Bayesian random-effects approach is well-suited for estimation of life expectancies in small areas.

Bayesian analysis; life expectancy; Monte Carlo method; small-area analysis

Abbreviations: MCMC, Markov chain Monte Carlo; RMSE, root mean square error; SMR, standardized mortality ratio.

Public health researchers are frequently required to compare mortality in different geographic areas. The standardized mortality ratio (SMR), which measures the excess or deficit of mortality compared with a chosen standard population, is commonly used for this purpose. However, inter-SMR comparisons are valid only if the areas compared have identical population structures (1–3). Accordingly, 2 or more areas can only be compared via their SMRs if they have (at least) very similar population age structures. This condition is often violated, particularly in small-area studies, where local differences are not averaged out as much as in larger geographic areas.

Life expectancies, in contrast, allow for comparisons between geographic areas to be made without having to assume a particular standard population. Even for areas with very diverse population structures, life expectancies are directly comparable, and, in contrast to SMRs, differences between areas are also more intuitive and straightforward to interpret (3–5). As such, life expectancies are preferable over SMRs as the outcome measure in studies concerning geographic mortality differences.

One of the major problems with life expectancies, however, is that the standard life-table method cannot be used to calculate life expectancy estimates for populations smaller than approximately 5,000 person-years at risk; below this size, the bias in the estimates, as well as the size of the accompanying standard errors, becomes too large for meaningful analysis (6–8). In virtually all small-area applications, a threshold of 5,000 person-years at risk implies that a substantial number of areas cannot be included in the analyses, even after pooling several years of data. Consequently, life expectancies are usually avoided in small-area analyses.

With a random-effects method, on the other hand, it should be possible to calculate accurate life expectancies

with sufficiently small standard errors for significantly smaller populations at risk. This can be achieved using a modeling approach that recognizes correlations (i.e., borrows strength) between different age groups, geographic areas, and sexes to stabilize the life expectancy estimates. These models are conveniently fitted using a Bayesian estimation approach. A further advantage of a Bayesian estimation approach is that it facilitates the calculation of standard errors and confidence intervals without having to rely on asymptotic normality assumptions that generally do not hold in small-area applications. It thereby produces standard errors that remain accurate for much smaller populations (9, 10).

Thus far, however, there has been no clear evidence that a Bayesian random-effects approach indeed results in more accurate life expectancy estimates than the traditional method. Neither is there an indication of a minimum required population size for estimating small-area life expectancies using a Bayesian random-effects approach. Thus, our aims in this study were to compare the relative efficiency of both methods (in terms of accuracy of the estimates and precision of the standard errors and confidence intervals) and to provide evidence that a Bayesian random-effects approach can indeed be used to calculate reliable life expectancies for smaller populations than with the traditional approach.

MATERIALS AND METHODS

Benchmark data

In our analysis, we use Monte Carlo simulations to evaluate the performance of traditional and Bayesian life expectancy calculations for (hypothetical) populations of varying size. A total of 6 simulations are performed, for populations of 500, 1,000, 2,000, 5,000, 10,000, and 25,000 person-years at risk. Instead of creating fully synthetic benchmark life-table data, with artificially imposed correlations between the different age groups, geographic areas, and sexes, the simulations are designed to mimic the exact male and female population age structures, age-specific mortality rates, and geographic locations of 33 European benchmark countries. For these countries, complete life-table data for the 2005–2006 period are obtained from either the Human Mortality Database (11) or the Eurostat Populat database (12). These data (based on millions of inhabitants) serve as the input for the simulations, implying that the simulations recreate Europe as if it were a city, with European countries as its neighborhoods.

The 33 European benchmark countries have very different life expectancies, ranging from close to 80 years in Scandinavia and the Mediterranean to approximately 60 years in Russia and Ukraine (see Figure 1). These differences are substantial and may be larger than typically encountered in small-area applications (e.g., Congdon (10) reports a range of 12 years between the highest and lowest life expectancies in wards in eastern England). On the other hand, these differences do allow for a meaningful comparison between the two methods without an inherent bias in

favor of the Bayesian random-effects approach due to the inclusion of very similar areas in the simulations.

Creating the hypothetical small-area data sets

Each of the 6 simulations starts with the scaling down of the European benchmark populations to the required population size (i.e., 500, 1,000, 2,000, 5,000, 10,000, or 25,000 person-years at risk). This is programmed in MATLAB (The MathWorks, Inc., Natick, Massachusetts) and performed deterministically in order for the hypothetical small-area populations to mimic the exact age distribution of the benchmark countries. Subsequently, the numbers of deaths are simulated randomly—10,000 times per simulation—and independently for all age groups and both sexes using draws from a Poisson distribution with means set to the age- and sex-specific mortality rates of the benchmark countries. After the input data are generated, the data are automatically aggregated into abridged life tables that are ready for life expectancy estimations. Based on the recommendations of Toson and Baker (6) and Eayres and Williams (7), standard 5-year abridged life tables with ≥85 years as the final age interval are used without any adjustment for age-specific death counts of zero within the life table. This results in 10,000 life-table data sets per simulation, each with a slight variation in the number of deaths but with average life expectancies exactly equal to those of the benchmark countries.

Life expectancy estimations

In the estimation phase, life expectancies and accompanying standard errors are calculated for all 10,000 life-table data sets per simulation using the traditional method (in MATLAB) and the Bayesian random-effects method (in OpenBUGS (http://www.openbugs.info/)). Similar to Toson and Baker (6), Eayres and Williams (7), and Williams et al. (13), the simulations focus on male life expectancies only (the difference in performance between the sexes is very small), and, following the recommendations in the same studies, all life expectancies are calculated using the Chiang life-table approach (14). This holds for both the traditional method and the Bayesian method, but an issue that pertains only to the traditional life-table calculations is the occasional occurrence of zero deaths in the final age intervals. Whereas the Bayesian random-effects approach does not break down, the mortality rate of the final age interval will be exactly zero for the traditional approach, resulting in an infinite mean length of survival (1/mortality rate) and consequently an infinite life expectancy. To avoid these problems, zero death rates in the final age interval of the traditional life-table calculations are replaced by the average observed sex-specific mortality rate of the final age groups.

Another computational problem that only concerns the traditional method is the occasional occurrence of age-specific mortality rates that are equal to or higher than 0.40. In larger populations, such mortality rates generally do not occur, but in smaller populations it occasionally happens that a (nonfinal) age group has, for example, a population of 5 and a number of deaths of 2. Here the Chiang formula
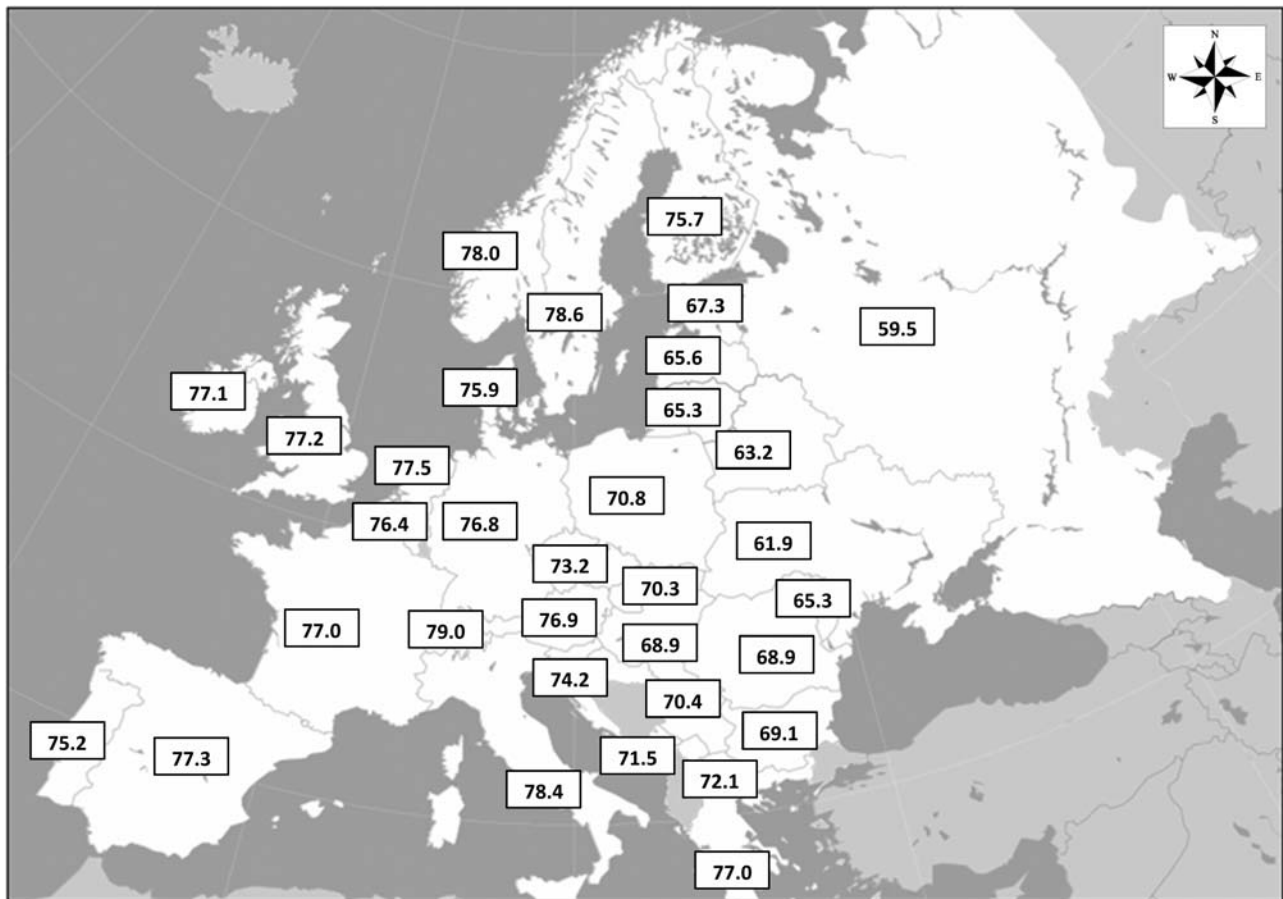
Figure 1.   Male life expectancy (in years) in the 33 European benchmark countries, 2005–2006.

for the conditional probability that persons who enter the age interval will survive the age interval (P_x) is zero, implying that the life-table calculations no longer include subsequent age groups and that the standard life expectancy formulas break down. In these cases, the life expectancies of subsequent age groups are set to zero to exclude them from the life-table calculations.

Bayesian modeling approach

For the Bayesian method, a relatively basic model and a more advanced model are evaluated in the simulations. Both models pool strength over sexes, age groups, and geographic areas using a random-effects method that includes multivariate (random) spatial effects that account for spatial clustering in mortality rates and multivariate (random) age effects that capture the usually high correlation between mortality rates for successive ages. The second model is more flexible than the first by allowing for age × area interactions, which means that the second model relaxes the assumption of a uniform age gradient per sex in the mortality patterns of the simulated areas. The model specifications are described in detail in the Appendix, and both models

are realistic examples of Bayesian life expectancy models. These models can also be fitted using a frequentist estimation approach. An advantage of the Bayesian estimation approach, however, is that it is easily extendible to more complicated models and is more convenient for obtaining reliable standard errors and confidence intervals, which are directly available from the posterior life expectancy distributions that are already estimated for the life expectancy calculations.

The Bayesian models are fitted in OpenBUGS using iterative Markov chain Monte Carlo (MCMC) sampling techniques. The estimations for model 1 and model 2 both start with 25,000 burn-in MCMC iterations to allow the chains to converge, followed by 100,000 MCMC iterations with a thinning interval of 10 to reliably approximate the posterior life expectancy distributions. Note that we use relatively good starting values and also use a very conservative burn-in period in order to avoid having to inspect convergence for all 120,000 Bayesian estimations. Instead, convergence for the first 10 iterations of each simulation is evaluated using the Gelman-Rubin criteria based on 2 parallel chains (15). Convergence was always obtained within 15,000 MCMC iterations, and, for all other estimations, convergence

Table 1.   Mean Error (Bias), Standard Error, and Root Mean Square Error for Male Life Expectancies (in Years), by Methodological Approach and Population Size[a]

Population Size	Bias			Standard Error			Root Mean Square Error		
	Traditional	Bayesian 1	Bayesian 2	Traditional	Bayesian 1	Bayesian 2	Traditional	Bayesian 1	Bayesian 2
500	1.1	0.8	0.1	6.2	3.8	3.8	6.3	3.9	3.8
1,000	0.7	0.5	0.0	4.4	3.0	2.8	4.5	3.0	2.8
2,000	0.6	0.2	0.0	3.1	2.2	2.0	3.1	2.3	2.0
5,000	0.3	0.1	0.0	2.1	1.8	1.5	2.2	1.8	1.5
10,000	0.2	0.1	0.0	1.5	1.5	1.1	1.5	1.5	1.1
25,000	0.1	0.0	0.0	0.9	1.2	0.8	0.9	1.2	0.8

[a] Results are based on 10,000 simulation iterations per population size.

was assumed after 25,000 MCMC iterations. The Bayesian life expectancy estimations are distributed on the Dutch Life Science Grid (16) to considerably reduce the required computation time for the large number of regressions required for the simulations.

Comparison between methods

After the life expectancy estimations are completed, the point estimates and calculated standard errors of the traditional method and the means, standard deviations, and 95% credibility intervals of the Bayesian posterior distributions are aggregated and summarized in MATLAB. Together, these estimates form distributions of life expectancies and standard error estimates from which reliable inferences can be made. The method that produces 1) the most accurate life expectancies and 2) the most accurate estimated standard errors and confidence intervals is preferred.

Regarding the first criterion, the accuracy of the life expectancy estimates is characterized by two properties: 1) the bias of the life expectancy estimates and 2) the standard error of the life expectancy estimates. These measures capture the systematic error and random sampling variability of the estimates, respectively. The overall accuracy of the life expectancy estimates is summarized by the root mean square error (RMSE), a measure that combines the bias and standard error of the estimates into a single composite measure of absolute fit. Smaller values indicate a closer fit to the benchmark life expectancies; accordingly, the method with the smallest RMSE is preferred.

Regarding the second criterion, the accuracy of the reported standard errors and confidence intervals is evaluated by comparing the average estimated standard error with the simulated random error and a tally of how often the benchmark life expectancies are located within the reported 95% confidence intervals. The latter is referred to as the coverage of the 95% confidence intervals. The first criterion gives an indication of whether the reported standard errors reflect the true random sampling variability of the estimates, while the second criterion also takes the systematic error of the estimates into account and provides a more direct indication of how reliable the reported 95% confidence intervals actually are: Values below 95% indicate

that the confidence intervals are too optimistic, and values greater than 95% indicate that they are too conservative.

Comparison in a real-life example

The methods are also compared in a real-life example. Similarly to the simulations, standard 5-year abridged life tables with ≥85 years as the final age group are used for the life expectancy calculations. Unlike the simulations, there is no benchmark that can be used to formally compare the quality of the life expectancy estimates. Instead, the real-life example is used to more tentatively substantiate the simulation results. Firstly, the example shows how both approaches handle real-life data and associated problems that often occur in real-life analyses. Secondly, the sizes of the reported standard errors of both approaches are compared, which should be in line with the sizes of the reported standard errors in the simulations. Thirdly, the relative size of the bias of both approaches is investigated by examining the ratios of the estimated life expectancies per neighborhood; these should concur with the relative bias as reported in the simulations.

The selected geographic area for the example is the city of Rotterdam, the second-largest city in the Netherlands, with a population of approximately half a million people. The required population and mortality data for Rotterdam are obtained from Statistics Netherlands and cover the 2008–2009 period. Rotterdam has 89 neighborhoods, of which several have little or no population (e.g., because they contain a public hospital, a city park, an airport, a zoo, or various industrial and harbor areas). Because it is not permittted by Statistics Netherlands to export life tables for populations smaller than 1,000 person-years at risk, 25 neighborhoods are excluded from the analysis. Together the excluded neighborhoods comprise 1% of the total population of Rotterdam.

RESULTS

Accuracy of the life expectancy estimates

Table 1 shows the bias, standard error, and RMSE of the traditional and Bayesian life expectancy calculations. As can be seen, the life expectancy estimates of the traditional

Table 2. Mean Estimated Standard Error and 95% Confidence Interval Coverage[a] for Male Life Expectancies (in Years), by Methodological Approach and Population Size[b]

Population Size	Estimated Standard Error			Coverage of 95% Confidence Interval		
	Traditional	Bayesian 1	Bayesian 2	Traditional	Bayesian 1	Bayesian 2
500	5.0	4.3	3.4	86	95	89
1,000	4.0	3.2	2.5	90	94	92
2,000	2.7	2.2	1.9	91	93	94
5,000	1.9	1.7	1.5	92	91	95
10,000	1.4	1.3	1.2	93	89	96
25,000	0.9	0.9	0.8	94	84	96

[a] Coverage denotes the percentage of benchmark life expectancies contained in the estimated 95% confidence intervals. Values below 95% indicate confidence intervals that are too optimistic, and values greater than 95% indicate confidence intervals that are too conservative.

[b] Results are based on 10,000 simulation iterations per population size.

method are increasingly (upwards) biased for smaller populations. In contrast, the bias of Bayesian model 1 is considerably lower and the bias of Bayesian model 2 is close to zero.

All methods have standard errors that decrease with larger population sizes. The standard errors of the Bayesian approach, however, are approximately 40% smaller than those of the traditional approach for population sizes of 5,000 or less. For larger populations, Bayesian model 2 still performs better than the traditional approach, but Bayesian model 1 has a less flexible specification that becomes increasingly restrictive when more data become available. As a result, Bayesian model 1 performs even worse than the traditional approach for population sizes of 25,000.

This is also reflected in the RMSE, which indicates that Bayesian model 1 has a lower RMSE than the traditional approach for population sizes smaller than 10,000 person-years at risk but a higher RMSE for population sizes of 25,000. Bayesian model 2, on the other hand, has the lowest RMSE regardless of population size and clearly performs best in terms of the accuracy of the life expectancy estimations. Compared with the traditional approach, it has an approximately 40% lower RMSE for populations below 5,000 and still a 17% lower RMSE for populations as large as 25,000.

Accuracy of the estimated standard errors and 95% confidence intervals

Table 2 shows the mean estimated standard errors and the coverage of the 95% confidence intervals for the traditional and Bayesian approaches. The estimated standard errors are highly correlated with the simulated standard errors in Table 1, with a correlation coefficient of 0.98. The traditional method, however, increasingly underpredicts the true variability of the life expectancy estimates for smaller populations. Bayesian model 1, on the other hand, overpredicts the true variability for smaller populations and underpredicts the true variability for larger populations, whereas Bayesian model 2 performs best with modest (but increasing) deviations for population sizes of 2,000 or less.

When one examines the coverage of the 95% confidence intervals, the results in Table 2 indicate that the traditional

approach performs increasingly worse for smaller populations, which reflects the bias in the estimates and the underprediction of the standard errors for smaller populations. In contrast, Bayesian model 1 performs well for smaller populations (where the small amount of bias mitigates the effect of the standard errors that are too large) but not for larger populations, where the size of the true standard errors is increasingly underpredicted. The more flexible Bayesian model 2 again performs best: It has a coverage of 94% for populations as small as 2,000 (note that the traditional approach requires populations of 25,000 to reach this level of accuracy), has a coverage of 95% for populations of 5,000, and even becomes somewhat conservative for populations of 10,000 and 25,000.

Performance in a real-life example

Figure 2 and Figure 3 summarize the results for the real-life example. Firstly, the ratios of estimated life expectancies in the city of Rotterdam, as depicted in Figure 2, substantiate the simulation results by showing that the traditional approach is slightly upwards biased relative to the Bayesian random-effects approach; this is indicated by the negatively sloped log-linear regression curve. Secondly, the ratios of estimated standard errors as shown in Figure 3 substantiate the simulation results by showing that the Bayesian approach produces smaller standard errors than the traditional approach, particularly for smaller populations.

Figures 2 and 3 also depict several significant deviations between the Bayesian and traditional approaches. These occur in a limited number of situations, which involve small populations with 1) zero populations at risk in the life table, 2) zero deaths in the life table, 3) zero deaths in the final age group, or 4) 1 or 2 deaths in the entire life table but for very young persons. Here the traditional approach has computational difficulties and generates extreme life expectancies with unrealistically small standard errors (or infinite life expectancies with infinite standard errors), whereas the Bayesian random-effects approach simply borrows more information and produces larger standard
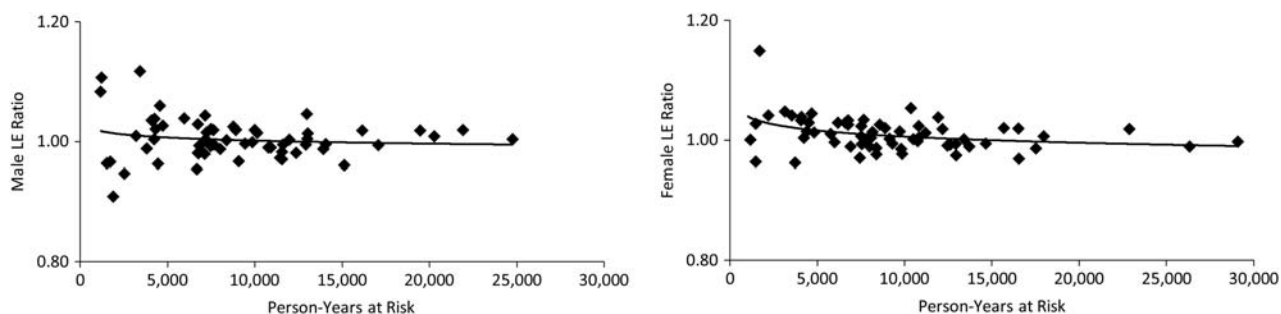
Figure 2. Ratios (traditional approach/Bayesian approach) of life expectancy (LE) among men (left) and women (right) for neighborhoods in Rotterdam, the Netherlands, 2008–2009. The solid line represents the fitted log-linear regression curve.

errors. Overall, the results of the Bayesian random-effects approach are more stable and seem more reliable for smaller populations, in concurrence with the results of the simulations.

Finally, when comparing the estimated standard errors in the simulations (Table 2) with those in the real-life example (Table 3), the standard errors of the Bayesian random-effects approach are 10%–25% smaller than in the simulations, whereas the standard errors of the traditional approach are very similar in size. This is not entirely unexpected, since a random-effects approach performs better when more and more similar areas are included in the estimations, and the range of estimated life expectancies with Bayesian model 2 is only 9.2 years (74.3–83.5) for males and 8.2 years (78.5–86.7) for females, which is significantly smaller than in the simulations.

DISCUSSION

The presented simulations indicate that the Bayesian random-effects approach can improve significantly upon the traditional life expectancy calculations, in terms of both the RMSE and the accuracy of the standard errors and confidence intervals. In particular, a Bayesian random-effects model that allows for age × area interactions performs significantly better than the traditional approach at all simulated population sizes (i.e., 500–25,000 person-years at risk).

The simulation results also provide evidence that the often recommended minimum population size of 5,000 person-years at risk for the traditional approach is rather optimistic and that substantially larger populations are advisable to obtain sufficiently accurate standard errors and confidence intervals. As shown, the coverage of the confidence intervals was only 92% for populations of 5,000 and remained below 95% for populations as large as 25,000. Based on observed deviations from normally distributed standard errors, Scherbov and Ediev (8) arrived at similar conclusions and recommended caution for population sizes up to 50,000 person-years at risk. In contrast, a minimum required sample size of 2,000 person-years at risk for the Bayesian random-effects approach seems warranted. At this point, Bayesian model 2 has a lower RMSE and considerably more accurate standard errors than the traditional approach at its recommended minimum sample size of 5,000.

The simulations are even likely to provide a somewhat conservative estimate of the relative performance of the Bayesian random-effects approach vis-à-vis the traditional approach. The first reason is that the implemented simulation approach does not introduce random variation in the age distributions of the hypothetical small areas when scaling down the benchmark populations. This is similar to previous simulation studies and is the correct method for validating the accuracy of the reported standard errors; yet it also avoids computational problems with the traditional
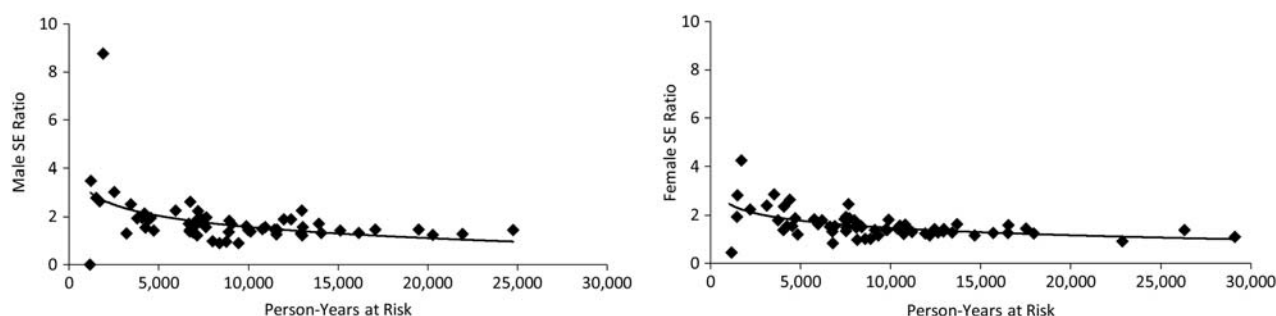

Figure 3. Ratios (traditional approach/Bayesian approach) of the estimated standard error (SE) of life expectancy among men (left) and women (right) for neighborhoods in Rotterdam, the Netherlands, 2008–2009. The solid line represents the fitted log-linear regression curve.

Table 3. Mean Estimated Standard Error for Life Expectancies (in Years) in Neighborhoods in Rotterdam, the Netherlands, by Methodological Approach, 2008–2009

Population Size	Men			Women		
	Average SE		No. of Observations	Average SE		No. of Observations
	Traditional	Bayesian 2		Traditional	Bayesian 2	
1,000–1,999	4.3	1.9	5	3.5	1.7	5
2,000–4,999	2.2	1.1	11	2.4	1.2	11
5,000–9,999	1.5	1.0	25	1.6	1.1	25
10,000–24,999	1.2	0.8	23	1.2	0.9	21
≥25,000	N/A	N/A	0	0.8	0.6	2

Abbreviations: N/A, not applicable; SE, standard error.

approach because it precludes zero populations at risk in the life tables. Whereas the Bayesian random-effects approach can adequately handle zero populations at risk in the life tables, the traditional approach breaks down and requires further aggregation or imputation of deaths to avoid computational problems. In real-life applications, where zero populations at risk frequently occur in small populations, the traditional approach is thus expected to perform somewhat worse than the simulations suggest.

The second reason for a conservative estimate of the relative performance of the Bayesian random-effects approach is that the simulations are based on benchmark countries with substantial differences in population structures and life expectancies. These differences are deliberately used to represent something of a "worst-case scenario" for the Bayesian random-effects approach, which derives its strength from similarities between the areas under analysis. Many real-life small-area life expectancy estimations will involve smaller differences between the areas and also include substantially more areas; in these cases, the Bayesian random-effects method performs better than as suggested in the simulations (whereas the performance of the traditional approach remains the same). This point is also illustrated in the real-life example, where the Bayesian standard errors are 10%–25% lower than as reported in the simulations.

Additionally, in a nonreported simulation (available from the first author upon request) where all geographic autocorrelation is removed by randomly shuffling the location of the benchmark countries in each individual regression while leaving the geographic structure of Europe intact, the results of the Bayesian random-effects approach remain unbiased and the RMSE remains 15%–25% lower than that in the traditional approach. Admittedly, such a scenario is quite unrealistic, but it does substantiate the robustness of the simulation results. Taking these considerations into account, a minimum sample size of 2,000 person-years at risk for the Bayesian random-effects approach seems prudent and is unlikely to be an artifact of the specific simulation approach.

Moreover, several improvements to the modeling approach can be envisaged that can further increase the performance of the Bayesian approach (in terms of bias and RMSE) and simultaneously have interesting public health applications. Firstly, due to computational constraints, the simulations in this paper do not incorporate more advanced spatial specifications and/or selection of random-effects models (10, 17). In real-life applications, where each model needs to be fitted only once, more flexible Bayesian models can be considered that are better suited to account for large differences in life expectancy between included areas. The estimation time for these models would increase from approximately 10–15 minutes for Bayesian model 2 to 20–30 minutes for more complicated models.

Secondly, it is also relatively straightforward to extend the Bayesian random-effects approach with a time dimension. By pooling strength over an additional dimension (in addition to sex, age group, and geographic area), the Bayesian results would further improve, and such a model could be especially useful for monitoring life expectancy in areas over time—for example, to detect atypical time trends or to evaluate the effect of interventions on the small-area level.

Thirdly, area-specific variables that are known to be correlated with geographic differences in life expectancy (e.g., the location of nursing homes and/or area deprivation scores) could be included to further improve the Bayesian life expectancy estimations. Such models are usually referred to as mixed models, and they can improve upon the results of purely random-effects models because the included predictors already explain part of the observed mortality differences between areas. Such an approach could, for example, be used to provide an elegant correction for the impact of nursing home deaths on small-area life expectancies (18).

Finally, public health researchers and health authorities may also be interested in exploring the causes of variation in estimated life expectancies. For this purpose, the Bayesian random-effects approach can easily be extended to analyze differences in estimated life expectancy in follow-up regressions that directly take the precision of the life expectancy estimations as well as the spatial configuration of areas into account.

In conclusion, the Bayesian random-effects approach is versatile and well-suited for small-area life expectancy estimations. It performs better than the traditional approach for all simulated population sizes and allows for the estimation of accurate life expectancies and accompanying confidence intervals for populations as small as 2,000 person-years at risk.

REFERENCES

1. Tsai SP, Wen CP. A review of methodological issues of the standardized mortality ratio (SMR) in occupational cohort studies. *Int J Epidemiol*. 1986;15(1):8–21.
2. Julious SA, Nicholl J, George S. Why do we continue to use standardized mortality ratios for small area comparisons? *J Public Health Med*. 2001;23(1):40–46.
3. Silcocks PB, Jenner DA, Reza R. Life expectancy as a summary of mortality in a population: statistical considerations and suitability for use by health authorities. *J Epidemiol Community Health*. 2001;55(1):38–43.
4. Tsai SP, Hardy RJ, Wen CP. The standardized mortality ratio and life expectancy. *Am J Epidemiol*. 1992;135(7):824–831.
5. Lai D, Hardy RJ, Tsai SP. Statistical analysis of the standardized mortality ratio and life expectancy. *Am J Epidemiol*. 1996;143(8):832–840.
6. Toson B, Baker A. *Life Expectancy at Birth: Methodological Options for Small Populations*. Norwich, United Kingdom: Office for National Statistics; 2003.
7. Eayres D, Williams ES. Evaluation of methodologies for small area life expectancy estimation. *J Epidemiol Community Health*. 2004;58(3):243–249.
8. Scherbov S, Ediev D. Significance of life table estimates for small populations: simulation-based study of estimation errors. *Demogr Res*. 2011;24(22):527–550.
9. Congdon P. Modelling trends and inequality in small area mortality. *J Appl Stat*. 2004;31(6):603–622.
10. Congdon P. Life expectancies for small areas: a Bayesian random effects methodology. *Int Stat Rev*. 2009;77(2):222–240.
11. University of California, Berkeley, and Max Planck Institute for Demographic Research. *The Human Mortality Database*. Berkeley, CA, and Rostock, Germany: University of California, Berkeley, and Max Planck Institute for Demographic Research; 2011. (http://www.mortality.org). (Accessed April 18, 2011).
12. Eurostat, European Commission. *Populat*. (Eurostat population database*)*. Brussels, Belgium: European Commission; 2011. (http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/). (Accessed April 18, 2011).
13. Williams ES, Dinsdale H, Eayres D, et al. *Technical Report: Calculating Life Expectancies in Small Areas*. Oxford, United Kingdom: South East Public Health Observatory; 2005.
14. Chiang CL. The life table and its construction. In: *Introduction to Stochastic Processes in Biostatistics*. New York, NY: John Wiley & Sons, Inc; 1968:189–214.
15. Gelman A, Carlin J, Stern H, et al. *Bayesian Data Analysis*. London, United Kingdom: Chapman & Hall Ltd; 1995.
16. SARA Computing and Networking Services. *Life Science Grid*. Amsterdam, the Netherlands: SARA Computing and Networking Services; 2011. (http://www.sara.nl/project/life-science-grid). (Accessed June–July 2011 and December 2011–January 2012).
17. Congdon P. A spatially adaptive conditional autoregressive prior for area health data. *Stat Methodol*. 2008;5(6):552–563.
18. Williams ES, Dinsdale H, Eayres D, et al. Impact of nursing home deaths on life expectancy calculations in small areas. *J Epidemiol Community Health*. 2004;58(11):958–962.

APPENDIX

Model 1

Let D_{six} and Pop_{six} denote deaths and populations at risk, classified by sex ($s = 1, 2$), area ($i = 1, \ldots, 33$), and age group ($x = 1, \ldots, 19$). Deaths are assumed to be Poisson-distributed:

$$D_{six} \sim \text{Poisson}(\text{Pop}_{six} \times m_{six}), \qquad (1)$$

with m_{six} denoting mortality rates specified in the same dimensions. For larger populations, a binomial distribution could also be specified; however, given our focus on small populations with few observed deaths, the Poisson distribution is considered more appropriate. A standard log-link function is used, with the first model specified as

$$\log(m_{six}) = \alpha_s + \beta_{1sx} + \beta_{2si}. \qquad (2)$$

This initial model contains:

1. Overall sex-specific mortality-level parameters α_s, which are assigned flat prior distributions from $-\infty$ to ∞ with the OpenBUGS "dflat()" distribution.
2. Parameters β_{1sx} that represent the age-sex mortality rates for age group x and sex s; these are assigned a multivariate conditional first-order random walk prior that takes correlations between adjacent age groups and the correlation between the mortality experience of males and females into account.
3. Area effects β_{2si} that represent spatially correlated mortality contrasts that are also allowed to be sex-differentiated; these are assigned a multivariate conditional autoregressive prior distribution that takes correlations between adjacent areas and the correlation between male and female mortality rates into account. Both β_1 and β_2 are estimated in OpenBUGS using the "MV.CAR" distribution with Wishart priors assigned to the precision matrices (both specified with 2 degrees of freedom and a 2-by-2 identity matrix as the inverse scale matrix).

Model 2

In model 1, the age effects β_{1sx} are assumed to operate independently of the area effects β_{2si}. This assumption results in a parsimonious model, but the actual mortality pattern may not conform to the simplifying assumption of a (sex-differentiated) uniform age gradient in all areas. Accordingly, model 2 allows for age × area interactions:

$$\log(m_{six}) = \alpha_s + \beta_{1sx} + \beta_{2si} \times \beta_{3sx}. \qquad (3)$$

The parameters α_s, β_{1sx}, and β_{2si} are given exactly the same priors as in model 1, whereas the additional β_{3sx} parameters are assigned gamma(1,1) priors. Together, the parameter combination $\beta_{2si} \times \beta_{3sx}$ provides a relatively parsimonious representation of age-sex-area mortality effects involving 104 parameters, thereby avoiding 1,254 (i.e., $2 \times 33 \times 19 = 1{,}254$) overall age-area interaction parameters β_{4six}. In further extensions, this full set of parameters could still be introduced to correct for remaining discrepancies, but this would come at the cost of model parsimony if not combined with an automatic selection mechanism such as that described, for example, by Congdon (10). Given the increase in computational time, these extensions are beyond the scope of the simulations presented in this paper. Finally, note that the β_{1sx} and β_{2si} parameters are constrained to sum to zero and that the β_{3sx} are constrained to sum to 1 for identification of the parameters. The OpenBUGS code for models 1 and 2 is available upon request.