PredictiveValues: A Program for Converting Information on Disease Prevalence and the Characteristics of One, Two or Three Independent Imperfect Diagnostic Tests into Information on Predictive Values

(version 2.1, September 2017)

1. Introduction

PredictiveValues easily converts information on disease prevalence and the characteristics of one, two or three independent diagnostic tests into information on predictive values.

When perfect diagnostic tests are available, a positive test always means the patient is disease positive, and a negative test rules out the disease. However, if imperfect tests are used, false positive and false negative cases arise. It is then of interest to know the probability that a subject with a given set of tests results is in fact positive or negative. This, in turn, depends on the sensitivity and specificity of the tests used. This program assumes that the sensitivity and specificity of each test used are only partially known, that is, these properties are known with uncertainty. These properties are entered into the program, along with the prior probability of disease, also assumed to be imperfectly known. The program then computes the probabilities of being truly positive for all possible combinations of tests results, as well as the uncertainty surrounding these probabilities.

2. How to use **PredictiveValues**

The inputs required for this program are:

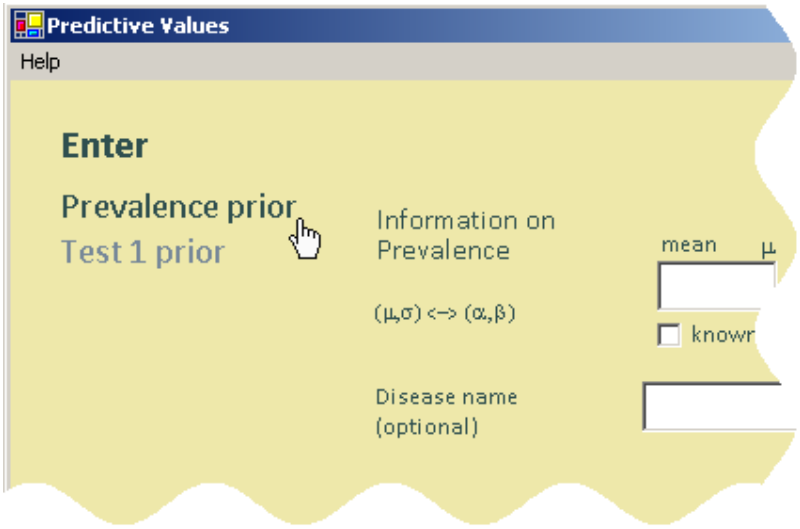
• information on disease prevalence in the form of point estimates and the standard deviation of these point estimates, and
• information on the sensitivity and specificity of each diagnostic test, also input as point estimates and the standard deviation of these point estimates.

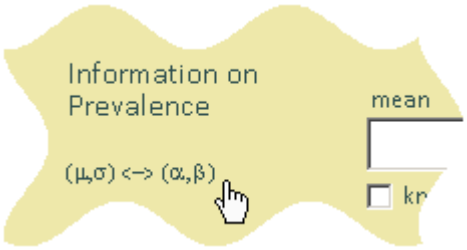
Internal to the program, information on the means and standard deviations as entered above for each parameter are converted to parameters from a Beta density, α and β[1]. The Beta distribution parameters can also be entered directly, but most users will probably find it easier to enter a mean and a standard deviation as this is how tests properties are usually expressed in the medical literature.

[1] Beta distribution with parameters (α,β) has a mean of $\alpha/(\alpha+\beta)$ and variance of $\alpha\beta/(\alpha+\beta)^2(\alpha+\beta+1)$.

After installation, open **PredictiveValues** by clicking on the Start button on your desktop and browse to **PredictiveValues** in the Programs menu. This brings up the Graphical User Interface below, where you will enter the your information on the prevalence of disease and diagnostic tests characteristics.

You may first enter your information on the prior probability of disease (typically called the prevalence of disease) for your population by clicking the **Prevalence prior** label on the left-hand side menu.





The (μ,σ) <–> (α,β) label allows you to express your information in terms of Beta distribution moments (α,β) rather than in terms of mean and s.d. (μ,σ), if you feel comfortable with this parameterization of the Beta distribution.

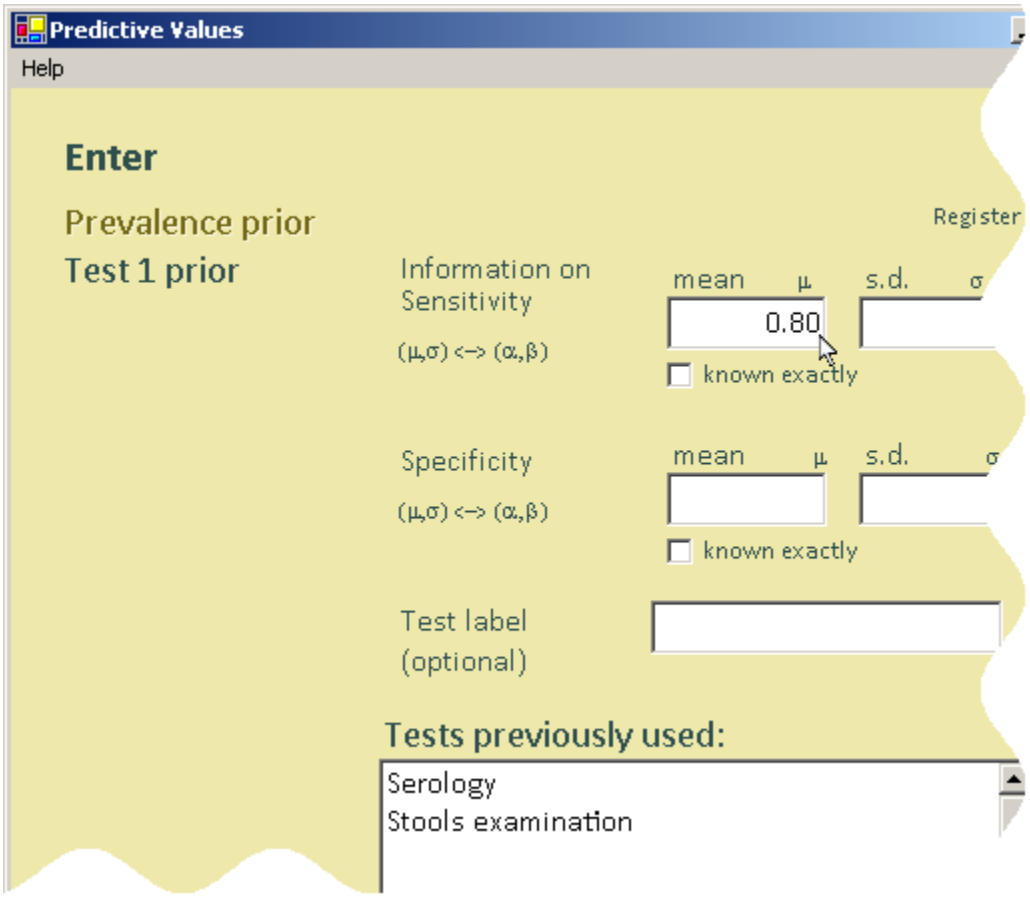
The tick box labelled *known exactly* allows calculating of predictive values under the assumption that the prevalence, the sensitivities or specificities are exactly known; although a parameter's value is hardly ever exactly known in practice, it may be of interest to run the program with some or all parameters considered as exactly known, later comparing the results to those from the more realistic case when all parameters are considered as being uncertain.

Entering a label in the *Disease name* cell will make the corresponding disease prevalence information only one click away the next time you run this program. Up to 100 prevalence values can be stored in this way, for quick and easy recall.

Click on the cells labelled mean and s.d. to enter your values, and click *Register* to proceed to the first diagnostic test description.

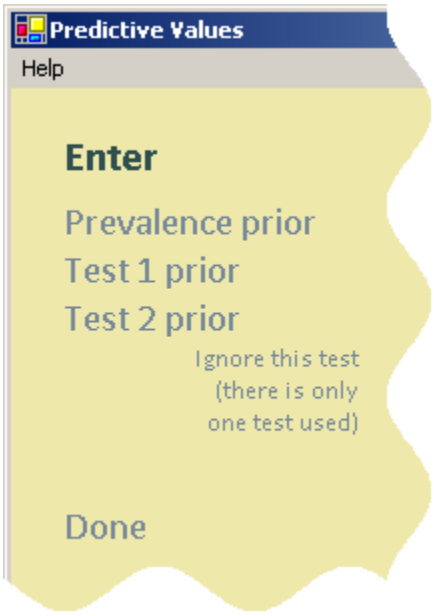
The click the **Test 1 prior** label to enter information on sensitivity and specificity of your first diagnostic test.

Fill in the four cells for mean and s.d. for both test sensitivity and test specificity.



Once again, clicking on the (μ,σ) <–> (α,β) label allows you to directly enter Beta parameters, if you prefer, and once again, the tick boxes allow you to assume perfect information on the test properties.

Entering a label in *Test label* cell will make this test information only one click away the next time you run this program. Up to 100 sets of tests properties can be saved for instant recall in this way.



Click **Register** to validate and record your parameter values.

A **Test 2 prior** label will now appear below the **Test 1 prior** label on the left-hand side menu.

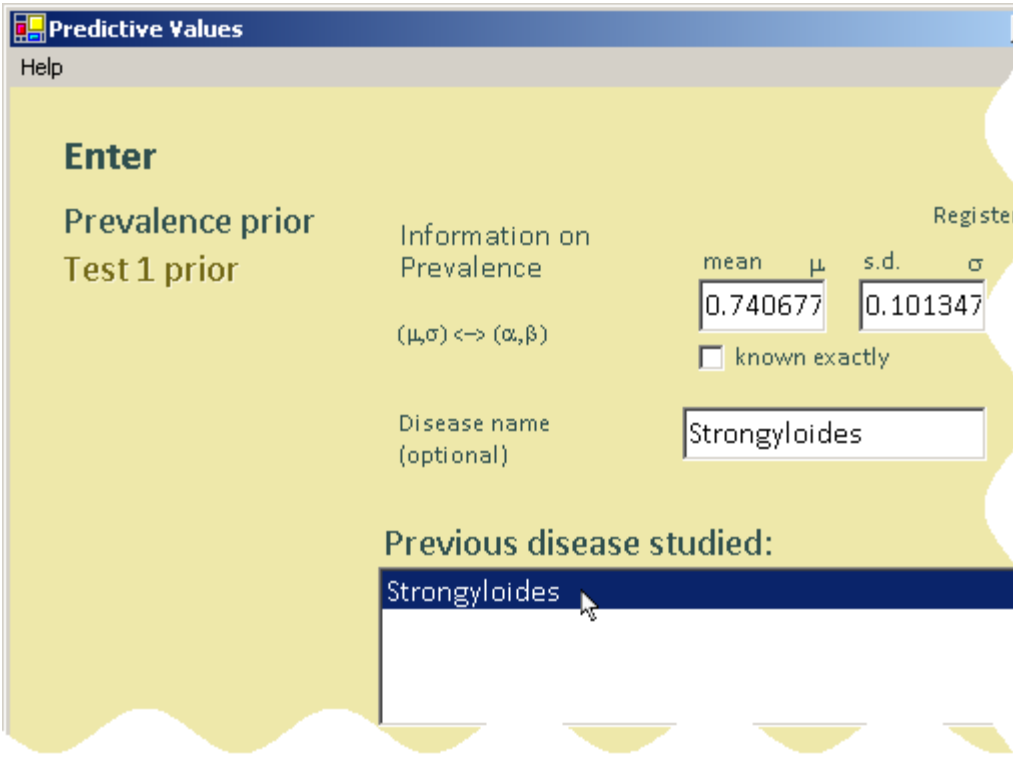
If you are using information from more than one test, click on the cell **Test 2 prior** label; if not, clicking **Done** will display a message encouraging you to review the parameters entered for information on prevalence and diagnostic tests used: this is done by clicking each of corresponding left-hand side menu items. Had you entered a parameter incorrectly in the earlier steps, this is your opportunity to change any incorrectly entered/mistyped value.

After reviewing all parameters descriptions, click the **Proceed to calculations** label. Calculations will vary in run times from a few seconds (most of the time) up to a few hours (very rarely), depending on the specific information entered. We will now illustrate these steps as well as show how to interpret the output with an example.

3. An Example of Running **PredictiveValues**

To illustrate use of this program, we will summarize the information on predictive values when using Serology and Stool Examination in the diagnosis of Strongyloides (Joseph L, Gyorkos T, Coupal L, 1995). Note that the information on prevalence of Strongyloides and Serology and Stool Examination sensitivities and specificities was included in the program to serve as a first example. Also note that no claim is made as to the medical accuracy of these inputs, this is an illustrative example only.

Upon clicking the **Prevalence prior** label, text boxes to enter your prior information appear as well as a list box presenting the diseases analysed previously: click *Strongyloides* in that box, labelled *Previous diseases studied.*



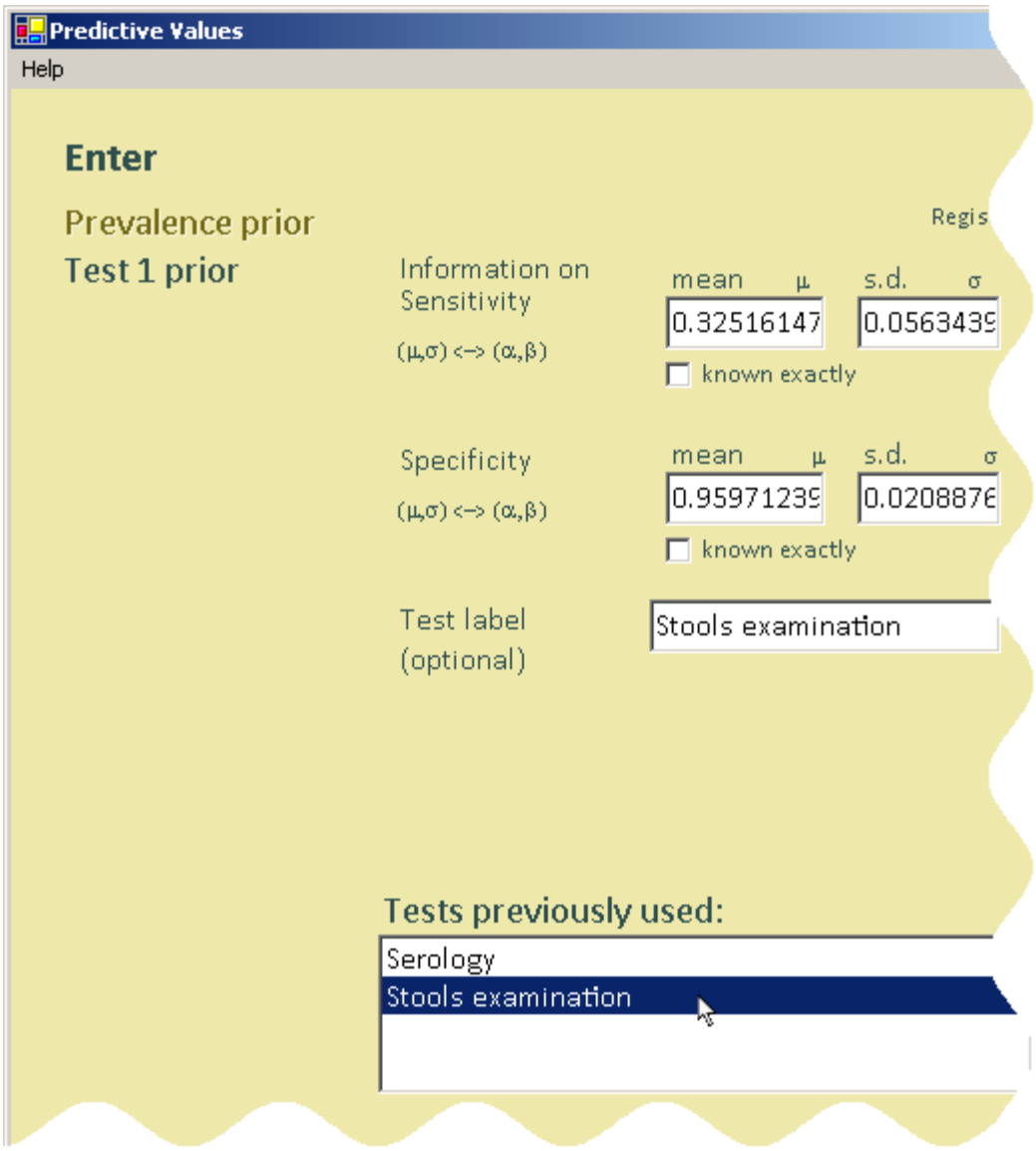
Note that the values are immediately inserted: with a mean of about 0.74 and a standard deviation of about 0.1, we are saying that the prior probability of disease is approximately[2] in the range of 0.54 to 0.94 (mean ± 2 s.d.).

Click **Register** to confirm prevalence prior description and then click the **Test 1 prior** label button to enter your information on first diagnostic test.

[2] The precision of the approximation depends on how close the chosen Beta distribution is to a normal distribution.

Next, select *Stools examination* in the lower list labelled *Tests previously used.*

With a mean of 0.325 and a s.d. of 0.056, sensitivity is expected to be in the range of 0.213 to 0.437, while specificity (with a mean ≈ 0.96 and a s.d. ≈ 0.02) is expected to be in the higher range of 0.92 to 0.999.

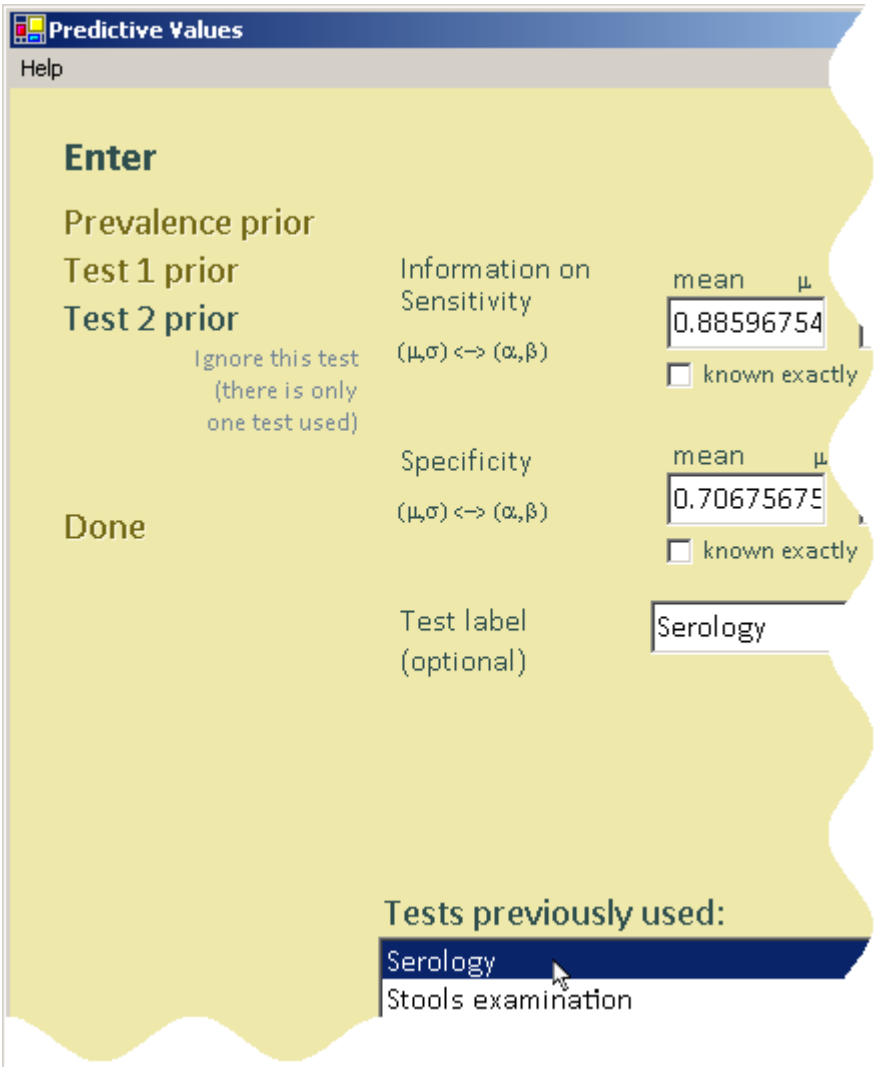


Click **Register** to confirm test 1 prior description. An item labeled **Test 2 prior** will appear on left-hand side menu; click it.

Select *Serology* in the list of tests; means and s.d.'s indicate that Serology's sensitivity is known to be in the range of 0.809 to 0.963, while its specificity is in the range of 0.39 to 0.999.

Click **Register** to confirm test 1 prior description. An item labelled **Test 3 prior** will appear on left-hand side menu, which we will ignore since we have only two diagnostic tests at hand for this example.

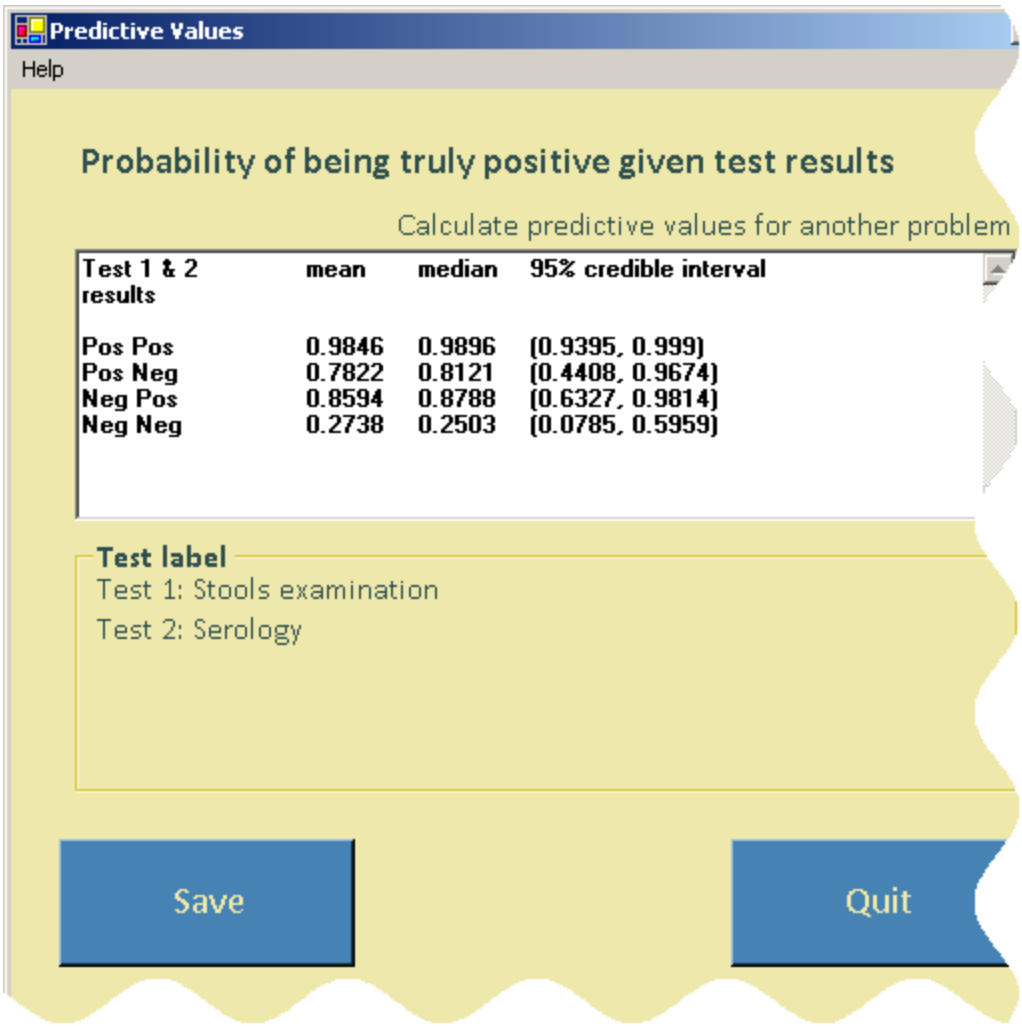
Click the **Done** label.



At this point, we encourage you to check all information and labels carefully by clicking again on each of the left-hand side menu items, which will redisplay the corresponding prior distribution parameters on screen. If all is correct, begin calculating predictive values by clicking the **Proceed to calculations** menu item.

Click the button "*It's all fine. Proceed to Predictive Values calculations >>"* to proceed when you have validated all of the information.

After the required computation time (progress bars will be displayed), information on the predictive values for each possible combination of tests results will be displayed.



Click **Save** to save results as an html or a text file (you will be prompted for the output file location) or quit without saving by clicking on the **Quit** button. You may want to click the label **Calculate predictive values for another problem** (top-right corner of results table) to run another problem.

As a result of the uncertainty in the prevalence and test sensitivities and specificities, predictive values are also uncertain. Uncertainty around predictive values is summarized in the results form. In the above example, positive results on both tests leads to a high probability of being truly positive, with an average estimated probability of 98.46%; we are 95% sure (cf. 95% credible interval) that this probability lies within the 93.95-99.9% range, and there is a 50% chance that the probability is higher than 98.96% (median). When only one test is positive, the average estimate of the probability of being truly positive goes down to 85.94% if Serology (test 2) is positive, while it goes down to 78.22% if Stools examination (test 1) shows the only positive result. Credible intervals are shifted down towards lower probabilities accordingly. It is interesting to note that even when a patient tests negative on both tests, the probability of being truly positive is still quite high, with an average estimate of 27.4% and a credible interval that includes values as high as 59.6%.

IMPORTANT: Results will slightly differ from one run to another, as results are based on Monte Carlo simulations (thus, subject to random variation). Also note that inputting very small s.d.'s might considerably lengthen run times; as an approximate workaround, one could obtain faster results by assuming that the parameter's value is exactly known. Nevertheless it is highly recommended that these approximate results be compared to those obtained from the inputs that more realistically incorporate any uncertainty.

4. Change log

Version 1.1 (September 2005)
You can easily access instructions file through top menu item *Help*.

Version 1.2 (February 2007)
Each parameter can be assumed to be known exactly.

Versions 1.3 and 1.3.1 (June 2008)
Earlier versions accepted commas in numeric inputs, which, depending on their placement, could have led to unintended inputs being used. If a comma is found, a pop-up box now asks you to remove it, eliminating all ambiguity.

Version 1.4 (March 2010)
It is now possible to enter your prior information in terms of (α, β) without first entering values for (μ, σ) in each entry form. Previously, you counter-intuitively had to enter values for (μ, σ) before being allowed to switch to the alternative (α, β) parametrization in the diagnostic tests form.

Versions 1.5 and 1.5.1 (October 2010)
Graphic presentation was slightly revisited.

Version 1.6 (November 2010)
It is now possible to save output results to an .html file.

Versions 1.7 and 1.7.1 (November 2010)
The Graphical User Interface now accepts either commas or periods as decimal symbols (in numeric inputs), depending on the value set for **Decimal symbol** in the **Customized Regional Options** form of the **Regional and Language Options** in the **Control Panel**. Both commas and periods cannot be used at the same time, you must use the option chosen for your computer.

Version 1.7.3 (April 2015)
Minor bug fix update: a potential installation problem was fixed.

Version 2.0 (February 2016)
Revised Graphical User Interface with some extra features.

Version 2.1 (September 2017)
PredictiveValues now works on Windows 8 & 10. Windows 7 users do not need to reinstall or upgrade.

5. Bibliography

Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard
Joseph L, Gyorkos T, Coupal L
American Journal of Epidemiology 1995; 141(3), pp. 263-272

Questions? Please send email to: Lawrence.Joseph@McGill.ca

Other Bayesian software packages are available at
http://www.medicine.mcgill.ca/epidemiology/Joseph.