Using the Kalman Filter to Improve Disparity Estimates for Rare Racial/Ethnic Minorities: An Application to American Indians / Alaska Natives and Chinese Americans using the National Health Interview Survey

1. PROBLEM

The United States has experienced dramatic increases in life expectancy and declines in rates of morbidity since 1900. Nevertheless, significant differences in health and mortality among subpopulations persist, even as the major causes of death and disease change over time.

While national health surveys collect substantial data on large racial/ethnic groups, little is known about the extent of disparities for smaller groups, such as American Indians/Alaska Natives (AI/AN) and Asian subgroups. In fact, major national health surveys have limited to no ability to adequately measure the health of major racial/ethnic subgroups smaller than Mexican-Americans (Waksberg et. al. 2000). Only the Census 2000 and the American Community Survey (ACS) provide sufficient sample sizes for accurate estimates for AI/AN and Asian subgroups,[1] but neither collects detailed health-related outcomes. As such, US policymakers are in dire need of a more thorough assessment of both the health status and the extent of health disparities in groups comprising 0.5-1.0% of the US population.

Current Data Limitations for Small Groups

Most major health surveys cannot support even simple distributional analysis for Hispanic or Asian subgroups and Native Americans. Only the National Health Interview Survey (NHIS) can generally support simple analyses for these subgroups, but it cannot support detailed analyses for subgroups other than Mexican-Americans. Table 1 describes these limitations.

NHIS is the principal source of information on the health of the US civilian non-institutionalized population. As the most promising data set to address questions of racial/ethnic disparities, we will focus on NHIS data as a means to evaluate our proposed approach. This does not imply that our approach is specifically designed to be used with the NHIS; it is intended to be applicable to other data sources as well.

[1] Sample sizes and design effects based on the survey sample designs as of 2000.

The NHIS is an annual multi-stage, national probability sample. Household interviews provide rich demographic and socioeconomic information, in addition to the wealth of health measurements. It is also worth noting that NHIS is a repeated annual cross-section of the US population: it is not a longitudinal measurement of the same individual(s).

Even though NHIS is the largest national health survey, it only includes 200 completes each for AI/AN and Chinese annually, and contains relative standard errors of 30% for such groups for several outcomes. These shortcomings mean that NHIS is unable to accurately estimate prevalences among smaller groups or determine the extent of disparities, and that health estimates using data from the current year only are very imprecise.

Table 1. Adequacy of Selected Health Data Sets (NCHS/CDC and AHRQ) With Acceptable Levels of Precision to Determine Prevalence Rates among Race/Ethnic Sub-Groups.

Race/Ethnicity	NHIS	NSFG	NIS	NHANES	MEPS
Mexican-American	A	C	B	C	B
Puerto Rican	B	D	C	D	C
Cuban	C	D	D	D	D
Central/South American	B	D	C	D	C
Other Hispanic	B	D	C	D	C
American Indian/Alaska Native	C	D	C	D	D
Chinese	C	D	C	D	D
Filipino	C	D	D	D	D
Japanese	C	D	D	D	D
Asian Indian	C	D	D	D	D
Korean	C	D	D	D	D
Vietnamese	C	D	D	D	D
Hawaiian	D	D	D	D	D
Other	C	D	D	D	D

Notes: Table is Adapted from Waksberg et al. 2000.
A=Detailed Cross-Classification Possible
B=Some Limited Cross-Classification Possible
C=Only Simple Distributions are Possible
D=No Analysis is Possible

NHIS=National Health Interview Survey; NSFG=National Survey of Family Growth; NIS=National Immunization Survey; NHANES=National Health and Nutrition Examination Survey; MEPS=Medical Expenditure Panel Study.

Approaches to Improving the Accuracy for Small Groups

Design and Analysis

Two general approaches for potential improvements in the accuracy of health estimates for small racial / ethnic groups are: (1) Design-based approaches: improvements in the effective sample size through targeted and efficient increases in sample sizes for these groups and; (2) Analytic approaches: more accurate inference from existing data through innovations in analysis.

These approaches are not mutually exclusive; one can both supplement the available sample size and further increase efficiency through analysis. In other work, we focus on design-based approaches. In this work, we focus on analytic approaches, which we will note are typically less expensive to implement.

Systematic Literature Review

During 2003-2004, we conducted a systematic, English-language literature review of documents relating to the estimation of prevalence of health outcomes for rare populations. We used a variety of search terms related to sample design and analysis. We employed conventional and unconventional searches: searching a variety of library databases, a snowball survey of experts that began with authors of review papers, and a review of documents from the authors' personal libraries. We hand-searched reference lists of documents for additional relevant citations and considered "gray literatures"—unpublished and limited-distribution documents. We screened titles and abstracts, ordered documents that appeared relevant, and abstracted data via a standardized form.

Our searches produced 1866 titles, including published articles, government reports, websites, technical reports, and other unpublished material. Most (84%) of the titles came from online databases, 14% from experts, and 3% from reference mining; this count contains duplicates of citations. Additional screening restricted us to 453 relevant documents, while twelve of which (3%) were unobtainable despite repeated attempts. In-depth screening further limited the total to 326 relevant documents, which were reviewed in detail. Many of these documents pertained to design work described elsewhere, but 175 pertained to analytic techniques.

Our specific focus

Our review of these 175 articles and discussion with the technical expert panel of this Office of Minority Health sponsored project (contract no. 282-00-0005), led to a conclusion that analytic techniques that pooled data over time were the most promising approach to improving power analytically for data with this repeated cross-sectional structure.

In order to have two test cases, we selected American Indians/Alaska Natives (AI/AN) and Chinese Americans. Both of these groups (1) represent ½-1% of the US population, (2) have been identified as distinct subgroups in NHIS data for a number of years, and (3) have about 200 observations annually in the NHIS adult sample in recent years.

These two groups also provide contrasts in characteristics such as health status (with AI/AN thought to be in generally in poorer health than non-Hispanic whites (NHW) and Chinese generally though to be in better health than NHW- see Loue, 1999a,b), generational status, and geographic location. The 1990 census indicated that 30 percent of AI/AN live on reservations (especially in the Western United States) or in blocks with a concentration of AI/AN greater than 60 percent (Massey et. al. 1993). By contrast, the 1990 census shows that three-fifths of Chinese lived in California or New York, predominantly in urban areas.

2. OBJECTIVE

Current Year Estimates in Disparities

We take as our objective obtaining a more accurate estimate for racial/ethnic disparities in the current year based on data from that current year combined with data from previous years. We focus on current year estimates as a means of providing policymakers with the most accurate estimates of disparities at a point in time. For measuring progress with respect to Healthy People 2010 goals, the National Center For Health Statistics (NCHS) defines disparities for a given group as the difference between a given group and a reference group that is the racial/ethnic group with the best level of that outcome among those racial/ethnic groups with a relative standard error of less than 10% for that outcome (Keppel et. al, 2005). This reasonable metric ensures that the reference group for disparity estimates is itself well-estimated. This implies that the accuracy of health estimates for rare racial/ethnic groups such as AI/AN and Chinese will be the limiting factor in making disparity estimates. Therefore that the goal of obtaining more accurate current year disparity estimates for these groups translates into a more specific objective of obtaining more accurate current year health estimates for AI/AN and Chinese, but in a way that does not bias estimates of disparities.

Mean-Squared Error (MSE) as a Metric

The precision of health estimates is usually discussed in terms of Standard Errors (SE, standard deviations of an estimate) or Relative Standard Errors (RSE, standard errors divided by their corresponding point estimates). Both of these terms are based on the variance of the estimator -- the expected squared deviation of the estimator from its own expected value. With conventional analytic approaches, the expected value of the estimator corresponds to the population value of the parameter being estimated, and the variance of an estimator is equivalent to its mean-squared error (MSE), the expected squared deviation of the estimator from the true population parameter.

Several of the analytic techniques we discuss improve accuracy in ways that result in expected values of estimators that differ from the population parameters. These measures require that we use the more general measure of MSE, which applies to both traditional and more innovative analytic approaches. For the purposes of this paper, we will follow the convention of using the term *accurate* to describe low MSE estimators and the term *precise* to describe low variance estimators. We will also refer to the *Root MSE* (RMSE) of an estimator, which is the square root of the MSE of an estimator, and which is analogous to the standard error. Finally, we will also consider the *Relative RMSE* of an estimator, which is the RMSE divided by the estimate, and which is analogous to the RSE.

Approaches to Pooling Data

Pooling data across racial/ethnic groups

One common approach to improving the accuracy for groups is to use small area estimation (Rao 2003) to “borrow strength” from larger groups. In the present application, that might correspond to shrinking estimates of health prevalences and means for AI/AN and Chinese towards an overall mean dominated by non-Hispanic Whites (NHW) and other larger groups. Such an approach would improve the MSE of current year health estimates for AI/AN and Chinese, but would systematically bias disparity estimates by shrinking health estimates towards the overall mean. This bias would be larger in magnitude for small groups and groups with true health status that differed greatly from the population average. For AI/AN in particular, such an approach would be likely to result in substantially understating the extent to which their health was poorer than the reference group.

For these reasons, we consider borrowing strength with respect to central tendency (means, proportions) across racial/ethnic groups to be incompatible with the objectives of the project. On the other hand, there may be opportunities to borrow strength to estimate other model parameters (e.g. variances, autocorrelation parameters, slopes with respect to time) in ways that do not threaten the validity of the resultant disparity estimates

Pooling data over time

<u>Limitations</u>

There are several limitations in pooling repeated cross-sectional data over time. The first is that because the data is not longitudinal (repeated measures within individuals), the various techniques that might take advantage of that structure to improve accuracy (Hedeker and Gibson, 2006) are not available. As will be discussed in greater detail below, this approach models a series of annual racial/ethnic means (group-year means) on the basis of data sampled within each group and year.

A second limitation is that for a given data set, one is restricted to a set of years over which the racial/ethnic subgroups of interest are broken out individually and defined consistently. One is further limited to outcome variables that are continuously available with an unchanging definition. As will be described below, our data take advantage of eight years of measurement over which a variety of outcomes remain unchanged and which includes 11 racial/ethnic subgroups (including AI/AN and Chinese) definitions that are available and consistently defined on the NHIS.

<u>Simple averaging</u>

The most straightforward pooling approach simply averages data from the *k* most recent years to estimate the current year. Other than a few minor issues as whether to weight years or observations equally, this is a very simple approach. If sample sizes are the same each year, and there is no autocorrelation in racial/ethnic group means over time, the variance would be divided by *k,* and the SE and RSE would be divided by the square root of *k*. In other words, pooling over four years would divide an RSE by two.

If one only wants to estimate the average health for Chinese or AI/AN (or the average disparity relative to a reference group) over the whole *k* year period, this approach is efficient and unbiased. As an estimator of current year health and health disparities, this approach may improve the accuracy of a direct estimate based on the current year alone. On the other hand, this approach has shortcomings that do not fully take advantage of the potential for pooled data to estimate health in the current year. These shortcomings are discussed next.

Linear trends and autocorrelated variance in racial/ethnic group means

First, to the extent that there is a linear trend (if the health indicator is improving or declining over the *k* years), then the unweighted *k*-year average estimator will be biased as an estimator of the current year. In particular, the simple *k*-year average will underestimate a rising outcome in the current year and overestimate a falling one. For a rapidly changing indicator, such as diabetes or obesity, this bias could be very large relative to the precision gains from pooling data over time, and could erode gains in accuracy from pooling even in very small groups.

Second, it is unlikely that group means for a given racial/ethnic group are independent over time for all health measures. A more likely state of affairs involves (a) a variance component at the level of the group mean, and (b) some positive autocorrelation in that group mean. A variance component in the group mean reflects the fact that there are likely to be some random annual factors that affect the health of all members in a given racial/ethnic group, in addition to a second set of random factors that individually affect the health of specific people within that group. The presence of positive autocorrelation in these means indicates a tendency for group means in subsequent years to regress less towards the mean than would be the case with independent observations. This reflects the fact that the health of population groups tends to move from its current state, rather than being created anew each year.

Taken together, autocorrelated group-year variance components result in a situation in which all observations are not equally valuable in predicting the current year. Older years are less predictive and therefore should be weighted less. Thus simple averaging is inefficient in the presence of structured (non-independent) variance of group-years means.

Our alternative: A modified Kalman Filter with a linear trend

In order to pool data across time and improve the MSE of racial/ethnic disparity estimates in the current year for small groups such as AI/AN and Chinese in the presence of linear trends in the health measures and autocorrelated variance components at the level of the group mean, we propose a modification of the Kalman Filter. The Kalman filter was developed by Rudolph Kalman (1960) as an iterative updating algorithm to "filter" out "noise" (as opposed to signal) in engineering applications. Since then, it has been recognized that it provides a very general set of tools that can be applied to a variety of settings far removed from this original application. In particular, it has been demonstrated that it can capture the variance structure of hierarchical data with autocorrelation at the higher of two levels (Blight & Scott, 1973; Binder & Dick, 1989; Lind 2005), which would include group means with positively autocorrelated variance components.

A Kalman filter typically assumes a stationary process, with no trends in means. Because we would like to capture possible trends in health means within our *k*-year window of available data, we implement a modified Kalman filter with a linear trend. We first use linear regression to estimate a linear trend over time within each racial/ethnic group and then apply the Kalman filter to the residuals of this model. We therefore produce current year estimates of health disparities for small racial/ethnic groups that are not only more accurate than the direct estimates, but which also account for both linear trends and structured cross-sections group year means in a manner that actually minimizes MSE under reasonable model assumptions. It should be noted that the gains from the MKF are likely to vary by outcome.

3. METHODS

Data

For the purposes of this paper we used data from eight years of the NHIS (1997-2004), restricting to the 258,279 cases in the adult sample 1997-2004 (approximately 31,000-36,000 cases each year). The adult sample includes a much broader set of health outcomes than the full (core) sample. Data from all respondents was included, regardless of race/ethnicity. The NHIS originally categorized race-ethnicity into approximately 20 categories, though these varied somewhat by year. Where necessary, we collapsed categories in order to achieve subgroup definitions that did not change over the eight years and which were consistent with OMB definitions. In other cases, we collapsed smaller subgroups that were not our focus in order to improve estimation. For example, 11 Hispanic subgroups were collapsed into four groups. This resulted in a total of 11 categories: NHW, Black, Mexican, Puerto Rican, Cuban, Other Hispanic, Chinese, Asian Indian, Filipino, AI/AN, and Other (includes "Other Asian," "Other Race," and "Multiple Race").

We selected a total of 18 health outcomes, a large proportion of all available variables for all eight years, omitting only six measures of pain and items regarding colds and stomach illness, as well as collapsing four indicators of heart disease into one indicator. We included a wide variety of outcomes that were continuous and dichotomous (with high and low prevalences), etc.: number of outpatient visits, number of inpatients days, number of workdays lost, number of functional limitations, indicators of having ever had specified chronic diseases (cancer, stroke, heart disease, diabetes, hypertension, asthma, kidney disease, emphysema, chronic bronchitis), indicators of recent episodes of acute illnesses (hay fever, sinusitis, bronchitis, ulcer), and substance use indicators (recent binge drinking, ever smoked).

Modified Kalman Filter (MKF)

Here the modified Kalman Filter models a total of 88 group-year means (8 years x 11 racial/ethnic groups) with independent intercepts and slopes for each racial/ethnic group, assuming a linear trend. It further assumes an AR(1) correlational structure, which means that the residual for a given year is independent of the past residuals after conditioning on the previous year's residual. While more complex autoregressive structures may exist, an AR(1) parameterization tends to be a good first approximation of more complex structures, and the number of group years means is insufficient to allows a more complex autoregressive structure to be well estimated (or even for the need for that structure to be established). As a further step to improve the estimation of this parameter, we estimate a single AR(1) parameter (rho) for all groups for any given outcome. We further assume that the group-year level variance does not differ by group for a given outcome.

Under the assumptions of a linear trend in group-year means with a true AR(1) autocorrelated group-year level variance that does not vary by groups or year within a given outcome, the MKF provides minimum MSE estimates for group means for each time period.

In what follows, we treat current nominal NHIS sample sizes as if they were effective sample sizes (design effect of 1), for the sake of simplicity. Because the MKF results in greater proportionate gains for groups with smaller annual sample sizes, this conservative assumption will understate the potential gains.

Parameters of the MKF

Below we (1) describe the individual-level model from which the group mean is derived, (2) describe the updating formula that models residuals of the group-year mean, (3) interpret key parameters, and (4) show how the MKF can be used to make predictions.

Individual-level model

The outcome for the jth, j=1,…,n_{it}, member of group i in period t is

$$y_{it[j]} = \mu_i + \gamma_{it} + \varepsilon_{it[j]}$$

where $\gamma_{it} = \rho\gamma_{it-1} + \xi_{it}$, $\xi_{it} \sim N(0, \tau^2)$ and $\varepsilon_{it[j]} \sim N(0, \sigma^2)$, reflecting AR(1) variance.

Here $\mu_{i\,t}$ is the linear trend in the group mean

The mean for each period can then be written as

$$\bar{y}_{it\bullet} - \mu_i = \gamma_{it} + \eta_{it},$$

where $\eta_{it} = \bar{\varepsilon}_{it\bullet}$ with variance $\nu_{it} = \sigma^2/n_{it\cdot}$, the mean of individuals' errors averaged within group by time. Here $\bar{y}_{it\bullet}$ is the year t deviation from the trend line for group i.

Kalman Filter Updating

The Kalman filter provides an updating formula for generating the minimum MSE estimators of the state variables (residuals). The MKF estimates starting values, then updates "state variables" (γ_{it-1}) which correspond to the residuals of the previous time period), thereby shrinking the current year's residual toward an estimate based on past data in the following recursive formula:

$$\hat{\gamma}_{it} = \lambda_{it}(\bar{y}_{it\bullet} - \mu_{it}) + (1-\lambda_{it})\rho\hat{\gamma}_{it-1}$$

where $\lambda_{\text{it}} = \delta_{it}/(\delta_{it} + \nu_{it})$, $\delta_{it} = \rho^2\omega_{it-1} + \tau^2$, and $\omega_{it} = \delta_{it}(1-\delta_{it}/(\delta_{it} + \nu_{it}))$.

The amount of shrinkage depends on the relative accuracy of the two estimators (current period and past). Future work may annually update slope as well as residuals, using such approaches as double exponential smoothing (Bowerman and O'Connell, 1993).

Kalman Filter Parameters

λ_{it} sets the weight of the current period. Note that this weight increases with the variance of the past relative to the variance of the current period. δ_{it} is the variance of the past. Note that this variance has both an autocorrelative portion and an annual group mean innovation. In the presence of a linear trend, positive autocorrelation implies a tendency for successive residuals (and group-year means) to stay on the same side of the regression line

Making Predictions

To predict the mean for a given racial/ethnic group in the current year, one simply adds the current year filtered residual to the regression prediction for the current year. The past has more influence on these predictions when innovation is low and when autocorrelation is high. The extent to which predictions based on the past differ from the direct estimate from the current year depends upon the magnitude of deviation of the current year residual from a linear regression line.

Fitting the Kalman Filter

Variance Estimation

There were several steps in fitting the MKF. First, we used PROC MIXED in SAS to estimate three values regarding the variance structure: ρ, τ^2, and σ^2. This estimation had to be decomposed into several steps to stay within the matrix limitations of SAS.

Estimates of the two variances were reasonably precise, but assuming an AR(1) parameter that did not vary by group was insufficient to estimate that parameter in this dataset with sufficient precision. We reduced the variance in that parameter estimate by borrowing strength across the 18 outcome models. In particular, we used Bayesian shrinkage based on a one-way random effects model (Carlin & Louis 1996), which shrinks the direct estimate of rho for each outcome toward the overall mean rho across outcomes (0.351). Note that this approach does not bias estimate of disparities.

Starting values

We also needed to estimate three starting values: μ_i, $\gamma_{i\,0}$, and ω_{i0}. We used $\bar{y}_{i\bullet\bullet}$, the overall mean for racial/ethnic group i, as the estimator of μ_i.. We assumed a starting value of zero for the initial residual $\gamma_{i\,0} = 0$, and estimated ω_{i0} as $\tau^2/(1-\rho^2)$, the variance of γ_{i1}. These last two starting values makes the first period residual estimate an empirical Bayes Stein estimator: the first period mean shrunk back toward zero in proportion to the MSE of the state variable and the noise.

4. RESULTS

We begin by considering four key parameters of the Modified Kalman Filter and follow this by examining for which outcomes, in the net, the Modified Kalman Filter (a) improves upon the MSE of a direct current year estimate, (b) substantially alters that direct current year point estimate, and (c) improves upon the MSE of a simple eight-year average.

Key Parameters by Outcome

One way to understand the Modified Kalman Filter in this application is to consider the four key parameters that determine the extent to which it improves upon traditional approaches such as a direct current-year estimate and a simple eight-year average. These four parameters are:

(1) The annual sample size for the target racial/ethnic group

(2) The standardized group-year mean innovation

(3) The autocorrelation parameter (rho)

(4) The standardized slope with respect to time

Annual Sample Size of target racial/ethnic group

The sample size available each year for a given target group is the main determinant of the accuracy of simple current year estimates. Because estimates are imprecise, for a given standardized group-year mean innovation and a given level of autocorrelation, smaller groups will rely more upon the past and stand the most to gain in accuracy from MKF. Our primary focus, as noted above, was Chinese Americans and AI/AN, each of which have about 200 completes in the NHIS sample each year. For comparison, we briefly consider NHW (about 20,000 completes annually) and Blacks (about 5,000 completes annually).

Standardized group-year mean innovation

A second factor influencing the contribution of the MKF is the size of the standardized group-year mean innovation. This ratio of variance components compares the annual variance that affects all members of a given racial/ethnic group to the variance at the individual level. If the group-year mean has a large independent annual variance component, the past cannot contribute much to estimating current values, even with small sample sizes in the current year, because its variance overwhelms any gains in precision that might otherwise result. For example, if this ratio of variance components exceeds the inverse of the annual (effective) sample size, the past cannot contribute whatsoever to estimation of the current year. Smaller values will result in greater gains in accuracy from MKF compared to direct estimates from the current year alone. Across the 18 outcomes we examined, the median value of this variance component ratio is 6 x 10^{-4} (1/1667). The largest value is 387 x 10^{-4} =1/26 for BMI, with the next largest values for t smoking and outpatient visits. These outcomes will stand to benefit very little from the MKF. The smallest ratios are less than 1 x 10^{-4} (stroke, followed by heart disease). These outcomes stand to benefit substantially from the MKF for small groups such as AI/AN and Chinese, and may even benefit from the MKF for groups with sample sizes as large as 10,000, such as Blacks.

Autocorrelation parameter (rho)

The MKF gives more weight to the past in its point estimate when autocorrelation is higher, as positive autocorrelation implies the past is more predictive of the current year. Positive autocorrelation also increases the contributions of the past to the accuracy of the current period estimate.

The shrunken estimates of rho have a mean of 0.32 across the 18 outcomes and a range of 0.00-0.68. Autocorrelation is highest for measures of utilization (inpatient and outpatient) and near zero for BMI, emphysema, diabetes, and smoking.

Standardized slope with respect to time

The standardized slope describes the rate of change in an outcome in individual-level standard deviations per year. The rate at which an outcomes changes over time is an important determinant of the viability of the simple eight-year average as an estimator of the current year mean. When the absolute standardized slope is large, the simple average over time is biased estimator of current year. If the absolute standardized slope is large then this bias can result in a worse MSE for an eight-year average than for a direct estimate based on a single year alone.

With (effective) sample sizes of 200 per year, slopes as small as 0.04 standard deviations per year can result in poorer MSEs from 8-year averages than for a direct current year estimate; with larger annual sample sizes, the sensitivity of 8-year averages to slopes is greater.

The median overall absolute values of standardized slope is 53 x 10^{-4} (1/200), with a largest overall value for diabetes (156 x 10^{-4}), followed by BMI and hypertension.. The smallest overall absolute slope is for heart disease (4 x 10^{-4}), followed by emphysema. Slopes were statistically significant for more than half of outcomes; we detrend linearly even when $p>0.05$.

For which outcomes does the Kalman Filter help?

For which outcomes does the Kalman Filter improve accuracy over direct current year (2004) estimates?

The MKF improves the MSE of a current year estimate when it gives non-trivial weight to the past. Smaller values of lambda in the last iteration of the Kalman filter updating equation indicate a greater weight given to the past and greater improvements in MSE from the MKF. In this sense, lambda summarizes the effects of small current year sample size and small innovation on increasing the contribution of the past to prediction of the current year.

For AI/AN and Chinese, the MKF gives weights of more than 1% to the past for 7 of 18 outcomes, as shown in Table 2. The MKF would notably improve the MSE of these measures, which span a wide range of areas. For NHW and Blacks, the weight of the past is <1% for virtually all outcomes.

Table 2. Weight Given to Past by MKF for Selected Outcomes

Outcome	Weight Given to Past	
	AI/AN	Chinese
Workdays lost	96%	95%

Inpatient days	88%	88%
Diabetes	47%	45%
Smoking	21%	20%
Ulcer	13%	12%
BMI	12%	13%
Emphysema	3%	3%

For which outcomes does the Kalman Filter change direct current year (2004) point estimates substantially, in addition to improving their accuracy?

It is possible for the MKF to reduce the MSE of the current year estimate without substantially changing its point estimate. The converse (substantial changes in point estimates without substantial reductions in MSE) will not be occur under our model assumptions, because if those assumptions are met, the MKF will not substantially alter the current year estimate unless the MSE of the current year estimate can be reduced by doing so.

It will change the point estimate based only on the current year when both (1) the weight of the past is fairly large and (2) the current year is a fairly large residual from the linear regression line. In these instances, the MKF pulls current year estimates in toward the regression line, which can be seem by examining differences of MKF point estimates from direct current-year estimates.

There were two instances in which the MKF point estimate differed from the current year direct point estimate by a relative difference of 5% or more for AI/AN:

- Inpatients days (MKF 5.67 mean days 2004 vs. .direct current year mean of 5.32 inpatient days)

- Diabetes (MKF 15.5% 2004 lifetime prevalence vs. direct current year mean of 16.9% diabetes).

These are also the two outcomes with the largest relative deviations of MKF estimates from the direct 2004 means for Chinese, though they only represent 3% relative deviations in the case of Chinese.

For AI/AN, the direct current year mean of 16.9% lifetime prevalence for diabetes is a notably above the prediction of a regression line with a strong positive slope. Under the model assumptions of the MKF, that is interpreted as a large current year residual that needs to be brought nearer the regression line. It is also possible that the underlying growth in diabetes is greater than linear (e.g. quadratic), a violation of our model assumptions. In that case, the MKF would be overly aggressive in pulling the current year in from its observed value.

For which outcomes are simple eight-year averages substantially biased?

As noted above, simple eight-year averages, a popular approach to improving precision, results in biased estimates of the current year in the presence of a linear trend. There may also be substantial differences between the eight-year averages and MKF estimates when lambda suggests weighting the present year in a manner that does not give it equal weight with past years.

Tables 3 and 4 illustrate the outcomes for which the MKF estimates have a relative difference of 20-26% or more from the eight-year averages.

Table 3. Differences of MKF Point Estimates from 1997-2004 Averages for Selected Outcomes: AI/AN

($p<0.05$ for all, differences exceed 20% of 1997-2004 average)

	MKF 2004 estimate	**1997-2004 average**

Inpatient Days	5.7d	11.5d
Ulcer	8.8%	13.6%
Cancer	7.7%	5.9%
Bronchitis	5.5%	7.7%
Emphysema	15.8%	19.5%
Diabetes	15.5%	12.6%

Table 4. Differences of MKF Point Estimates from 1997-2004 Averages for Selected Outcomes: Chinese

($p<0.05$ for all, differences exceed 26% of 1997-2004 average)

	MKF 2004 estimate	1997-2004 average
Stroke	2.1%	1.0%
Emphysema	<0.1%	0.1%
Bronchitis	.5%	1.4%
Sinusitis	4.2%	7.1%
Diabetes	5.5%	3.9%
Binge Drink	8.7%	6.7%
Heart Disease	3.7%	5.2%

5. DISCUSSION AND CONCLUSIONS

For a substantial subset of health outcomes, the MKF appears to be a powerful tool to improve the accuracy of health and health disparity estimates for small racial/ethnic groups, such as AI/AN and Chinese. In particular, the MKF is likely to improve the MSE of direct 2004 estimates for 7 of 18 outcomes examined, some substantially:

-Past year mean inpatient days and workdays missed

-Lifetime incidence of diabetes, emphysema, and smoking

-Past year episodes of ulcers

-Current BMI

For other outcomes, the MKF is not likely to improve accuracy substantially. Although not the focus of this approach, the MKF could improve the MSE somewhat for even groups as large as Blacks and NHW for outcomes such as mean inpatient days and workdays missed.

While the current practice of using unweighted averages as estimators of the average of a number of years is inherently valid, such an approach is subject to severe bias for about half of all outcomes examined when used as an estimator of the current period. Such an approach cannot be recommended.

We recommend that the MKF be used for the substantial subset of items for which it improves the MSE of direct current year estimates for small racial/ethnic groups such as AI/AN or Chinese. For other outcomes, MKF estimates will resemble direct current year estimates in accuracy and value. Where the MKF does help, synergistic gains can be achieved by combining this approach with targeted increases in sample size for these subgroups through changes in sample design, although these gains would be less than fully multiplicative, since the proportionate gains from the MKF decrease somewhat as sample sizes increase.

Finally, longitudinal continuity of racial/ethnic subgroup definitions and outcome variable definitions and availability in NHIS and similar data sources will support the use and effectiveness of this approach. We encourage those designing and administering these surveys to maintain this continuity where possible, as it may lead to substantially more precise information regarding health disparities for small racial/ethnic groups.

6. FUTURE WORK UNDER DEVELOPMENT

We are developing detailed simulations that compare the MSE of the Modified Kalman Filter to several alternatives across the parameter space defined by the outcomes we modeled here. We will also investigate the extent to which double exponential methods allow more flexible fitting of group mean trends.

REFERENCES

Binder DA and Dick JP (1989) “Modeling and Estimation for Repeated Surveys” Survey Methodology, 15: 29-45.

Blight BJN and Scott AJ (1973) “A Stochastic Model for Repeated Surveys” JRSS-B 35:61-66

Bowerman, BJ. and O’Connell RT (1993). *Forecasting and Time Series: An Applied Approach*. Duxbury Thomson
Learning.

Carlin BP and Louis TA (1996). *Bayes and Empirical Bayes Methods for Data Analysis, 2nd Ed.*. Chapman & Hall: New York..

Hedeker, D and Gibbons, R D (2006). *Longitudinal Data Analysis.* John Wiley & Sons: New York.

Kalman R (1960) “A New Approach to Linear Filtering and Prediction Problems” Transactions of the ASME--Journal of Basic Engineering, 82: 35-45

Keppel, K., Pamuk, E., Lynch, J., Carter-Pokras, O., Kim, I., Mays, V.M., Pearcy, J., Schoenbach, V., Weissman, J.S. (2005). Methodological issues in measuring health disparities. National Center for Health Statistics. Vital and Health Statistics 2, 141, 1-16.

Lind JT (2005). “Repeated Surveys and the Kalman Filter” Econometrics Journal, 8:418-427.

Loue, S. (1999a). Asian and Pacific Islander Health. Pp. 101-16 in *Gender, Ethnicity and Health Research*. New York, N.Y.: Kluwer Academic.
.
Loue, S. (1999b) The Health of Native Americans. Pp. 133-45 in *Gender, Ethnicity and Health Research*. New York, N.Y.: Kluwer Academic.)

Massey JT, Judkins D, Waksberg J (1993). Collecting health data on minority populations in a national survey. Paper presented at: Proceedings of the Survey Research Methods Section, American Statistical Association; 1993, 75-84 .
Available from: http://www.amstat.org/Sections/Srms/Proceedings/papers/1993_009.pdf [accessed November 2004].

Rao, J.N.K. 2003. *Small Area Estimation.* New York, John Wiley and Sons.

Waksberg J, Levine D, Marker. D (2000) . Assessment of Major Federal Data Sets for Analyses of Hispanic and Asian or Pacific Islander Subgroups and Native Americans: Inventory of Selected Existing Federal Databases. Washington, DC: US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation..