Bayesian Approach to "Healthy Worker Hire Effect" in Standardized Mortality Ratio Analysis

Igor Burstyn, PhD, Ghassan B. Hamra, PhD, and J. Morel Symons, PhD

Objectives: In this study, we address the healthy worker hire effect that arises when people with greater than average health are recruited to work in industrial jobs. **Methods:** Epidemiologists have used both general and working population reference rates to gauge influence of healthy worker hire effect on the standardized mortality ratio. We propose a Bayesian procedure that uses information derived from general and working population reference rates to calculate standardized mortality ratio. **Results:** The procedure is illustrated in the context of heart disease and lung cancer mortality analyses of a cohort of workers from a fluoropolymer production facility. **Conclusions:** Application of our method should allow for fuller discussions of the healthy worker effect when one of its components, the healthy worker hire effect, is evaluated quantitatively. Our method can be utilized to improve risk estimates for a cohort with occupational exposure.

The concept of a healthy worker effect (HWE) has been traced to the foundations of the fields of occupational medicine and mortality registry statistics.[1] A notable early description of the phenomenon has been attributed to Ogle in 1885 who commented, "...some occupations may repel, while others attract, the unfit at the age of starting work..." (as cited in [1]). The most cited explanation of the HWE has been attributed to a series of studies authored by McMichael et al[2] that involved mortality studies of a cohort of rubber workers. Although the HWE is invoked as a potential bias that may mask the effects of an occupational hazard among workers in a particular industry, Monson[3] has commented that the potential of the HWE to bias risk estimates is limited relative to the ability to detect a true causal effect of an occupational exposure on health.

The healthy worker hire effect (HWHE) has been identified as a component of the overall HWE that involves the initial entry of healthy individuals into an occupation. Selected workers are described as "healthy" based on the assumption that their underlying disease risk at the time of hire is less than the corresponding risk of the general population.[4] Contributing factors to the lowered risk profile include the self-selection of individuals who apply for employment and the condition that hiring often depends on passing both a pre-employment medical examination and illegal drug screening test. Moreover, these individual selection factors are affected by secular trends such as the unemployment rate that may influence competition for available jobs.[5] The structure of this bias has been illustrated as a selection bias problem.[6]

Our work focuses on bias that arises from preferentially hiring persons with differential baseline health, notably the potential for developing a specific disease from a particular enterprise. We refer specifically to the HWHE to differentiate this component from other biases that fall under the broad umbrella of the HWE including the healthy worker survivor effect.[7] The HWHE causes bias in epidemiological analyses that rely on risk comparisons to a reference population that is external to the selected cohort due to use of inappropriate comparison group.[8]

Mortality rate information from general populations is typically used to generate expected numbers for outcomes in standardized mortality ratio (SMR) calculations based on the assumption that health and lifestyle characteristics are shared by both the reference population and the cohort save the potential for occupational exposure. This assumption provides a valid estimate of risk because of occupational exposure in a cohort selected if the workers who are exposed are a random sample from the general population used for the reference comparison; however, an occupational cohort is not a random sample of individuals from the general population, and therefore an SMR may be a biased estimate of risks because of occupational exposure within a cohort. This typically is noted as a deficit of deaths observed among working populations over follow-up compared with expected numbers derived from the general population rate. The classical solution to this selection bias is to conduct internal comparisons among members of the same working population who are assumed to be unexposed to the agent of interest; this assumes that the selective pressure is the same for the exposed and unexposed.[9] This approach is typically limited by the fact that a comparable occupational cohort with no exposure to the agent of interest can be difficult to define. Even if this limitation is overcome, the reduced size of available unexposed cohort populations measurably diminishes the power of the comparison relative to that achieved by an SMR calculation wherein the reference group is presumed to have no sampling variability. Furthermore, it is ideal to have a reference group with no exposure of interest whatsoever or negligible likelihood of exposure (eg, null chance of occupational exposure to beryllium in a sample of industrial workers who do not handle this metal).

One alternative is to calculate the expected number of outcomes from external industrial cohorts (working populations) known to experience the outcome of interest and hiring selection mechanisms but are free of occupational exposure to known or studied causes of the outcome.[10] If reference rates are derived from large working populations, then they can be as precise as those obtained from general population. This approach leaves investigators with two SMRs derived from different reference rates. If the two estimates largely agree, this can be taken to mean that HWHE is not at play; however, this is not necessarily the case. Because both estimates contain valuable information about disease rates, it is appealing to numerically reconcile them based on knowledge about the likely strength and form of selection into workforce without making strong assertions that the same selection mechanism applies to all members of the cohort (eg, fitter-than-average people recruited into armed forces relative to less-fit-than-average people recruited into low-paid manual labor). We set out to accomplish this through a Bayesian procedure that is easy to implement. Such a procedure quantifies the degree to which investigators believe that HWHE is at play. It then uses this knowledge to skew SMR calculation either toward general population or working population reference rates,

From the Department of Environmental and Occupational Health (Drs Burstyn, Hamra); Department of Epidemiology and Biostatistics (Dr Burstyn), School of Public Health, Drexel University, Philadelphia, Pennsylvania; and DuPont Epidemiology Program—Integrated Health Services (Dr Symons), Newark, Delaware.

The authors report no conflicts of interest.

Supplemental digital contents are available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.joem.org).

Address correspondence to: Igor Burstyn, PhD, School of Public Health, Drexel University, Nesbitt Hall, Room 614, 3215 Market Street, Philadelphia, PA 19104 (igor.burstyn@drexel.edu).

DOI: 10.1097/JOM.0000000000000556

while reflecting the degree of uncertainty about which reference rate is best. The procedure is numerically equivalent to making many plausible guesses about which reference rate to use and then averaging them to obtain an estimate of SMR that reflects both random and systematic sources of error as is done in Monte Carlo sensitivity analyses. We will examine the archetypal case when the HWHE arises from selection of healthier persons into industrial cohorts. In this context, we propose a Bayesian method for calculating the SMR and associated credible intervals. It is important to note that the method is equally applicable to standardized incidence ratio analysis. Our work encompassed accounting for uncertainty about which comparison group is ideal.

METHODS

SMR is defined as observed/expected disease events, given the person-years (PY) structure of an observed cohort. We posit that some combination of the expected outcomes based on the two reference groups is defensible as a numerical summary of the expected cases in absence of selection bias, and we define ω as parameter that governs how the two expected counts are combined. In other words, the expected rate of disease in the absence of a specific occupational exposure is informed by both the rate in the general population wherein very few persons are potentially "exposed," and by the rate for an "unexposed" population of workers who have been selected for employment into similar jobs to those of the cohort members but do not share a similar occupational exposure profile. In Bayesian terms, we can view the following conditions: (1) data as observed outcomes (O) based on the age-sex-period structure and associated PY, (2) posterior distribution as the expected number (E) of outcomes given the observed age-sex-period structure in presence of the HWHE governed by parameter ω, and (3) required prior probabilities for ω, as well as the two counts of cases obtained when reference rates from either general or working population are applied to the observed PY.

The count of outcomes follows a Poisson distribution such that the observed data has distribution $[O]\sim Pois(\lambda_o)$, the expected counts of outcomes derived from application of general population reference rates to the observed PY follow the distribution $[g]\sim Pois(\lambda_g)$, and the expected counts of outcomes derived from application working population reference rates to the observed PY follow the distribution $[w]\sim Pois(\lambda_w)$. We define $[E]\sim Pois(\lambda_e)$ where λ_e is the average number of outcomes in an unobserved population that is identical to the exposed group in all confounders and selection pressures, such that when there is no "exposure", that is, the SMR $= 1$. As the SMR is the ratio of O and E counts, we need to specify how the two expected values of counts w and g are combined to derive a combined expected value of E.

The values of ω are constrained between 0 and 1, and the larger values favor use of general population reference rates in the SMR calculations, whereas the smaller values favor the use of working population reference rates in the SMR calculations. We note that computationally, it is easier to work with expected numbers of cases, that is, estimated parameters of Poisson distribution λ_g and λ_w, rather than rates, because in this approach the weight ω is not applied to every calculation that invokes a rate of event in multiple strata (Supplemental Digital Content 1, http://links.lww.com/JOM/A216). Thus, the estimation of SMR becomes:

$$\widehat{SMR} = \frac{\widehat{\lambda}_o}{\widehat{\lambda}_e}, \quad where\ \widehat{\lambda}_e = \omega\widehat{\lambda}_g + (1-\omega)\widehat{\lambda}_w \qquad (1)$$

This is straightforward to implement because we work with $\widehat{\lambda}_g$ (estimates of expected counts of outcomes derived from application of general population reference rates), and $\widehat{\lambda}_w$ (expected counts of outcomes derived from application working population reference rates), both derived in the conventional manner for SMR calculations. These estimated counts are provided whenever SMRs are reported in the literature, making the method applicable to post-publication analysis of data without the need to obtain tables used to estimate the expected counts from the authors.

We propose to use the beta distribution for the prior on ω: $[\omega]\sim$beta(α,β). The beta distribution is a frequency distribution of a random variable that takes on values between 0 and 1, inclusive. It is described by two parameters, α and β, that take on positive values. The mean of beta distribution is $\alpha/(\alpha+\beta)$. When $\alpha=\beta=1$, a random variable has equal chance to take any value between 0 and 1, that is, a uniform distribution, with mean 1/2. When α is less than β, the distribution is skewed toward zero, and when α is more than β, distribution is skewed toward 1. The variance of the beta distribution decreases as α and β become large; for example, two random variables $x_1\sim$beta(1,1) and $x_2\sim$beta(10, 10) have the same mean (1/2) but variance of x_1 is 7 times larger than the variance of x_2. Parameters of beta distribution can be derived from any two percentiles. In our application, if ω takes on value of 0, it is equivalent to claiming that the general population reference rates are inappropriate and if ω takes on value of 1, it is equivalent to claiming the working population reference rates are inappropriate.

We now present the problem in explicit Bayesian framework: our data are O,g,w, that is, the O and E counts for a mortality outcome derived under different reference rates in the conventional manner. Therefore, the likelihood is $[O,g,w|\lambda_o,\lambda_g,\lambda_w,\omega]$, that is, the probability of observing given counts of observed and expected, and the posterior distribution is $[\lambda_o,\lambda_g,\lambda_w,\omega|O,g,w]$. The only additional priors that need to be specified are $[\lambda_o]$, $[\lambda_g]$, and $[\lambda_w]$. We can leave these as flat and use a conjugate prior of the Poisson distribution: $\Gamma(k=0.01, \theta=0.01)$. (It may be more natural to use non-conjugate prior Uniform(0, N), where N is the size of cohort: any number of counts of outcomes is equally likely, from no occurrences at all to all members of the cohort having the outcome; however, it is also sensible to try to ensure that posterior distribution is proper. In any case, the numerical results from using either prior, in examples that we examined, are indistinguishable for all practical purposes.) In specific applications, these priors can be made to be informative based on knowledge of the prevalence of the outcome of interest. It seems instructive for now to make no assumptions about the magnitude of relative health of general and working populations with respect to the given outcome and focus on the strength of selection from the general to the working population. An alternative presentation of the model is to view the likelihood as $[O,g,w|\lambda_o,\lambda_e]$, the posterior distribution as $[\lambda_o,\lambda_e|O,g,w]$, and priors that are required as $[\lambda_o]$ and $[\lambda_e]$, where the prior on λ_e is induced by expression,[1] and where $[\omega]$ informs how the prior on λ_e is influence by $[\lambda_g]$ and $[\lambda_w]$. We do not place an explicit prior on the SMR but calculate it based on the posterior distributions of λ_o and λ_e, which themselves are obtained using identical flat priors. Thus, we let the empirical data guide the calculations except for determining the choice of general population versus working population reference rates. Heuristic presentation of the calculations is detailed in Supplemental Digital Content 2, http://links.lww.com/JOM/A216. The implementation of this calculation is straightforward in WinBUGS[11] and requires an analyst to supply only the values of O, g, w, α, and β, following conventional calculation of g and w and enumerating of O from the cohort in question (Supplemental Digital Content 3, http://links.lww.com/JOM/A216).

Prior Elucidation

We propose an elucidation procedure for the prior strength of the HWHE that is transparently derived from guesses of two percentiles of distribution of ω by a team of experts and averaging the independent evaluations of each percentile. This approach is favorably viewed upon evaluation by Wu et al[12] and was previously used in occupational epidemiology.[13] The key step in such prior elucidation is phrasing a question about the potential health profile of people selected into workforce relative to the general population. There are alternate ways to elucidate priors that may be considered in a given context if our approach appears unworkable in a specific setting,[14] including visual aids for creating the beta distribution for ω.[15] The questions we propose to use are detailed in Supplemental Digital Content 4, http://links.lww.com/JOM/A216.

Motivating Examples

We illustrate the application of our method to calculate the SMR for two causes of death for which it is reasonable to assume a different degree of HWHE: heart disease and lung cancer with strong and weak selective pressures with respect to future risk, respectively. Data come from a study of mortality at a fluoropolymer production facility located in West Virginia that has been previously analyzed with published SMRs estimated from both the US general population and reference rates derived from a company-based registry.[10] The working population reference rates covered the period 1955 to 2009, and included 67,294 male and 19,404 female workers in plants in the Appalachian region. These were considered to be workers similar to the fluoropolymer production employees with respect to general health characteristics. The working population reference rates were calculated after exclusion of workers in the cohort of interest. For heart disease, the SMR relative to general population was 0.68 (95% confidence interval [CI]: 0.60, 0.77) and relative to the working population reference was 0.97 (0.86, 1.09). For lung cancer, the SMR relative to general population was 0.60 (95% CI 0.48, 0.74) and relative to the working population reference was 0.78 (0.62, 1.64). Both estimates exhibit classic evidence of the HWHE, particularly for lung cancer mortality. It is informative that the selected cohort is more similar to the working population reference than to the US general population with respect to heart disease mortality.

We perform SMR calculations under three different beliefs about the degree of selection, that is, about the strength of HWHE. This is accomplished by employing three different priors for [ω], each implying different assumptions for how the results using general and working population reference rates are pooled. We chose these priors for the motivating example to illustrate the sensitivity of the calculations to a variety of plausible priors that reflect equally certain but qualitatively different beliefs about the degree of health-based selection into workforce. First, we consider the situation where the working or general population reference rates are equivalent such that the mode of [ω] is 0.5. Next, we consider a relatively weak influence of the HWHE for which the mode of [ω] is 0.95. Lastly, we consider a relatively strong HWHE for which the mode of [ω] is 0.15. In all three cases, the second percentile was chosen such that variances among the three priors are the same: this avoids artifacts in precision that arise when the samples from posterior distributions of SMRs are compared; however, there is no such general requirement application of our method when sensitivity to different priors is examined. The beta distributions selected to represent the three of beliefs described above were beta($\alpha = 3.2618$, $\beta = 3.2618$), beta($\alpha = 3.4834$, $\beta = 1.1307$), and beta($\alpha = 1.5502$, $\beta = 4.1177$), respectively. Please refer to the earlier description of the properties of the beta distribution. Three Markov Monte-Carlo Chains with Gibbs sampler were employed, each producing 1000 samples from the posterior distribution; there was good mixing and no auto correlation in the chains even without burn-in. The posterior distributions are described in Table 1.

We note that, within the first decimal place of the SMR, when the assumption of a weak influence of the HWHE effect is invoked, the means of posterior distributions of the SMRs are the same as for the traditional analyses that use general population reference rates, for example, 0.64 versus 0.60 for lung cancer. In the lung cancer mortality analysis when a strong HWHE is assumed, the posterior distribution of the SMR is centered, on average, at a value of 0.7 with 95% of samples from the posterior falling between 0.6 and 1. The data do not allow us to determine whether there is, contrary to expectation, a strong HWHE for lung cancer that reduces the SMR, or that there is indeed a deficit of lung cancer deaths.

For the heart disease analysis, when a strong HWHE is assumed, the posterior distribution of the SMR is centered on average at a value of 0.9 with 95% of samples from the posterior falling between 0.7 and 1. This is indicative of the removal of bias because of the HWHE that is similar to that achieved when the working population reference rates are used, but without placing a strong bet on general population rates being completely unsuitable.

The calculations allow investigators to be uncertain about the presence of the HWHE, that is, selecting a mode of [ω] equal to 0.5. This results in SMR estimates that are more aligned with the expectation that the general population reference rates produce biased effects indicating a deficit of mortality outcomes in the occupational cohort. For example, the 95% credible interval of the posterior distribution of the SMR for lung cancer is 0.5 to 0.9, which is more equivocal as evidence of reduced mortality compared with evaluation of 95% CI of 0.5 to 0.7 for the SMR based on the general population reference rate. It is noteworthy that the variability of the risk estimate is increased as a result of Bayesian calculation. This is reasonable because in the traditional SMR calculation, we make a strong assumption about the certainty of which reference rates to use; however, in reality there is doubt; thus, the practice of calculating SMR under different competing reference rates better reflects the uncertainty in what the true reference rate may be. This doubt must be reflected in reduced precision of the estimate of the SMR. In other words, the traditional approach to the calculation of an SMR results in overly precise and possibly biased risk estimates, whereas allowing for uncertainty about selection bias to be reflected in the calculations should produce less biased yet less certain estimates.

TABLE 1. Bayesian Calculations of SMR Under Varying Degrees of Presumed Strength of Health-Based Selection into Workforce (HWHE) as Reflected by Assumption About Suitable Reference Rates

Cause of Death	Assumption About Strength of HWHE*	Mean (Standard Deviation) of Posterior of ω	Mean of Posterior Distribution of SMR	95% Credible Interval of SMR
Heart disease	Unsure	0.50 (0.18)	0.80	0.67, 0.96
	Weak	0.76 (0.18)	0.74	0.61, 0.90
	Strong	0.27 (0.17)	0.87	0.72, 1.06
Lung cancer	Unsure	0.50 (0.18)	0.68	0.51, 0.89
	Weak	0.76 (0.18)	0.64	0.49, 0.83
	Strong	0.27 (0.17)	0.73	0.55, 0.96

HWHE, healthy worker hire effect; SMR, standardized mortality ratios.

We are certain that in this case, wherein the empirical SMR is 0.6 for lung cancer, that it is biased away from the null, which clearly indicates the influence of the health worker hire effect. It is tempting to propose that the prior that reflects a lack of certainty be adopted as a preferred default, but it is more appropriate to develop priors that reflect beliefs about any selective pressures that can be elicited during study design and not after data are collected.

The main utility of risks estimated by the SMRs in Table 1 is that they do not force the investigator to choose a specific reference group but allow for borrowing information from two comparison groups to yield SMR estimates that reflect assumptions about the plausible degree of the HWHE as a selection factor. This can be helpful for stating qualitative influences on the potential excess of occupational risk. It is also valid for applying epidemiologic results to a risk assessment where the expression of uncertainty of risk calculations is desired.[16] The conceptual utility of a Bayesian SMR for risk assessment is that the posterior distributions of risk do not reflect only sampling variability as in frequentist calculations of CIs, but also quantify systematic errors and provide direct measure of true distribution of risk.

We note that as the number of parameters of interest exceeds the number of data points in our approach, the problem is not identifiable from the frequentist perspective; however, this does not preclude a useful Bayesian inference because we make no claim of convergence of the posterior distribution on a "true value" in probability; for proof-of-principle of this argument see the study by Gustafson.[17] We have no data on the strength of HWHE to reconcile with the sample from the prior of ω. If a guess about the strength of HWHE is poor, there is no data to alert us to this fact, unlike for other parameters such as the counts of observed and expected outcomes. Therefore, our procedure is only suitable for the estimation of the SMR given certain assumptions about selection because of hiring practices.

In conclusion, we advocate for a quantitative expression via suitably articulated priors estimating the influence of the HWHE. Our method can be utilized to produce more applicable risk estimates for a cohort with occupational exposure. This should allow for fuller discussions of the HWE when one of its components, the HWHE, is dealt with quantitatively instead by subjective speculation. A necessary condition for application of our method is the availability of reference rates for selected working populations that can provide stable reference rates for SMR calculations.

REFERENCES

1. Weed DL. Historical roots of the healthy worker effect. *J Occup Med*. 1986;28:343–347.
2. McMichael AJ. Standardized mortality ratios and the "healthy worker effect": scratching beneath the surface. *J Occup Med*. 1976;18:165–168.
3. Monson RR. Observations on the healthy worker effect. *J Occup Med*. 1986;28:425–433.
4. Baillargeon J. Characteristics of the healthy worker effect. *Occup Med*. 2001;16:359–366.
5. Shah D. Healthy worker effect phenomenon. *Indian J Occup Environ Med*. 2009;13:77–79.
6. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
7. Arrighi HM, Hertz-Picciotto I. The evolving concept of the healthy worker survivor effect. *Epidemiology*. 1994;5:189–196.
8. Wang JD, Miettinen OS. Occupational mortality studies. Principles of validity. *Scand J Work Environ Health*. 1982;8:153–158.
9. Li CY, Sung FC. A review of the healthy worker effect in occupational epidemiology. *Occup Med (Lond)*. 1999;49:225–229.
10. Steenland K, Woskie S. Cohort mortality study of workers exposed to perfluorooctanoic acid. *Am J Epidemiol*. 2012;176:909–917.
11. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS: a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*. 2000;10:325–337.
12. Wu Y, Shih WJ, Moore DF. Elicitation of a beta prior for Bayesian inference in clinical trials. *Biom J*. 2008;50:212–223.
13. Liu J, Gustafson P, Cherry N, Burstyn I. Bayesian analysis of a matched case-control study with expert prior information on both the misclassification of exposure and the exposure-disease association. *Stat Med*. 2009;28:3411–3423.
14. Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Feldman BM. Methods to elicit beliefs for Bayesian priors: a systematic review. *J Clin Epidemiol*. 2010;63:355–369.
15. Johnson SR, Tomlinson GA, Hawker GA, Granton JT, Grosbein HA, Feldman BM. A valid and reliable belief elicitation method for Bayesian priors. *J Clin Epidemiol*. 2010;63:370–383.
16. Burns CJ, Wright JM, Pierson JB, et al. Evaluating uncertainty to strengthen epidemiologic data for use in human health risk assessments. *Environ Health Perspect*. 2014;122:1160–1165.
17. Gustafson P. Bayesian inference for partially identified models. *Int J Biostat*. 2010;6:Article 17.