Comments from Sandra Greenland: Here you go:

The usual goal of adjustment was never variance reduction, but rather to get an unbiased estimate or prediction of what the crude rate in some population (which supplies the "standard" distribution) would be (or would have been) if the covariate-specific rates were different than what they actually were. The goal is to allow comparisons unconfounded by the adjustment covariates. The final sections of Ch. 3 and Ch. 4 of Rothman and Greenland (ME2) talks about this (they should read those) and there is an enormous literature on the topic going back a century at least.

I think that there are several issues here that need to be clearly articulated so that the problem addressed is clear. In the BRFSS, there is a sample of a population. The statistical context is sampling, and there is a target parameter (the crude rate in population). (Note that for example, if smoking rates vary by gender, the crude rate is an weighted average of the gender specific rates. Some may argue that the crude rate isn’t interesting here, since the gender specific rates differ. This is another issue.). I would disagree with Sandra that the goal is ‘to allow comparisons unconfounded by the adjustment covariates’. I would say the goal is to develop the most accurate estimate of the crude rate.

Yates (JRSS 1934) is a classic and had it right (he even pointed out how two SMRs were not comparably standardized). There was a very nice article on standardization and its relation regression and prediction by Lane and Nelder in Biometrics 1982, which they should download from JSTOR and read.

I didn’t go back and look at Yates paper, but the focus of comparing populations is different from the focus of estimating a parameter in a population. This is the problem that we are addressing here.

The paper by Lane and Nelder seems like a similar context to the one we consider- where we want to estimate what would happen if we knew the entire population. In section 2, they describe a problem in terms of a randomized block experiment. This is a more complicated problem, and is not the same as the simple estimation problem since there is randomization. It is possible, but it would take some work, to see if the simplification of the problem they address really matches the setting we are considering. I don’t think this paper is a good place to start.

Another less common goal of standardization, seen in demography, sociology, and policy, is to see what the crude rate would be if the distribution of the covariates were different (they are addressing situations in which a planned intervention changes the covariates rather than the covariate-specific rates). Presumably that isn't the goal in what they were examining.

I agree, this isn’t the goal we are addressing.
The bias issue comes down to parameters: inequality of the crude and adjusted parameters = bias in the crude estimator when the adjusted parameter is the target, or estimand.

I believe Sandra’s comments here are helpful. First, he says ‘inequality of the crude and adjusted parameters’. One question is ‘what is the adjusted parameter?’ In our context, the adjusted parameter is equal to the crude parameter. Here, bias is defined in terms of non-stochastic constants. I believe part of the confusion is that there is no accounting for sampling. I believe that we may need to illustrate the difference in ideas.

To address precision, one could set out the goal of adjustment clearly, as the above citations do, and then discuss asymptotically unbiased estimators under a given model.

I do believe the ‘goal’ matters here, as Sandra indicates. The goal in the references cited does not match the goal we have, so I’m not sure they are very relevant.


Model based approaches offer a very different starting point, since the parameters are defined by the models, not by the population. In our example, we know that some model based approaches result in the same predictor. This doesn’t change the problem, but it does agree with the comment above.

If one discards unbiasedness in favor of mean-squared error (MSE) reduction, one could then ask whether the crude or adjusted estimator has smaller MSE for the target parameter (which is always an adjusted parameter), and that depends on bias and variance (and hence sample size). One then finds both the crude and the usual adjusted estimators are suboptimal, and that some compromise may minimize MSE. An example is Greenland, S. (1991). Reducing mean squared error in the analysis of stratified epidemiologic studies./ Biometrics/, *47*, 773-775 which can also be gotten off JSTOR.

The presumption in the previous paragraph is that there is bias. Since there is no bias in the setting we consider, I don’t think this comment if very relevant.

This is closely related to empirical-Bayes estimation, which has been used to improve standardized rates and ratios, for example see Greenland, S. and Robins, J. M. (1991). Empirical-Bayes adjustments for multiple comparisons are sometimes useful./ Epidemiology/, *2*, 244-251.
which reproduces an example from Efron and Morris (JASA 1975).

This is a nice note, since it relates to empirical bayes estimation. There is a similarity between the resulting predictors in the two approaches (empirical bayes and finite population prediction approaches), but the starting points and assumptions are different. This can be brought out in the discussion.

Hope this helps.
Best Wishes,
Sander
Wenjun, I like this draft much better. I've made bunch of comments, and would be happy to discuss them more. Ed

Covariate Adjustment May Not Be Better: Thresholds of Relative Risk when Controlling for Confounders

Wenjun Li
Biostatistics Research Group, Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, Worcester, MA

Edward J. Stanek III
Department of Public Health, University of Massachusetts, Amherst, MA

Co-authors (??)

Jenny Kelsey??
Liz Bertone ??

Table of Contents

Abstract .................................................................................................................................... 5
Introduction .......................................................................................................................... 6
Background .......................................................................................................................... 7
  Crude rates ....................................................................................................................... 9
  Covariate-adjusted rates ............................................................................................... 9
Simulation methods ........................................................................................................ 13
Results ............................................................................................................................... 14
Discussion ......................................................................................................................... 16
Reference ........................................................................................................................... 20

Deleted:
2
3
4
6
6
10
10
9
11
13
12
17
15
J:\Projects\Cluster\report\liw_06_adjusted_rates-ejs.doc

Deleted:
C:\Documents and Settings\Edward\My Documents\cluster\report\liw_06_adjusted_rates.doc

Deleted: 3/13/2006

Insert:
J:\Projects\Cluster\report\wenjun\liw_06_adjusted_rates-ejs.doc

3/21/2006
Abstract

In theory, covariate-adjusted rates have smaller variance when regression coefficients are known. In practice, regression coefficients needed to achieve the adjustment are unknown and their sample estimators are used instead. We develop guidelines to determine when adjusted rates have smaller variance than unadjusted crude rates, illustrating the results with an example of adjusting for gender when estimating cigarette smoking prevalence. We show that in many practical settings, adjustment for gender when estimating smoking prevalence rate results in less accuracy. The variance reductions due to covariate adjustment depend on sample sizes, gender ratios and male-to-female prevalence rate ratios (RRs). In populations with balanced gender ratios and prevalence rates of 35%, the adjusted smoking prevalence estimates were more accurate when RRs were above 1.6, 1.5, 1.4, 1.3 and 1.2 for sample sizes of 25, 50, 100, 150 and 200, respectively. In populations with unbalanced gender ratios or lower prevalence rates, the RR thresholds were higher. In sum, adjustment for covariates will not result in more accurate estimates of smoking prevalence in all settings, and in particular, not when both RRs and sample sizes are small and prevalence rates are low. When reporting smoking rates at the municipal level based on Massachusetts BRFSS data, gender adjustment is not recommended when sample size is less than 200.

Key Words: Epidemiologic method, survey sampling, finite population, rate adjustment, variance components

Deleted: with known variance components.
Deleted: variance components
Deleted: empirical
Deleted: in
Deleted: should not be conducted.
Deleted: rates had smaller variances
Deleted: ed rates should not be used
Introduction

Epidemiologic studies often involve random sampling of subjects from a population defined in space and time. For example, the Behavioral Risk Factors Surveillance Systems (BRFSS) of the Centers for Disease Control and Prevention (CDC) conducts annual telephone surveys on adults living in households in the 50 states of the United States. BRFSS surveys cover many behavioral risk factors, such as cigarette smoking, sexual behaviors, and drunk driving. When the sample data are analyzed, adjustments are usually made to account for possible imbalance of covariates (such as gender and age) in the study samples, which is also commonly referred as “control for confounding”. In theory, estimators can be made more accurate (I was thinking that accurate may be the better term. The reason is that if you don’t do adjustment, people may think of the estimator as being biased. I don’t think it is biased, but if the thinking is like that, accuracy would reflect it) precise by adjusting for known auxiliary information such as the population gender proportion and average age of the study population (Särndal et al. 1992). Adjusted rate estimators have smaller variance when relevant variance components are known. However, in practice, variance components are unknown and replaced by their sample estimator. The uncertainty introduced by using variance component estimates has implications for whether controlling for a confounder effectively reduces the size of the confidence interval for prevalence.

In many epidemiologic or public health studies relying on random samples, relevant variance components are unknown. To our knowledge, there is no established guideline with regard to when the covariate-adjustment should or should not be made in popular epidemiology textbooks, such as Rothman and Greenland (1998) and Fleiss, Levin and Paik (2003). We illustrate the problem using a simple example of adjustment...
for gender in estimating smoking prevalence, and suggest an empirical guideline on when such adjustment should not be used in epidemiologic reports.

Background

We use a simple hypothetical example to introduce the problem and define notation. Suppose a survey of cigarette smoking prevalence rates among adult residents living in a small town in Massachusetts is conducted by random digit dialing following the procedures similar to the BRFSS (write out first time) (CDC 2004). The town list of residents, updated annually by the town government as mandated by law, provides telephone numbers of adult residents. The total number of adults ($N$) and the proportion of male adult residents $\pi_x$ are known. (I'd suggest changing this section.

Maybe say, “In practice, RDD has many problems (i.e., answering machines, cell phones, caller id etc) that result in a sample that is not a simple random sample (include references). While these problems are practically important, for simplicity we assume that the sampling process results in a simple random sample (SRS) with all adults having an equal probability of being interviewed. In total, a sample of $n$ subjects is selected. We represent the smoking status for the $i$-th sample subject with an indicator random variable $Y_i$ (one if the subject is a smoker, and zero otherwise), and the auxiliary variable with an indicator random variable, $X_i$, of male gender (one if the subject is male, and zero otherwise). We assume that smoking status and gender are recorded for all the subjects in the sample. The goal of the survey is to estimate the prevalence rate of cigarette smoking $\pi_y$. We summarize the data from the sample and information on the number of male and female residents in Table 1. We estimate smoking rates using the crude and the gender-adjusted rates, both of which are unbiased (Cochran 1977; Sämdal et al. 1992).
Table 1 Number of subjects in a simple random sample / a hypothetic population by gender and smoking status

<table>
<thead>
<tr>
<th>Gender</th>
<th>Smoking status</th>
<th>Sample</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Smoker</td>
<td>Nonsmoker</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>$n_{11}$</td>
<td>$n_{10}$</td>
<td>$n_*$</td>
</tr>
<tr>
<td>Female</td>
<td>$n_{21}$</td>
<td>$n_{20}$</td>
<td>$n_*$</td>
</tr>
<tr>
<td>Total</td>
<td>$n_*$</td>
<td>$n_*$</td>
<td>$n_*$</td>
</tr>
</tbody>
</table>

Notation:

$n_*$ = number of smokers in the sample
$n_{11}$ = number of male smokers in the sample
$n_{21}$ = number of female smokers in the sample
$n_*$ = number of nonsmokers in the sample
$n_{10}$ = number of male nonsmokers in the sample
$n_{20}$ = number of female nonsmokers in the sample
$n = n_1 + n_0 = n_* + n_*$ = sample size

$N_*$ = number of males in the population
$N_2*$ = number of females in the population
$N$ = population size

$\pi_1 = N_*/N$ proportion of males in the population
Crude rates

The crude rates are straightforward estimators of the prevalence rate, i.e., sample mean of the indicator variable for smoking status $Y$. The crude rate and its variance are given by $\hat{\pi}_y = Y = n_y / n$ and $\var(\hat{\pi}_y) = (1 - f)\sigma^2_y / n$, where $f = n / N$, and $\sigma^2_y = \pi_y (1 - \pi_y)$ is variance of $Y$.

Covariate-adjusted rates

The covariate-adjusted rate estimator corresponds to a simple weighting of the gender specific rates according to gender proportions of the population. With the data in Table 1, the gender-adjusted rate estimator is calculated as (you use the same notation for two different things. Introduce different notation)

$$\hat{\pi}_y = \frac{N_{1y}}{N} \left( \frac{n_{1y}}{n_{1*}} \right) + \frac{N_{2y}}{N} \left( \frac{n_{2y}}{n_{2*}} \right).$$

(1)

This estimator, routinely used in epidemiological and public health reports, is the direct adjusted estimator formed by multiplying population weights by sample rates. In the finite sampling literature, it is referred as the poststratified estimator which can be developed from either a calibration (Deville and Särndal 1992) or a model-assisted approach (Särndal et al. 1992).

Although the adjustment is intuitive and seemingly straightforward, the underlying variability resulting from sampling both $y$ and $x$ results in variability that may overshadow the gains made by the adjustment. To examine this issue, it is valuable to consider more general statistical frameworks that give rise to this solution.

Various approaches have been proposed in finite sampling literature to obtain covariate-adjusted estimates of rates based on simple random samples. For example,
methods of improving estimation with auxiliary information have been discussed in a model-based approach (Bolfarine and Zacks 1992; Valliant et al. 2000), model-assisted (Cassel et al. 1977; Cochran 1977; Särndal et al. 1992) or calibration methods (Särndal and Wright 1984; Deville and Särndal 1992), and design-based prediction approach (use your thesis as a ref here) (Li and Stanek III 2006, under review). For the simple scenarios concerned in this paper, the above approaches result in very similar estimators, which vary slightly according to what variance components are used in calculation (how do they vary? This suggests that you’ll indicate this. If they do vary, why is this not important. If they don’t vary, could this be a more minor remark that the same results can be obtained using other frameworks.). For example, with design-based prediction approach by Stanek and Singer (2005) and Li and Stanek (2006, under review), the gender-adjusted estimator is derived as (this notation is the same as for the crude estimate. You should use different notation. by doing so, you can present this as a crude estimate plus an adjustment. I don’t think the first form of the equation is relevant to your discussion.)

\[ \hat{\pi}_y = \frac{1}{N} \left( \sum_{i=1}^{n} Y_i + \sum_{i=1}^{N} \left( \bar{Y} - \frac{1}{1-f} \beta (\bar{X} - \pi_x) \right) \right), \]

\[ = \bar{Y} - \beta (\bar{X} - \pi_x) \]

and its variance

\[ \text{var}(\hat{\pi}_y) = \left(1 - \rho^2 \right) \left(1 - f \right) \sigma^2 \frac{1}{n}, \]

where \( f = n/N \), \( \beta = \sigma_x \left/ \sigma^2 \right. \), \( \rho = \sigma_{xy} \left/ \sigma_x \sigma_y \right. \) is the correlation coefficient of smoking status (\( Y \)) on gender (\( X \)), and \( \bar{Y} \) is the sample proportion of cigarette smokers in the sample and \( \bar{X} \) is the proportion of male subjects in the sample, respectively. (would it be worthwhile to discuss that var(x) is known, since we know the gender of everyone in
the population. This means that to get an estimate, we only need to estimate the covariance. You could then say that even though it makes sense to use the known variance, if you would use the estimated variance, you’ll get the post-stratified estimator (and show how it equals 2).

In practice, these quantities can be estimated directly from the sample. For example, with the data in Table 1, \( \bar{Y} = n_y/n \), \( \bar{X} = n_x/n \), and \( \sigma_x^2 = N_x N_y / N(N - 1) \). For \( \sigma_y^2 \) and \( \sigma_{xy} \), we apply “method of moments” and replace \( \sigma_y^2 \) and \( \sigma_{xy} \) with their sample estimates, \( \hat{\sigma}_y^2 = \frac{n_y n_{2y}}{n(n - 1)} \) and \( \hat{\sigma}_{xy} = \frac{n_x n_{2x}}{n(n - 1)} \left( \frac{n_{11} - n_{21}}{n_x n_{2x}} \right) \), and obtain

\[
\hat{\beta}_{xy} = \left( \frac{N(N - 1)}{n(n - 1)} \right) \left( \frac{n_y n_{2y}}{N_x N_y} \right) \left( \frac{n_{11} - n_{21}}{n_x n_{2x}} \right).
\]

(4)

Alternatively, if \( \hat{\sigma}_y^2 \), \( \hat{\sigma}_{xy} \) and \( \hat{\sigma}_x^2 \) are all replaced with their sample estimates \( \hat{\sigma}_y^2 \), \( \hat{\sigma}_{xy} \) and \( \hat{\sigma}_x^2 = \frac{n_x n_{2x}}{n(n - 1)} \), then

\[
\hat{\beta}_{xy} = \frac{n_{11} - n_{21}}{n_x n_{2x}},
\]

(5)

and consequently, we obtain the post-stratified estimator (see equation (1)).

(It seems to me like this next paragraph is more a discussion point, and may not be critical to the results. I’d be inclined to leave it out- I don’t quite understand why interpreting this as a shrinkage estimator makes sense- are we shrinking gender specific estimates to the overall population mean estimate? The way the idea is presented, it seems different)

The gender-adjusted rate estimator in equation (2) has interesting interpretations. The first expression is divided into the sum of random variables in the observed sample and a predictor of the random variables in the unobserved subjects of the population.
Since the weighted total of the sample and the unobserved random variables equals the population mean, it is clear that for a realized sample, we simply substitute the values observed for the sample random variables in the estimator. The adjusted rates, in either the first or the second expressions, include the sample prevalence rate, \( \bar{Y} \), and an adjustment term for gender. The adjustment ‘shrinks’ the observed difference in male proportions between the sample and the population. For example, if the value observed for the male proportion in the sample, \( \bar{X} \), exceeds the population male proportion, \( \pi_x = \frac{N_x}{N} \), and the cigarette smoking status and gender random variables are positively correlated, we may anticipate that the sample smoking prevalence, \( \bar{Y} \), will also exceed the population smoking prevalence \( \pi_y \). The adjusted rate compensates for the over estimation of \( \pi_y \) by the sample random variables, and the adjustment is proportional to the difference in male proportions, \( \bar{X} - \pi_x \), and weighted by the regression coefficient, \( \beta \). As shown in formula (3), the adjustment results in variance reduction over the simple sample proportion, with the percent reduction given by \( (1 - \rho^2) \times 100\% \).

The adjusted estimator (2) is a function of the regression coefficient, which, in turn, is a function of variance components in the population. In practice, although \( \sigma_x^2 \) may be known, \( \sigma_{xy} \) will need to be estimated. Särndal et al. (1992, p229) recommends estimating both terms, a recommendation that we evaluate via simulation. The estimate of \( \beta \) is given by \( h = \hat{\sigma}_{xy} / \hat{\sigma}_y \) (see formula (5)) (I’d use a hat notation for b1 to indicate an estimate), and the resulting adjusted rate estimator is the well-known post-stratified estimator (1) (you already said this). Li and Stanek (2006, under review) showed that
use of known population variance of $X$ to estimate $\beta$ by

$$b_2 = \hat{\sigma}_y / \sigma_x$$

(formula (4))

(once again, use a hat notation) generally (what is meant by ‘generally’? can you be more specific) result in smaller variance of the adjusted rates, especially when sample sizes are relatively small. In this analysis (which analysis? Does this mean that you won’t use the predictor with $b_1$? gender-adjusted rates were made by using $b_2$. I think you need to be clear about what the point is of the paper. The main point is that you shouldn’t make adjustments when $rr$ is small. A second point may be that you should use $b_2$ instead of $b_1$. Am I right? The way it is now, I’m not sure what the simulation is designed to show.

For the adjusted estimators, we examine how the uncertainty of estimating variance components influence the precision of rate estimates (using $b_2$?). To do so, we conducted a series of Monte Carlo simulations to compare the variances of the crude and the adjusted rates ($b_2$).

**Simulation methods**

The simulation study generated a series of hypothetical populations of sizes 200, 400, 800, 1,600 and 3,200; each with male proportions ranging from 30% to 70% ($\pi_x = 0.30, 0.35, 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, 0.70$), and with prevalence rates of cigarette smokers ranging from 15% to 35% ($\pi_y = 0.15, 0.20, 0.25, 0.30$ and 0.35). The prevalence rate ratios (RR) of men to women ranged from 1.0 to 4.0 (RR = 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 2.0, 2.2, 2.4, 3.0, 3.5, 4.0). According to data from 2004 BRFSS data of Massachusetts (http://apps.nccd.cdc.gov/brfss/), the prevalence of current cigarette smokers is 18.4%; male-to-female RR is 1.13, RR of adults with low annual per ca income (< $15,000) to those with high annual per ca income (> $50,000) is 2.1; the RR of individuals with less than high school education to college graduates is
2.6; and the RR of those aged 18 to 25 to those aged 65 or older is 3.1. The ranges of gender ratios and smoking prevalence rates are similar to those in 1999-2003 BRFSS data of Massachusetts.  (good)  (Can you comment on the sample size in Mass for the BRFSS?  Also, what are sampling fractions?  Do the sampling fractions matter?)

We evaluated the rate estimators by comparing the average variance over 10,000 independent simple random samples for each scenario. The sample sizes ranged from 25 to 200 (n = 25, 50, 100, 150, 200, n < N), corresponding to a sampling fraction from 1.5% - 75%. In total, 15,120 scenarios were evaluated.

To compare the variance of the rate estimators, we recorded whether the variance of gender-adjusted rates was smaller than the variance of the crude rates. Ratios of variances of the adjusted rates (AR) to variances of the crude rates (CR) were calculated. The variance ratios were then plotted against percentage of males in population, by population smoking prevalence rate, population size and sample size. Thresholds of RRs, above which adjusted rates have smaller expected variances, were estimated graphically.

**Results**

We present results graphically in terms of a set of threshold relative risk of smoking by gender (Male/Female). The estimated thresholds for prevalence rate ratios are presented in Figure 1. Results for scenarios with population size of 3200 are not included due to their similarities to those with population size of 1600.

As shown in Figure 1, the variance reductions due to gender-adjustment depend on sample sizes, gender ratios, prevalence rate ratios, and overall prevalence rate of the population. Using the empirical estimators based on sample estimates of covariance between smoking status and male gender (\(\hat{\sigma}_{xy}\)), a reduction in the variance will occur...
only when the male-to-female prevalence rate ratios are sufficiently large, or the association between smoking status and the covariate are sufficiently strong.

In populations with balanced gender ratios and overall prevalence rates of 35%, the adjusted rates have smaller variances when RRs are above 1.6, 1.5, 1.4, 1.3 and 1.2 for sample sizes of 25, 50, 100, 150 and 200, respectively. In populations with unbalanced gender ratios, the RR thresholds are higher.

When the population prevalence rates are lower, for example, at 15% in gender balanced populations, the RR thresholds are much higher, i.e., 2.6, 2.1, 1.8, 1.6 and 1.5 for sample sizes of 25, 50, 100, 150 and 200, respectively.

The RR thresholds are much higher in populations with high male proportions, in particular, when sample sizes are under 150. For example, at 70% males in a population of size 1,600 with prevalence rate of 15%, the RR thresholds were above 4, 3, 2.3, 1.6, and 1.5 for sample sizes of 25, 50, 100, 150 and 200, respectively.

Sampling fractions influence the variance reduction as well, with higher sampling fraction related to slightly higher RR thresholds. For example, with a sample of size 150 from a population with balanced gender ratio and prevalence of 15%, the RR thresholds are 1.7, 1.55, 1.45 and 1.4 for populations of sizes 200, 400, 800 and 1,600 (corresponding to sampling fractions of 75%, 38%, 19%, and 9%). According to these data, in a study of a small town of 1600 adults, with a sample of 200, if the RR of smoking with gender matches the Massachusetts population RR (1.13), gender-adjustment should not be used.
Discussion

In this analysis, we illustrate that adjustment for confounding, while designed to make estimators more precise, may actually make the estimators worse. Although this study was limited to smoking and gender, the results have broad implications for other settings in epidemiological applications. The basic idea is that adjustments for confounding involve estimating regression coefficients from the sample. If the relationship between the confounder and the response variable is not strong, the added variability due to estimating the coefficients can overshadow the gain. (I think you could expand on the ‘bias’ interpretation. For example, if we fix the number of males and females as the number realized in a sample, then repeated sampling (with the same numbers) will result in unadjusted rates that are biased. This may be a person’s intuition, since they may view these sample sizes as fixed. I don’t know if this is common and some discussion with Jenny and Liz would help here. In any case, I think this is a separate point), usually viewed as resulting in a reduction in bias. The implication is that wider confidence intervals will result for the prevalence estimates if control for confounding is used. In fact, when the RR is lower than the thresholds indicated in Figure 1, gender-adjusted rates are less precise than un-adjusted rates.

We simulated a large number of scenarios that are similar to common situations in epidemiologic sample surveys, where sample sizes are small and prevalence rates are relatively low. In certain situations, a much stronger relationship is needed for the adjustment for confounding to be warranted. For example, with a sample size of 100 and prevalence rate of 20%, the variance reduction will occur only if the prevalence rate ratio is greater than 1.7 in a population with balanced gender proportions, or over 2.5 in a population with 70% males and 30% females.
You may want this paragraph earlier. Covariate-adjusted rates, in theory, do lead to rate estimates that are more precise than crude rates when the regression coefficient $\beta$ or the relevant variance components ($\sigma_x$, and $\sigma_y^2$) are known (Särndal et al. 1992).

The remainder of this paragraph seems redundant... In practice, these parameters are rarely known, leading in some cases to more variable results (sentence seems redundant with earlier paragraph). Small sample sizes ($n < 200$) are often encountered when analyzing small-scale sample surveys or in subgroup analysis of larger epidemiologic datasets. In such scenarios, the simulation studies provide guidelines for whether or not adjustment is desirable. Based on analysis of a large number of scenarios simulated to mimic the practical situations, we have provided empirical guidelines with regard to situations where covariate-adjusted rates should not be used in place of crude rates. Such guidelines are presented in Figure 1 for many settings. Mention this in the results section: A program (in Stata 9.0 (working in progress)) is available on our web site that can be used to evaluate whether confounding should be controlled in other settings.

In this paper, we illustrated very simple scenarios where there is only one covariate (gender) with two classes (male versus female). More complicated yet common scenarios, such as those with multiple subgroups (e.g., race-ethnicity), or multiple categorical covariates (e.g., gender and race-ethnicity), or with mixture of categorical covariate (gender) and continuous covariate (age). I think that with age, the key is whether the relationship of the category with the response rate is linear. If it is, the regression coefficient should be like the one you use with two groups. Is this right? However, the linear assumption would have to hold. If it is categorical, then will the
adjustments be like many 2-category adjustments?, should be further investigated. Can we say something about this? what happens with 3 groups? You can always reduce 3 groups into 2-group lumpings. In particular, the impacts of the uncertainty of variance component estimation on the variances of covariate-adjusted rate estimators should be evaluated by Monte Carlo simulations. We are currently developing simulation results that address these settings. (are you doing this? I wonder if you could speculate on what you think will happen, and say that you’ll be studying it more).

The findings of our analysis may have important implications on the reporting methods of the public health sample survey data, such as BRFSS surveys, in particular, how the municipal (county, town or city)-level statistics should be reported. For example, among the 351 communities (towns or cities) of the Commonwealth of Massachusetts, 49 have a sample size ≥30, and only 7 have a sample size ≥100 in the 2003 BRFSS survey data. If the smoking prevalence rates of the 49 municipals with sample sizes ≥30 are to be reported based on direct estimates from the sample, our empirical guidelines suggest that gender-adjusted rates should not be used in place of crude rates. (I think you could expand on this, as it is very interesting. Suppose you stratify the 351 communities by sample size. Can you then list what should be adjusted for? For example, you have commented not only on gender, but also income, education, and age (young vs old).

In summary, we (can we be more pro-active. Can we recommend practical guidelines for not adjusting.) caution about automatically including control for confounding without reviewing evidence of a relationship between the response variable and the confounder. Covariate-adjusted rates should not be used in all settings, and in particular, not when prevalence rates, prevalence rate ratios and sample sizes are small. In such settings, covariate-adjustment will lead to worse precision of the rate estimates.
and only an illusion of statistical control. We anticipate that this problem will be aggravating by inclusion of multiple covariates, but this problem warrants further investigation.
Reference


Legend

Figure 1 Relative risk threshold for variance reduction due to gender adjustment by prevalence rate, population and sample sizes
The diagram illustrates the relationship between Relative Risk Threshold and % of Males in population, with different prevalence rates and numbers of cases. The prevalence rates range from 15% to 35%, and the numbers of cases range from 200 to 1600. The graph shows how the Relative Risk Threshold changes with different prevalence rates and numbers of cases, indicating the impact on the % of Males in population.

**Prevalence Rate**
- 15%
- 20%
- 25%
- 30%
- 35%

**Relative Risk Threshold**
- 200, 100
- 400, 100
- 800, 100
- 1600, 100
- 200, 150
- 400, 150
- 800, 150
- 1600, 150
- 400, 200
- 800, 200
- 1600, 200

**% of Males in population**
- 30, 40, 50, 60, 70

---

**Deleted:** C:\Documents and Settings\Edward\My Documents\cluster\report\liw_06_adjusted_rates.doc
**Inserted:** C:\Documents and Settings\Edward\My Documents\cluster\report\liw_06_adjusted_rates-ejs.doc
**Deleted:** 3/13/2006

---

**Deleted:** J:\Projects\Cluster\report\liw_06_adjusted_rates-ejs.doc