



Published in final edited form as:

J Health Serv Res Policy. 2009 October ; 14(4): 226–233. doi:10.1258/jhsrp.2009.008186.

A Bayesian method for the synthesis of evidence from qualitative and quantitative reports: the example of antiretroviral medication adherence

Corrine I Voils, PhD [Research Health Science Specialist and Associate Professor]

Health Services Research and Development Service, Veteran Affairs Medical Center (152), 508 Fulton Street, Durham, NC 27705, USA

Department of Medicine, Duke University Medical Center, Durham

Vic Hasselblad, PhD [Professor of Biostatistics]¹, Jamie L Crandell, PhD [Assistant Professor]², YunKyung Chang, PhD [Social Research Specialist]², EunJeong Lee, MPH [Doctoral Student]², and Margarete Sandelowski, PhD [Boshamer Distinguished Professor]²

¹Department of Medicine, Duke University Medical Center, Durham

²University of North Carolina at Chapel Hill School of Nursing, Chapel Hill, North Carolina, USA

Abstract

Objectives—Bayesian meta-analysis is a frequently cited but very little-used method for synthesizing qualitative and quantitative research findings. The only example published to date used qualitative data to generate an informative prior probability and quantitative data to generate the likelihood. We developed a method to incorporate both qualitative and quantitative evidence in the likelihood in a Bayesian synthesis of evidence about the relationship between regimen complexity and medication adherence.

Methods—Data were from 11 qualitative and six quantitative studies. We updated two different non-informative prior distributions with qualitative and quantitative findings to find the posterior distribution for the probabilities that a more complex regimen was associated with lower adherence and that a less complex regimen was associated with greater adherence.

Results—The posterior mode for the qualitative findings regarding more complex regimen and lesser adherence (using the uniform prior with Jeffreys' prior yielding highly similar estimates) was 0.588 (95% credible set limits 0.519, 0.663) and for the quantitative findings was 0.224 (0.203, 0.245); due to non-overlapping credible sets, we did not combine them. The posterior mode for the qualitative findings regarding less complex regimen and greater adherence was 0.288 (0.214, 0.441) and for the quantitative findings was 0.272 (0.118, 0.437); the combined estimate was 0.299 (0.267, 0.334).

Conclusions—The utility of Bayesian methods for synthesizing qualitative and quantitative research findings at the participant level may depend on the nature of the relationship being synthesized and on how well the findings are represented in the individual reports.

Introduction

Among the most frequently cited but as yet very little-used methods for synthesizing qualitative and quantitative research findings is Bayesian meta-analysis.¹⁻³ A Bayesian analysis begins with a prior distribution describing plausible potential values for parameters that will be estimated. This distribution may be informative, taking account of previous data or expert beliefs, or it may be non-informative, allowing equitably for any of a wide range of parameter values to be true. The distribution of the observed data is then described by writing the likelihood as a function of parameter values. In the final step, the prior distribution and likelihood distribution are multiplied and normalized to create a posterior distribution for each parameter. The mean, median or mode of the posterior distribution can be taken as the point estimate of the parameter, and credible set limits (a Bayesian analog to a confidence interval) can be calculated to describe the uncertainty in the estimate.

The only paper referred to as an example of Bayesian meta-analysis of both qualitative and quantitative evidence and, to our knowledge, the only paper to date showing the use and results of this synthesis method is a report by Roberts and colleagues⁴ of a study of factors affecting the uptake of childhood immunizations.

A quick perusal of CINAHL (Cumulative Index of Nursing and Allied Health Literature), Google Scholar, Thompson ISI Web of Science, PsycInfo, PubMed and ScienceDirect shows that of the 44 unique references to this article (as of 19 November 2008), 73% were solely to identify it as an example of Bayesian meta-analysis and the remainder, to support an argument concerning childhood immunization. We know of no-one other than Roberts *et al.* who has used a Bayesian approach to synthesize qualitative and quantitative findings, nor anyone who critically analysed what Roberts *et al.* did. In this article, we first examine the Roberts *et al.* paper, then we describe the results of a Bayesian analysis we conducted of qualitative and quantitative findings on antiretroviral adherence and how this analysis diverged from the one Roberts *et al.* described.

The Roberts *et al.* report

The stated goal of the Roberts *et al.*⁴ study was to identify the factors that predict uptake of childhood immunizations. To begin, the investigators listed, ranked and weighed factors they believed affected uptake of childhood immunizations. They then reviewed 11 qualitative studies regarding factors that affect uptake and updated their beliefs based on that evidence. Their updated beliefs were combined to estimate the prior probability that each factor was associated with immunization uptake. The quantitative evidence was then used to generate the likelihood, which updated their expert beliefs to create a posterior probability.

Their analysis had two characteristics that we directly address in this article. First, the qualitative and quantitative findings were not treated equally. The qualitative evidence was used to refine the prior probabilities, but only the quantitative evidence was considered to constitute the data in the Bayesian analysis, and it alone contributed to the likelihood. This is only one possible way of using the Bayesian paradigm to synthesize qualitative and quantitative findings. Data from both qualitative and quantitative studies could enter the analysis in the same phase. Second, Roberts *et al.* were not able to construct credible intervals quantifying the uncertainty in their probability estimates (although their approach could be expanded to include credible intervals by assigning a distribution to the probabilities corresponding to each factor and/or using a different estimation method, such as Markov Chain Monte Carlo). The estimate of the probability of an effect is more meaningful when a measure of the variability of the estimate is provided.

Our goal was to apply a Bayesian method for synthesis of qualitative and quantitative findings that allowed the qualitative and quantitative evidence to affect the results in the same way, and that produced a measure of uncertainty in the estimate of the probability of an effect. Due to the complexity of the method, we synthesized the findings on only the effect of one factor and its converse for illustrative purposes. We show how we applied Bayesian methods to synthesize a set of qualitative and quantitative research findings on the relationship between regimen complexity and adherence to antiretroviral medications in HIV-positive women.

Methods

Data for the synthesis

The data for the analysis consisted of qualitative and quantitative reports of empirical studies of antiretroviral adherence in HIV-positive women of any race/ethnicity, class or nationality living in the USA. These reports are part of a larger collection retrieved in the course of a study aimed at developing methods for synthesizing qualitative and quantitative research findings. They were retrieved between June 2005 and January 2006, with updates in June 2006 and October 2007, using 40 databases housing citations to literature across the health, behavioral and social sciences. Reports were eligible for inclusion if they were of qualitative and quantitative observational studies of adherence to antiretroviral therapy conducted with HIV-positive women and published in or after 1997 (when protease inhibitors became part of the standard of care) to 2007. The research question asked of this set of reports was: ‘Which factors are (positively or negatively) associated with medication adherence?’

For the purposes of clearly demonstrating our application of Bayesian methods for combining qualitative and quantitative findings, we focused on a set of findings pertaining to a relationship addressed in both qualitative and quantitative studies, namely, the association between adherence and regimen complexity. This relationship was the most frequently addressed in both types of reports: 11 qualitative and 10 quantitative reports. Common to all of these studies was the reduction of the interview or focus group data collected in ways that remained close to those data as given to the researcher, that is, with minimal interpretation or transformation, or what Sandelowski and Barroso referred to as qualitative survey findings.⁵ The quantitative studies reviewed included observational studies of factors associated with adherence.

Regimen complexity is a broad category of factors commonly appearing in systematic reviews of research on antiretroviral adherence and consisting of an array of diverse factors. In systematic reviews, it is usually necessary to combine factors seen to be similar and, therefore, treatable as the same because otherwise, there would be few findings available to synthesize, no matter what the method of synthesis used. We included in the medication regimen category any finding addressing dosing frequency, size of pills, timing of medications, availability of medication refills, medication side-effects, ease or difficulty of incorporation of pill-taking into daily routine, dietary requirements of drugs and regimen changes.

Determining the prior and posterior distributions

Our first step was to determine what would constitute the prior distribution. Prior distributions can be broadly classified into two types: informative and non-informative. Informative prior distributions are derived from a source, such as previous data or expert opinion. Roberts *et al.*⁴ used the qualitative findings, combined with their own expert opinions, to generate a prior probability. In contrast, we wanted to treat the qualitative research findings as evidence contributing to the posterior probability, which is more in keeping with the calls to give qualitative findings equal weight with quantitative findings in evidence-based syntheses. We could have sought expert opinion to help create an informative prior distribution, but to allow for the posterior distributions to be dominated by the likelihoods of the qualitative and

quantitative evidence, we used non-informative prior distributions. Such distributions have no population basis and, as such, have a minimal effect on the posterior distribution, which is informed primarily by the observed data used to construct the likelihood. There are different possibilities for prior distributions. For example, the uniform prior is a perfectly non-informative prior, giving every possible outcome the exact same prior probability. Frequently, the uniform prior is undesirable, so Jeffreys' prior⁶ is used because its mathematical properties make it fairly non-informative.⁷ As explained further below, we conducted the analyses both ways – using the uniform and Jeffreys' prior distributions – to ensure that our choice had no measurable effect on the posterior point estimate.

Deciding how to synthesize the findings

Before the qualitative and quantitative findings were to be combined to update the prior and generate the posterior distribution, we had to find a common metric. This proved challenging, as qualitative findings are usually represented at the study level, with thematic and interpretive lines typically prevailing over frequency counts, and within-participant or between-thematic lines comparisons typically prevailing over between- and cross-participant comparisons. Partly a consequence of open-ended and minimally structured modes of data collection resulting in non-comparable data-sets, in reports of qualitative studies, no one may have mentioned an association (which could be coded as 0), or ≥ 1 participants may have mentioned the association (which could be coded as 1). In contrast, quantitative findings are represented as group-level statistics (e.g. odds ratio) based on participant-level information.

Given these differences, we broadly defined two options for finding a common metric whenever the qualitative and quantitative findings will be used to calculate the likelihood. One option is to make the quantitative findings more like the qualitative findings by summarizing them at the study level. Thus, we would be estimating the probability that a study was linked to a finding. One way to do this might be to assign a value of 0 or 1 to each study to indicate whether the finding was absent or present, respectively. This could be accomplished by using a predetermined p value or odds ratio as a threshold. Thus, for example, if the odds ratio in a study were greater than the predetermined threshold, the study would receive a value of 1; otherwise, it would receive a 0. The findings could then be combined with the study-level qualitative findings. This solution is limited because it does not reflect the proportion of participants for whom, for example, a more complex regimen was associated with lesser adherence, thus resulting in a loss of information.

The alternative option is to make the qualitative findings more like the quantitative findings by summarizing them at the participant level. Thus, we would be estimating the probability that a participant was linked to a finding. This could be accomplished by extracting from the qualitative reports all information pertaining to the frequency of a finding and sample size and then using that information to estimate the proportion of participants for whom the finding was true. The corresponding proportion could then be calculated for the quantitative studies.

From an initial reading of the qualitative and quantitative reports, we found that there were two possible associations to summarize: more complex regimen associated with lesser adherence, or less complex regimen associated with greater adherence. In the quantitative reports, these two categories were mutually exclusive (i.e. statistically independent) because each participant was categorized as more adherent or less adherent in relation to a more complex regimen or a less complex regimen. In contrast, in the qualitative reports, participants fell into one, both, or neither category. For example, a woman reporting that she did not adhere because the regimen was too complex could not be assumed to be adherent with a less complex regimen. A woman who reported decreased adherence after a switch to a more complex regimen would fit into both categories.

If the categories had been statistically independent (i.e. represented by different participants) for both the qualitative and quantitative data, then we could have analysed them together. Because of the nature of the qualitative data in the studies reviewed, however, we could only synthesize findings concerning one component of the association between regimen complexity and adherence at a time because otherwise we would be counting some participants twice. Thus, we synthesized separately the association between a more complex regimen and lesser adherence, and the association between a less complex regimen and greater adherence.

Calculating the proportion of participants associated with a finding

Qualitative findings—While reviewing the qualitative reports, it became apparent to us that determining the number of participants indicating that a more complex regimen was associated with lesser adherence or that a less complex regimen was associated with greater adherence would be difficult, as the findings in these reports were typically presented with verbal count adjectives such as ‘few’, ‘many’ and ‘most’. As described more fully elsewhere,⁸ this is not a deficiency of qualitative research *per se*, but rather a reflection of the differences between the purposeful sampling and analytic imperatives associated with qualitative research and the probability sampling and analytic imperatives associated with quantitative research. Moreover, this style of reporting findings is characteristic of qualitative survey research, which characterizes the qualitative studies in our data-set. Because qualitative survey findings are similar in form and interpretive level to quantitative findings, they are more amenable to quantitative transformation than highly interpretive qualitative findings.⁹ We, therefore, developed a method to transform these verbal counts into numbers. The details of this method are fully described elsewhere.⁸ Briefly, we used the results of a quantitative survey we conducted to ascertain which specific numbers, or ranges of numbers, respondents had in mind when using seven verbal count terms (couple, few, majority, many, most, several and some) for each of four sample sizes common in qualitative studies: 5, 10, 20 and 50. For each sample size, we constructed a range for each of the seven words using the smallest and largest number reported. To avoid the influence of outliers on these ranges, if the smallest or largest number for a word was reported by only one respondent, we discarded that number and used the second-smallest or the second-largest reported value. To apply the results of the survey to the actual sample sizes in the 11 qualitative reports, we estimated the mathematical relationship between sample size and the number represented by each word by fitting a regression equation for the lower and upper limit of the range for each of the seven words. Table 1 shows the results of this method for both findings for each of the 11 qualitative reports.

Quantitative findings—To calculate the proportion of participants associated with each finding in the quantitative reports, we extracted all bivariate (unadjusted) findings, as it is impossible to calculate the percentage of people reporting a finding from multivariable (adjusted) results. We excluded two reports for which the outcome was mean adherence, as there was no way to calculate percentages, and two other reports that did not provide sufficient information for calculating percentages, even after contacting the authors. As a result of these exclusions, we were able to calculate percentages for 6 of the 10 quantitative reports (Table 2). Some of these reports contain findings from more than one variable that could be classified as medication regimen. To meet the independence assumption (i.e. participants may contribute to only one finding) for any quantitative synthesis method, we extracted the variable from each report that was most similar to the variables from other reports. In five of the six reports that we utilized, that variable was regimen type (HAART versus pre-HAART, or combination therapy versus monotherapy); in the remaining study, the variable was having too many pills to take.

Analysis

We applied Bayesian methods to determine the probabilities of a participant reporting that a more complex medication regimen was associated with lesser adherence and of a participant reporting that a less complex regimen was associated with greater adherence. The following methods were used to calculate each probability separately. First, for each finding, we synthesized the qualitative and quantitative findings individually. As they are typically classified as different methods, all of the qualitative studies could be seen as estimating one probability, and all of the quantitative studies could be seen as estimating another probability. We do not mean to invoke stereotypes of qualitative and quantitative research, as the dichotomy is often false. Rather, we argue that researchers should examine each set of studies critically to determine whether and how the two types of studies differ. If both types of studies estimate the same probability, the posterior distributions will be similar and we can combine them to estimate the posterior probability. If the posterior distributions are very different (as indicated by nonoverlapping 95% credible sets), this may be an indication that the two study types are not estimating the same probability and, arguably, that the distributions should not be combined.

We analysed the data using two different prior distributions: (1) the uniform prior, allowing all probabilities of an effect to be equally likely; and (2) the standard Jeffreys' prior⁶ (Figure 1). We wrote the likelihood for each study using the observed data, then combined the prior distribution and the likelihood to form a posterior distribution. The likelihoods were combined by multiplying them together using the Mathematica software.¹⁰

Each quantitative study can be thought of as a draw from a binomial distribution, where the possible number of 'successes' (i.e. times the effect is observed) is equal to the total sample size. The probability that a participant is linked to the finding is represented by probability p . To test whether the quantitative studies all had the same underlying value of p , we calculated an index of heterogeneity. For the association between more complex regimen and lesser adherence, Cochran's Q was $\chi^2(5, n = 6) = 17.846$, $p = 0.003$, and for the association between less complex regimen and greater adherence, it was $\chi^2(5, n = 6) = 138.36$, $p < 0.001$. This suggests the presence of heterogeneity and that a fixed effects model (i.e. a single value of p) is inappropriate. Therefore, we used the beta-binomial random effects model.¹¹

In this model, the parameter, p , is randomly varying across studies according to a beta distribution with parameters α, β . The resulting distribution is no longer called a binomial distribution, but rather a beta-binomial distribution:

$$P[x=j] = \binom{N}{j} \frac{B(\alpha+j, \beta+N-j)}{B(\alpha, \beta)}, \quad j=0, 1, 2, \dots, N,$$

where $B(\alpha, \beta)$ is the beta function.¹² Equivalently, the distribution can be written as:

$$P[x=j] = \frac{1}{C} \frac{\alpha(\alpha+1)\dots(\alpha+j-1)}{j!} \times \frac{\beta(\beta+1)\dots(\beta+N-j-1)}{(N-j)!}, \quad j=0, 1, 2, \dots, N,$$

$$\text{where } C = \binom{\alpha+\beta+N-1}{N}.$$

The parameter estimates were calculated using the modified Gauss-Newton method.¹³ The posterior mean and 95% credible set limits were computed using the FAST*PRO Software.¹⁴

Results

Association between a more complex medication regimen and lesser adherence

The resulting posterior probability for the qualitative studies is shown on the right side of Figure 2 (in all cases, the same shape results, regardless of which prior was used). Using the uniform prior, the posterior mode for the probability of adherence in the qualitative studies was 0.588 (95% credible set limits 0.519, 0.663). Using Jeffreys' prior changed the estimate very slightly, with a posterior mode of 0.589 (0.519, 0.664). Further examination of the results revealed that the following studies had a minimal effect on the posterior mode (i.e. the credible set limits and posterior mode remain the same to 3 decimal places): Misener and Sowell,¹⁵ Remien *et al.*,¹⁶ Richter *et al.*,¹⁷ Siegel and Gorey,¹⁸ Siegel *et al.*,¹⁹ and Wood *et al.*²⁰ This is because the estimated ranges were wide relative to the study n , leading to a large amount of imprecision.

The resulting posterior probability for the quantitative studies is shown on the left side of Figure 2. Using the uniform prior, the posterior mode for the probability of adherence in the quantitative studies was 0.224 (0.203, 0.245). Using Jeffreys' prior, the posterior mode was 0.222 (0.182, 0.268).

Because the point estimates for the qualitative and quantitative studies were vastly different, with non-overlapping credible sets, it is possible that each set of findings is estimating a different probability. Therefore, we did not combine them into a single point estimate, and no statistically meaningful conclusion could be drawn about the relationship between a more complex regimen and lesser adherence.

Association between a less complex medication regimen and greater adherence

The resulting posterior probability for the qualitative studies is shown in Figure 3. Using the uniform prior, the posterior mode for the probability of adherence in the qualitative studies was 0.288 (0.214, 0.441). Using Jeffreys' prior changed the estimate very slightly, with a posterior mode of 0.272 (0.118, 0.437). Using the uniform prior, the posterior mode for the probability of adherence in the quantitative studies was 0.297 (0.258, 0.339). Using Jeffreys' prior, the posterior mode was 0.299 (0.260, 0.341).

Because the point estimates for the qualitative and quantitative studies were very similar, with overlapping credible sets, we combined them into a single point estimate. Using the uniform prior, the posterior mode was 0.300 (0.267, 0.334). Using Jeffreys' prior, the posterior mode was 0.299 (0.267, 0.334). As shown in Figure 3, in this example, the combined posterior is more similar to the posterior from the quantitative studies. This is because the posterior from the qualitative studies is relatively flat over the region where the posterior from the quantitative studies has appreciable mass.

Discussion

To our knowledge, Roberts and colleagues⁴ were the first to demonstrate the use of Bayesian methods to synthesize qualitative and quantitative research findings. Our approach differed from theirs in a few key ways. For one, Roberts *et al.* used the qualitative research findings to inform the prior probability, which was then updated with quantitative evidence to create a posterior probability. In contrast, we used two different non-informative prior distributions, with the goal of updating them with the qualitative and quantitative research findings. Our approach also differed from Roberts *et al.* in that we calculated 95% credible set limits to indicate the degree of certainty about the point estimates (although this could be added easily

to their approach). Finally, we synthesized findings of two relationships to determine if this affected our ability to synthesize findings or the estimated posterior probabilities.

Our conclusions differed according to which relationship was synthesized. When synthesizing the relationship between a more complex regimen and lesser adherence, the qualitative data yielded a much greater estimate of the probability than did the quantitative data, and the two estimates had non-overlapping 95% credible sets. Moreover, the distribution for the quantitative findings was narrower and more precise (i.e. less variability) than the distribution for the qualitative findings, likely owing to the use of ranges rather than precise probability estimates. Although combining the two distributions into a mixture distribution is mathematically possible, the resulting estimate may not be meaningful because the studies were drawn from two different, known populations of studies (i.e. qualitative and quantitative). The probabilities should only be combined if they were assumed to be drawn from a single population of studies.

In contrast, when synthesizing the relationship between a less complex regimen and greater adherence, the posterior modes were very similar, with overlapping credible sets; therefore, we were able to combine the estimates into an overall probability estimate. As before, wider 95% credible sets were obtained for the qualitative, in comparison to the quantitative, estimates.

Another way of viewing these findings is that the posterior modes for the two complementary quantitative relationships were similar, whereas for the qualitative relationships, they were quite different. Inspection of columns 3 and 4 of Table 1 indicates that the derived ranges were wider, and, in some cases, unknown (thus not contributing to the synthesis), for the relationship between less complex regimen and greater adherence than for the relationship between more complex regimen and lesser adherence. This is sensible because in many studies, participants were asked what made it difficult to adhere to the regimen, as opposed to what made it easy or a non-valence question about adherence in general. Even subtle differences in the way questions are worded, to obtain information about ostensibly the same phenomenon, has long been understood to have a profound impact on responses.²¹ Moreover, it cannot be assumed that asking an individual what makes it hard to adhere yields information about what makes it easy; they are not symmetrically opposite questions.²²

One limitation of our approach is that the posterior probability estimates for the qualitative studies may be inaccurate due to the method we used to calculate the range estimates. Due to the way that the qualitative findings were reported, we did not have information on the number of participants for each finding, which led us to the effort to calculate ranges. Although obtaining estimates is not the goal of qualitative research, in order to synthesize qualitative and quantitative research findings mathematically, this information will sometimes be needed. Our calculations were contingent on authors choosing words that have shared meaning. For example, if the authors reported that ‘many’ participants reported a finding to denote that 20–30% did, but ‘many’ is commonly understood as >50%, as we found previously,⁸ then the ranges and, thus, the probabilities that we calculated will be inflated. Although the ranges provided us with more precision than would have been available otherwise (i.e. the only information provided was that the finding was true for 1 to n), they were still relatively large. This contributed to wider credible intervals for the qualitative findings. This method for estimating ranges and calculating probabilities is only possible when the qualitative data are presented as qualitative survey findings, which are largely untransformed and remain close to data as given. More interpretive qualitative findings (e.g. in the form of grounded theories or ethnographic explanations) would not lend themselves to the synthesis method presented here because their goal is not to generate lists of ideas or occurrences that are, in turn, more amenable to counting.

Another limitation is that the regimen complexity category was defined differently in qualitative and quantitative studies. We chose regimen complexity because it was available in (some of) both types of reports. Within this category, we were only able to synthesize some of the quantitative findings because of differences in measurement (continuous vs. dichotomous outcome) and because some reports contained insufficient information for calculating proportions. Nonetheless, in five of the six quantitative reports, the same variable was used and thus synthesized (HAART vs. non-HAART regimen). In contrast, in the qualitative reports, several different aspects of regimen complexity were mentioned, even within a single study. For example, in a case in which the minimum n was 2, one person may have mentioned pill size whereas the second person may have mentioned availability of medication refills. Moreover, the most frequently mentioned aspect of regimen complexity in the qualitative reports was side-effects. Thus, the components of regimen complexity were different for the qualitative and quantitative studies. This limitation will be difficult to circumvent when conducting a synthesis of qualitative and quantitative findings. As mentioned previously, combining factors that can be seen as similar is usually necessary in any effort to synthesize findings, otherwise there would be too few findings to synthesize. If we were only willing to combine factors that were defined and operationalized the same way across studies, we would not have been able to combine any association.

Although Bayesian methods have been promoted for synthesizing qualitative and quantitative research findings, they may have limited utility.³ At issue is whether both types of research findings can be treated as evidence, factoring into the calculation of the likelihood, and whether the findings can be represented at the participant level and/or the study level. Our attempt to synthesize the relationship between regimen complexity and adherence at the participant level revealed that the conclusion may depend on the nature of the relationship and on the ability to obtain precise frequencies associated with each finding. Future research should investigate whether Bayesian methods to synthesize qualitative and quantitative research findings at the study level (i.e. finding present or absent in each study) prove more fruitful because this information is readily available in both types of reports. Indeed, we have taken this approach in a different analysis, but because of the amount of detail required to describe it, we will report our results elsewhere.

Acknowledgements

The method study referred to here, 'Integrating qualitative and quantitative research findings', is funded by the National Institute of Nursing Research, National Institutes of Health (5R01NR004907, 3 June 2005–31 March 2010). This material is the result of work also supported with resources and the use of facilities at the Durham Veterans Affairs Medical Center. The views expressed in this article are those of the authors and do not necessarily represent the Department of Veterans Affairs.

References

1. Dixon-Woods M, Agarwal S, Jones D, et al. Synthesising qualitative and quantitative evidence: A review of possible methods. *J Health Serv Res Policy* 2005;10:45–53. [PubMed: 15667704]
2. Mays N, Pope C, Popay J. Systematically reviewing qualitative and quantitative evidence to inform management and policy-making in the health field. *J Health Serv Res Policy* 2005;10(Suppl 1):6–20. [PubMed: 16053580]
3. Boaz A, Ashby D, Denyer D, et al. A multitude of syntheses: A comparison of five approaches from diverse policy fields. *Evidence & Policy* 2008;2:479–502.
4. Roberts KA, Dixon-Woods M, Fitzpatrick R, et al. Factors affecting uptake of childhood immunisation: A Bayesian synthesis of qualitative and quantitative evidence. *Lancet* 2002;360:1596–9. [PubMed: 12443615]
5. Sandelowski, M.; Barroso, J. *Handbook for Synthesizing Qualitative Research*. Springer; New York, NY: 2007.

6. Jeffreys, H. *Theory of Probability*. Oxford Press; New York, NY: 1961.
7. Gelman, A.; Carlin, J.; Stern, H., et al. *Bayesian Data Analysis*. Vol. 2nd edn.. Chapman & Hall/CRC; Boca Raton, FL: 2004.
8. Chang Y, Voils C, Sandelowski M, et al. Transforming verbal counts in reports of qualitative descriptive studies into numbers. *West J Nurs Res*. in press
9. Sandelowski M, Barroso J, Voils CI. Using qualitative metasummary to synthesize qualitative and quantitative descriptive findings. *Res Nurs Health* 2007;30:99–111. [PubMed: 17243111]
10. Wolfram, S. *Mathematica*. Wolfram Research, Inc.; Champaign, IL: 1996.
11. Kleinman J. Proportions with extraneous variance: Single and independent samples. *J Am Stat Assoc* 1973;68:46–54.
12. Abramowitz, M.; Stegun, I. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications; New York, NY: 1972.
13. Hartley H. The modified Gauss Newton method for the fitting of nonlinear regression functions by least squares. *Technometrics* 1961;3:269–80.
14. Eddy, D.; Hasselblad, V. *Fast*Pro: Software for meta-analysis by the confidence profile method*. Academic Press; San Diego, CA: 1992.
15. Misener TR, Sowell RL. HIV-infected women's decisions to take antiretrovirals. *West J Nurs Res* 1998;20:431–47. [PubMed: 9686522]
16. Remien RH, Hirky AE, Johnson MO, et al. Adherence to medication treatment: A qualitative study of facilitators and barriers among a diverse sample of HIV + men and women in four US cities. *AIDS Behav* 2003;7:61–72. [PubMed: 14534391]
17. Richter DL, Sowell RL, Pluto DM. Attitudes toward antiretroviral therapy among African American women. *Am J Health Behav* 2002;26:25–33. [PubMed: 11795602]
18. Siegel K, Gorey E. HIV-infected women: Barriers to AZT use. *Soc Sci Med* 1997;45:15–22. [PubMed: 9203266]
19. Siegel K, Lekas HM, Schrimshaw EW, et al. Factors associated with HIV-infected women's use or intention to use AZT during pregnancy. *AIDS Educ Prev* 2001;13:189–206. [PubMed: 11459356]
20. Wood SA, Tobias C, McCree J. Medication adherence for HIV positive women caring for children: In their own words. *AIDS Care* 2004;16:909–13. [PubMed: 15385246]
21. Willis, G. *Cognitive interviewing: A tool for improving questionnaire design*. Sage; Thousand Oaks, CA: 2005.
22. Holleman B. The meanings of 'yes' and 'no'. An explanation for the forbid/allow asymmetry. *Qual Quant* 2006;40:1–38.
23. Abel E, Painter L. Factors that influence adherence to HIV medications: perceptions of women and health care providers. *J Assoc Nurses AIDS Care* 2003;14:61–9. [PubMed: 12953613]
24. Gant L, Welch L. Voices less heard: HIV-positive African American women, medication adherence, sexual abuse, and self-care. *J HIV AIDS Soc Serv* 2004;3:67–91.
25. Powell-Cope GM, White J, Henkelman EJ, et al. Qualitative and quantitative assessments of HAART adherence of substance-abusing women. *AIDS Care* 2003;15:239–49. [PubMed: 12856345]
26. Roberts KJ, Mann T. Barriers to antiretroviral medication adherence in HIV-infected women. *AIDS Care* 2000;12:377–86. [PubMed: 11091770]
27. Schrimshaw EW, Siegel K, Lekas HM. Changes in attitudes toward antiviral medication: a comparison of women living with HIV/AIDS in the pre-HAART and HAART eras. *AIDS Behav* 2005;9:267–79. [PubMed: 16088368]
28. Durante AJ, Bova CA, Fennie KP, et al. Home-based study of anti-HIV drug regimen adherence among HIV-infected women: feasibility and preliminary results. *AIDS Care* 2003;15:103–15. [PubMed: 12655838]
29. Phillips KD, Moneyham L, Murdaugh C, et al. Sleep disturbance and depression as barriers to adherence. *Clin Nurs Res* 2005;14:273–93. [PubMed: 15995155]
30. Schuman P, Ohmit SE, Cohen M, et al. Prescription of and adherence to antiretroviral therapy among women with AIDS. *AIDS Behav* 2001;5:371–8.

31. Stone VE, Hogan JW, Schuman P, et al. Antiretroviral regimen complexity, self-reported adherence, and HIV patients' understanding of their regimens: survey of women in the HER study. *J Acquir Immune Defic Syndrome* 2001;28:124–31.
32. Wilson TE, Barron Y, Cohen M, et al. Adherence to antiretroviral therapy and its association with sexual behaviour in a national sample of women with human immunodeficiency virus. *Clin Infect Dis* 2002;34:529–34. [PubMed: 11797182]
33. Wilson TE, Ickovics JR, Fernandez MI, et al. Self-reported zidovudine adherence among pregnant women with human immunodeficiency virus infection in four US states. *Am J Obstet Gynecol* 2001;184:1235–40. [PubMed: 11349194]

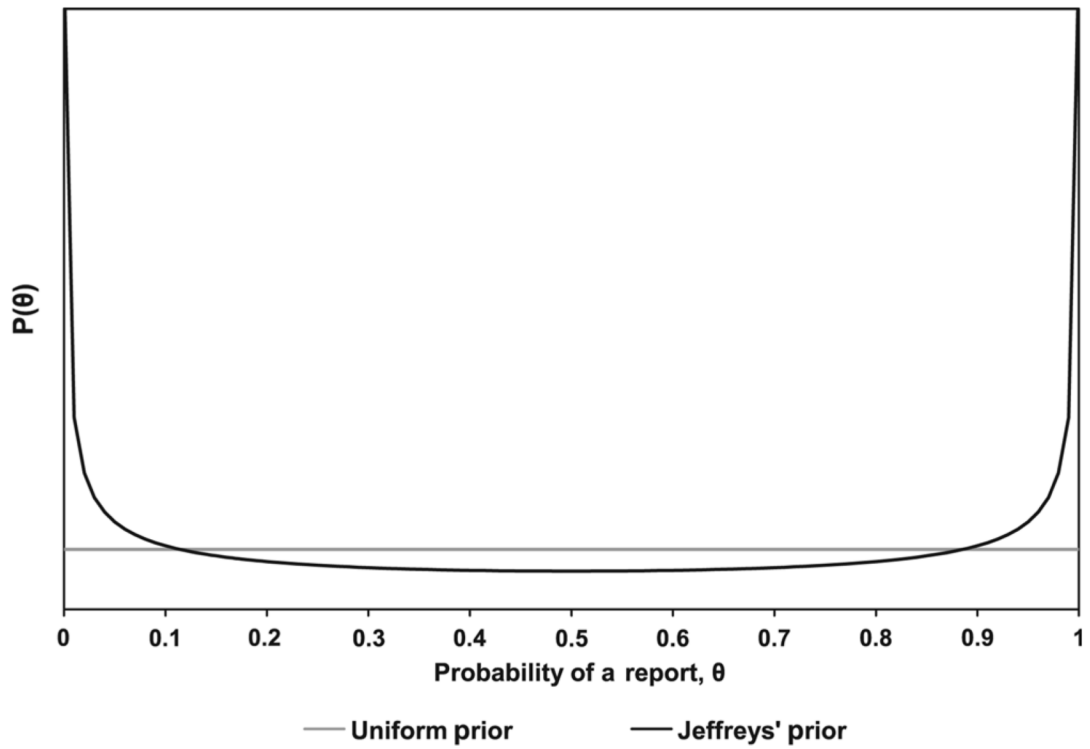


Figure 1.
Prior distribution of the probability of a participant reporting a relationship

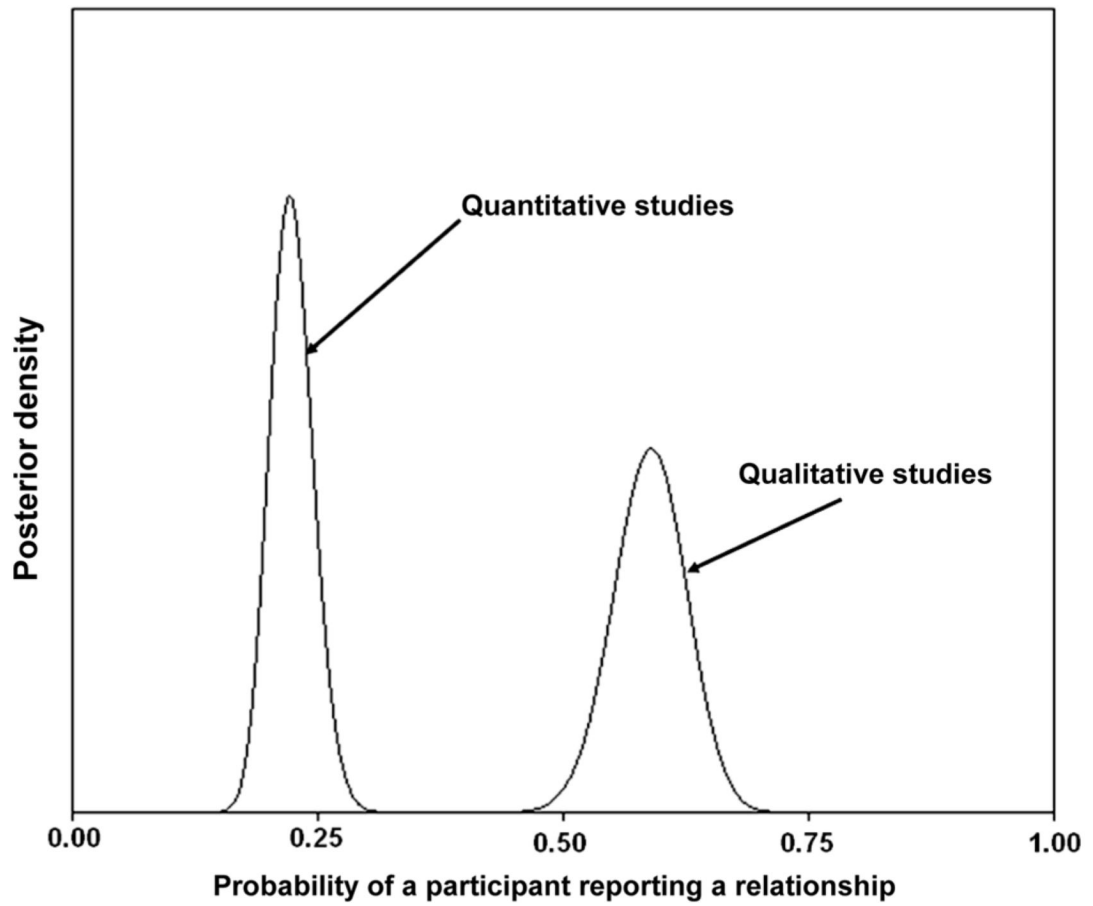


Figure 2.
Posterior distribution of the probability of a participant reporting the relationship between a more complex regimen and lesser adherence

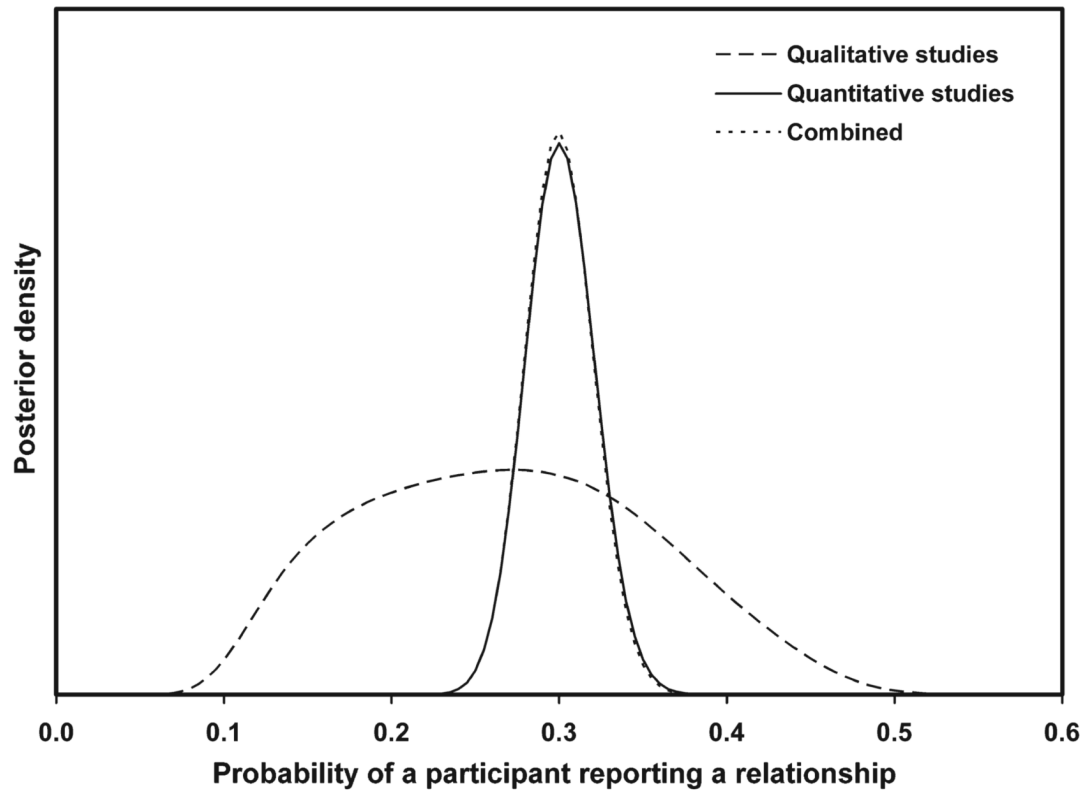


Figure 3. Posterior distribution of the probability of a participant reporting the relationship between a less complex regimen and greater adherence

Table 1

Findings from 11 qualitative reports on the relationship between a more complex medication regimen and lesser adherence and a less complex regimen and greater adherence

Report	Total study (n)	Calculated range of findings in more complex regimen/ lesser adherence*	Calculated range of findings in less complex regimen/ greater adherence*
Abel & Painter, 2003 ²³	6	2–6	1–6
Gant and Welch, 2004 ²⁴	30	3–15	2–12
Misener and Sowell, 1998 ¹⁵	22	6–22	3–11
Powell-Cope <i>et al.</i> , 2003 ²⁵	24	13–24	1–24
Remien <i>et al.</i> , 2003 ¹⁶	110	56–110	5–103
Richter <i>et al.</i> , 2002 ¹⁷	33	2–33	2–13
Roberts and Mann, 2000 ²⁶	20	15–20	NF
Schrimshaw <i>et al.</i> , 2005 ²⁷	158	100–142	16–79
Siegel and Gorey, 1997 ¹⁸	71	15–71	NF
Siegel <i>et al.</i> , 2001 ¹⁹	51	11–51	NF
Wood <i>et al.</i> , 2004 ²⁰	36	2–36	NF

NF, no findings

* Calculated ranges

Table 2

Findings from six quantitative reports on the relationship between a more complex medication regimen and lesser adherence and between a less complex medication regimen and greater adherence

Report	Total (n)	Non-adherence (n) ^a	Adherence (n) ^b
Durante <i>et al.</i> ²⁸	48	12	12
Phillips <i>et al.</i> ²⁹	125	21	8
Schuman <i>et al.</i> ³⁰	326	57	18
Stone <i>et al.</i> ³¹	186	59	71
Wilson <i>et al.</i> ³²	562	136	159
Wilson <i>et al.</i> ³³	132	25	74

^aNumber of women for whom a more complex regimen was associated with less adherence

^bNumber of women for whom a less complex regimen was associated with more adherence