Multiple Imputation to Account for Missing Data in a Survey: Estimating the Prevalence of Osteoporosis

Andrew Kmetic,[1,2] *Lawrence Joseph,*[1,2] *Claudie Berger,*[2] *and Alan Tenenhouse*[3]

Background. Nonresponse bias is a concern in any epidemiologic survey in which a subset of selected individuals declines to participate.

Methods. We reviewed multiple imputation, a widely applicable and easy to implement Bayesian methodology to adjust for nonresponse bias. To illustrate the method, we used data from the Canadian Multicentre Osteoporosis Study, a large cohort study of 9423 randomly selected Canadians, designed in part to estimate the prevalence of osteoporosis. Although subjects were randomly selected, only 42% of individuals who were contacted agreed to participate fully in the study. The study design included a brief questionnaire for those invitees who declined further participation in order to collect information on the major risk factors for osteoporosis. These risk factors (which included age, sex, previous fractures, family history of osteoporosis, and current smoking status) were then used to estimate the missing osteoporosis status for nonparticipants using multiple imputation. Both ignorable and nonignorable imputation models are considered.

Results. Our results suggest that selection bias in the study is of concern, but only slightly, in very elderly (age 80+ years), both women and men.

Conclusions. Epidemiologists should consider using multiple imputation more often than is current practice. (EPIDEMIOLOGY 2002;13:437–444)

Key words: Bayesian, imputation, osteoporosis.

Virtually every large epidemiologic study has at least some degree of missing data, arising from either the refusal of some subjects to participate in the study or missing items from subjects who otherwise participate. Because participating subjects may differ in important ways from nonparticipants, parameter estimates may be biased. For example, the Canadian Multicentre Osteoporosis Study (CaMos) is a randomly selected, prospective cohort study, with a total of 9423 full participants enrolled between 1995 and 1997.[1] One of the major objectives of CaMos is to estimate the prevalence of osteoporosis in Canada. Because the study design was a random selection of Canadians, a simple estimate would be the observed proportion of participants (or subgroups thereof) with osteoporosis. However, only 42% of individuals who were contacted agreed to fully participate in CaMos, creating the potential for nonresponse bias. For example, people with direct concerns about osteoporosis may be more likely to participate, which would imply that nonrespondents would be less likely to have osteoporosis compared with respondents. Alternatively, less well individuals, especially among the elderly, may refuse to participate, as they may remain home more than healthier persons. This implies the possibility of higher rates of osteoporosis among the nonrespondents. Therefore, it is interesting to note that even the expected direction of potential nonresponse bias was not clear when CaMos was being designed.

We used CaMOS data to illustrate a Bayesian approach to account for missing data, called multiple imputation. Originally proposed by Rubin,[2] multiple imputation is based on forming "complete" datasets by simulating the

unknown missing data from each nonresponding subject, given a model that relates observed data about the subjects to the missing data items. The "complete" data sets are then analyzed using standard methods that would have been used had there been no missing data. Several "complete" data sets are simulated and analyzed concurrently, so that the uncertainty in predicting the missing data for each subject is included in the final inferences.

Although the initial motivation was Bayesian, papers by Little and Rubin[3] and by Rubin[4] have extensively evaluated the frequentist properties of multiple imputation. Most frequentist uses of multiple imputation simply create two or more complete datasets, as discussed above, and run the appropriate frequentist complete data analysis on each. The results from each data set are then combined, with overall point estimates taken as the means of the individual point estimates in each data set, and the variances of the point estimates taken to be the sum of the within- and between-imputation variances.[3,4]

The above frequentist method is easy to implement, and is often sufficient in practice. Bayesian methods of multiple imputation, however, can offer several important advantages, including the ability to incorporate prior information into the estimation procedure, and to include the uncertainty of the imputation model itself into the final estimates. Many good introductions to Bayesian inference are available.[5–9] Briefly, Bayesian analysis operates as follows: First, the parameter of interest, θ, is identified. This parameter, which is usually vector valued, includes all unknown quantities of interest, such as prevalences, odds ratios, a set of regression coefficients, or missing data. Next, a prior distribution for θ, $f(\theta)$, is specified. This distribution summarizes what is known about θ before the collection of new data. If there is little prior information, diffuse or noninformative prior distributions can be used, which in practice means that the data drive the final inferences. The distribution of the data, x, given the parameter value θ, is then specified in a likelihood function, $f(x|\theta)$. The posterior distribution, $f(\theta|x)$, is determined via Bayes Theorem, which states that $f(\theta|x) \propto f(x|\theta)f(\theta)$, that is, the posterior distribution is proportional to the prior times the likelihood. The posterior distribution summarizes the knowledge about the unknown parameter, θ, given the information contained in the data (as represented by the likelihood function) and the prior information. Note that in a Bayesian approach, all unknown quantities are given a probability distribution, which represents the uncertainty about their values. Therefore, in missing data problems, the missing data items are simply considered as additional unknown parameters to be estimated.

In practice, the posterior distribution is usually a complex multidimensional function that can be accurately approximated by modern Bayesian Monte Carlo algorithms such as the Gibbs sampler. Being an iterative algorithm, the Gibbs sampler is ideally suited to estima-tion in missing data problems, as it iterates between two basic steps. First, the missing data is imputed, by drawing from the predictive distribution for the missing data, y, given the observed data, x, and the unknown parameters. Next, given that the missing values have now been "filled in", the usual Bayesian complete data methods can be applied to derive posterior estimates of the unknown parameters of interest, such as the prevalence and the parameters of the imputation model. The next imputed dataset uses these updated estimates of the model parameters, so that final inferences incorporate the uncertainty in the parameter estimates as well.

Multiple imputation is a generic technique that can be applied to virtually any missing data situation. Aside from missing data in surveys, which we discuss in detail here, recent examples have included missing covariate data in regression,[10,11] latent data,[12] survival analysis,[13] and interval censored data.[14] Excellent introductions to multiple imputation include Little and Rubin,[3] Rubin,[4] and Schafer.[15]

Methods

Participants in the CaMos study were recruited via randomly selected, residential telephone listings for nine urban centers. Determination of peak bone mass (PBM) for the CaMos cohort is described by Tenenhouse *et al.*[16] For our purposes, a CaMos participant was considered as osteoporotic if she or he had osteoporosis at either the femoral neck or lumbar spine (L1-L4) sites, defined by a bone mineral density value at least 2.5 standard deviations below PBM.

Missing data can be classified as either ignorable or nonignorable for the estimation of a given parameter. When the pattern of missing data, given the observed data, does not supply any information about the parameters of interest, inferences can be made based solely on the observed data. This condition is called missing at random. If, in addition, the parameters that define the missing data process are independent of the parameters used to model the observed data, then the missing data are ignorable.[4] Very roughly speaking, this means that if investigators are able to collect sufficient information about nonresponders in order to model the missing data in a reasonable way, multiple imputation would provide valid inferences adjusted for nonresponse bias. Nonignorability implies a systematic difference between responders and nonresponders even after accounting for all observed data. In this case, there is no choice but to propose plausible differences between the missing data items in the nonresponders and the observed data in the respondents, and then to check the robustness of any inferences made across a reasonable range of these modeling assumptions. Given that in any real data situation

one cannot empirically verify whether missing data are ignorable or not, both possibilities are considered here.

To adjust the crude prevalence estimates for selection bias using multiple imputation, one must be able to estimate the probability that each nonrespondent has osteoporosis. To enable prediction of osteoporosis, a refusal questionnaire (RQ) was administered to those individuals declining to participate fully. The RQ collected information on major risk factors for osteoporosis, including age, gender, race, fracture history (yes/no and, if yes, before or after age 50), family history of osteoporosis (yes/no, osteoporosis of family members including the individual responding), and current cigarette smoking status (yes/no). Information on subjects declining to complete the RQ was limited to geographic location, the number of telephone calls made to establish contact, and, in some cases, age, gender, and number of household members older than 25 years of age. As a result there were two main groups of nonrespondents, based on their level of participation and available information: the nonresponders who completed the RQ (hereafter called the RQ group), and the total refusers, the TR group, composed of those individuals who refused to participate at all. The TR group was further subdivided into two subgroups. The first group, TR1, comprises those individuals who provided information on age, gender, and number of household members older than 25 years of age. The second group, TR2, comprises those individuals who provided no information at all, so that the only information available was on urban center and the number of telephone trials required to establish initial contact. Up to 12 calls were made to each household.

The monotone pattern[4] of missing data implies that multiple imputation could have easily been used throughout. Because a number of the TR group were missing data on age and gender, arguably the two most important predictors of osteoporosis, it is difficult to model osteoporosis prevalence accurately in this group. Rather than discard these data records, a single imputation was used to impute the missing demographic data, which, given the limited amount of available information, was as effective as performing multiple imputation on this variable. Gender was imputed by making draws from a Bernoulli distribution for the probability of being female conditional on the observed data, using logistic regression. Age was then imputed via linear regression, including at least study center and number of telephone trials. Therefore, there are two levels of imputations for these subjects, an initial single imputation for age and gender, followed by multiple imputations for osteoporosis. All analyses were stratified on gender.

Analysis Assuming Ignorable Nonresponse for Total Refusers

Multivariate hierarchical logistic regression analyses[8] were performed using the information available from the full responders to determine the best predictors of osteoporosis among those variables available in the various groups of nonrespondents. A random effects term was used for the center variable, in order to "borrow strength" from all centers in estimating the center-specific effects.[8] Final imputation models were selected from the list of all possible combinations of the potential predictor variables listed in Table 1 using the Bayesian Information Criterion (BIC).[17] This criterion was selected because of its optimality properties when predicting the dependent variable in future subjects.[18] These final models were then used to create the imputed "complete" data sets within each age and sex group. Different models were used for each of these groups.

Analysis Assuming NonIgnorable Nonresponse for Total Refusers

Several main risk factors for osteoporosis were collected via the RQ, so that adjustments assuming ignorability for the RQ group of nonresponders should be reasonable. For example, if persons with close relatives with osteoporosis chose to participate in higher proportions compared with those without such relatives, then the ignorable model should be able to adjust for this, thus lowering the estimate of the prevalence of osteoporosis compared with the crude estimate. However, there is more concern about the TR group, where information on major risk factors for osteoporosis is missing, so that nonignorable nonresponse cannot easily be ruled out. Essentially, the missing data of interest (osteoporosis status, in this case) may depend on items not observed. In this situation, the estimation of osteoporosis in the TR group is dominated by assumptions made about the relation between osteoporosis prevalence in the respondents and the nonrespondents. In general, these assumptions can be based on findings from other studies or the substantive knowledge of researchers familiar with the problem, and the pattern of nonresponse. This knowledge is rarely precise, and it is important to check the robustness of the conclusions across a range of plausible assumptions. We begin by using the refusal questionnaire to provide insight into the prevalence of osteoporosis in the partial (RQ group) responders. The extent of nonresponse bias attributable to this group was estimated by considering the difference between the prevalence in the full participants and the estimated (imputed) prevalence in the RQ group.

We then considered three possibilities for the distribution of the prevalence of osteoporosis in the TR groups, depending on the calculated difference in mean prevalence between the responders and the estimated

TABLE 1. Comparison of Major Risk Factors for Osteoporosis in Female and Male Canadian Multicentre Osteoporosis Study Responders, Refusal Questionnaire Nonresponders, and Total Refusal Nonresponders

Risk Factor	Females						Males					
	Responders N = 5699		Refusal Questionnaire Nonresponders N = 3859		Total Refusal Nonresponders N = 3209		Responders N = 2565		Refusal Questionnaire Nonresponders N = 2212		Total Refusal Nonresponders N = 2011	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Age (mean ± SD)	61.9 ± 12.4		67.1 ± 13.8		62.6 ± 14.2		58.9 ± 14.1		61.2 ± 16.1		58.6 ± 15.5	
Race (Caucasian)	5454	95.7	3624	93.9	n.a.		2396	93.4	2002	90.5	n.a.	
Fracture history (ever)	2354	41.3	1154	29.9	n.a.		1295	50.5	728	32.9	n.a.	
Family history of osteoporosis	849	14.9	294	7.6	n.a.		208	8.1	64	2.9	n.a.	
Current smoking	810	29.7	566	17.5			474	27.3	379	21.7		
Number of telephone calls required to establish contact (mean ± SD)	2.1 ± 1.9		1.9 ± 1.8		2.2 ± 2.3		2.2 ± 2.2		2.1 ± 2.0		2.5 ± 2.4	
Number of household members 25 years of age or older (mean ± SD)	1.7 ± 0.7		1.7 ± 0.8		1.7 ± 0.7		1.9 ± 0.6		1.9 ± 0.7		1.8 ± 0.7	
City												
Calgary	707	12.4	424	11.0	400	12.5	310	12.1	236	10.7	270	13.4
Halifax	632	11.1	493	12.8	372	11.6	281	11.0	257	11.6	206	10.3
Hamilton	627	11.0	399	10.0	236	7.3	297	11.6	217	9.8	129	6.4
Kingston	569	10.0	325	8.4	274	8.5	248	9.7	156	7.1	213	10.6
Quebec	637	11.2	309	8.0	89	2.9	277	10.8	186	8.4	40	2.0
Saskatoon	696	12.2	295	7.7	473	14.7	310	12.1	146	6.6	307	15.3
St. Johns	633	11.1	437	11.3	355	11.0	267	10.4	263	11.9	211	10.5
Toronto	499	8.8	530	13.8	687	21.4	254	9.9	394	17.8	392	19.5
Vancouver	699	12.3	647	16.8	323	10.1	321	12.5	357	16.1	243	12.1

n.a. indicates not available.

* Some missing smoking data lead to denominators that are less than total numbers shown in first row of table, as follows: female responders (2727), refusal questionnaire nonresponders (3234), male responders (1736), and refusal questionnaire nonresponders (1746).
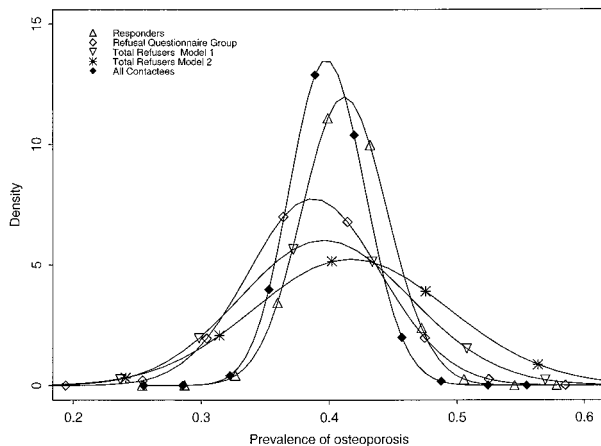
FIGURE 1. Posterior distribution of prevalence of osteoporosis for responders (crude), refusal questionnaire nonresponders, total refusers, and overall (adjusted) in CaMos women 80+ years, assuming ignorable missing data.

mean prevalence in the RQ group. Our motivation for the choices below was to create a plausible range of possible degrees of bias, based on the estimated degree of bias in the RQ group. For example, suppose the imputed prevalence in the RQ group is 2% greater than the prevalence in the full responders. The first possibility is that the prevalence of the TR group is similar to that of the RQ group. In this case, we assume that the mean of the distribution of the prevalence of osteoporosis for the TR group will also be 2% greater than the mean for the full responders. The second possibility is that the TR group differs from the full responders in the same direction as the RQ group, but with twice as large a nonresponse bias. Under this assumption, the mean of the

distribution of the prevalence of osteoporosis for the TR group will be 4% greater than the mean for the full responders. The third possibility is that the TR group is so dissimilar to the RQ group that the bias is in the opposite direction. Here, the mean of the distribution of the prevalence of osteoporosis for the TR group will be 2% *less* than the mean for the full responders. To summarize, we assume that the bias in the TR group compared with the bias in the RQ group is similar, twice as large, or similar in magnitude but opposite in direction. By calculating the overall prevalence of osteoporosis across this range of assumptions, we can check the sensitivity of the estimates to three plausible selection biases in the TR group.

All inferences were carried out via Gibbs sampling using BUGS software.[19] In our analysis, 5000 iterations of the sampling process were performed in BUGS resulting in 5000 imputed "complete" datasets. To ensure convergence of the algorithm, we examined the criteria of Raftery and Lewis[20] and Gelman's R statistic.[21] Under both ignorable and nonignorable assumptions, the imputed missing data were combined with the observed data to estimate the prevalence of osteoporosis in each sex/age grouping.

Results

As expected, the crude prevalence rates were lower in the younger age groups, at just under 1%, and increased to over 20% for men and over 40% for women in those 80 years of age and older (see Tables 2 and 3).

The concern is that the crude estimates may be affected by nonresponse bias, especially because some of the main risk factors for osteoporosis are differently distributed in responders and nonresponders (Table 1). Of all women contacted, 45% agreed to participate as full responders, 30% agreed to complete the RQ, 12% fell into the TR1 group, and 13% fell into the TR2 group. For men, the breakdown was 38%, 33%, 14% and 16%, respectively. In both men and women, the RQ group is substantially less likely to have had a previous bone fracture, to have a family history of osteoporosis, or to be current smokers than the full responder group. With fewer risk factors among nonresponders, one may expect that the crude prevalence rates may overestimate the true prevalence, but the degree of bias depends on the strength of the risk factors and their distribution across age/sex groupings. We used multiple imputation to provide adjusted prevalence estimates.

Generally, similar variables were selected for predicting osteoporosis in men and women. For RQ women, the variables selected were age, race (as a Caucasian yes/no variable, as per the original questionnaire), fracture history, family history of osteoporosis, the number of telephone calls made to establish contact, and the study
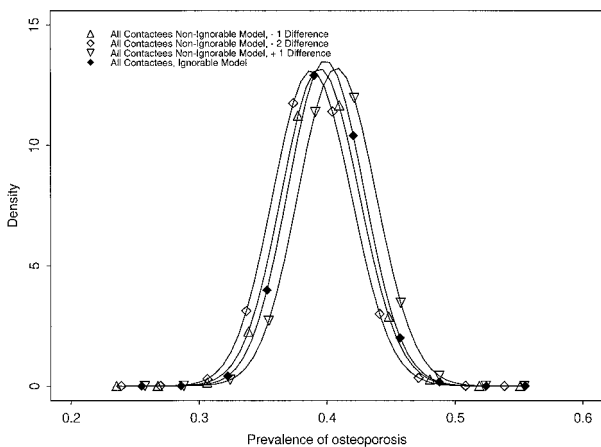

FIGURE 2. Posterior distribution of prevalence of osteoporosis adjusted assuming nonignorability of missing data. Three models shown, based on the difference in prevalence between responders and refusal questionnaire nonresponders. The overall (adjusted assuming ignorability) posterior distribution is also shown for comparative purposes.

TABLE 2. Crude and Multiple Imputation Adjusted Prevalence of Osteoporosis in Female Participants in the Canadian Multicentre Osteoporosis Study

	Age Range (Years)					
	50–59	60–69	70–79	80+	50+*	25+*
Responder group						
N	1273	1841	1356	313	4783	5699
Prevalence	0.060	0.183	0.269	0.413	0.188	0.082
95% CI	0.048–0.074	0.165–0.201	0.246–0.293	0.359–0.467	0.177–0.199	0.076–0.088
Refusal questionnaire group						
N	497	1056	1228	654	3435	3859
Prevalence	0.064	0.185	0.253	0.390	0.183	0.080
95% CI	0.034–0.100	0.148–0.224	0.213–0.296	0.309–0.474	0.162–0.205	0.070–0.090
Total refuser group, model 1						
N	226	435	320	231	1212	1518
Prevalence	0.064	0.178	0.266	0.398	0.185	0.081
95% CI	0.024–0.116	0.127–0.235	0.196–0.340	0.295–0.507	0.155–0.217	0.067–0.095
Total refuser group, model 2						
N	473	520	323	153	1469	1691
Prevalence	0.062	0.182	0.271	0.419	0.189	0.082
95% CI	0.032–0.098	0.135–0.233	0.201–0.344	0.298–0.545	0.161–0.219	0.069–0.097
All contactees (ignorable)						
N	2469	3852	3227	1351	10899	12767
Prevalence	0.062	0.182	0.262	0.399	0.185	0.080
95% CI	0.050–0.075	0.166–0.199	0.241–0.284	0.352–0.448	0.175–0.195	0.075–0.086
All contactees (nonignorable $\pm \Delta$)						
N	2469	3852	3227	1351	10899	12767
Prevalence	0.061	0.184	0.260	0.395	0.184	0.080
95% CI	0.050–0.075	0.167–0.201	0.238–0.281	0.347–0.445	0.174–0.195	0.075–0.086
All contactees (nonignorable $\pm 2\Delta$)						
N	2469	3852	3227	1351	10899	12767
Prevalence	0.062	0.184	0.256	0.389	0.183	0.080
95% CI	0.050–0.075	0.168–0.201	0.235–0.278	0.341–0.439	0.173–0.193	0.075–0.085
All contactees (nonignorable $\pm \Delta$)						
N	2469	3852	3227	1351	10899	12767
Prevalence	0.060	0.183	0.266	0.408	0.186	0.081
95% CI	0.048–0.072	0.166–0.200	0.244–0.287	0.360–0.458	0.176–0.197	0.076–0.087

* Weighted to the 1996 Canadian Census.
95% CI = 95% credible interval.
Δ = the mean difference in observed and imputed prevalence in the RQ group.

center. For the RQ group of men, all the same variables were selected, with the exception of the race variable, which was eliminated by the BIC criterion. For TR groups of women and men, the variables selected were the number of telephone calls made to establish contact, the number of household members older than 25 years of age, and the study center. The logistic regression coefficients in all of the above models varied by sex and age groups.

To illustrate the effect of adjusting for selection bias, the results for women 80 years of age and older are presented in Figures 1 and 2. The posterior distributions for the prevalence of osteoporosis in the various groups of responders and nonresponders, assuming ignorability, are shown in Figure 1. The posterior density representing the observed or crude prevalence of osteoporosis has a mean of 0.413 (95% credible interval (CI) = 0.359–0.467). The estimated prevalence in the RQ group has a mean of 0.390 (95% CI = 0.309–0.474), which is shifted approximately 2% to the left from the crude estimate. The wider credible interval indicates the

greater degree of uncertainty regarding the imputed prevalence in the RQ group compared with that in the full participant group. The mean prevalences in the TR1 and TR2 groups are 0.398 (95% CI = 0.295–0.507) and 0.419 (95% CI = 0.298–0.545), respectively. The very large credible intervals for the TR groups indicate the lack of good risk factor information. The overall adjusted prevalence distribution, obtained by combining the observed and imputed posterior densities over all four response groups, has a mean of 0.399 (95% CI = 0.352–0.448). Thus, taking into account all groups of nonresponders shows that selection bias may have a small but noticeable effect in this age group, given that the assumption of ignorability is valid. This assumption is questionable, however, given the lack of good risk factor information in the TR groups.

Figure 2 shows the adjusted prevalence of osteoporosis, assuming nonignorability, in women 80 years of age and older. Given that the crude prevalence of osteoporosis is approximately 2% greater than that imputed in the RQ group, we first assume that the bias in the TR

TABLE 3.   Crude and Multiple Imputation Adjusted Prevalence of Osteoporosis in Male Participants in the Canadian Multicentre Osteoporosis Study

	Age Range (Years)					
	50–59	60–69	70–79	80+	50+*	25+*
Responder group						
N	568	699	522	123	1912	2565
Prevalence	0.046	0.046	0.095	0.208	0.068	0.035
95% CI	0.030–0.065	0.031–0.062	0.072–0.122	0.142–0.284	0.057–0.080	0.028–0.045
Refusal questionnaire group						
N	375	516	527	262	1680	2212
Prevalence	0.036	0.040	0.084	0.160	0.056	0.031
95% CI	0.012–0.070	0.016–0.072	0.046–0.129	0.084–0.255	0.040–0.076	0.021–0.042
Total refuser group, model 1						
N	270	154	170	102	696	922
Prevalence	0.049	0.043	0.091	0.203	0.067	0.035
95% CI	0.016–0.095	0.008–0.098	0.035–0.162	0.092–0.333	0.042–0.096	0.022–0.051
Total refuser group, model 2						
N	294	306	170	86	856	1089
Prevalence	0.043	0.041	0.090	0.170	0.061	0.033
95% CI	0.013–0.084	0.013–0.079	0.034–0.161	0.062–0.308	0.039–0.087	0.021–0.047
All contactees (ignorable)						
N	1507	1675	1389	573	5144	6788
Prevalence	0.042	0.041	0.088	0.176	0.061	0.033
95% CI	0.029–0.057	0.029–0.055	0.068–0.111	0.128–0.231	0.052–0.070	0.026–0.040
All contactees (nonignorable ± Δ)						
N	1507	1675	1389	573	5144	6788
Prevalence	0.039	0.042	0.088	0.169	0.060	0.032
95% CI	0.026–0.054	0.030–0.056	0.067–0.110	0.119–0.226	0.051–0.069	0.026–0.040
All contactees (nonignorable ± 2Δ)						
N	1507	1675	1389	573	5144	6788
Prevalence	0.036	0.040	0.085	0.154	0.056	0.031
95% CI	0.023–0.051	0.028–0.055	0.065–0.107	0.104–0.210	0.047–0.066	0.025–0.038
All contactees (nonignorable ± Δ)						
N	1507	1675	1389	573	5144	6788
Prevalence	0.046	0.045	0.093	0.201	0.067	0.035
95% CI	0.032–0.063	0.032–0.060	0.072–0.117	0.148–0.260	0.057–0.077	0.029–0.043

* Weighted to the 1996 Canadian Census.
95% CI = 95% credible interval.
Δ = the mean difference in observed and imputed prevalence in the RQ group.

group is either 2% or 4% in the same direction. This results in estimates slightly less than the crude prevalence, with adjusted estimates of 0.395 (95% CI = 0.347–0.445) and 0.389 (95% CI = 0.341–0.439), respectively. The third option, assuming that the bias in the TR group is similar in magnitude but opposite in direction compared with the RQ group, yields an adjusted prevalence of 0.408 (95% CI = 0.360–0.458), very similar to the crude estimate. For comparative purposes, the adjusted prevalence distribution, assuming ignorability, is also shown in Figure 2. This figure suggests that, although some nonresponse bias is present regardless of the ignorability assumption, the overall effect remains small.

Tables 2 and 3 present the crude and adjusted prevalence rates of osteoporosis, assuming ignorability and the various nonignorability models, for men and women of all age groups across the different groups of participants and nonparticipants. The degree of selection bias ranges from negligible in the younger age groups to a small but noticeable effect in the older age groups for both men and women. For both men and women, nonresponse bias is very slight when collapsing the age groups into those 25 years and older and those 50 years and older, each weighted by the 1996 Canadian census data.

Discussion

Application of multiple imputation to the CaMos data suggests little nonresponse bias, except perhaps in the very elderly. It seems that contacts decided to participate in CaMos for reasons largely unrelated to their osteoporosis status, or at least that differences in known major risk factors were not large enough to seriously affect the estimates of the prevalence of osteoporosis.

Nevertheless, our estimates depend largely on having collected all of the important risk factor differences for osteoporosis in the RQ. It is possible that there are other osteoporosis-related differences not captured by the questionnaire. Furthermore, our nonignorability assumptions also depend on the RQ, because the magnitude of

nonresponse bias assumed in the TR groups depends on the adjustment in the RQ group.

Multiple imputation has at least four main advantages over other missing data techniques. First, once the missing values have been imputed, standard complete data analysis techniques can be used rather than more complex missing data techniques. Therefore, there is no need to derive a new methodology for each type of analysis. This is an important factor in CaMos, given multiple investigators in this multicenter study with interest in analyzing the data. Second, "complete" datasets can be distributed to other users. If two or more data sets with different imputed values are distributed, then the simple addition of between- and within-dataset variations provide nonexpert users with a means to adjust for nonresponse in a valid way.[4] Third, multiple levels of missing data, as exist here, can easily be accommodated. Finally, the methodology is relatively easy to implement, as all that is required is a model for the missing data items, a random number generator to impute the missing items according to the model, and a calculation of the within- and between-imputation variances. Further, freely available software such as BUGS can be used, which largely automates the procedure.[22] Other commercial packages, such as SAS[22] and SPSS,[23] have also recently added versions of multiple imputation algorithms to their software.

References

1. Kreiger N, Tenenhouse A, Joseph L, *et al*. The Canadian multicentre osteoporosis study (CaMos): background, rationale, methods. *Can J Aging* 1999;18:376–387.

2. Rubin D. Inference and missing data. *Biometrika* 1976;63:581–590.

3. Little RJ, Rubin D. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons, 1987.

4. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons, 1987.

5. Goodman SN. Toward evidence-based medical statistics. 2: the Bayes factor. *Ann Intern Med* 1999;130:1005–1013.

6. Brophy JM, Joseph J. Placing trials in context using Bayesian analysis. *JAMA* 1995;273:871–875.

7. Congdon P. *Bayesian Statistical Modelling*. New York: John Wiley & Sons, 2001.

8. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. London: Chapman & Hall, 1995.

9. Berry D, Stangl D, eds. *Bayesian Biostatistics*. New York: Marcel Dekker, 1996.

10. Wu H, Wu L. A multiple imputation method for missing covariates in non-linear mixed-effects models with application to HIV dynamics. *Stat Med* 2001;20:1755–1769.

11. Sheppard L, Levy D, Norris G, Larson TV, Koenig JQ. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987–1994. *Epidemiology* 1999:10:23–30.

12. Barnard J, Meng XL. Applications of multiple imputation in medical studies: from AIDS to NHANES. *Stat Methods Med Res* 1999;8:17–36.

13. Bebchuk JD, Betensky RA. Multiple imputation for simple estimation of the hazard function based on interval censored data. *Stat Med* 2000;19:405–419.

14. Carabin H, Gyorkos TW, Joseph L, Payment P, Soto JC. Comparison of methods to analyse imprecise faecal coliform count data from environmental samples. *Epidemiol Infect* 2001;126:81–190.

15. Schafer JL. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall, 1997.

16. Tenenhouse A, Joseph L, Kreiger N, *et al*. Estimation of the prevalence of low bone density in Canadian women and men using a population-specific DXA reference standard: the Canadian multicentre osteoporosis study (CaMos). *Osteoporos Int* 2000;11: 897–904.

17. Raftery A. Bayesian model selection in social research (with discussion by Gelman A, Rubin DB, Hauser RM). In: Marsden PV, ed. *Sociological Methodology*. Oxford, U.K.: Blackwells, 1995;111–196.

18. Kass R, Raftery A. Bayes factors. *J Am Stat Assoc* 1995;90:773–795.

19. Spiegelhalter D, Thomas A, Best N, Gilks W. BUGS Bayesian inference using Gibbs sampling. BUGS Version 0.50. MRC Biostatistics Unit, Cambridge, U.K. (http://www.Mrc-Bsu.Cam.Ac.UK/Bugs/Mainpage.html); 1995.

20. Raftery A, Lewis S. How many iterations in the Gibbs sampler? In: Bernardo J, *et al.*, eds. *Bayesian Statistics 4*. Oxford, U.K.: University Press, 1992;763–773.

21. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences (with discussion). *Stat Sci* 1992;7:457–511.

22. SAS. *SAS/STAT software: changes and enhancements release 8.2*. Cary, NC: SAS Institute Inc., (www.sas.com/newversion), 2001.

23. SPSS. *Base 10 User's Guide*. Chicago: SPSS Inc., 2001.