# Common before-after accident study on a road site: a low-informative Bayesian method

Thierry Brenac

Published in 2009 in the *European Transport Research Review* 1(3): 125-134.
The original publication is available at *www.springerlink.com/content/1867-0717*
DOI: 10.1007/s12544-009-0015-4

# Common before-after accident study on a road site: a low-informative Bayesian method

Thierry Brenac*

*INRETS (National Institute for Transport and Safety Research)*
*Chemin de la Croix Blanche F-13300 Salon de Provence, France*

**Abstract:**
*Purpose:* This note aims at providing a Bayesian methodological basis for routine before-after accident studies, often applied to a single road site, and in conditions of limited resources in terms of time and expertise. *Methods*: A low-informative Bayesian method is proposed for before-after accident studies using a comparison site or group of sites. As compared to conventional statistics, the Bayesian approach is less subject to misuse and misinterpretation by practitioners. The low-informative framework seems appropriate in situations of limited expertise. The proposed approach gives the possibility of correcting for regression to the mean. Examples illustrate the application of this method. *Results and conclusions*: It is shown that a relatively simple method, based on the Jeffreys's rule prior considered as a "reasonable standard", can be implemented without major difficulties. Posterior distributions are proper. The numerical calculation of posterior probabilities can be done without using Monte-Carlo simulations nor specialised software tools.
**Keywords** Road safety, controlled before-after study, odds-ratio, low-informative prior, Bayes

## 1. Introduction

It is common that road sites are modified in order to achieve improvements from various points of view (traffic conditions, better integration of various uses and users of the road and public space, reduction of noise and air pollution, traffic safety, etc.). A few years after a site has been modified, local engineers generally have to study the effects of this road change, regarding various aspects including road safety. Thus, a retrospective before-after accident study is often needed.

In such routine situations, resources are limited in terms of time and expertise, and the risk of misuse of conventional statistical methods is increased. Even among people who are more experienced in statistics, like researchers, erroneous uses of conventional methods are common: misuse of tests of significance, erroneous understanding of *p*-values, misinterpretation of confidence intervals (as pointed out by many authors [15, 18, 19, 24, 27, 32]; see also [5, 11, 28]). For example, the *p*-value is often erroneously regarded as the probability that the null hypothesis is true, and the 95% confidence interval obtained is wrongly assumed to contain the true parameter with a 95% chance. The Bayesian approach to statistics is more in accordance with the expectations and intuitions of non-specialists. In particular, the posterior distribution can be legitimately used to give the probabilities that the parameter of interest is contained in various regions of the parameter space (a 95% credible interval, for example), or exceeds a particular value, given the data observed and prior knowledge. Some authors consider that teaching Bayesian statistics is easier than teaching frequentist statistics [10, 31]. Nevertheless, aids to practitioners are necessary to implement Bayesian methods, since the calculations in these approaches are sometimes complex.

*E-mail address: thierry.brenac@inrets.fr

In this paper, we will not deal with studies based on large samples of sites and using multivariate modelling, for which Bayesian approaches were proposed in the recent period [4, 30, 35, 37]. Bayesian methods adapted to meta-analyses or to overviews of several studies (see, for example, [12]) will not be considered here. We will focus on methods applicable to a single site and transferable to engineers for common practice.

In the case we deal with here (routine evaluation, single site), the methods currently used and recommended are conventional statistical methods (see, for example, [23]), even though they sometimes make use of empirical Bayes estimates of the expected accident number on the treated site in order to cope with 'regression to the mean' bias. The principle of a 'full' Bayesian approach was described by Hauer [21, 22] for studying the index of effectiveness $\theta$ of a road measure: the prior probability density function of the parameter $\theta$, reflecting the prior knowledge concerning this parameter, is combined with the likelihood function (probability of the data given the parameter) to obtain the posterior probability density function. The posterior probabilities reflect the revised knowledge about the parameter, given previous knowledge and the data analysed. The method proposed by Hauer, however, is an informative (subjective) Bayes method and presupposes expertise or previously formalised knowledge: the prior probabilities are based on the "elicitation of prevailing opinion about the effectiveness of a treatment" ([22], p. 289), or possibly on the results of previous studies or meta-analyses. Road safety expertise is limited, however, in the routine situations we consider here, since the study is often carried out by a local road engineer, and not by a road safety specialist. Moreover, the site modification is often singular and not generic (it may combine several treatments, for example: redesigning of islands, resurfacing and marking at a junction site). Therefore, it may be difficult to make use of results from previous meta-analyses. A method coping with this problem was described by Al Masaied *et al*. [3]: prior probabilities were estimated using a part of the accident data, for both the before and the after periods. In the case of a single site, however, this may lead to very small accident numbers for each data subset. Another way is to use the 'objective' or 'low-informative' Bayesian framework [6, 7, 17, 25, 26] where the prior probabilities are chosen in order to be neutral in some way as regards the possible parameter values, reflecting the lack of previous knowledge. Besides, it can be argued that results based on low-informative approaches are generally easier to communicate to a diverse or uninitiated audience, since, as mentioned by Box and Tiao [13], they represent "what someone who a priori knew very little about an unknown parameter should believe in light of the data" (p. 22).

In before-after accident studies, it is important to be able to control for regression to the mean bias, which can be done by incorporating some limited information into the prior distribution concerning one component of the vector of parameters (see section 4). Besides, although such studies are retrospective and not experimental, one should seek to control for the confusing influence of factors other than the road change. To this end, it can be useful to take into consideration a comparison group of similar sites, for example. The method described by Hauer [22] uses a comparison sample, but the calculations are based on approximations which presuppose that the accident counts in the comparison sample are large. The method proposed by Al-Masaied *et al*. [3] is a simple before-after method without comparison sites.

In this methodological note, we describe a low-informative Bayesian method adapted to the current practice of before-after accident studies concerning a single treated site (or a group of sites considered as a whole). A comparison site (or group of sites) is used in order to control for factors other than the road modification. Practical means of calculation, for a commonly available spreadsheet software package, will also be provided on the author's webpage (http://www.inrets.fr/ur/ma/Brenac.html).

## 2. Data structure and parameters for the before-after study with comparison sites

When a comparison site or group of sites is used, the basic data take the form of a 2×2 contingency table (table 1) containing the observed accident counts $x_i$ . These counts are considered as observations of independent Poisson variables $X_i$ with expected values $\mu_i$ (unknown).

**Table 1** Usual form of the basic accident data

| | Treated site | Comparison site (or group of sites) |
|---|---|---|
| Period I (Before) | $x_1$ | $x_3$ |
| Period II (After) | $x_2$ | $x_4$ |

Under the assumption of a strong similarity between the treated site and the comparison site, and if the evolution of traffic does not differ between them, the effect of the measure can be represented by the odds ratio

$$\theta = \mu_2 / [\mu_1 \times (\mu_4 / \mu_3)] = \mu_2 \, \mu_3 / (\mu_1 \, \mu_4) \qquad (1)$$

$\theta$ expresses the ratio of *the 'accidentality' on the treatment site during period II* (after modification) to *what this 'accidentality' would have been during the same period II, had the site not been modified* — here we use the term 'accidentality' in the somewhat unusual sense of the expected value of the accident count. From a practical point of view, an odds ratio of 0.8, for example, would mean that the effect of the treatment is a 20 % reduction in accidentality. The ratio $\eta = \mu_4 / \mu_3$ reflects the effect of other factors on the evolution from period I to period II, assumed to be common to both the treated and comparison sites ($\eta$ can be considered as a trend parameter). In other terms, $\mu_2$ and $\mu_4$ can be expressed as follows: $\mu_2 = \mu_1 \, \theta \, \eta$ and $\mu_4 = \mu_3 \, \eta$ .

Thus, we are in the presence of a problem with four observations $x_1$, $x_2$, $x_3$, $x_4$ from four independent random Poisson variables $X_1$, $X_2$, $X_3$ and $X_4$, and four unknown parameters $\theta$, $\eta$, $\mu_1$ and $\mu_3$ with the following relationships:

$$X_1 \sim \text{Poiss}(\mu_1), \; X_2 \sim \text{Poiss}(\mu_1 \theta \, \eta), \; X_3 \sim \text{Poiss}(\mu_3), \; X_4 \sim \text{Poiss}(\mu_3 \eta) \qquad (2)$$

## 3. The Bayesian framework

According to the Bayesian approach to statistics, the unknown parameters $\theta$, $\eta$ , $\mu_1$ and $\mu_3$ are considered as instantiations of variables $\Theta$, $H$, $M_1$ and $M_3$, treated as random variables, but which in fact reflect our uncertainty about the values of these parameters. Given the observed data $\mathbf{x} = (x_1, x_2, x_3, x_4)$, given the likelihood function $L(\mathbf{x} \mid \theta, \eta, \mu_1, \mu_3)$[1] and the joint prior probability density function of the parameters $\pi(\theta, \eta, \mu_1, \mu_3)$, the application of Bayes' theorem leads to the joint posterior distribution

$$p(\theta, \eta, \mu_1, \mu_3 \mid \mathbf{x}) = \frac{L(\mathbf{x} \mid \theta, \eta, \mu_1, \mu_3) \times \pi(\theta, \eta, \mu_1, \mu_3)}{\int\int\int\int L(\mathbf{x} \mid \theta, \eta, \mu_1, \mu_3) \pi(\theta, \eta, \mu_1, \mu_3) \mathrm{d}\theta \mathrm{d}\eta \mathrm{d}\mu_1 \mathrm{d}\mu_3} \qquad (3)$$

The joint prior distribution $\pi(\theta, \eta, \mu_1, \mu_3)$ represents our previous assumptions or knowledge (or lack of knowledge) regarding the parameters (see section 4). The joint posterior distribution represents our revised knowledge about the parameters, after the observations are

[1] This notation means: probability of the data $\mathbf{x} = (x_1, x_2, x_3, x_4)$ given the parameter values $\theta$, $\eta$ , $\mu_1$, $\mu_3$.

taken into account. The likelihood function can be easily derived from the problem formulation given in section 2:

$$L(\mathbf{x} \mid \theta, \eta, \mu_1, \mu_3) = \frac{\exp(-\mu_1)\mu_1^{x_1}}{x_1!} \frac{\exp(-\mu_1\theta\eta)(\mu_1\theta\eta)^{x_2}}{x_2!} \frac{\exp(-\mu_3)\mu_3^{x_3}}{x_3!} \frac{\exp(-\mu_3\eta)(\mu_3\eta)^{x_4}}{x_4!} \quad (4)$$

The parameter of interest is $\theta$. Its posterior probability density function can be obtained by integrating the joint posterior distribution with respect to the three other parameters:

$$p(\theta \mid \mathbf{x}) = \int\int\int p(\theta, \eta, \mu_1, \mu_3 \mid \mathbf{x}) \mathrm{d}\eta \mathrm{d}\mu_1 \mathrm{d}\mu_3 \quad (5)$$

From a practical point of view, however, the most useful result is the posterior cumulative distribution function of $\Theta$,

$$F_\Theta(t \mid \mathbf{x}) = \Pr(\Theta < t \mid \mathbf{x}) = \int_{\theta=0}^{\theta=t} \int_{\eta} \int_{\mu_3} \int_{\mu_1} p(\theta, \eta, \mu_1, \mu_3 \mid \mathbf{x}) \mathrm{d}\mu_1 \mathrm{d}\mu_3 \mathrm{d}\eta \mathrm{d}\theta \quad (6)$$

This cumulative distribution function makes it possible to calculate credible intervals and the probability that the effect studied is lower or higher than a particular value, given the data and prior probabilities.

## 4. Low-informative prior distributions

In this paper we assume a lack of previous knowledge or sufficient expertise regarding the parameters. Thus, the prior distributions should be low-informative or neutral as regards these parameters. This choice also tends to "let the data speak for themselves", giving a higher importance to the likelihood function in the calculation of posterior probabilities. Two situations should be distinguished, however, according to whether regression to the mean bias is likely or not. Regression to the mean (see, for example, [1]) occurs when the site was chosen for treatment in consideration of a high accident record. In this case, the count $x_1$ gives only biased information on the expected value $\mu_1$, and a low-informative prior distribution for $\mu_1$ would lead to biased results, overestimating the treatment effect. In this situation, other data or information are needed and should be taken into account in the prior distribution of $\mu_1$ (see point 4.2).

### 4.1. Case where regression to the mean bias is unlikely

In many circumstances, regression to the mean bias is unlikely: for example, when the site modification was not decided for safety reasons, but for other purposes (really independent from accident counts). In this case, a low-informative joint prior distribution can be chosen for the four parameters $\theta$, $\eta$ , $\mu_1$ and $\mu_3$. The way of selecting low-informative priors (also called non-informative, objective, default or reference priors) is widely discussed in Bayesian statistics (see the review by Kass and Wassermann [29]; see also [2, 8, 9, 17, 20, 25, 26, 39]). We will not enter this debate here since, as mentioned by Ghosh *et al*. [17], "even though there is no unique objective prior, the posteriors will usually be very similar even with a modest amount of data" (p. 147). In the present paper, for the sake of simplicity, we will only consider the prior obtained by the Jeffreys's general rule[2] [26], which is widely accepted as a

[2] This rule can be justified from several points of view, in particular: invariance by re-parameterisation, uniformity, in the sense of equiprobability of regions of same size in the parameter space with a Riemannian metric, and minimisation of information (the Jeffreys's rule prior can be considered as a special case of the Bernardo-Berger prior). For developments of these arguments, see for example Ghosh *et al*. [17] and Kass and Wassermann [29].

"reasonable standard" [29]. For a vector of parameters $\xi$, the Jeffreys's rule prior is proportional to the square root of the determinant of the Fisher information matrix:

$$\pi(\xi) \propto [\det(\mathbf{I}(\xi))]^{1/2} \qquad (7)$$

where $\propto$ denotes proportionality. In this expression, $\mathbf{I}(\xi)$ is the Fisher information matrix defined by $I(\xi)_{i,j} = E\left(-\frac{\partial^2 l}{\partial\xi_i \partial\xi_j}\right)$ where $l$ is the log-likelihood. Applied to our problem, using equation 4, this rule leads to the joint prior

$$\pi(\theta, \eta, \mu_1, \mu_3) \propto (1/\theta)^{1/2} \qquad (8)$$

Like many non-informative priors, this prior is improper since it does not integrate to a finite value over the parameter space. In Bayesian statistics, however, this is not regarded as a problem, provided that the posterior distribution is proper (*i.e.*, the integral in the denominator of equation 3 converges to a finite value).

4.2. Case where regression to the mean bias is likely

In this situation, conventional methods correct for regression to the mean by considering that the site is taken from a population of comparable sites and extracting complementary information from a sample of such sites[3]. Each of the accident counts $x_{1j}$ at these sites, during period I, is considered as an observation from a Poisson variable with mean $\mu_{1j}$ . The $\mu_{1j}$ are assumed to be distributed like a Gamma variable with shape parameter $\alpha$ and scale parameter $\lambda$ (some empirical justifications can be found in the literature [1, 34]). This Poisson-Gamma structure leads to a negative binomial distribution of the counts $x_{1j}$ among this sample of sites. Based on the mean $m$ and variance $s^2$ of this distribution, estimated from the $x_{1j}$, it is possible to estimate[4] $\alpha$ and $\lambda$: $\alpha = m^2/(s^2-m)$ and $\lambda = m/(s^2-m)$. Conventional evaluation methods then replace $x_1$, the usual estimate of $\mu_1$, by the empirical Bayes estimate $\mu_1^* = m^2/s^2 + x_1(s^2-m)/s^2 = (\alpha+x_1)/(1+\lambda)$ for the calculation of the odds ratio [16, 23, 36]. This technique has proved to be effective for correcting for regression to the mean bias [38].

The equivalent in the 'full' Bayes approach consists in taking the Gamma($\alpha$,$\lambda$) prior distribution for the parameter $\mu_1$:

$$\pi(\mu_1) = \frac{\lambda^\alpha \mu_1^{\alpha-1} \exp(-\lambda\mu_1)}{\Gamma(\alpha)} \qquad (9)$$

In this situation, a joint prior distribution can be obtained by calculating $\pi(\theta, \eta, \mu_3)$ with the Jeffreys's rule applied while holding $\mu_1$ fixed (see [29]), which gives $\pi(\theta, \eta, \mu_3) \propto (1/\theta)^{1/2}$ and leads to

$$\pi(\theta, \eta, \mu_1, \mu_3) \propto (1/\theta)^{1/2}\, \mu_1^{\alpha-1} \exp(-\lambda\mu_1) \qquad (10)$$

(Constants are not taken into account since they would be cancelled anyway as common factors in the denominator and numerator of equation 3). This prior is also improper. The estimates of $\alpha$ and $\lambda$ are drawn from accident data at a sample of similar sites (independent from the group of comparison sites), or from an accident model, as described above for conventional methods using empirical Bayes estimates. Although this joint distribution

[3] Alternatively, when a general accident model is available (see, for example, the models published by the Transport Research Laboratory in the United Kingdom), it can be applied for obtaining complementary information, instead of using a sample of the population of comparable sites [23, 38].

[4] Instead, if an accident model is available, it can give the mean $m$ and variance $s^2$ of the accident counts on a virtual population of sites with the same characteristics as the site of interest [23, 38]. The parameters $\alpha$ and $\lambda$ are then also obtained by calculating $\alpha = m^2/(s^2-m)$ and $\lambda = m/(s^2-m)$.

(equation 10) uses some prior information concerning $\mu_1$ (through $\alpha$ and $\lambda$), it remains, however, low-informative in a relative sense, since no prior knowledge is used concerning the parameter of interest $\theta$ and the two other parameters $\eta$ and $\mu_3$.

## 5. Posterior probabilities

5.1. Case where regression to the mean bias is unlikely

Applying the likelihood function (equation 4) and the joint prior distribution given in equation 8 to the calculation of the joint posterior distribution (equation 3) gives the following expression, after simplification (cancelling of factors present both at the numerator and the denominator):

$$p(\theta,\eta,\mu_1,\mu_3 \mid \mathbf{x}) = \frac{1}{C}\exp(-(1+\theta\eta)\mu_1)\mu_1^{\,x_1+x_2}\exp(-(1+\eta)\mu_3)\mu_3^{\,x_3+x_4}\theta^{x_2-½}\eta^{x_2+x_4} \qquad (11)$$

where $C = \int\int\int\int \exp(-(1+\theta\eta)\mu_1)\mu_1^{\,x_1+x_2}\exp(-(1+\eta)\mu_3)\mu_3^{\,x_3+x_4}\theta^{x_2-½}\eta^{x_2+x_4}\mathrm{d}\theta\mathrm{d}\eta\mathrm{d}\mu_1\mathrm{d}\mu_3$.

This latter integral converges to a finite value even when some (or all) of the $x_i$ equal zero. Therefore, a proper posterior distribution can always be obtained. The terms in $\mu_1$ and $\mu_3$ are proportional to Gamma density functions, which makes it possible to integrate the expression given in equation 11 with respect to $\mu_1$ and $\mu_3$, leading to the joint posterior of $\theta$ and $\eta$

$$p(\theta,\eta \mid \mathbf{x}) = \frac{1}{K}\frac{\theta^{x_2-½}}{(1+\theta\eta)^{x_1+x_2+1}}\frac{\eta^{x_2+x_4}}{(1+\eta)^{x_3+x_4+1}} \qquad (12)$$

with $K = \int_\theta\int_\eta \frac{\theta^{x_2-½}}{(1+\theta\eta)^{x_1+x_2+1}}\frac{\eta^{x_2+x_4}}{(1+\eta)^{x_3+x_4+1}}\mathrm{d}\eta\mathrm{d}\theta = \mathrm{B}(x_2+½, x_1+½)\mathrm{B}(x_4+½, x_3+½)$ where B denotes the Beta function[5]. The posterior cumulative distribution function of $\Theta$ is then given by

$$F_\Theta(t \mid \mathbf{x}) = \frac{1}{K}\int_{\theta=0}^{\theta=t}\int_\eta \frac{\theta^{x_2-½}}{(1+\theta\eta)^{x_1+x_2+1}}\frac{\eta^{x_2+x_4}}{(1+\eta)^{x_3+x_4+1}}\mathrm{d}\eta\mathrm{d}\theta \qquad (13)$$

The calculation of this integral is generally not possible by analytical means. We describe in the appendix of this paper a way of calculating it numerically.

5.2. Case where regression to the mean bias is likely

For the prior given in equation 10, the same kind of calculations as those described in point 5.1 lead to the following expression for the posterior cumulative distribution of $\Theta$:

$$F_\Theta(t \mid \mathbf{x}) = \frac{1}{K'}\int_{\theta=0}^{\theta=t}\int_\eta \frac{\theta^{x_2-½}}{(1+\lambda+\theta\eta)^{x_1+x_2+\alpha}}\frac{\eta^{x_2+x_4}}{(1+\eta)^{x_3+x_4+1}}\mathrm{d}\eta\mathrm{d}\theta \qquad (14)$$

where $K' = \frac{\mathrm{B}(x_1+\alpha-½, x_2+½)\mathrm{B}(x_3+½, x_4+½)}{(1+\lambda)^{x_1+\alpha-½}}$. The calculation leading to this result is not valid if $x_1+\alpha \le ½$ (which is unlikely: $\alpha$ is a positive parameter and we are in a situation where the treated site was chosen in consideration of a high accident count $x_1$). For the numerical calculation of this integral, see the appendix.

[5] In the expression of *K*, the term in $\theta$ is proportional to a three-parameter Beta-prime distribution, which makes it possible to integrate with respect to $\theta$ over [0,+∞). The integration with respect to $\eta$ is then possible, over [0,+∞).

5.3. Practical uses of the posterior cumulative distribution function of $\Theta$

From a practical point of view, various useful results can be obtained using the function $F_\Theta(t \mid \mathbf{x})$. For example, the lower limit $\theta_{LL}$ and upper limit $\theta_{UL}$ of a 95% symmetrical credible interval are defined by $F_\Theta(\theta_{LL} \mid \mathbf{x}) = 0.025$ and $F_\Theta(\theta_{UL} \mid \mathbf{x}) = 0.975$; the probability, given the data, that $\theta$ is contained in this interval is 95%. The median $\theta_{med}$ defined by $F_\Theta(\theta_{med} \mid \mathbf{x}) = 0.5$ gives a point estimate of the odds ratio for which the posterior risks of overestimation and underestimation are equal. The value $F_\Theta(1 \mid \mathbf{x})$ represents the posterior probability that $\theta$ is lower than 1, *i.e.* the probability that the treatment is beneficial to safety, given the data and initial assumptions (see section 2).

## 6. Particular cases

6.1 Group of comparison sites instead of a single comparison site

In this situation, the group of $q$ comparison sites is considered as a whole, with $x_3 = \Sigma\, x_{3k}$ and $x_4 = \Sigma\, x_{4k}$ (where $x_{3k}$ and $x_{4k}$ are the accident counts during periods I and II on each comparison site $k$, with $k = 1$ to $q$). The aggregated counts $x_3$ and $x_4$ are observations from random variables $X_3$ and $X_4$ which are Poisson variables (since they are obtained by summing the independent Poisson variables $X_{3k}$ and $X_{4k}$) with means $\mu_3 = \Sigma\, \mu_{3k}$ and $\mu_4 = \Sigma\, \mu_{4k}$. The calculations described in sections 3 to 5 are then applied by simply using the aggregated counts $x_3$ and $x_4$ and the aggregated means $\mu_3$ and $\mu_4$ . The low informative joint prior is given by equation 8 or 10. The posterior cumulative distribution function of $\Theta$ is then given by equation 13 or 14 (with $x_3 = \Sigma\, x_{3k}$ and $x_4 = \Sigma\, x_{4k}$ ).

6.2 Multiple treated sites

The general case of several treated sites, considered independently, with possibly different odds ratios $\theta_i$ due to heterogeneity in the treatment effect is beyond the purpose of this paper and will be the subject of further publications. Nevertheless, in the simpler situation where a group of treated sites is considered as a whole (with a focus on the overall effect of treatment), the methods described above can be easily adapted.

Let us consider $n$ treated sites with accident counts $x_{1i}$ and $x_{2i}$ ($i = 1$ to $n$) during periods I and II, with corresponding means $\mu_{1i}$ and $\mu_{2i}$, and $q$ comparison sites with accidents counts $x_{3k}$ and $x_{4k}$ ($k = 1$ to $q$) during periods I and II, with corresponding means $\mu_{3k}$ and $\mu_{4k}$.

When regression to the mean bias is unlikely, and if we consider the treated sites as a whole (and the comparison sites as a whole), the calculations and results described in sections 3 to 5 can be applied by simply using the aggregated counts $x_1 = \Sigma\, x_{1i}$, $x_2 = \Sigma\, x_{2i}$, $x_3 = \Sigma\, x_{3k}$, $x_4 = \Sigma\, x_{4k}$ and the corresponding aggregated means, with the prior given in equation 8. In this case, the parameter $\theta$ represents the overall effect of the programme of treatment. The posterior probabilities are given by equation 13.

When regression to the mean bias appears likely, if the same prior Gamma($\alpha,\lambda$) distribution can be assumed for the mean $\mu_{1i}$ of each treated site $i$, the prior distribution of the overall mean $\mu_1 = \Sigma\mu_{1i}$ is a Gamma($n\alpha,\lambda$) distribution (using the classical property of the sum of independent Gamma variables with same scale parameter $\lambda$). Considering the treated sites as a whole (and the comparison sites as a whole), and considering $\theta$ as the overall effect, the joint prior distribution becomes

$$\pi(\theta, \eta, \mu_1, \mu_3) \propto (1/\theta)^{½}\, \mu_1^{n\alpha-1} \exp(-\lambda\mu_1) \qquad (15)$$

where $\mu_3 = \Sigma\mu_{3k}$ .

The posterior cumulative distribution function of $\Theta$ is then

$$F_{\Theta}(t \mid \mathbf{x}) = \frac{1}{K''} \int_{\theta=0}^{\theta=t} \int_{\eta} \frac{\theta^{x_2 - ½}}{(1+\lambda+\theta\eta)^{x_1+x_2+n\alpha}} \frac{\eta^{x_2+x_4}}{(1+\eta)^{x_3+x_4+1}} \mathrm{d}\eta \mathrm{d}\theta \qquad (16)$$

where $K'' = \frac{\mathrm{B}(x_1 + n\alpha - ½, x_2 + ½)\mathrm{B}(x_3 + ½, x_4 + ½)}{(1+\lambda)^{x_1+n\alpha-½}}$ and where $n$ is the number of treated sites, $x_1 = \Sigma\, x_{1i}$, $x_2 = \Sigma\, x_{2i}$, $x_3 = \Sigma\, x_{3k}$ and $x_4 = \Sigma\, x_{4k}$.

## 7. Examples of application

### 7.1. Example 1: Safety effect of redesigning an urban road section

We describe here the case of an urban section of road where the infrastructure was largely modified in order to enhance the quality of local urban life. Raised median islands, small roundabouts, speed humps and raised tables were implemented in 2000 on this section of a main urban road in a town of 40,000 inhabitants (length of the treated section: 700 meters). All the unmodified sections of the main roads in this town were taken as a comparison group of sites. The comparability between the treated site and the comparison group of sites was verified by comparing the yearly injury accident counts for the 1989-1999 period. The 'before' period is the five-year period from 1995 to 1999. The 'after' period is the five-year period from 2001 to 2005. The presence of regression to the mean bias was considered to be unlikely for the following reasons: this project was not decided for safety reasons, and the proportion of accidents during the 1995-1999 period relative to 1989-1999 was not higher on the treated site as compared to all the unmodified sections of main roads in this town. For the 'before' period, 16 injury accidents occurred on the treated site and 61 injury accidents occurred on the comparison group of sites. For the 'after' period, 3 injury accidents occurred on the treated site, and 46 injury accidents occurred on the comparison group of sites.

The calculations applied to these data ($x_1 = 16$, $x_2 = 3$, $x_3 = 61$, $x_4 = 46$) with the low-informative prior based on the Jeffreys's rule (equation 8) give the following results based on the posterior cumulative distribution function of $\Theta$:

| | |
|---|---|
| 95% symmetrical credible interval: | 0.062 to 0.815 |
| Median: | 0.259 |
| Posterior probability that $\theta < 1$: | 0.990 |

Figure 1 shows an example of spreadsheet screen for the posterior probability calculation (see the appendix), applied to these data.

These results suggest a beneficial effect on safety. They can be compared to the results that would be obtained by conventional statistical methods. Nevertheless, as mentioned in the introduction, Bayesian and non-Bayesian concepts (like credible interval and confidence interval) can not be interpreted in the same way[6]. In this example, the usual unconditional maximum likelihood estimator of the odds ratio, with the related approximate 95% confidence interval (Woolf interval), would lead to the following results:

| | |
|---|---|
| $\theta_{ML}$* = 0.249 | |
| Woolf 95% confidence interval: | 0.068 to 0.904 |

In this example, a practitioner would probably conclude in favour of a positive effect on safety, from both these Bayesian and non-Bayesian results.

[6] A correct interpretation of a classical (non-Bayesian) 95% confidence interval is: if we could indefinitely repeat the same "experiment" with the same parameter value, 95% of the confidence intervals thus obtained would contain this value.

**Fig. 1** Example of spreadsheet screen for the posterior probability calculation

Before-after accident study, with comparison sites — Low-informative Bayesian approach
Posterior probability calculator

**CALCULATION OF** Pr(Θ ≤ t | data)

Probability that the "index of effectiveness" (represented here by the odds ratio Θ)
is less than a particular value t, given the data and assumptions (posterior probability)

| **Enter the total accident counts** | Treated site | Comparison site(s) |
|---|---|---|
| "Before" period | 16 | 61 |
| "After" period | 3 | 46 |

**Select a prior distribution**

PRIOR 1: Jeffreys's rule joint prior for (θ, η, μ1, μ3)
PRIOR 2: Modified joint prior : prior Gamma(α,λ) for μ1 and Jeffreys's rule prior for (θ, η, μ3)
(this modified prior is recommended if regression-to-the-mean is likely)

| | | |
|---|---|---|
| To use PRIOR 1, enter α = 0 and λ = 0 | α = | 0.000 |
| To use PRIOR 2, enter the values (>0) of α and λ | λ = | 0.000 |
| **Enter a value (>0) for t** | t = | 1.000 |

| **RESULT** | Pr(Θ ≤ t \| data) = | 0.990 |
|---|---|---|

NB: The intermediate calculations (numerical integration) are provided in another sheet.

7.2. Example 2: Safety effect of a rural crossroads modification

This example deals with a priority intersection on a main rural two-lane road. This crossroads was modified in 1986 (installation of median raised islands, marking) for safety reasons. Therefore, regression to the mean is likely to occur. At this junction, 14 injury accidents occurred during the three-year period before the treatment. During the three-year period following the treatment, 4 injury accidents occurred.

This evolution was compared to the evolution observed at a set of 11 similar intersections on main rural two-lane roads in the same region, used as a comparison group of sites. At these sites, considered as a whole, 33 injury accidents occurred during the before period and 22 injury accidents occurred during the after period.

The calculations applied to these data ($x_1 = 14$, $x_2 = 4$, $x_3 = 33$, $x_4 = 22$), using the low-informative prior distribution given by equation 8 (Jeffreys's rule prior), would lead to the following results based on the posterior cumulative distribution function of $\Theta$:

95% symmetrical credible interval: 0.117 to 1.389
Median: 0.439
Posterior probability that $\theta < 1$: 0.917

Due to the high risk of regression to the mean in this case, however, these results are probably biased. In order to correct for this regression to the mean bias, it is necessary to use a more 'informed' prior, concerning the parameter $\mu_1$ (see section 4). To this end, the parameters $\alpha$ and $\lambda$ of a prior Gamma distribution for $\mu_1$ have to be estimated. By applying an accident model (which was established at a national level [14]) to the characteristics of this junction (traffic volumes, number of arms, number of traffic lanes), as mentioned in section 4.2., it is possible to obtain the mean $m = 3.55$ and variance $s^2 = 15.90$ of the accident counts for a virtual population of similar sites during the same period I. On this basis, we can calculate the

estimates $\alpha = 1.02$ and $\lambda = 0.29$. The joint prior given by equation 10 is then precisely defined and leads to the following results, in terms of posterior probabilities:

| | |
|---|---|
| 95% symmetrical credible interval: | 0.151 to 1.789 |
| Median: | 0.566 |
| Posterior probability that $\theta < 1$: | 0.828 |

These results show that, in this case, the safety effect is in reality smaller than indicated by the biased results obtained with the low-informative prior given by equation 8. The median of the posterior distribution (0.566) can be used as a point estimate of the odds ratio (where the posterior probabilities of overestimation and underestimation are equal). This value corresponds to an accident reduction of approximately 43%. The 95% credible interval, however, is large and the beneficial effect of the treatment remains uncertain.

Using the same data, a more conventional approach would lead, for example, to the maximum likelihood estimate $\theta_{ML}^* = 0.429$ (without controlling for regression to the mean), or to a corrected estimate of 0.515 based on the empirical Bayes estimate of $\mu_1$ [16, 36].

7.3. Example 3: Safety effect of resurfacing on main roads

This example is based on some of the data published in an article by Leden *et al.* [33], dealing with the effect of resurfacing on friction, speeds and safety on main roads in Finland. The treated sites are all sections on main roads (in the south of Finland) which were resurfaced in 1991. The comparison sites are all the untreated main roads in the same region. Due to the particular road conditions in winter in Finland, only the effects on the non-winter period (from April 1 to September 30) are studied. Regression to the mean bias is considered to be unlikely, since "sections were selected for treatment on a routine maintenance base" [33] (p. 82) and not for safety reasons. We consider the treated sites as a whole, and the comparison sites as a whole. The parameter $\theta$ thus represents the overall effect of the treatment on the set of sites. The following data concern the 'before' period from April to September 1990 and the 'after' period from April to September 1992. Before and after injury accident counts are $x_1 = 80$ and $x_2 = 74$ on the treated sites, and $x_3 = 931$ and $x_4 = 779$ on the comparison sites. Based on the Jeffreys's rule prior, the results are as follows, in terms of posterior probabilities:

| | |
|---|---|
| 95% symmetrical credible interval: | 0.794 to 1.537 |
| Median: | 1.106 |
| Posterior probability that $\theta < 1$: | 0.275 |

One can note the proximity of these results with the following results which would be obtained with a conventional frequentist approach:

| | |
|---|---|
| $\theta_{ML}^* = 1.105$ | |
| Woolf 95% confidence interval: | 0.794 to 1.537 |

This proximity is not surprising: posterior credible intervals based on the Jeffreys's rule prior are frequently close to frequentist confidence intervals in large sample conditions [17, 40] although they do not have the same meaning.

Based on these results, the posterior median estimate of $\theta$ would suggest a slight detrimental effect on safety (increase of accidentality of approximately 11%), but no certain conclusion can be drawn since the 95% credible interval is large. Based on the posterior probability that $\theta < 1$ (approximately 28%), however, one could say that the probability that the treatment increases the accidentality, given the data and assumptions, is 72%. No equivalent result from a conventional statistical analysis could lead to this kind of

interpretation, except if one wrongly interprets a *p*-value as a posterior probability. A possible increase of accidentality could be explained by the fact that resurfacing tends to increase the average speeds, at least when the road is dry, as shown by Leden *et al*. [33].

## 8. Discussion and conclusion

In this note, we described a low-informative Bayesian method for before-after accident studies, using a comparison site or group of sites, and giving the possibility of correcting for regression to the mean bias. The aim was to provide a methodological basis for routine evaluation studies, often applied to a single treated site, and in conditions of limited resources in terms of time and expertise. As compared to conventional statistics, the Bayesian approach is less subject to misuse and misinterpretation by practitioners with limited statistical experience. The low-informative or objective Bayesian methods seem appropriate in routine evaluation studies, where expertise or previous knowledge are often limited or hard to formalise. As shown in sections 2 to 6, a relatively simple method, based on the Jeffreys's rule prior considered as a "reasonable standard", can be implemented without major difficulties. Posterior distributions are proper. The numerical calculation of posterior probabilities can be done without using Monte-Carlo methods nor specialised software tools. The examples given in section 7 show that the results can be analysed in a direct way, without the high risk of misinterpretation involved in the analysis of frequentist results.

Further developments, however, are still needed. Although this method seems to be transferable to engineers for common practice, further work is necessary in order to provide a simple, didactic description of the Bayesian line of reasoning, with minimal use of mathematical formalisms, appropriate for communicating this approach to practitioners. Concerning the practical means for calculating the posterior probabilities, the spreadsheet mentioned in the appendix (for a common spreadsheet software package) will be made available on our website.

The proposed method has limitations, of course. Retrospective before-after studies are not randomised experiments and the validity of their results is based on the assumption that the treated and comparison sites are similar. Before-after studies based on multivariate generalised linear models make it possible to better control for the influence of differences between treated and comparison sites. But such methodologies would generally involve a thorough data collection and analysis on a large sample of sites, which seems hard to implement by practitioners in the routine situations we considered in this paper. The comparability of treated and comparison sites, however, can be checked by examining their accident history, when accident data are available for a long period before the treatment (see [23]). A Bayesian approach to this subject could be studied. Besides, other developments could contribute to extending the field of application of the proposed method: in this paper, we only dealt with the case of a single treated site (or a group of sites treated as a whole, with a focus on the overall effect of the programme of treatment), with a comparison site or group of sites. The case of multiple treated sites considered independently and with possibly different odds ratios remains to be dealt with. However, this would involve an increased complexity and more difficulties for practitioners.

We hope this methodological note will contribute to an increased use of the Bayesian approach, which is more in accordance with the expectations and intuitions of non-statisticians, in the current practice of before-after accident studies.

## Acknowledgements

The author would like to thank Sylvie Després (Université Paris-Nord) and three anonymous referees for their helpful comments.

## Appendix: Numerical calculation of the posterior cumulative distribution function of $\Theta$

In the case where regression to the mean bias is unlikely, this function is given by equation 13. It can be written in a more generic form:

$$F_\Theta(t \mid \mathbf{x}) = \frac{1}{\mathrm{B}(a,b)\mathrm{B}(c,d)} \int_{\theta=0}^{\theta=t} \int_{\eta} \frac{\theta^{a-1}}{(1+\theta\eta)^{a+b}} \frac{\eta^{a+c-1}}{(1+\eta)^{c+d}} \mathrm{d}\eta \mathrm{d}\theta \qquad (17)$$

where the parameters $a$, $b$, $c$ and $d$ take the values $a = x_2+½$, $b = x_1+½$, $c = x_4+½$, $d = x_3+½$ for the Jeffreys's rule prior (equation 8), and the values $a = \Sigma x_{2i}+½$, $b = \Sigma x_{1i}+½$, $c = \Sigma x_{4k}+½$, $d = \Sigma x_{3k}+½$ in case of multiple sites considered as a whole (with the Jeffreys's rule prior; see section 6), for example.

If we successively use the three changes of variable $z(\eta) = \eta/(1+\eta)$, $\omega(\theta) = \theta z/((\theta-1)z + 1)$ and lastly $u(z) = \mathrm{Betacdf}_{c,d}(z)$, where $\mathrm{Betacdf}_{c,d}$ denotes the cumulative distribution function of the Beta distribution with parameters $c$ and $d$, we obtain from equation 17:

$$F_\Theta(t \mid \mathbf{x}) = \int_{u=0}^{u=1} \mathrm{Betacdf}_{a,b}\left( \frac{t \times \mathrm{Betacdf}_{c,d}^{-1}(u)}{1+(t-1)\times \mathrm{Betacdf}_{c,d}^{-1}(u)} \right) \mathrm{d}u \qquad (18)$$

where $\mathrm{Betacdf}_{a,b}$ denotes the cumulative distribution function of the Beta distribution with parameters $a$ and $b$, and $\mathrm{Betacdf}_{c,d}^{-1}$ denotes the inverse function of $\mathrm{Betacdf}_{c,d}$. The Beta cumulative distribution function and its inverse are commonly available in spreadsheet software, and this integral can be calculated without using specialised tools (see below).

In the case where regression to the mean bias is suspected, the posterior cumulative distribution function of $\Theta$ is preferably calculated from equation 14. This equation can be written in the following form:

$$F_\Theta(t \mid \mathbf{x}) = \frac{(1+\lambda)^b}{\mathrm{B}(a,b)\mathrm{B}(c,d)} \int_{\theta=0}^{\theta=t} \int_{\eta} \frac{\theta^{a-1}}{(1+\lambda+\theta\eta)^{a+b}} \frac{\eta^{a+c-1}}{(1+\eta)^{c+d}} \mathrm{d}\eta \mathrm{d}\theta \qquad (19)$$

where the parameters $a$, $b$, $c$ and $d$ take the values $a = x_2+½$, $b = x_1+\alpha-½$, $c = x_4+½$, $d = x_3+½$ for the prior given by equation 10, and the values $a = \Sigma x_{2i}+½$, $b = \Sigma x_{1i}+n\alpha-½$, $c = \Sigma x_{4k}+½$, $d = \Sigma x_{3k}+½$ in case of multiple sites considered as a whole, with the prior given in equation 15. After the three successive changes of variable $z(\eta) = \eta/(1+\eta)$, $\omega(\theta) = \theta z/((\theta-1-\lambda)z + 1 + \lambda)$ and lastly $u(z) = \mathrm{Betacdf}_{c,d}(z)$, this integral becomes

$$F_\Theta(t \mid \mathbf{x}) = \int_{u=0}^{u=1} \mathrm{Betacdf}_{a,b}\left( \frac{t \times \mathrm{Betacdf}_{c,d}^{-1}(u)}{1+\lambda+(t-1-\lambda)\times \mathrm{Betacdf}_{c,d}^{-1}(u)} \right) \mathrm{d}u \qquad (20)$$

We will propose (on our website) a spreadsheet calculating $F_\Theta(t \mid \mathbf{x})$ for any given value of $t$, given the accident counts and the prior choice (among the priors mentioned in this paper), for a common spreadsheet software package. The calculation of the integrals of equations 18 and 20 is based on a classical trapezoidal quadrature method, with an increase of the number of calculation points in the vicinity of 0 and 1. The reliability of the results was carefully checked by comparing them to the results obtained with a more powerful software tool using other quadrature methods (adaptive Simpson quadrature and Lobatto quadrature), for a large range of possible values for accident counts $x_1$, $x_2$, $x_3$, and $x_4$.

*NB*: Simplified calculations can be equivalently used in the special case where the counts $x_3$ and $x_4$ are very large, since the trend parameter $\eta$ can then be considered as non-random and approximately equal to $x_4 / x_3$. The problem then reduces to a two-parameter problem ($\theta$ and $\mu_1$), with two random observations $x_1$ and $x_2$. In this situation, the Jeffreys's rule prior is again $\pi(\theta, \mu_1) \propto (1/\theta)^{½}$. In the case where regression to the mean should be taken into account, the method described in section 4 leads to the following prior: $\pi(\theta, \mu_1) \propto (1/\theta)^{½}\, \mu_1^{\alpha-1}\exp(-\lambda\mu_1)$. With the Jeffreys's rule prior, the same kind of calculations as those described in section 3 to 6 lead to the posterior cumulative distribution function

$$F_\Theta(t \mid x_1, x_2) = \int_{\theta=0}^{\theta=t} \frac{(1/\eta)^{x_1+½}}{\mathrm{B}(x_2+½, x_1+½)} \frac{\theta^{x_2-½}}{(\theta+1/\eta)^{x_2+x_1+1}} \mathrm{d}\theta \qquad (21)$$

This is a Beta-prime distribution function with three parameters $x_2$+½, $x_1$+½ and $1/\eta$. After a change of variable $z(\theta) = \theta/(\theta + 1/\eta)$, this integral can be written in the form of a Beta cumulative distribution function:

$$F_\Theta(t \mid x_1, x_2) = \text{Betacdf}_{a,b}\left(\frac{t}{t+1/\eta}\right) \qquad (22)$$

where $a = x_2$+½ and $b = x_1$+½. In the case where the prior is $\pi(\theta, \mu_1) \propto (1/\theta)^{½}\, \mu_1^{\alpha-1}\exp(-\lambda\mu_1)$, the calculation leads to the following result:

$$F_\Theta(t \mid x_1, x_2) = \text{Betacdf}_{a,b}\left(\frac{t}{t+(1+\lambda)/\eta}\right) \qquad (23)$$

where $a = x_2$+½ and $b = x_1+\alpha$–½.

## References


1. Abbess C, Jarrett D, Wright CC (1981) Accidents at blackspots: estimating the effect of remedial treatment, with special reference to the 'regression-to-mean' effect. Traffic Eng Control 22:535-542
2. Agresti A, Hitchcock DB (2005) Bayesian inference for categorical data analysis. Stat Methods Appl 14:297-330
3. Al-Masaied HR, Sinha KC, Kuczek T (1993) Evaluation of safety impact of highway projects. Transp Res Rec 1401:9-16
4. Aul N, Davis G (2006) Use of propensity score matching method and hybrid Bayesian method to estimate crash modification factors of signal installation. Transp Res Rec 1950:17-23
5. Belia S, Fidler F, Williams J, Cumming G (2005) Researchers misunderstand confidence intervals and standard errors. Psychol Methods 10:389-396
6. Berger J (1985) Statistical decision theory and Bayesian analysis. Springer, New York
7. Berger J (2006) The case for objective Bayesian analysis. Bayesian Anal 1:385-402
8. Berger JO, Bernardo JM (1992) On the development of the reference prior method. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds) Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting. Clarendon Press, Oxford, pp 35-60
9. Bernardo JM (1979) Reference posterior distributions for Bayesian inference. J R Stat Soc Series B Stat Methodol 41:113-147
10. Berry DA (1995) Decision analysis and Bayesian methods in clinical trials. Cancer Treat Res 75:125-154
11. Berry DA (1997) Teaching elementary Bayesian statistics with real applications in science. Am Stat 51:241-246
12. Bin Ibrahim K, Metcalfe AV (1993) Bayesian overview for evaluation of mini-roundabouts as a road safety measure. Statistician 42:525-540
13. Box GEP, Tiao GC (1973) Bayesian inference in statistical analysis. Addison-Wesley, Reading
14. Brenac T (1994) Accidents en carrefour sur routes nationales, modélisation du nombre d'accidents prédictible sur un carrefour et exemples d'applications. INRETS report 185. INRETS, Arcueil (France)
15. Cohen J (1994) The earth is round (p<0.05). Am Psychol 49:997-1003
16. De Brabander B, Nuyts E, Vereeck L (2005) Road safety effects of roundabouts in Flanders. J Saf Res 36:289-296
17. Ghosh JK, Delampady M, Samanta T (2006) An introduction to Bayesian analysis, theory and methods. Springer, New York
18. Goodman SN (2005) Introduction to Bayesian methods, I: measuring the strength of evidence. Clin Trials 2:282-290
19. Haller H, Krauss S (2002) Misinterpretations of significance: A problem students share with their teachers? Methods Psychol Res 7:1-20
20. Hasofer AM (1970) On the representation of ignorance in Poisson processes. J R Stat Soc Series B Stat Methodol 32:268-271
21. Hauer E (1983a) Reflections on methods of statistical inference in research on the effect of safety countermeasures. Accid Anal Prev 15:275-285
22. Hauer E (1983b) An application of the likelihood/Bayes approach to the estimation of safety countermeasure effectiveness. Accid Anal Prev 15:287-298
23. Hauer E (1997) Observational before-after studies in road safety. Pergamon, Oxford
24. Hauer E (2004) The harm done by tests of significance. Accid Anal Prev 36:495-500
25. Jaynes ET (1968) Prior probabilities. IEEE Trans Syst Sci Cybern 4:227-241

26. Jeffreys H (1961) Theory of probability, 3rd edn. Oxford University Press, London
27. Johnson DH (1999) The insignificance of statistical significance testing. J Wildl Manag 63:763-772
28. Kadane JB (1995) Prime time for Bayes. Control Clin Trials 16:313-318
29. Kass RE, Wasserman L (1996) The selection of prior distribution by formal rules. J Am Stat Assoc 91:1343-1370
30. Lan B, Persaud B, Lyon C, Bhim R (2008) Validation of a full Bayes methodology for observational before-after road safety studies and application to evaluation of rural signal conversions. Transportation Research Board annual meeting, Washington, 2008
31. Lecoutre B (1999) Beyond the significance test controversy: Prime time for Bayes? Bull Int Stat Inst LVIII(2):205-208
32. Lecoutre B, Lecoutre MP, Poitevineau J (2001) Uses, abuses and misuses of significance tests in the scientific community: Won't the Bayesian choice be unavoidable? Int Stat Rev 69:399-418
33. Leden L, Hämäläinen O, Manninen E (1998) The effect of resurfacing on friction, speeds and safety on main roads in Finland. Accid Anal Prev 30:75-85
34. Maher MJ (1987) Fitting probability distributions to accident frequency data. Traffic Eng Control 28:356-357
35. Miranda-Moreno LF, Fu L (2007) Traffic safety study: empirical Bayes or full Bayes? Transportation Research Board annual meeting, Washington, 2007
36. Mountain L, Fawaz B, Sineng L (1992) The assessment of changes in accident frequencies on link segments: a comparison of four methods. Traffic Eng Control 33:429-431
37. Pawlovich MD, Li W, Carriquiry A, Welch T (2006) Iowa's experience with road diet measures, use of Bayesian approach to assess impacts on crash frequencies and crash rates. Transp Res Rec 1953:163-171
38. Persaud B, Lyon C (2007) Empirical-Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. Accid Anal Prev 39:546-555
39. Robert CP (2007) The Bayesian choice, from decision-theoretic foundations to computational implementation, 2nd edn. Springer, New York
40. Welch BL, Peers HW (1963) On formulae for confidence points based on intervals of weighted likelihoods. J R Stat Soc Series B Stat Methodol 25:318-329