Bayesian Statistics: A Walkthrough with a Simulated Dental Dataset

Eldon Sorensen, MS¹/Chandler Pendleton, MS²/Xian Jin Xie, PhD²

When a clinician sees a patient with a complication, they often go through a Bayesian style of logic, most likely without even knowing it. They assess whether they have seen the complication before, provide an intervention based on historical knowledge of what leads to improvement, and then later assess how the intervention is performed. This process, which is routine in clinical practice, can be mathematically extended into an alternative way of performing statistical analyses to assess clinical research. However, this process is contrary to the most common statistical methods used in dental research: frequentist statistics. Though powerful, frequentist methods come with advantages and disadvantages. Bayesian statistics are an alternative method, one that mirrors how we as researchers think and process new information. In this primer, a walkthrough of Bayesian statistics is performed by constructing priors, defining the likelihood, and using the posterior result to draw conclusions on parameters of interest. The motivating example for this walkthrough was a Bayesian analog to logistic regression, fit using a simulated dental-related dataset of 50 patients who received a dental implant—classified as either within or outside normal limits—from practitioners who did or did not receive a training course in implant placement. The results of the Bayesian and traditional frequentist logistic regression models were compared, resulting in very similar conclusions regarding which parameters seemed to be strongly associated with the outcome. Int J Oral Maxillofac Implants 2022;37:1095–1099. doi: 10.11607/jomi.10210

In dental research, some of the most common statistical methods that you will see are t tests, chi-square tests, ANOVA, linear regression, and logistic regression. All these tests fall under one general branch of statistics: frequentist statistics. While these familiar tools are incredibly useful methods for testing researchers’ hypotheses, frequentist statistics are only one side of statistics and come with both advantages and disadvantages. This primer seeks to introduce you to another branch of statistics that could be utilized in dental research, known as Bayesian statistics.

Frequency (or Frequentist) Statistics

To start, frequentist methods are based on the idea that any one random experiment performed is one instance of a potentially infinite number of independent repetitions of that same experiment.¹⁻⁴ Drawing conclusions from random experiments therefore becomes a process of assessing the plausibility of our observed result in comparison to what is expected across infinite repetitions.¹⁻²,⁴,⁵ This leads to recognizable metrics such as:

$P$ value: The probability of seeing a result equal to or more extreme than what is witnessed in one performed experiment if the experiment were performed repeatedly, assuming that a null hypothesis is true. For example, if you get a $P$ value of .05, you would expect five experiments out of one hundred ($5/100 = 0.05$) to have a result that is either equally extreme or more so than your single experiment if the null hypothesis is true, based on chance alone.¹⁻²,⁴⁻⁶

95% Confidence interval: A range of plausible values for a true parameter of interest, constructed in such a way that if these confidence intervals were made across many repetitions of an experiment, 95% of them would contain the true value of the parameter of interest. For example, if we performed an experiment 100 times and constructed a confidence interval for each one, we would expect 95 of those intervals to contain the true parameter value.²,⁴,⁶

Any reader, upon seeing these definitions, might think they seem incredibly strange, unintuitive, and clunky. They would be right, to a degree. $P$ values and confidence intervals are typically misunderstood in practice and often used arbitrarily to decide whether a yes/no result is significant.⁶ The $P$ value < .05 cutoff is one example.⁶

Frequentist statistics are no doubt a powerful statistical methodology with incredible capabilities. However, it would be wrong to presume that it is perfect, or the only way statistics can be performed.

Bayesian Statistics

Bayesian methods are best described through the idea of probability as a degree of belief about some true event or parameter.¹⁻²,⁴,⁶,⁷ For instance, if you were asked

¹Department of Biostatistics, University of Iowa College of Public Health, Iowa City, Iowa, USA.
²Division of Biostatistics and Computational Biology, University of Iowa College of Dentistry and Dental Clinics, Iowa City, Iowa, USA.

Correspondence to: 801 Newton Road, N419, Iowa City, IA, 52242. Email: xianjin-xie@uiowa.edu

Submitted August 16, 2022; accepted August 24, 2022. ©2022 by Quintessence Publishing Co Inc.
about the chances of seeing rain today, your answer would be a direct case of a Bayesian probability. You might say 0%, 50%, 100%, or some other percentage, and that would represent your degree of belief about whether it will rain today.

Thus, with this line of thinking, researchers perform experiments so that new data can be gathered to inform how to update our beliefs on some event or parameter. For example, if a pharmaceutical company has developed a new drug to treat some condition, they will most likely go through the process of a clinical trial to gather new data to show it works.5,7

This primer aims to go through a formal walkthrough of Bayesian analysis using a simulated dental-related dataset.

### Simulated Data Description

Imagine a periodontist created a new continuing education (CE) course that provided training on implant placement. They believed that dentists who attended this CE course would be more likely to place implants within normal limits compared to those who did not attend this training course. Table 1 provides a description of the variables in our simulated dataset. Since our outcome is binary, the frequentist approach would fit a logistic model to our data. The variable denoting if the dentist placed the implant within normal limits would be our outcome variable, and the variable denoting if the dentist attended this CE course would be the explanatory variable. The patient and dentist demographic variables would be entered into the model to control for any confounding. Table 2 shows the results of fitting the data to this traditional model.

Now, we will walk through how to approach this analysis from a Bayesian framework.

#### Bayesian Walkthrough: The Prior

Let’s say the following question was asked about the odds ratio for an implant placed within normal limits between providers who did or did not receive the training course: What do you believe is the “true” odds ratio? No direct experimental data have been provided, nor has useful information been given to you. You must go off your first instinct. Whatever probability you have focused on is formally known as the prior probability. This is your belief about the true parameter or event of interest before any experiment is performed.1,2,4,5

A common term in Bayesian statistics is the prior distribution. This is because in analyses, prior beliefs are represented through probability distributions, where the shape of the distribution informs what prior beliefs you hold.1,2,4

Let us define the following notation:

- $\theta$: The odds ratio described earlier.
- $f(\theta)$: The prior distribution of $\theta$. In other words, a function that describes your degree of belief about $\theta$.

Figures 1a through 1d detail four potential priors that resemble what one may choose to represent beliefs about $\theta$.

Qualitatively, here is what each prior tells regarding the beliefs about $\theta$:

- Prior A: The highest density is on 1, but covers the range from 0 to 15. This represents someone who believes that the most likely true odds ratio is 1, but it is very possible it could be less than 1 or greater than 1.
- Prior B: The density peaks around 4.5 and spreads around the range from 1 to 15. This represents someone who believes that the odds ratio is most likely around 4.5, and could possibly be anywhere between 1 and 15. Notably, they strongly believe the odds ratio is greater than 1, as the density to the left of the dashed line is very small.
- Prior C: The density peaks around 0.25 and spreads around the range from 0 to 1. This represents someone who believes that the odds ratio is most likely around 0.25, and could possibly be anywhere between 0 and 1. Notably, they strongly believe the odds ratio is less than 1, as the density to the right of the dashed line is very small.

### Table 1 Variable Descriptions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>ID number</td>
</tr>
<tr>
<td>Outcome</td>
<td>Indicator of whether the implant was placed within normal limits</td>
</tr>
<tr>
<td>Patient age</td>
<td>Patient age at the time of placement</td>
</tr>
<tr>
<td>Patient sex</td>
<td>Patient sex</td>
</tr>
<tr>
<td>Provider taken course</td>
<td>Indicator of whether the dentist took the CE course</td>
</tr>
<tr>
<td>Practicing years</td>
<td>Years the dentist has spent practicing dentistry</td>
</tr>
</tbody>
</table>

### Table 2 Frequentist Logistic Regression Results

<table>
<thead>
<tr>
<th></th>
<th>OR estimate</th>
<th>95% Confidence interval</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.079</td>
<td>(0.003, 2.352)</td>
<td>.133</td>
</tr>
<tr>
<td>Patient age</td>
<td>1.005</td>
<td>(0.968, 1.044)</td>
<td>.796</td>
</tr>
<tr>
<td>Patient sex</td>
<td>0.875</td>
<td>(0.233, 3.284)</td>
<td>.839</td>
</tr>
<tr>
<td>Provider taken course</td>
<td>4.472</td>
<td>(1.167, 17.142)</td>
<td>.025</td>
</tr>
<tr>
<td>Years practicing</td>
<td>1.122</td>
<td>(1.021, 1.232)</td>
<td>.014</td>
</tr>
</tbody>
</table>

OR: odds ratio.
Prior D: The density peaks around 1 and spreads around the range 0.75 to 1.25. This represents someone who strongly believes that the odds ratio is around 1, with relatively little belief it is outside the range 0.75 to 1.25.

Try to picture what your prior distribution might be. Do you align perhaps with prior A, where the odds ratio could be anything? Or do you fall in line somewhere with B, C, or D, where you believe in a most likely value or range for the odds ratio?

Prior(s) for Analysis
Since the analysis being performed is a Bayesian analog to logistic regression, a prior must be placed on all the parameters present in the intended model (patient age, patient sex, provider taken course, and provider practicing years). For simplicity, all parameters in the model were given the same prior on the logarithm of the odds ratio visualized in Fig 2.

Bayesian Walkthrough: The Likelihood
Let’s say you have decided on your prior beliefs regarding the true odds ratio. In response to this, the data described in “Simulated Data Description” are presented to you. The summary statistics for that data are shown in Table 3.

Most likely, right now you are weighing whether the summary statistics are enough for you to change your prior beliefs. Of note is the fact that the counts in Table 3 under “Provider Taken Course” equate to an empirical odds ratio of 3.78, suggesting that the odds of an implant placed within normal limits are higher for providers who took the new training course compared to providers who did not take it. For some, the summary statistics might make them question their

### Table 3 Summary Statistics Mean (SD); n (%)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Outside normal limits, n = 25</th>
<th>Within normal limits, n = 25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient age (y), mean (SD)</td>
<td>71.72 (15.75)</td>
<td>70.76 (18.54)</td>
</tr>
<tr>
<td>Patient sex, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>13 (52)</td>
<td>15 (60)</td>
</tr>
<tr>
<td>Female</td>
<td>12 (48)</td>
<td>10 (40)</td>
</tr>
<tr>
<td>Provider taken course, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Course not taken</td>
<td>16 (64)</td>
<td>8 (32)</td>
</tr>
<tr>
<td>Course taken</td>
<td>9 (36)</td>
<td>17 (68)</td>
</tr>
<tr>
<td>Provider practicing years, mean (SD)</td>
<td>10.04 (6.73)</td>
<td>15.76 (8.19)</td>
</tr>
</tbody>
</table>
prior beliefs. For others, seeing the results may make them more confident in prior beliefs.

In Bayesian statistics, the likelihood is formally defined as “the probability of seeing the results of an experiment given the prior beliefs are true.”1,2,4,5 For the present analysis, the likelihood is the probability of seeing the results presented above, given that whatever you initially believed about the odds ratio is true. Depending on how your prior beliefs align with the data, the likelihood could be small or large.

Note that, like the prior, the likelihood is usually designated using a probability distribution.

Bayesian Walkthrough: The Posterior

Now let’s say you are asked this last concluding question: Given the result of the experiment, what do you now believe is the true odds ratio?

Most likely your answer to this question depends on your prior beliefs and the result of the experiment. If you initially believed the true odds ratio was less than or equal to 1, the results may make you believe that the true odds ratio is in the range of 1 to 5. Or, if you initially believed the true odds ratio was greater than 1, the results may reinforce your belief that the odds ratio is in the range of 3 to 5.

In Bayesian statistics, the posterior is formally defined as your belief about the probability of some event or parameter given the result of an experiment performed.1,2,4,5 In this walkthrough, the posterior is your belief about the true odds ratio given that you witnessed the data collected. Thus, in Bayesian analyses, finding the posterior analysis (eg, odds ratio) is the tool through which inference is performed.

Like the prior and likelihood, the posterior is defined as a probability distribution. The posterior distribution is used to calculate the posterior mean estimate of the true parameter or event of interest, as well as the 95% credible intervals. The posterior mean estimate is the mean value of the posterior distribution and represents a point estimate of the true parameter or event of interest after accounting for the results of the experiment and one’s prior beliefs.1,2,4,5 The 95% credible interval is an interval calculated from the posterior distribution such that there is exactly a 95% probability the true value of the parameter or event of interest is within the bounds of the interval.1,2,4,5 Credible intervals are one of the biggest advantages of Bayesian statistics, as their definition is a straightforward probability of the true parameter being in some calculated range.

Posterior Analysis

Using the priors described previously and the data, the rstan® package in R was used to fit the Bayesian model and do the following:

- Estimate the posterior distribution.
- Get posterior mean estimates for all the odds ratios.
- Get 95% credible intervals for all the odds ratios.

Figure 3 shows the prior and posterior distributions for the logarithm of the odds ratio comparing the odds of an implant placed within normal limits between providers who did or did not receive a new training course in implant placement. Overall, while the prior set was wide, the posterior is now concentrated in a much smaller range. This behavior suggests that the data have provided evidence of the parameter being in a specific smaller range.

Table 4 provides posterior means and 95% credible intervals for all the odds ratios. Since the primary variable of interest is the odds ratio for “provider taken course,” the results are interpreted for that parameter with the understanding that the interpretation is generalizable to all the parameters.

The posterior mean estimate is 4.487. Thus, given the prior beliefs specified and adjusting for all other parameters in the model, it is estimated that the odds of an implant placed within normal limits for the patient are 4.487 times higher for those whose provider did not have the new training course compared to those patients whose provider did not have the new training course. In addition, the 95% credible interval (1.286, 16.954) suggests that there is a 95% probability that the true odds ratio is within the range of 1.286 to 16.954. Since this credible interval does not contain 1, there is at least a 95% probability the true odds ratio is greater than 1. In fact, we

![Fig 3](image-url)
can calculate that the exact estimated probability of the true odds ratio being greater than 1 is 99.1%.

**Comparison with Frequentist Logistic Model**

In Table 5, the posterior means and 95% credible intervals from the Bayesian model are compared to the estimates and 95% confidence interval from the frequentist model. For the odds ratio, the Bayesian and frequentist methods produced estimates and intervals that were all very close to one another. Furthermore, the Bayesian and frequentist models generally suggest the same broad conclusion that, after adjusting for all the other variables in the model, the odds of an implant placement within normal limits for providers with the new training course are higher than the odds of an implant placement within normal limits for providers without the new training course.

**CONCLUSIONS**

Bayesian statistics represent an alternative paradigm in the context of performing statistics. With the simulated dental dataset, a Bayesian analog to logistic regression was implemented that enabled an estimation of posterior mean estimates and 95% credible intervals for the odds ratios for all parameters in the model. These estimates all closely lined up with the traditional frequentist implementation of logistic regression.

While frequentist methods are very predominant and established in research and publication practices, Bayesian methods offer some advantages that are worth highlighting. As established in this primer, interpretations of the 95% credible interval can be much more intuitive compared to the traditional 95% confidence interval. Furthermore, Bayesian statistics can lend itself very well to working with complex data. For instance, Bayesian hierarchical modeling has become increasingly popular for many reasons, including its capability to model correlated, repeated, and/or multi-level data.9-11 In addition, Bayesian methods have seen usage in clinical trials.5,7,12 This is especially true regarding adaptive clinical trial designs.12

However, it is important to note the potential disadvantages of Bayesian analyses. Probably the most well-known and discussed weakness is the reliance on specification of the prior distribution.13 In certain cases, proper specification of the prior is imperative to obtain valid estimates of the posterior. In addition, Bayesian methods can be complex and computationally expensive to implement, and are often reliant on software such as WinBUGS or STAN, which assume an in-depth knowledge of Bayesian statistics.8,14 Thus, it is imperative that Bayesian methods are undertaken by someone who knows fully how to construct, implement, and assess a Bayesian model.

**Table 5 Bayesian vs Frequentist Model Results**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Frequentist</th>
<th>Bayesian</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR: Intercept</td>
<td>0.079 [0.003, 2.352]</td>
<td>0.121 [0.006, 2.217]</td>
</tr>
<tr>
<td>OR: Patient age</td>
<td>1.005 [0.968, 1.044]</td>
<td>0.999 [0.963, 1.034]</td>
</tr>
<tr>
<td>OR: Patient sex</td>
<td>0.875 [0.233, 3.284]</td>
<td>0.823 [0.224, 3.09]</td>
</tr>
<tr>
<td>OR: Provider practicing years</td>
<td>1.122 [1.021, 1.232]</td>
<td>1.126 [1.03, 1.24]</td>
</tr>
</tbody>
</table>

Overall, Bayesian methods represent an alternative methodology for performing statistics. For dental researchers, it is worth knowing that Bayesian methods exist and represent a valid alternative to traditional frequentist methods. Depending on the experiment you may wish to analyze, Bayesian methods may be the optimal path forward.

**REFERENCES**