ABSTRACT

Data from existing administrative databases and ongoing surveys or surveillance methods may prove indispensable after mass traumas as a way of providing information that may be useful to emergency planners and practitioners. The analytic approach, however, may affect exposure prevalence estimates and measures of association. We compare Bayesian hierarchical modeling methods to standard survey analytic techniques for survey data collected in the aftermath of a terrorist attack. Estimates for the prevalence of exposure to the terrorist attacks of September 11, 2001, varied by the method chosen. Bayesian hierarchical modeling returned the lowest estimate for exposure prevalence with a credible interval spanning nearly 3 times the range of the confidence intervals (CIs) associated with both unadjusted and survey procedures. Bayesian hierarchical modeling also returned a smaller point estimate for measures of association, although in this instance the credible interval was tighter than that obtained through survey procedures. Bayesian approaches allow a consideration of preexisting assumptions about survey data, and may offer potential advantages, particularly in the uncertain environment of postterrorism and disaster settings. Additional comparative analyses of existing data are necessary to guide our ability to use these techniques in future incidents. (Disaster Med Public Health Preparedness. 2008;2:119–126)

Key Words: Bayesian; survey; surveillance; data
measurement issues such as the lack of uniform definitions among the multiple sources of health information, correlation among questions and data sources, and the need to frame questions in appropriate geographic contexts that account for varying levels of exposure. Finally, in the context of complex surveys, weighting and extrapolation of postdisaster data to a target population may become difficult when only a subset of the total survey sample is interviewed about disaster-related outcomes.

The September 11, 2001, terrorist attacks in New York City were in many respects the prototypical large-scale disaster. They were an unanticipated (and unprecedented) event, which shocked the city and the country and which paralyzed normal function in New York City for weeks. Although a number of original research studies were launched in the aftermath of the attacks, public health authorities quickly recognized that extant surveillance systems may offer unique insights into the health of New Yorkers after these events. To that end, the Centers for Disease Control and Prevention and the New York State Department of Health added questions to address the general population–level behavioral and mental health impacts of the September 11, 2001, terrorist attacks to their ongoing Behavioral Risk Factor Surveillance System (BRFSS) telephone survey. Some results from this survey have been reported in aggregate with similar efforts in New Jersey and Connecticut.

It is the purpose of this article to use this BRFSS effort as a model to introduce and examine some aspects of the challenges involved in analyzing and interpreting existing surveillance data, or new modules appended to existing surveillance efforts, to provide information that is useful in the postdisaster context. We are interested, in particular, in how analytic approach may affect exposure prevalence assessment and measures of association. We compare frequentist survey analytic techniques with Bayesian approaches and examine how such outcomes may affect disaster response and resource allocation and suggest directions for future efforts.

There have been few reports on how best to analyze and interpret such surveys for the purposes of informing postdisaster emergency response efforts. This article should serve as an initial illustration of the implications of different modeling decisions for the purposes of informing future analyses of existing surveillance data in postdisaster contexts, and to highlight the effect analytic approaches may have on the interpretation of such efforts.

Specifically, we compare Bayesian hierarchical modeling methods to standard survey analytic techniques for survey data collected in the aftermath of a terrorist attack. We present results on exposure prevalence and on the outcome of a cross-sectional logistic regression for a possible mental health outcome, and discuss the implications of statistical methodology on efforts to guide disaster relief. We are not aware of published reports to date that have considered a Bayesian approach to postterrorism or disaster survey data.

**METHODS**

**Data Sources and Definitions**

The BRFSS is a yearly random-digit–dialed telephone survey of noninstitutionalized adult civilians conducted by each state under the auspices of the US Centers for Disease Control and Prevention. The objective is to collect state-level data on health-related behaviors to guide preventive health practices.

Following the terrorist attacks of September 11, 2001, New York state added 17 questions to their ongoing BRFSS to measure the behavioral and emotional effects of the attacks. During October, November, and December 2001, 1168 respondents of the total 3899-member year-long sample were asked these additional questions.

We restricted our analyses to those 1168 respondents asked questions about the terrorist attacks of September 11, 2001. We created an exposure variable based on the question, “Did you get help with problems you have experienced since the attack?” Control variables consisted of an individual’s age (imputed from date of birth), sex, marital status (married vs 5 nominal unmarried states), education (6 ordinal levels), and race (non-Hispanic white, non-Hispanic black, non-Hispanic other, non-Hispanic multiracial, Hispanic).

We compared and contrasted survey analytic techniques that use Taylor linearization methods to Bayesian hierarchical models for the prevalence of exposure and for the odds ratio (OR) for association of the downtown Manhattan exposure variable with the seeking help outcome variable controlling for age, sex, marital status, education, and race.
Survey Methods
We used SAS version 9.1.16 to conduct survey analyses. SAS uses the Taylor linearization method to account for complex survey designs and estimate variances. These procedures require the user to specify the first-level or primary sampling unit as well as any strata from which the primary sampling unit was drawn and any clustering variables. Publicly available BRFSS data sets clearly identify the primary sampling unit (essentially the individual chosen from a household to participate in a survey) and intrastate strata; there are generally no clustering variables.

A weight variable also must be specified for the SAS survey procedures. The weight for each individual in the sample is essentially the inverse of the selection probability for that individual and reflects poststratification adjustments, nonresponse, and unequal or oversampling. The sum of the weights for a survey sample reflects the population for which inferences are being made.

We adjusted the weights assigned to the 1168 post-9/11 respondents to reflect the total population of New York state by multiplying each individual by a weight of 3.343446. We arrived at this number by dividing the sum of the weights for the World Trade Center (WTC) attack respondents into the sum of the weights for the total sample (14,512,463.00/4,340,570.67). The sum of the adjusted WTC sample weights was 14,512,463.65 compared with the 14,512,463.00 sum of the weights for the entire year-long sample. We assessed the validity of extrapolating from the postattack respondents to the entire New York state population by comparing the pre- and postattack respondents through chi square tests for categorical variables and t tests for continuous variables.

Bayesian Methods
In a Bayesian approach, our 2 main sources of information about parameters of interest ($\theta$) are our prior beliefs or the prior distribution of the parameter (Pr[$\theta$]) and the likelihood of observing the data given the parameter (Pr[y | $\theta$]). Our prior distribution indicates how we believe the parameter would behave if we had no data upon which to base our judgments. The likelihood informs about $\theta$ via the data themselves. When we have a lot of data, the likelihood predominates, and our results will essentially be the maximum likelihood estimate. When we have fewer data, the prior distribution has greater influence.17,18 The result of combining the prior distribution and the likelihood is called the posterior distribution and follows Bayes theorem:

$$\text{Pr}[\theta | y] \propto \text{Pr}[y | \theta] \times \text{Pr}[\theta],$$

In a hierarchical (multilevel or mixed) model, we are interested in making inferences on some number of parameters ($\theta_1, \ldots, \theta_k$) measured on N units (eg, individuals, subsets, geographic areas, time periods, published studies) that are somehow related or connected by the nature of the problem. We could assume that there is 1 underlying parameter that arises from a single underlying population, in which case the units from which the data arose are uninformative. We could, at the opposite extreme, assume that each unit is a distinct population and should be analyzed separately, leading essentially to a series of descriptive analyses. If, however, we believe that knowing something about some of the units tells us something about the others (eg, data from 9 contiguous geographic areas tell us something about the 10th), then we may assume that the unit data are being drawn from some prior population with unknown parameters much like a random effects model in meta-analysis. At a spatial level, by specifying how the units vary across the population, we can help account for irregular groupings, autocorrelation, and the effects of extreme values in any particular unit.

In our analyses, we specified the subsets or geographic areas that constitute the N units as the New York counties from which samples were drawn. We assumed a normal likelihood for the data in each county $x_i \sim \text{N}(\theta_i, \sigma^2_{\text{e}_i}) i = 1, n_z, z = 1–20$ counties, and assumed $z$ distinct but related mean parameters, $\theta_i$. We took $\sigma^2_{\text{e}_i}$ to be unknown and assumed a vague gamma prior for the prevalence estimate and a more informative normal prior for the hierarchical logistic regression model based on maximum likelihood estimation runs of the data.

We used WinBUGS software19 to run 2 parallel Monte Carlo Markov chains with overdispersed initial values for 22,000 iterations. The first 2000 iterations were discarded as a burn-in, and our inferences were based on the final 20,000 iterations. We assessed convergence by examining trace histories and our inferences were based on the final 20,000 iterations. We assessed convergence by examining trace histories for parallel chains, and we used R software20 to conduct the Brooks, Gelman and Rubin, and Geweke convergence diagnostics, as well as the Heideberger and Welch stationarity test. We present these results at median values for the coefficients:

### TABLE 1

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Comparison of Pre–World Trade Center Attack to Post–World Trade Center Attack Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-WTC Sample, %</td>
</tr>
<tr>
<td>Married</td>
<td>47.9 (1308/2731)</td>
</tr>
<tr>
<td>Female</td>
<td>58.7 (1603/2731)</td>
</tr>
<tr>
<td>High school graduate</td>
<td>36.4 (995/2731)</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>70.74 (1932/2731)</td>
</tr>
</tbody>
</table>

Data from the New York state 2001 Behavioral Risk Factor Surveillance Survey.

Disaster Medicine and Public Health Preparedness

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
coefficients with their associated 95% equal-tailed Bayesian credible intervals and their kernel density graphs.

**Missing Data**

Geographic county of residence was missing for 258 of the 1168 (22.1%) records. WinBUGS software uses the same Gibbs sampling approach to simulate best guesses for missing values as it does for estimating the model, basing the guesses on the conditional distributions of all of the other model parameters. For the Bayesian analyses, records missing geographic data were coded to a separate, additional stratum. We based this decision on the assumption that for the demonstrative purposes of this analysis, coding them in this way allowed us to make greater use of the geographic units from which the data arose in a way that survey procedures did not, thus highlighting differences between the 2 approaches.

Eighty-four of 1168 (7.2%) records were missing data on exposure and were imputed to the more conservative “non-exposed” category. Race was missing from 22 records (1.9%) and was imputed to “white.” The 7 missing responses for marital status (0.6%) were imputed to “married.” Complete codes for the models are presented in the Appendix. The study was approved by the Columbia University institutional review board, reference number AAAB0209.

**RESULTS**

The 1168 (29.96% of total year-long sample) respondents who were asked questions about the World Trade Center in October, November, and December 2001 did not differ in any significant way from the 2731 who were sampled up to the time of the attacks. The 2 groups were similar for county of residence ($\chi^2 = 10.5668; df = 21; P = .9567$) and age (mean 45.0 pre-WTC, 44.9 post-WTC; $t = -0.15; P = .8786$).

Table 1 presents additional comparisons. Approximately 48% of respondents were married, 58% were female, 71% were white non-Hispanic, and about one third of the respondents were college graduates.

The results for exposure prevalence and the association of exposure with seeking help under each of the 2 approaches are presented in Table 2 and Figures 1 and 2. Estimates for the prevalence of exposure to the terrorist attacks of September 11, 2001, varied by the method chosen. Bayesian hierarchical modeling returned the lowest estimate (7.8%), with a credible interval spanning nearly 3 times the range of the CIs associated with both unadjusted and survey procedures. Approximately 15% (0.1502) of the variance in the Bayesian hierarchical modeling approach was related to geographic county (95% credible interval 0.0796–0.236).

Survey procedures returned a point estimate that was one half of a percentage point higher than the Bayesian estimate with a narrower CI. Unadjusted analyses returned the highest point estimate for prevalence at nearly 10%.

Bayesian approaches to survey data may offer potential advantages, particularly in the uncertain environment of postterrorism and disaster settings.

**DISCUSSION**

Accurate information in postdisaster settings is essential to guide relief efforts and public health interventions. It has been said that the effort “required to collect the information necessary to provide apt and well-directed aid is more than justified by the improved results.” Our results indicate that choice of statistical method may affect the interpretation of such information.

**TABLE 2**

<table>
<thead>
<tr>
<th>Comparison of Unadjusted, Survey, and Bayesian Methods for Prevalence of Exposure and Association of Exposure With Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prevalence Estimate (95% CI) for Exposure, %</strong></td>
</tr>
<tr>
<td>Bayesian hierarchical modeling</td>
</tr>
<tr>
<td>Survey</td>
</tr>
<tr>
<td>Unadjusted</td>
</tr>
</tbody>
</table>

Modeling for % of respondents indicating they were in “downtown Manhattan” during the September 11, 2001, terrorist attacks (column 1) and odds ratio (OR) results for association of this exposure with seeking “help for problems,” controlling for age, sex, race, education, and marital status (column 2). Data from the New York state 2001 Behavioral Risk Factor Surveillance Survey.
The implications of such differences may be significant. The 6.5% lower bound of the exposure prevalence obtained from a survey analysis approach translates to a minimum of nearly 1 million people in downtown Manhattan on the day of the terrorist attacks. By one estimate, on a typical workday the entire 1.5 million person residential population of Manhattan rises to 2.1 million; the 50,000 employees population of the WTC complex itself was augmented by 100,000 visitors. Consistent with this estimate, reports put the number of direct victims of this attack at 164,000. The 2.5% lower bound of the Bayesian credible interval for the prevalence of exposure translates to a minimum of 350,000 people in downtown Manhattan and appears more in line with the likely actual number of individuals exposed.

The wider credible interval from the Bayesian approach may better reflect the uncertainty of the prevalence estimates. By contrast, the tighter credible interval for the OR for association between exposure and seeking help illustrates both an advantage and a difficulty of a Bayesian approach: the choice of a prior distribution. We chose a noninformative (gamma) distribution for the prior in our prevalence calculations to highlight our prior assumption that such an estimate was fraught with uncertainty. For the OR calculation we assumed, based on prior studies of postdisaster pathology, a fairly informative normally distributed prior distribution. Because the posterior distribution in a Bayesian analysis is a weighted average of the likelihood and the prior distribution, placing a restrictive prior distribution on the analysis will necessarily return a tighter credible interval. Sensitivity analyses using less restrictive prior distributions may be helpful in analyzing postterrorism and disaster survey data.

The hierarchical Bayesian approach allowed us to take geography into account and (through the variance partition coefficient) estimate that only about 15% of the probability of exposure was due to geographic residence. This is in line with New York being a commuter city, and could very well be important information in planning and allocating public health resources. We did not calculate a variance partition coefficient for the noncontinuous, non-normal logit-based OR. Approximations exist but are analytically complex to implement and interpret. Sparse cells precluded a hierarchical spatial approach with survey techniques entirely.

Despite these potential advantages, the present study is at best an initial examination of Bayesian approaches to postdisaster population health and these techniques. The reliability and validity of survey methods such as Taylor linearization have been vetted over a long period of time and are widely accepted. As illustrated, Bayesian approaches require choices, particularly about appropriate prior distributions, that may
vary from investigator to investigator and that may mean-
ingfully affect the results. An additional example of the
use of Bayesian modeling in postdisaster settings may be
informative.

In contrast to the evidence on psychopathology, little consensus
exists about changes in substance use after mass traumas.25,26
Although some studies have shown an increase in substance use
and misuse after mass traumas,26,27 others have not.25 One study
used Bayesian hierarchical modeling to look at the relation
between residential proximity to the terrorist attacks of Septem-
ber 11, 2001, and substance use–related diagnoses in a popula-
tion of New York City Medicaid enrollees.28

The Bayesian hierarchical analysis drew inferences on the
statistical significance of model coefficients for the distance
in miles of New York City ZIPcode tabulation areas from the
WTC site, controlling for median household income, age,
sex, and race. Results indicated that prevalence did indeed vary,
but in a predictable fashion, with each 2-mi increment
away from the WTC site resulting in 18% more substance
abuse diagnoses in the latter part of 2001. This was in
contrast to 2000, when each 2-mi increment farther from the
WTC resulted in 11% fewer such diagnoses.

CONCLUSIONS
Much remains to be learned about Bayesian approaches to
survey data in general29 and postdisaster data in particular.
Although it offers potential advantages, the approach has yet
to be rigorously tested in the setting of postdisaster popula-
tion health. A Bayesian approach has been called well suited to
the “real-world complexities” of clinical medicine, in
which information is scarce, observations are clustered, and
results should be intuitive to nonstatisticians.30 The same
may be said for disaster public health.

Disaster care practitioners and planners should be aware that
statistical approach may affect point estimates and intervals.
Bayesian approaches to survey data may offer potential ad-
vantages, particularly in the uncertain environment of post-
terrorism and disaster settings, but cannot yet be recom-
mended unreservedly. Additional comparative analyses of
existing data will help guide our ability to use these tech-
niques for future incidents. The epidemiological information
obtained could help improve and inform disaster relief efforts.

APPENDIX: MODEL CODE

1. SAS Code for Exposure Prevalence

   proc Surveymeans Data=nybrfss Mean Clm;
   strata _ststr;
   cluster _psu;
   weight _wt;
   var Exposed; /* 0, 1, Indicator Variable */
   run;

2. Winbugs Code For Exposure Prevalence:

   model {
     for(i In 1:Nobs) {
       exp[i] dnorm(theta[zone[i]], tau.e) # likelihood for observed data
     }
     for(z in 1:Nzone) {
       theta[z] Dnorm(mu, Tau.z) # Zone-specific Means (random Effects)
     }
   # Priors On Random Effects Mean And Variance
   mu dnorm(0, 0.000001)
   tau.z Dgamma(0.001, 0.001)
   sigma2.z <-1/tau.z # random effects variance (between-zone variance of mean)
   tau.e Dgamma(0.001, 0.001)
   sigma2.e <-1/tau.e # residual error variance
   vpc <-Sigma2.z / (sigma2.z + Sigma2.e) # Variance Partition Coefficient
   }

3. SAS Code For Logistic Regression Model

   proc surveylogistic data=brfssdat.nybrfss_wtc;
   strata_ststr;
   cluster_psu;
   weight New_wt;
   class exposed(reference=last)
   sex(reference=first)
   educa(reference=first)
   married(reference=first)
   race(reference=first)
   /Param=reference;
   model probhelp (event=first) = exposed age sex educa married race;
   run;

4. WinBUGS Code for Logistic Regression Model

   Model
   # likelihood for observed data where y = ‘probhelp’
   { For(i In 1 : Nobs ) {
     y[i] dbern(p[i]) # outcome Bernoulli distributed
   }
# Specification Of Logistic Regression Model
\[
\text{logit}(p[i]) = \alpha_0 + \alpha_1 \exp[i] + \alpha_2 \text{sex}[i] + \\
\alpha_3 \text{age}[i] + \\
\alpha_4 \text{race}[i] + \alpha_5 \text{married}[i] + \alpha_6 \text{educ}[i] + \\
b[cty[i]] \] # term for county-level random effects
\]

for (j in 1:Ncty){
\[b[j] \sim \text{Dnorm}(0.0, \tau_b)\]
}

# Priors For Coefficients
\[\alpha_0 \sim \text{dnorm}(-2.5, 0.8)\]
\[\alpha_1 \sim \text{Dnorm}(0, 0.001)\]
\[\alpha_2 \sim \text{dnorm}(0.7, 0.06)\]
\[\alpha_3 \sim \text{dnorm}(-0.01, 0.0001)\]
\[\alpha_4 \sim \text{dnorm}(0.1, 0.07)\]
\[\alpha_5 \sim \text{Dnorm}(0.3, 0.04)\]
\[\alpha_6 \sim \text{dnorm}(0.5, 0.5)\]

# Prior For Random Effects Term
\[\tau_b \sim \text{dunif}(0.01, 1.0)\]

# Variance Random Effect Term
\[\sigma^2_{b} < -1 / \tau_b\]

### About the Authors

Dr DiMaggio is with the Department of Epidemiology and Dr Abramson is with the National Center for Disaster Preparedness, Columbia University Mailman School of Public Health; and Dr Galea is with the Department of Epidemiology and Population Health, University of Michigan.


### Authors’ Disclosures

The study was supported by Centers for Disease Control and Prevention Health Protection Research Initiative Grant K01 CE000494.

The authors report no conflicts of interest.

### REFERENCES

2. Norris FH, Friedman MJ, Watson PJ. 60,000 disaster victims speak: II.

