

# A software tool for exploring the relation between diagnostic accuracy and measurement uncertainty

Theodora Chatzimichail

Hellenic Complex Systems Laboratory

Aristides T. Hatjimihail (✉ [ath@hcsl.com](mailto:ath@hcsl.com))

Hellenic Complex Systems Laboratory <https://orcid.org/0000-0002-0502-7996>

---

## Software

**Keywords:** diagnostic accuracy measures, ROC curve, measurement uncertainty, diagnostic tests, screening tests, risk

**Posted Date:** January 14th, 2020

**DOI:** <https://doi.org/10.21203/rs.2.20819/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on August 19th, 2020. See the published version at <https://doi.org/10.3390/diagnostics10090610>.

# A software tool for exploring the relation between diagnostic accuracy and measurement uncertainty

***Theodora Chatzimichail, MD<sup>a</sup>, Aristides T. Hatjimihail, MD, PhD<sup>a</sup>***

<sup>a</sup> Hellenic Complex Systems Laboratory, Kostis Palamas 21, Drama 66131, Greece

Theodora Chatzimichail, MD

Address: Hellenic Complex Systems Laboratory, Kostis Palamas 21, Drama 66131, Greece

Phone: +30 2521058227

Mobile: +44 7745991039

Email: [tc@hcsl.com](mailto:tc@hcsl.com)

## Corresponding author:

Aristides T Hatjimihail, MD, PhD

Address: Hellenic Complex Systems Laboratory, Kostis Palamas 21, Drama 66131, Greece

Phone: +30 2521058227

Mobile: +30 6937017876

Email: [ath@hcsl.com](mailto:ath@hcsl.com)

# ***ABSTRACT***

---

## **BACKGROUND**

Screening and diagnostic tests are used to classify people with and without a disease. Diagnostic accuracy measures are used to evaluate the correctness of a classification in clinical research and practice. Although the correctness of a classification based on a measurand depends on the uncertainty of measurement, there has been limited research on their relation. The objective for this work is to develop an exploratory tool for the relation between diagnostic accuracy measures and measurement uncertainty, as diagnostic accuracy is fundamental to clinical decision making, while measurement uncertainty is critical to quality and risk management in laboratory medicine.

## **RESULTS**

For this reason, a freely available interactive program has been developed, written in Wolfram Language. The program provides four modules for calculating, optimizing, plotting and comparing various diagnostic accuracy measures and the corresponding risk of diagnostic or screening tests measuring a normally distributed measurand, applied at a single point in time in non-diseased and diseased populations. This is done for differing prevalence of the disease, mean and standard deviation of the measurand, diagnostic threshold, standard measurement uncertainty of the tests and expected loss.

The application of the program is illustrated with a case study of glucose measurements in diabetic and non-diabetic populations, that demonstrates the relation between diagnostic accuracy measures and measurement uncertainty.

## **CONCLUSION**

The presented interactive program is user-friendly and can be used as a flexible educational and research tool in medical decision making, to explore the relation between diagnostic accuracy measures and measurement uncertainty.

***Keywords:*** diagnostic accuracy measures; ROC curve; measurement uncertainty;

diagnostic tests; screening tests; risk

# 1 INTRODUCTION

---

An increasing number of in vitro screening and diagnostic tests are extensively used as binary classifiers in medicine, to classify people in the non-overlapping classes of populations with and without a disease, which are categorized as quantitative and qualitative. The quantitative and many qualitative screening or diagnostic tests are based on measurements. There is a joint probability distribution of the measurements in the diseased and non-diseased populations. To classify the patients with and without a disease using a test based on a measurement, a diagnostic threshold or cutoff point is defined. If the measurement is above the threshold the patient is classified as test-positive, otherwise as test-negative (or vice versa) (Fig 1). The possible test results are summarized in Table 1.

**Table 1.** A 2x2 contingency table

		populations	
		nondiseased	diseased
test results	negative	true negative ( <i>TN</i> )	false negative ( <i>FN</i> )
	positive	false positive ( <i>FP</i> )	true positive ( <i>TP</i> )

## 1.1 DIAGNOSTIC ACCURACY MEASURES

From the large number of diagnostic accuracy measures (DAM) appearing in literature, only a few are used for evaluating the diagnostic accuracy in clinical research and practice (1).

These include:

1. Sensitivity ( $Se$ ), specificity ( $Sp$ ), overall diagnostic accuracy ( $ODA$ ), diagnostic odds ratio ( $DOR$ ), likelihood ratios for positive or negative result ( $LR+$  and  $LR-$  respectively), that are defined conditionally on the true disease status (2) and are prevalence invariant.
2. Overall diagnostic accuracy ( $ODA$ ), that is defined conditionally on the true disease status and is prevalence dependent.
3. Positive predictive and negative predictive values ( $PPV$  and  $NPV$ ), that are defined conditionally on the test outcome and are prevalence dependent.

The natural frequency and the equivalent probability definitions and of the diagnostic accuracy measures derived from Table 1 and analyzed by the program, are presented in Table 2. The symbols are explained in Section 12 (Notation).

**Table 2:** Natural frequency and probability definitions of diagnostic accuracy measures

measure	natural frequency definition	probability definition
Se	$\frac{TP}{FN + TP}$	$Pr(T D)$
Sp	$\frac{TN}{TN + FP}$	$Pr(\bar{T} \bar{D})$
PPV	$\frac{TP}{FP + TP}$	$Pr(D T)$
NPV	$\frac{TN}{TN + FN}$	$Pr(\bar{D} \bar{T})$
ODA	$\frac{TN + TP}{TN + FN + TP + FP}$	$Pr(D) Pr(T D) + Pr(\bar{D}) Pr(\bar{T} \bar{D})$
DOR	$\frac{TN TP}{FN FP}$	$\frac{Pr(T D)}{Pr(\bar{T} \bar{D})} \cdot \frac{Pr(\bar{T} \bar{D})}{Pr(T D)}$
LR+	$\frac{TP(FN + TN)}{FP(FN + TP)}$	$\frac{Pr(T D)}{Pr(\bar{T} \bar{D})}$
LR-	$\frac{FN(FN + TN)}{TN(FN + TP)}$	$\frac{Pr(\bar{T} \bar{D})}{Pr(T D)}$
J	$\frac{TN TP - FN FP}{(TN + FP)(FN + TP)}$	$Pr(T D) + Pr(\bar{T} \bar{D}) - 1$
ED	$\sqrt{\left(\frac{FN}{FN + TP}\right)^2 + \left(\frac{FP}{TN + FP}\right)^2}$	$\sqrt{Pr(\bar{T} \bar{D})^2 + Pr(T D)^2}$
CZ	$\frac{TN TP}{(TN + FP)(FN + TP)}$	$Pr(T D) Pr(\bar{T} \bar{D})$
R	$l_0 + \frac{l_{TN}TN + l_{FN}FN + l_{TP}TP + l_{FP}FP}{TN + FN + TP + FP}$	$l_0 + l_{TN}Pr(\bar{D})Pr(\bar{T} \bar{D}) + l_{FN}Pr(D)Pr(\bar{T} \bar{D}) + l_{TP}Pr(D)Pr(T D) + l_{FP}Pr(\bar{D})Pr(T D)$

The symbols are explained in Section 12 (Notation).

Receiver operating characteristic (*ROC*) curves have been also used for the evaluation of the diagnostic performance of a screening or diagnostic test (3). *ROC* curves are plots of  $Se$  vs  $1-Sp$  of the test.

A related summary measure of diagnostic accuracy is the area under a *ROC* curve (*AUC*) (4, 5). The area over a *ROC* curve (*AOC*) has been proposed as a complementary summary measure of the diagnostic inaccuracy (6).

Recently, the predictive receiver operating characteristic (*PROC*) curves have been also proposed. *PROC* curves are plots of  $PPV$  vs  $1-NPV$  of the test (2).

For the optimization of binary classifiers, objective or loss functions have been proposed. They are based on diagnostic accuracy measures that can be maximized or minimized by finding the optimal diagnostic threshold. These measures include Youden's index  $J$ (7), Euclidean distance of a *ROC* curve point from the point (0, 1) (*ED*) (8) and the concordance probability measure (*CZ*) (9). The above-mentioned measures are defined conditionally on the true disease status and are prevalence invariant. Their respective probability and natural frequency definitions are presented in Table 2.

## 1.2 RISK

The risk of a diagnostic or screening test is related to its diagnostic accuracy and is defined as its expected loss. Therefore, it depends upon (Table 2):

1. The expected loss for the testing procedure, for a true negative result, for a false negative result, for a true positive result and for a false positive result.
2. The probabilities for a true negative result, for a false negative result, for a true positive result and for a false positive result.

### **1.3 MEASUREMENT UNCERTAINTY**

As there is inherent variability in any measurement process, there is measurement uncertainty, which is defined as a “parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand” (10). The parameter may be the standard measurement uncertainty ( $u$ ), expressed as a standard deviation and estimated as described in “Expression of Measurement Uncertainty in Laboratory Medicine” (11). Bias may be considered as a component of the standard measurement uncertainty (12).

The uncertainty is gradually replacing the total analytical error concept (13).

### **1.4 RELATION BETWEEN DIAGNOSTIC ACCURACY AND MEASUREMENT UNCERTAINTY**

Although the estimation of measurement uncertainty is essential for quality assurance in laboratory medicine (11), its effect on clinical decision making and consequently on clinical outcomes is rarely quantified (14). As direct outcome studies are very complex, a feasible first step is exploring the effect of measurement uncertainty on misclassification (15) and subsequently on diagnostic accuracy measures and the corresponding risk. Exploring this relation could assist the process of estimation of the optimal diagnostic threshold or the permissible measurement uncertainty.

## 2 IMPLEMENTATION

---

### 2.1 COMPUTATIONAL METHODS

For the calculation of the diagnostic accuracy measures it is assumed that:

1. Either the values of the measurand or their transforms are normally distributed in each of the diseased and non-diseased populations.
2. Or the measurement uncertainty is normally distributed and uncorrelated with the distribution of the measurand.

Thereafter, we use the term measurand to describe either the normally distributed value of a measurand or its normally distributed applicable transform.

Consequently, if  $\sigma$  is the standard deviation of the measurements of a screening or diagnostic test applied in a population  $P$ ,  $u$  the standard measurement uncertainty and  $\sigma_p$  the standard deviation of the measurand in the population, then

$$\sigma = \sqrt{\sigma_p^2 + u^2} \quad 1$$

The definitions of the diagnostic accuracy measures can be expressed in terms of sensitivity ( $Se$ ) and specificity ( $Sp$ ). These definitions are derived from table 2 and presented in Table 3.

**Table 3.** Definitions of diagnostic accuracy measures versus sensitivity and specificity

measure	definition
$PPV$	$\frac{Se v}{Se v + (1 - Sp)(1 - v)}$

<i>NPV</i>	$\frac{Sp(1-v)}{Sp(1-v) + (1-Se)v}$
<i>ODA</i>	$Sev + Sp(1-v)$
<i>DOR</i>	$\frac{\frac{Se}{1-Se}}{\frac{1-Sp}{Sp}}$
<i>LR +</i>	$\frac{Se}{1-Sp}$
<i>LR -</i>	$\frac{1-Se}{Sp}$
<i>J</i>	$Se + Sp - 1$
<i>ED</i>	$\sqrt{(1-Se)^2 + (1-Sp)^2}$
<i>CZ</i>	$Se Sp$
<i>R</i>	$l_0 + l_{TN}Sp(1-v) + l_{FN}(1-Se)v + l_{TP}Sev + l_{FP}(1-Sp)(1-v)$

The symbols are explained in Section 12 (Notation).

The functions of sensitivity (*Se*) and specificity (*Sp*), hence the functions of all the above diagnostic accuracy measures, can be expressed in terms of the cumulative distribution function of the normal distribution, therefore of the error function and the complementary error function.

The error function  $erf(x)$  is defined as:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt, x \geq 0 \quad 2$$

while the complementary error function  $erfc(x)$  is defined as:

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt, x \geq 0 \quad 3$$

Following the definition of the sensitivity and specificity of a test (Table 2), the respective functions versus diagnostic threshold  $d$  are calculated as:

$$se(d, \mu_D, \sigma_D, u) = 1 - \Psi\left(d, \mu_D, \sqrt{\sigma_D^2 + u^2}\right) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{-d + \mu_D}{\sqrt{2(\sigma_D^2 + u^2)}} \right) \right) \quad 4$$

$$sp(d, \mu_{\bar{D}}, \sigma_{\bar{D}}, u) = \Psi\left(d, \mu_{\bar{D}}, \sqrt{\sigma_{\bar{D}}^2 + u^2}\right) = \frac{1}{2} \operatorname{erfc} \left( \frac{-d + \mu_{\bar{D}}}{\sqrt{2(\sigma_{\bar{D}}^2 + u^2)}} \right) \quad 5$$

Then, the sensitivity function of a test versus its specificity  $z$  is calculated as:

$$se_{sp}(z, \mu_D, \sigma_D, \mu_{\bar{D}}, \sigma_{\bar{D}}, u) = 1 - \Psi\left(\Psi^{-1}\left(z, \mu_{\bar{D}}, \sqrt{\sigma_{\bar{D}}^2 + u^2}\right), \mu_D, \sqrt{\sigma_D^2 + u^2}\right) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{\mu_D - \mu_{\bar{D}} + \sqrt{2(\sigma_{\bar{D}}^2 + u^2)} \operatorname{erfc}^{-1}(2z)}{\sqrt{2(\sigma_D^2 + u^2)}} \right) \right), 0 \leq z \leq 1 \quad 6$$

where  $\Psi$  denotes the cumulative distribution function of a normal distribution.

The specificity function of a single test versus its sensitivity  $y$  is calculated as:

$$sp_{se}(y, \mu_D, \sigma_D, \mu_{\bar{D}}, \sigma_{\bar{D}}, u) = \Psi\left(\Psi^{-1}\left(1 - y, \mu_D, \sqrt{\sigma_D^2 + u^2}\right), \mu_{\bar{D}}, \sqrt{\sigma_{\bar{D}}^2 + u^2}\right) = \frac{1}{2} \operatorname{erfc} \left( \frac{-\mu_D + \mu_{\bar{D}} + \sqrt{2(\sigma_D^2 + u^2)} \operatorname{erfc}^{-1}(2 - 2y)}{\sqrt{2(\sigma_{\bar{D}}^2 + u^2)}} \right), 0 \leq y \leq 1 \quad 7$$

Following the Table 3 and the equations 4-7, the diagnostic accuracy measures of a test are defined equivalently as functions of either its diagnostic threshold, or sensitivity, or specificity. Consequently, the derived parametric equations, defining each measure, can be used to explore the relations between any two measures.

Following the definition of the *ROC* curves and assuming a normal probability density function of the measurands of each of the diseased and non-diseased populations, the *ROC* function is calculated as:

$$roc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = S\left(S^{-1}\left(t, \mu_{\bar{D}}, \sqrt{\sigma_{\bar{D}}^2 + u^2}\right), \mu_D, \sqrt{\sigma_D^2 + u^2}\right), 0 \leq t \leq 1 \quad 8$$

where  $S$  denotes the survival function of normal distribution.

Consequently:

$$roc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = \frac{1}{2} \operatorname{erfc}\left(\frac{-\mu_D + \mu_{\bar{D}} + \sqrt{2(\sigma_{\bar{D}}^2 + u^2)} \operatorname{erfc}^{-1}(2t)}{\sqrt{2(\sigma_D^2 + u^2)}}\right), 0 \leq t \leq 1 \quad 9$$

The function of the area under the *ROC* curve is defined as:

$$auc(\mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = \int_0^1 roc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) dt \quad 10$$

and is calculated as:

$$auc(\mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = \frac{\Phi(\mu_D - \mu_{\bar{D}})}{\sqrt{\sigma_{\bar{D}}^2 + \sigma_D^2 + 2u^2}} \quad 11$$

where  $\Phi$  denotes the cumulative distribution function of the unit normal distribution.

The function of the area over the *ROC* curve is defined as:

$$aoc(\mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = 1 - auc(\mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) \quad 12$$

Another *ROC* curve related quantity is the Euclidean distance (*ED*) of a *ROC* curve point

$(t, roc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u))$  from the point  $(0, 1)$  or equivalently the Euclidean distance of

the point  $(Se, Sp)$  from the point  $(1, 1)$  of perfect diagnostic accuracy. The respective

function is defined as follows:

$$ed(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = \sqrt{t^2 + (1 - roc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u))^2} \quad 13$$

The predictive *ROC* (*PROC*) curve relation is defined as follows (2):

$$proc(t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) = ppv(npv^{-1}(1 - t, \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u), \mu_{\bar{D}}, \mu_D, \sigma_{\bar{D}}, \sigma_D, u) \quad 14$$

This relation cannot be expressed in terms of elementary or survival functions.

## **2.2 THE PROGRAM**

To explore the relation between diagnostic accuracy measures or the corresponding risk and measurement uncertainty, an interactive program written in Wolfram Language (16) was developed in Wolfram Mathematica<sup>®</sup>, ver. 12.0 (17). This program was designed to provide four modules and six submodules for calculating, optimizing, plotting and comparing various diagnostic accuracy measures and the corresponding risk of two screening or diagnostic tests, applied at a single point in time in non-diseased and diseased populations (Fig 2). The two tests measure the same measurand, for varying values of the prevalence of the disease, the mean and standard deviation of the measurand in the populations and the standard measurement uncertainty of the tests. The two tests differ in measurement uncertainty. It is assumed that the measurands and the measurement uncertainty are normally distributed.

Parts of this program have been presented in a series of Demonstrations, at Wolfram Demonstration Project of Wolfram Research (6, 18-24).

The program is freely available as a computable document (.cdf) (Supplementary file Relation.cdf). It can be run on Wolfram Player<sup>®</sup> or Wolfram Mathematica<sup>®</sup> (see Section 6).

## **3 RESULTS**

---

### **3.1 INTERFACE OF THE PROGRAM**

The modules and the submodules of the program include panels with controls which allow the interactive manipulation of various parameters, as described in detail in Section 13 (Appendix). These are the following:

### **3.1.1 ROC plots**

The receiver operating characteristic (*ROC*) curves or the predictive receiver operating characteristic (*PROC*) curves of the two tests are plotted.

A table with the respective *AUC* and *AOC* and their relative difference is also presented with the *ROC* curves plot (Fig 3).

### **3.1.2 Diagnostic accuracy measures plots**

It includes the following submodules:

#### ***3.1.2.1 Diagnostic accuracy measures versus standard measurement uncertainty plots***

The values of the diagnostic accuracy measures or the corresponding risk of a test are plotted versus the standard measurement uncertainty of the test (Fig 4).

#### ***3.1.2.2 Diagnostic accuracy measures versus sensitivity or specificity plots***

The values of the diagnostic accuracy measures or the corresponding risk of the two tests, their partial derivatives with respect to standard measurement uncertainty, their difference, relative difference and ratio are plotted versus either the sensitivity or the specificity of each test (Fig 5).

#### ***3.1.2.3 Diagnostic accuracy measures versus sensitivity and specificity plots***

The values of the diagnostic accuracy measures or the corresponding risk of the two tests or their partial derivatives, with respect to standard measurement uncertainty, are plotted versus the sensitivity and the specificity of each test in three-dimensional line plots (Fig 6).

#### ***3.1.2.4 Diagnostic accuracy measures versus diagnostic threshold plots***

The values of the diagnostic accuracy measures or the corresponding risk of the two tests, their partial derivatives with respect to standard measurement uncertainty, their difference, relative difference and ratio are plotted versus the diagnostic threshold of each test (Fig 7).

### ***3.1.2.5 Diagnostic accuracy measures versus prevalence plots***

The values of the diagnostic accuracy measures or the corresponding risk of the two tests, their partial derivatives with respect to standard measurement uncertainty, their difference, relative difference and ratio are plotted versus the prevalence of the disease (Fig 8).

### ***3.1.2.6 Diagnostic accuracy measures relations***

As any two of the diagnostic accuracy measures can be expressed as functions of their sensitivities, their respective parametric equations are plotted to show the relations between the values of the two measures of each test (Fig 9).

### **3.1.3 Diagnostic accuracy measures calculator module**

The values of various diagnostic accuracy measures and the corresponding risk of each of the two tests and their respective relative differences, at a selected diagnostic threshold, are calculated and presented in a table (Fig 10).

### **3.1.4 Optimal diagnostic accuracy measures calculator**

An optimal diagnostic threshold for each test is calculated according to a selected objective or loss function. Then the values of various diagnostic accuracy measures and the corresponding risk of each of the two tests, at the respective optimal threshold, are presented in a table (Fig 11).

## **3.2 ILLUSTRATIVE CASE STUDY**

The program was applied to a bimodal joint distribution of log-transformed blood glucose measurements in non-diabetic and diabetic Malay populations, during an oral glucose tolerance test (OGTT) (25). Briefly, after the ingestion of 75 g glucose monohydrate, the 2-h postprandial blood glucose of 2667 Malay adults, aged 40 - 49 years, was measured with reflectance photometry. The estimated prevalence of diabetes was 0.067. To estimate the

distribution of the measurand in the diabetic and non-diabetic populations it was assumed that the measurement coefficient of variation and bias were equal to 4% and 2% respectively. The log-transformed measurands of each population were normally distributed, as shown in Fig 1. In this case study, the normalized log-transformed measurand means and standard deviations in the diseased and non-diseased populations, the standard measurement uncertainty and the diagnostic threshold were expressed in units equal to the standard deviation of the log-transformed measurand in the non-diseased population. The normalized log-transformed diagnostic threshold 2.26 corresponds to the American Diabetes Association (ADA) diagnostic threshold for diabetes of the 2-h postprandial glucose during OGTT, that is equal to 11.1 mmol/l (26). The normalized log-transformed standard measurement uncertainties 0.023 and 0.23 of the two tests in the Fig 3 and 5-17, correspond to standard measurement uncertainties equal to 1% and 10% of the mean of the measurand of the non-diabetic population.

The results of the illustrative case study are presented:

1. In the plots of Fig 3-9 and 12-16.
2. In the tables of Fig 10 and 12.
3. In table 5.

The parameter settings of Fig 12-17 and Table 5 are presented in Table 4.

**Table 4.** The parameter settings of Fig 12-17 and Table 5

settings	Fig 12	Fig 13	Fig 14	Fig 15	Fig 16	Fig 17	Table 5
$\mu_D$	2.99	2.99	2.99	2.99	2.99	2.99	2.99
$\sigma_D$	0.75	0.75	0.75	0.75	0.75	0.75	0.75
$\mu_{\bar{D}}$	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\sigma_{\bar{D}}$	1.0	1.0	1.0	1.0	1.0	1.0	1.0
$\nu$	0.067	0.067	-	0.067	0.067	0.067	0.067

$d$	2.26	2.26	2.26	-	-	-	
$u_a$	-	-	0.023	0.023	0.023	0.023	0.023
$u_b$	-	-	0.23	0.23	0.23	0.23	0.23
$l_0$	-	-	1	-	1	1	1
$l_{TN}$	-	-	0	-	0	0	0
$l_{FN}$	-	-	100	-	100	100	100
$l_{TP}$	-	-	0	-	0	0	0
$l_{FP}$	-	-	76	-	76	76	76

The symbols of the settings column are explained in Section 12 (Notation).

In this case, the measurement uncertainty has relatively little effect on the *ROC* and *PROC* curves, on *AUC*, sensitivity, specificity, overall diagnostic accuracy, positive predictive value, negative predictive value, Euclidean distance and concordance probability of the test, in accordance with previous findings (27, 28). Measurement uncertainty has relatively greater effect on diagnostic odds ratio, on likelihood ratio for a positive or negative result, Youden's index and risk.

As a result, the measurement uncertainty has relatively little effect on the optimal diagnostic thresholds maximizing the Youden's index or the concordance probability or minimizing the Euclidean distance. Conversely, it has a relatively greater effect on the optimal diagnostic thresholds minimizing risk (Table 5).

**Table 5.** Optimal diagnostic thresholds

			optimal diagnostic threshold		
			first test	second test	relative difference
<b>optimizing DAM</b>	<b>Youden's index</b>	$J$	1.637	1.623	0.009
	<b>Euclidean distance</b>	$ED$	1.676	1.663	0.008
	<b>concordance probability</b>	$CZ$	1.640	1.627	0.008
	<b>Risk</b>	$R$	2.258	2.290	-0.014

The optimal diagnostic thresholds with the respective parameters in Table 4.

## 4 DISCUSSION

---

The purpose of this program is to explore the relation between diagnostic accuracy measures and measurement uncertainty, as diagnostic accuracy is fundamental to clinical decision making, while defining the permissible measurement uncertainty is critical to quality and risk management in laboratory medicine. There has been extensive research on either diagnostic accuracy or measurement uncertainty, however very limited on both subjects (14, 29, 30).

This program demonstrates the relation between the diagnostic accuracy measures and the measurement uncertainty for a screening or diagnostic test measuring a single measurand (Fig 12, 13, 3-11 and 14-19). This relation depends on the population parameters, including the prevalence of the disease (Fig 8 and 14), and on the diagnostic threshold (Fig 5, 14 and 15). In addition, measurement uncertainty affects the relation between any two of the diagnostic accuracy measures (Fig 9 and 17).

As the program provides plots of the partial derivative of the diagnostic accuracy measures with respect to the standard measurement uncertainty, it offers a more detailed insight (Fig 16). In antithesis to the complexity of the relation, the program simplifies its exploration with a user-friendly interface.

Furthermore, it provides calculators for the calculation of the effects of measurement uncertainty on the diagnostic accuracy measures and corresponding risk (Fig 10) and for calculating the diagnostic threshold optimizing the objective and loss functions of Section 1.1 (Fig 11).

The counterintuitive finding that the measurement uncertainty has relatively little effect on the *ROC* and *PROC* curves, on *AUC*, sensitivity, specificity, overall diagnostic accuracy,

positive predictive value, negative predictive value, Euclidean distance and concordance probability suggests that we should reconsider their interpretation in medical decision making. However, further research is needed to explore the effect of measurement uncertainty on diagnostic accuracy measures with different clinically and laboratory relevant parameter settings.

Compared to risk measure, a shortcoming of Youden's index  $J$ , Euclidean distance of a *ROC* curve point from the point (0, 1) and concordance probability as objective functions, is that they do not differentiate the relative significance of a true negative and a true positive test result or equivalently of a false negative and a false positive test result. Accordingly, in the case study, the optimal diagnostic thresholds maximizing the Youden's index or the concordance probability or minimizing the Euclidean distance, are considerably less than the ADA diagnostic threshold for diabetes of the 2-h postprandial glucose during OGTT (Table 5). Nevertheless, the optimal diagnostic threshold minimizing the risk can be close to the ADA threshold, with specific loss factors settings (Fig 11). Although risk assessment is evolving as the preferred method for optimization of medical decision making (31) and for quality assurance in laboratory medicine (32), the estimation of the loss factors for each test result (Tables 2, 3) is still a complex task. In the future, as the potential of the data analysis will increase exponentially, appropriate loss factors could be estimated using evidence-based methods.

A limitation of the program is that it does not calculate confidence intervals, therefore, it is not intended to analyze samples of measurements. We are currently developing a program for the calculation of the confidence intervals of diagnostic accuracy measures.

All major general or medical statistical software packages (Matlab<sup>®</sup>, NCCS<sup>®</sup>, R, SAS<sup>®</sup>, SPSS<sup>®</sup>, Stata<sup>®</sup> and MedCalc<sup>®</sup>) include routines for the calculation and plotting of various diagnostic accuracy measures and their confidence intervals. The program presented in this work

provides 269 different types of plots of diagnostic accuracy measures (Fig 2), many of which are novel. To the best of our knowledge, no one of the above-mentioned programs or any other software provides this range of plots without advanced statistical programming.

## 5 CONCLUSION

---

The program developed for this work demonstrates clearly various respects of the relation between diagnostic accuracy measures and measurement uncertainty and can be used as a flexible, user-friendly, interactive educational or research tool in medical decision making, to explore and analyze this relation.

## 6 AVAILABILITY AND REQUIREMENTS

---

**Project name:** Relation

**Project home page:** <https://www.hcsl.com/Tools/Relation/>

**Operating systems:** Microsoft Windows, Linux, Apple iOS

**Programming language:** Wolfram Language

**Other software requirements:** Wolfram Player®, freely available at:

<https://www.wolfram.com/player/> or Wolfram Mathematica®

**System requirements:** Intel® Pentium™ Dual-Core or equivalent CPU and 2GB of RAM

**License:** Attribution - NonCommercial - ShareAlike 4.0 International Creative Commons

License

**Any restrictions to use by non-academics:** No other restrictions

## **7 LIST OF ABBREVIATIONS**

---

DAM: diagnostic accuracy measure

OGTT: oral glucose tolerance test

ADA: American Diabetes Association

## **8 DECLARATIONS**

---

### **8.1 ETHICS APPROVAL AND CONSENT TO PARTICIPATE**

Not applicable.

### **8.2 CONSENT FOR PUBLICATION**

Not applicable.

### **8.3 AVAILABILITY OF DATA AND MATERIALS**

All data of the case study are included in this published article (Subsection Sample in Research Design and Methods, and Table 3):

Lim TO, Bakri R, Morad Z, Hamid MA. Bimodality in blood glucose distribution: is it universal?

Diabetes Care. 2002;25(12):2212-7. Epub 2002/11/28. doi: 10.2337/diacare.25.12.2212.

PubMed PMID: 12453963.

## **8.4 COMPETING INTERESTS**

The authors declare that they have no competing interests.

## **8.5 FUNDING**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## **8.6 AUTHORS' CONTRIBUTIONS**

TC substantially contributed to the conception and design of the research project, analyzed and interpreted the case data, substantially contributed to the creation of the software and drafted the article; she has approved the submitted version; she has agreed both to be personally accountable for her own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which she was not personally involved, are appropriately investigated, resolved and the resolution documented in the literature.

ATH supervised the research project, substantially contributed to the creation of the software and substantially revised the draft of the article; he has approved the submitted version; he has agreed both to be personally accountable for his own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which he was not personally involved, are appropriately investigated, resolved and the resolution documented in the literature.

## 9 ACKNOWLEDGEMENTS

---

Not applicable.

## 10 REFERENCES

---

1. Šimundić A-M. Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC*. 2009;19(4):203-11.
2. Shiu S-Y, Gatsonis C. The predictive receiver operating characteristic curve for the joint assessment of the positive and negative predictive values. *Philos Trans A Math Phys Eng Sci*. 2008;366(1874):2313-33.
3. McNeil BJ, Hanley JA. Statistical approaches to the analysis of receiver operating characteristic (ROC) curves. *Med Decis Making*. 1984;4(2):137-50.
4. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. 1982;143(1):29-36.
5. Hilden J. The area under the ROC curve and its competitors. *Med Decis Making*. 1991;11(2):95-101.
6. Hatjimihail AT. The Area Over a Receiver Operating Characteristic (ROC) Curve as an Index of Diagnostic Inaccuracy. *Wolfram Demonstrations Project*; 2011.
7. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3(1):32-5.
8. Hajian-Tilaki K. The choice of methods in determining the optimal cut-off value for quantitative diagnostic test evaluation. *Stat Methods Med Res*. 2018;27(8):2374-83.
9. Liu X. Classification accuracy and cut point selection. *Stat Med*. 2012;31(23):2676-86.
10. Evaluation of measurement data — Guide to the expression of uncertainty in measurement: Joint Committee for Guides in Metrology; 2008 2008.

11. Kallner A, Boyd JC, Duewer DL, Giroud C, Hatjimihail AT, Klee GG, et al. Expression of Measurement Uncertainty in Laboratory Medicine; Approved Guideline: Clinical and Laboratory Standards Institute; 2012 2012.
12. White GH. Basics of estimating measurement uncertainty. Clin Biochem Rev. 2008;29 Suppl 1:S53-60.
13. Oosterhuis WP, Theodorsson E. Total error vs. measurement uncertainty: revolution or evolution? Clin Chem Lab Med. 2016;54(2):235-9.
14. Smith AF, Shinkins B, Hall PS, Hulme CT, Messenger MP. Toward a Framework for Outcome-Based Analytical Performance Specifications: A Methodology Review of Indirect Methods for Evaluating the Impact of Measurement Uncertainty on Clinical Outcomes. Clin Chem. 2019;65(11):1363-74.
15. Ceriotti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, et al. Criteria for assigning laboratory measurands to models for analytical performance specifications defined in the 1st EFLM Strategic Conference. Clin Chem Lab Med. 2017;55(2):189-94.
16. Wolfram S. An Elementary Introduction to the Wolfram Language. Second ed2017. 340 p.
17. Wolfram Research I. Mathematica, Version 12.0. Champaign, IL2019.
18. Hatjimihail AT. Receiver Operating Characteristic Curves and Uncertainty of Measurement. Wolfram Demonstrations Project; 2007.
19. Hatjimihail AT. Uncertainty of Measurement and Areas Over and Under the ROC Curves. Wolfram Demonstrations Project; 2009.
20. Hatjimihail AT. Uncertainty of Measurement and Diagnostic Accuracy Measures. Wolfram Demonstrations Project; 2009.
21. Chatzimichail T. Analysis of Diagnostic Accuracy Measures. Wolfram Demonstrations Project; 2015.

22. Chatzimichail T. Calculator for Diagnostic Accuracy Measures. Wolfram Demonstrations Project; 2018.
23. Chatzimichail T. Correlation of Positive and Negative Predictive Values of Diagnostic Tests. Wolfram Demonstrations Project; 2018.
24. Chatzimichail T, Hatjimihail AT. Calculation of Diagnostic Accuracy Measures. Wolfram Demonstrations Project; 2018.
25. Lim TO, Bakri R, Morad Z, Hamid MA. Bimodality in blood glucose distribution: is it universal? *Diabetes Care*. 2002;25(12):2212-7.
26. American Diabetes A. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2019. *Diabetes Care*. 2019;42(Suppl 1):S13-S28.
27. Kupchak P, Wu AHB, Ghani F, Newby LK, Ohman EM, Christenson RH. Influence of imprecision on ROC curve analysis for cardiac markers. *Clin Chem*. 2006;52(4):752-3.
28. Kroll MH, Biswas B, Budd JR, Durham P, Gorman RT, Gwise TE, et al. Assessment of the Diagnostic Accuracy of Laboratory Tests Using Receiver Operating Characteristic Curves; Approved Guideline—Second Edition: Clinical and Laboratory Standards Institute; 2011 2011.
29. Theodorsson E. Uncertainty in Measurement and Total Error: Tools for Coping with Diagnostic Uncertainty. *Clin Lab Med*. 2017;37(1):15-34.
30. Padoan A, Sciacovelli L, Aita A, Antonelli G, Plebani M. Measurement uncertainty in laboratory reports: A tool for improving the interpretation of test results. *Clin Biochem*. 2018;57:41-7.
31. Aggarwal R. Risk, Complexity, Decision Making, and Patient Care. *JAMA Surg*. 2018;153(3):208.
32. Hatjimihail AT. Estimation of the optimal statistical quality control sampling time intervals using a residual risk measure. *PLoS One*. 2009;4(6):e5770.

# 11 ADDITIONAL FILES

---

**File name:** Relation.cdf (It can be run on Wolfram Player®, available at:

<https://www.wolfram.com/player/>)

**Description of the file:** It is the software tool presented in the article.

# 12 NOTATION

---

## 12.1 POPULATIONS

$\bar{D}$ : nondiseased population

$D$ : diseased population

## 12.2 TEST OUTCOMES

$\bar{T}$ : negative test result

$T$ : positive test result

$TN$ : true negative test result

$TP$ : true positive test result

$FN$ : false negative test result

$FP$ : false positive test result

## 12.3 DIAGNOSTIC ACCURACY MEASURES

*Se*: sensitivity

*Sp*: specificity

*PPV*: positive predictive value

*NPV*: negative predictive value

*ODA*: overall diagnostic accuracy

*DOR*: diagnostic odds ratio

*LR+*: likelihood ratio for a positive test result

*LR-*: likelihood ratio for a negative test result

*J*: Youden's index

*ED*: Euclidean distance of a *ROC* curve point from the point (0,1)

*CZ*: concordance probability

*R*: risk

*ROC*: receiver operating characteristic curve

*AUC*: area under the *ROC* curve

*AOC*: area over the *ROC* curve

*PROC*: predictive receiver operating characteristic curve

## 12.4 PARAMETERS

$\mu_P$ : mean of the measurand of a single test in the population  $P$

$\sigma_P$ : standard deviation of the measurand of a single test in the population  $P$

$\nu$ : prevalence of the disease

$d$ : diagnostic threshold of a single test

$u$ : standard measurement uncertainty of a single test

## 12.5 EXPECTED LOSS

$l_0$ : expected loss for the testing procedure

$l_{TN}$ : expected loss for a true negative result

$l_{FN}$ : expected loss for a false negative result

$l_{TP}$ : expected loss for a true positive result

$l_{FP}$ : expected loss for a false positive result

## 12.6 FUNCTIONS AND RELATIONS

$se(d, \dots)$ : sensitivity function of a single test versus its diagnostic threshold  $d$

$sp(d, \dots)$ : specificity function of a single test versus its diagnostic threshold  $d$

$se_{sp}(z, \dots)$ : sensitivity function of a single test versus its specificity  $z$

$sp_{se}(y, \dots)$ : specificity function of a single test versus its sensitivity  $y$

$roc(\dots)$ : receiver operator characteristic function of a screening or diagnostic test

$auc(\dots)$ : function of the area under the receiver operator characteristic curve

$aoc(\dots)$ : function of the area over the receiver operator characteristic curve

$proc(\dots)$ : predictive receiver operator characteristic relation of a screening or diagnostic test

$ed(t, \dots)$ : Euclidean distance function of the ROC curve point  $(t, roc(t, \dots))$  from the point  $(0, 1)$

$\Phi(x)$ : cumulative distribution function of the unit normal distribution, evaluated at  $x$

$\Psi(x, \mu, \sigma)$ : cumulative distribution function of a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , evaluated at  $x$

$S(x, \mu, \sigma)$ : survival function of a normal distribution with mean  $\mu$  and standard deviation  $\sigma$ , evaluated at  $x$

$erf(x)$ : error function, evaluated at  $x$

$erfc(x)$ : complementary error function, evaluated at  $x$

$Pr(a)$ : probability of an event  $a$

$Pr(a|b)$ : probability of an event  $a$  given the event  $b$

$F^{-1}(\dots)$ : The inverse function  $F$

# 13 APPENDIX

---

## 13.1 ABOUT THE PROGRAM CONTROLS

The numerical settings are defined by the user with menus or sliders. Sliders can be finely manipulated by holding down the *alt* key or *opt* key while dragging the mouse. They be even more finely manipulated by also holding the *shift* and/or *ctrl* keys.

Dragging with the mouse rotates the three-dimensional plots, while dragging with the mouse while pressing the *ctrl*, *alt*, or *opt* keys zooms in or out.

## 13.2 RANGE OF THE INPUT PARAMETERS

$v$ : 0.00000001 – 0.99

$\mu_D, \mu_{\bar{D}}$ : 0 - 6

$\sigma_D, \sigma_{\bar{D}}$ : 0.01 - 6

$d$ : Minimum of  $(\mu_D - 3.5\sigma_D, \mu_{\bar{D}} - 3.5\sigma_{\bar{D}})$  – Maximum of  $(\mu_D + 3.5\sigma_D, \mu_{\bar{D}} + 3.5\sigma_{\bar{D}})$

$u$ : 0 - 6

$l_0$ : 0 - 1000

$l_{TN}$ : 0 - 1000

$l_{FN}$ : 0 - 1000

$l_{TP}$ : 0 - 1000

$l_{FP}$ : 0 – 1000

*Initial plot points per axis*: 10 – 50

## 13.3 RANGE OF THE COORDINATES OF THE PLOTS

There are two options for the range of coordinates to be included in each plot:

- 1) Full: All the calculated coordinate points are plotted.
- 2) Partial: The distribution of coordinate values is found and any points sufficiently far out in the distribution are not considered.

## **13.4 INPUT AND OUTPUT**

The program provides in four modules plots and tables of diagnostic accuracy measures of two screening or diagnostic tests for a single measurand, differing in measurement uncertainty, applied at a single point in time in a diseased and a non-diseased population.

Singularity points are excluded from the plots.

Indeterminate results of the calculation modules represent numerical quantities whose magnitudes cannot be determined, because they are either too small or too large.

### **13.4.1 ROC curves**

#### ***Input***

The user defines:

- 1) The prevalence of the disease.
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The standard measurement uncertainty of each test.

#### ***Output***

- 1) a) The plots of the *ROC* curves of the two tests and  
b) A table with the areas under and over the *ROC* curves (*AUC* and *AOC* respectively),  
or
- 2) The plots of the *PROC* curves of the two tests.

### 13.4.2 Diagnostic accuracy measures plots

The module presents in six submodules plots of diagnostic accuracy measures of the tests.

#### 13.4.2.1 Diagnostic accuracy measures versus uncertainty plots

##### Input

The user defines:

- 1) The prevalence of the disease.
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The diagnostic threshold.
- 4) The expected loss for:
  - i) Testing,
  - ii) A true positive result,
  - iii) A true negative result,
  - iv) A false positive result and
  - v) A false negative result.
- 2) The diagnostic accuracy measure function to be plotted:
  - a) Sensitivity ( $Se$ ),
  - b) Specificity ( $Sp$ ),
  - c) Overall diagnostic accuracy ( $ODA$ ),
  - d) Positive predictive value ( $PPV$ ),
  - e) Negative predictive value ( $NPV$ ),
  - f) Diagnostic odds ratio ( $DOR$ ),
  - g) Likelihood ratio for a positive test result ( $LR+$ ),
  - h) Likelihood ratio for a negative test result ( $LR-$ ),
  - i) Youden's index ( $J$ ),

- j) Euclidean distance ( $ED$ ),
  - k) Concordance probability ( $CZ$ ),
  - l) Risk ( $R$ ).
- 3) The range of the coordinate points to be plotted:
- i) Full,
  - ii) Partial.

### **Output**

Plots of the values of the measure of each test versus uncertainty.

#### ***13.4.2.2 Diagnostic accuracy measures versus sensitivity or specificity plots***

### **Input**

The user defines:

- 1) The prevalence of the disease
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The standard measurement uncertainty of each test.
- 4) The expected loss for:
  - a) Testing,
  - b) A true positive result,
  - c) A true negative result,
  - d) A false positive result and
  - e) A false negative result.
- 2) The diagnostic accuracy measure function to be plotted by selecting:
  - a) The measure:
    - i) Overall diagnostic accuracy ( $ODA$ ),
    - ii) Positive predictive value ( $PPV$ ),

- iii) Negative predictive value ( $NPV$ ),
  - iv) Diagnostic odds ratio ( $DOR$ ),
  - v) Likelihood ratio for a positive test result ( $LR+$ ),
  - vi) Likelihood ratio for a negative test result ( $LR-$ ),
  - vii) Youden's index ( $J$ ),
  - viii) Euclidean distance ( $ED$ ),
  - ix) Concordance probability ( $CZ$ ),
  - x) Risk ( $R$ ).
- b) The domain of the function:
- i) The sensitivity ( $Se$ ) of the test,
  - ii) The specificity ( $Sp$ ) of the test.
- c) The codomain of the function, that is
- i) The value of the measure,
  - ii) The partial derivative of the measure with respect to standard measurement uncertainty,
  - iii) The difference between the measures of the two tests,
  - iv) The relative difference between the measures of the two tests,
  - v) The ratio of the measures of the two tests.
- 3) The range of the coordinate points to be plotted:
- a) Full,
  - b) Partial.

### **Output**

Plots of:

- 1) The values of the measure of each test,

- 2) The partial derivatives of the measure of each test with respect to standard measurement uncertainty ( $u$ ),
  - 3) The difference between the measures of the two tests,
  - 4) The relative difference between the measures of the two tests,
  - 5) The ratio of the measures of the two tests,
- versus the respective domain variable.

### ***13.4.2.3 The diagnostic accuracy measures versus sensitivity and specificity plots***

#### **Input**

The user defines:

- 1) The prevalence of the disease.
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The standard measurement uncertainty of each test.
- 4) The expected loss for:
  - c) Testing,
  - d) A true positive result,
  - e) A true negative result,
  - f) A false positive result and
  - g) A false negative result.
- 4) The diagnostic accuracy measure function to be plotted by selecting:
  - a) The measure:
    - i) Overall diagnostic accuracy ( $ODA$ ),
    - ii) Positive predictive value ( $PPV$ ),
    - iii) Negative predictive value ( $NPV$ ),
    - iv) Diagnostic odds ratio ( $DOR$ ),

- v) Likelihood ratio for a positive test result ( $LR+$ ),
  - vi) Likelihood ratio for a negative test result ( $LR-$ ),
  - vii) Youden's index ( $J$ ),
  - viii) Euclidean distance ( $ED$ ),
  - ix) Concordance probability ( $CZ$ ),
  - x) Risk ( $R$ ).
- b) The codomain of the function, that is
    - i) The value of the measure,
    - ii) The partial derivative of the measure with respect to standard measurement uncertainty ( $u$ ).
- 5) The range of the coordinate points to be plotted:
- a) Full,
  - b) Partial.

### **Output**

Three-dimensional line plots of the values or the derivatives of the measure of each test versus its sensitivity ( $Se$ ) and specificity ( $Sp$ ).

#### ***13.4.2.4 Diagnostic accuracy measures versus diagnostic threshold plots***

### **Input**

- 1) The user defines:
  - a) The prevalence of the disease.
  - b) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
  - c) The standard measurement uncertainty of each test.
  - d) The expected loss for:
    - i) Testing,

- ii) A true positive result,
  - iii) A true negative result,
  - iv) A false positive result and
  - v) A false negative result.
- 2) The diagnostic accuracy measure function to be plotted, by selecting:
- a) The measure:
    - i) Sensitivity ( $Se$ ),
    - ii) Specificity ( $Sp$ ),
    - iii) Overall diagnostic accuracy ( $ODA$ ),
    - iv) Positive predictive value ( $PPV$ ),
    - v) Negative predictive value ( $NPV$ ),
    - vi) Diagnostic odds ratio ( $DOR$ ),
    - vii) Likelihood ratio for a positive test result ( $LR+$ ),
    - viii) Likelihood ratio for a negative test result ( $LR-$ ),
    - ix) Youden's index ( $J$ ),
    - x) Euclidean distance ( $ED$ ),
    - xi) Concordance probability ( $CZ$ ),
    - xii) Risk ( $R$ ).
  - b) The codomain of the function, that is
    - i) The value of the measure,
    - ii) The partial derivative of the measure with respect to standard measurement uncertainty ( $u$ ),
    - iii) The difference between the measures of the two tests,
    - iv) The relative difference between the measures of the two tests,
    - v) The ratio of the measures of the two tests.
  - c) The range of the coordinate points to be plotted:

- i) Full,
- ii) Partial,

### **Output**

Plots of:

- 1) The values of the measure of each test,
- 2) The partial derivatives of the measure of each test with respect to standard measurement uncertainty ( $u$ ),
- 3) The difference between the measures of the two tests,
- 4) The relative difference between the measures of the two tests,
- 5) The ratio of the measures of the two tests,

versus diagnostic threshold ( $d$ ).

#### ***13.4.2.5 Diagnostic accuracy measures versus prevalence plots***

### **Input**

The user defines:

- 1) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 2) The diagnostic threshold.
- 3) The standard measurement uncertainty of each test.
- 4) The expected loss for:
  - a) Testing,
  - b) A true positive result,
  - c) A true negative result,
  - d) A false positive result and
  - e) A false negative result.

- 2) The diagnostic accuracy measure function to be plotted, by selecting:
  - a) The measure:
    - i) Sensitivity ( $Se$ ),
    - ii) Specificity ( $Sp$ ),
    - iii) Overall diagnostic accuracy ( $ODA$ ),
    - iv) Positive predictive value ( $PPV$ ),
    - v) Negative predictive value ( $NPV$ ),
    - vi) Diagnostic odds ratio ( $DOR$ ),
    - vii) Likelihood ratio for a positive test result ( $LR+$ ),
    - viii) Likelihood ratio for a negative test result ( $LR-$ ),
    - ix) Youden's index ( $J$ ),
    - x) Euclidean distance ( $ED$ ),
    - xi) Concordance probability ( $CZ$ ),
    - xii) Risk ( $R$ ).
  - b) The codomain of the function, that is
    - i) The value of the measure,
    - ii) The partial derivative of the measure with respect to standard measurement uncertainty ( $u$ ),
    - iii) The difference between the measures of the two tests,
    - iv) The relative difference between the measures of the two tests,
    - v) The ratio of the measures of the two tests.
- 3) The range of the coordinate points to be plotted:
  - a) Full,
  - b) Partial.

## Output

Plots of:

- 1) The values of the measure of each test,
- 2) The partial derivatives of the measure of each test with respect to standard measurement uncertainty ( $u$ ),
- 3) The difference between the measures of the two tests,
- 4) The relative difference between the measures of the two tests,
- 5) The ratio of the measures of the two tests,

versus prevalence ( $v$ ).

### ***13.4.2.6 Diagnostic accuracy measures relations***

## Input

The user defines:

- 1) The prevalence of the disease.
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The standard measurement uncertainty of each test.
- 4) The expected loss for:
  - c) Testing,
  - d) A true positive result,
  - e) A true negative result,
  - f) A false positive result and
  - g) A false negative result.
- 5) The relation between the diagnostic accuracy measures to be plotted, by selecting any two of the following:

- a) Overall diagnostic accuracy ( $ODA$ ),
  - b) Positive predictive value ( $PPV$ ),
  - c) Negative predictive value ( $NPV$ ),
  - d) Diagnostic odds ratio ( $DOR$ ),
  - e) Likelihood ratio for a positive test result ( $LR+$ ),
  - f) Likelihood ratio for a negative test result ( $LR-$ ),
  - g) Youden's index ( $J$ ),
  - h) Euclidean distance ( $ED$ ),
  - i) Concordance probability ( $CZ$ ),
  - j) Risk ( $R$ ).
- 6) The range of the coordinate points to be plotted:
- a) Full,
  - b) Partial.

### **Output**

Plots of the respective parametric equations of any two of the above diagnostic accuracy measures.

### **13.4.3 Diagnostic accuracy measures calculator**

#### ***Input***

The user defines:

- 1) The prevalence of the disease.
- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The diagnostic threshold.
- 4) The standard measurement uncertainty of each test.
- 5) The expected loss for:

- a) Testing,
- b) A true positive result,
- c) A true negative result,
- d) A false positive result and
- e) A false negative result,.

### ***Output***

A table of the values and the relative differences of the following diagnostic accuracy measures of the two tests for the selected diagnostic threshold:

- b) Sensitivity ( $Se$ ),
- c) Specificity ( $Sp$ ),
- d) Overall diagnostic accuracy ( $ODA$ ),
- e) Positive predictive value ( $PPV$ ),
- f) Negative predictive value ( $NPV$ ),
- g) Diagnostic odds ratio ( $DOR$ ),
- h) Likelihood ratio for a positive test result ( $LR+$ ),
- i) Likelihood ratio for a negative test result ( $LR-$ ),
- j) Youden's index ( $J$ ),
- k) Euclidean distance ( $ED$ ),
- l) Concordance probability ( $CZ$ ),
- m) Risk ( $R$ ).

#### **13.4.4 Optimal diagnostic accuracy measures calculator**

### ***Input***

The user defines:

- 1) The prevalence of the disease.

- 2) The means and the standard deviations of the measurand in the diseased and non-diseased populations.
- 3) The diagnostic threshold.
- 4) The expected loss for:
  - a) Testing,
  - b) A true positive result,
  - c) A true negative result,
  - d) A false positive result and
  - e) A false negative result.
- 5) The objective or loss function to find the diagnostic threshold for the following optima:
  - a) Maximum Youden's index ( $J$ ),
  - b) Minimum Euclidean distance ( $ED$ ),
  - c) Maximum concordance probability ( $CZ$ ),
  - d) Minimum risk ( $R$ ).

### ***Output***

- 1) The optimal diagnostic thresholds for the two tests.
- 2) A table of the values of the following diagnostic accuracy measures of the two tests at the respective optimal diagnostic thresholds and of their relative differences:
  - a) Sensitivity ( $Se$ ),
  - b) Specificity ( $Sp$ ),
  - c) Overall diagnostic accuracy ( $ODA$ ),
  - d) Positive predictive value ( $PPV$ ),
  - e) Negative predictive value ( $NPV$ ),
  - f) Diagnostic odds ratio ( $DOR$ ),
  - g) Likelihood ratio for a positive test result ( $LR+$ ),

- h) Likelihood ratio for a negative test result ( $LR^-$ ),
- i) Youden's index ( $J$ ),
- j) Euclidean distance ( $ED$ ),
- k) Concordance probability ( $CZ$ ),
- l) Risk ( $R$ ).

## 14 FIGURE LEGENDS

---

**Fig 1. Probability density function plots.** The probability density functions plots of a measurand in a non-diseased and diseased population.

**Fig 2. Program flowchart.** The flowchart of the program with the number of the input parameters and of the output types for each module or submodule.

**Fig 3. ROC curves module screenshot.** ROC curves plots with the settings shown at the left.

**Fig 4. DAM plots module, DAM versus uncertainty submodule screenshot.** Overall diagnostic accuracy ( $ODA$ ) vs standard measurement uncertainty ( $u$ ) curve plot with the settings shown at the left.

**Fig 5. DAM plots module, DAM versus sensitivity or specificity submodule screenshot.** Relative difference of the diagnostic odds ratio ( $DOR$ ) of the two screening or diagnostic tests versus specificity ( $Sp$ ) curve plot, with the settings shown at the left.

**Fig 6. DAM plots module, DAM versus sensitivity and specificity submodule screenshot.** Likelihood ratio for a positive test result ( $LR^+$ ) versus sensitivity ( $Se$ ) and specificity ( $Sp$ ) curves plot, with the settings shown at the left.

**Fig 7. DAM plots module, DAM vs threshold submodule screenshot.** Relative difference of the risk ( $R$ ) of the two screening or diagnostic tests vs diagnostic threshold ( $d$ ) curve plot, with the settings at the left.

**Fig 8. DAM plots module, DAM vs prevalence submodule screenshot.** Ratio of the positive predictive values ( $PPV$ ) of the two screening or diagnostic tests versus prevalence ( $v$ ) of the disease curve plot, with the settings at the left.

**Fig 9. DAM plots module, DAM relations submodule screenshot.** Positive predictive value ( $PPV$ ) versus negative predictive value ( $NPV$ ) curves plot, with the settings at the left.

**Fig 10. DAM calculator module screenshot.** Calculated diagnostic accuracy measures and their relative differences, with the settings at the left.

**Fig 11. Optimal DAM calculator module screenshot.** Calculated diagnostic accuracy measures minimizing risk ( $R$ ) and their relative differences, with the settings at the left.

**Fig 12. DAM versus uncertainty plots.** Plots of (A) Sensitivity ( $Se$ ), (B) specificity ( $Sp$ ), (C) positive predictive value ( $PPV$ ) and (D) negative predictive value ( $NPV$ ) versus standard measurement uncertainty ( $u$ ) curves, with the respective parameters in Table 4.

**Fig 13. DAM versus uncertainty plots.** Plots of (A) diagnostic odds ratio ( $DOR$ ), (B) risk ( $R$ ), (C) likelihood ratio for a positive result ( $LR+$ ) and (D) likelihood ratio for a negative result ( $LR-$ ) versus standard measurement uncertainty ( $u$ ) curves, with the respective parameters in Table 4.

**Fig 14. DAM relative differences versus prevalence plots.** Plots of the relative difference of the (A) positive predictive value ( $PPV$ ), (B) negative predictive value ( $NPV$ ), (C) overall diagnostic accuracy ( $ODA$ ) and (D) risk ( $R$ ) of two diagnostic or screening tests, measuring

the same measurand with different uncertainties, versus prevalence ( $\nu$ ) curves, with the respective parameters in Table 4.

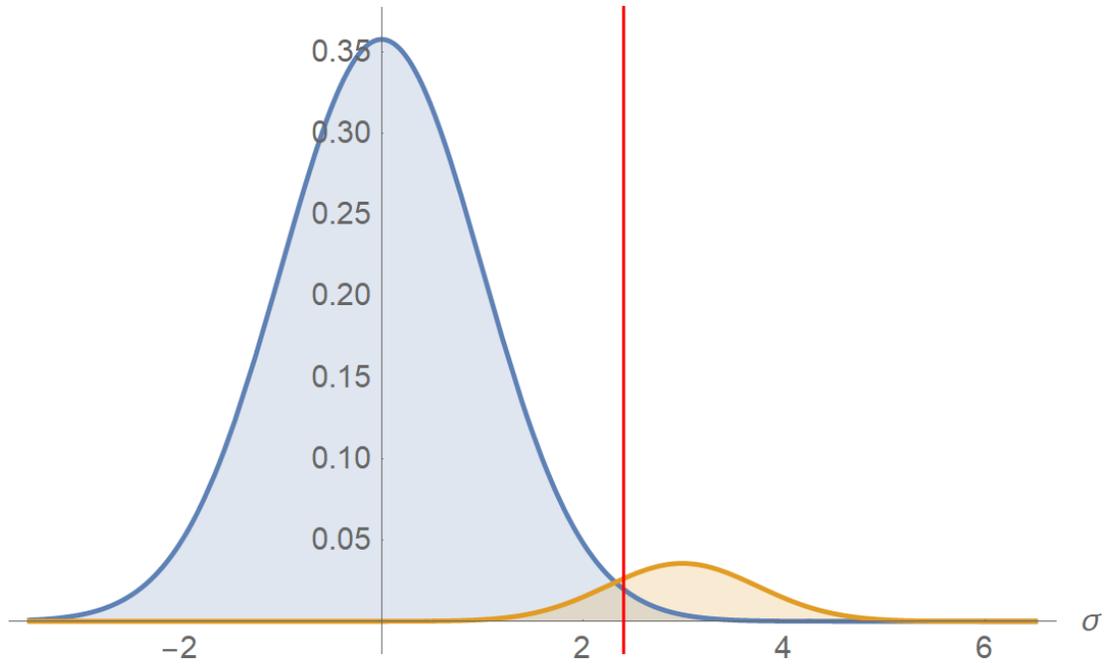
**Fig 15. DAM relative differences versus diagnostic threshold plots**

Plots of the relative difference of the (A) likelihood ratio for a positive result ( $LR+$ ), (B) likelihood ratio for a negative result ( $LR-$ ), (C) diagnostic odds ratio ( $DOR$ ) and (D) Youden's index ( $J$ ) of two screening or diagnostic tests, measuring the same measurand with different uncertainties, versus diagnostic threshold ( $d$ ) curves, with the respective parameters in Table 4.

**Fig 16. DAM partial derivatives versus diagnostic threshold plots.** Plots of partial derivatives of (A) overall diagnostic accuracy ( $ODA$ ), (B) Youden's index ( $J$ ), (C) positive predictive value ( $PPV$ ) and (D) risk ( $R$ ), with respect to measurement uncertainty, of two tests, measuring the same measurand with different uncertainties, versus diagnostic threshold ( $d$ ) curves, with the parameters in Table 4.

**Fig 17. DAM relations plots.** Plots of the relations between (A) negative predictive value ( $NPV$ ) and overall diagnostic accuracy ( $ODA$ ), (B) positive predictive value ( $PPV$ ) and Youden's index ( $J$ ), (C) likelihood ratio for a negative result ( $LR-$ ) and risk ( $R$ ) and (D) Euclidean distance ( $ED$ ) and diagnostic odds ratio ( $DOR$ ), of two tests, measuring the same measurand with different uncertainties, with the respective parameters in Table 4.

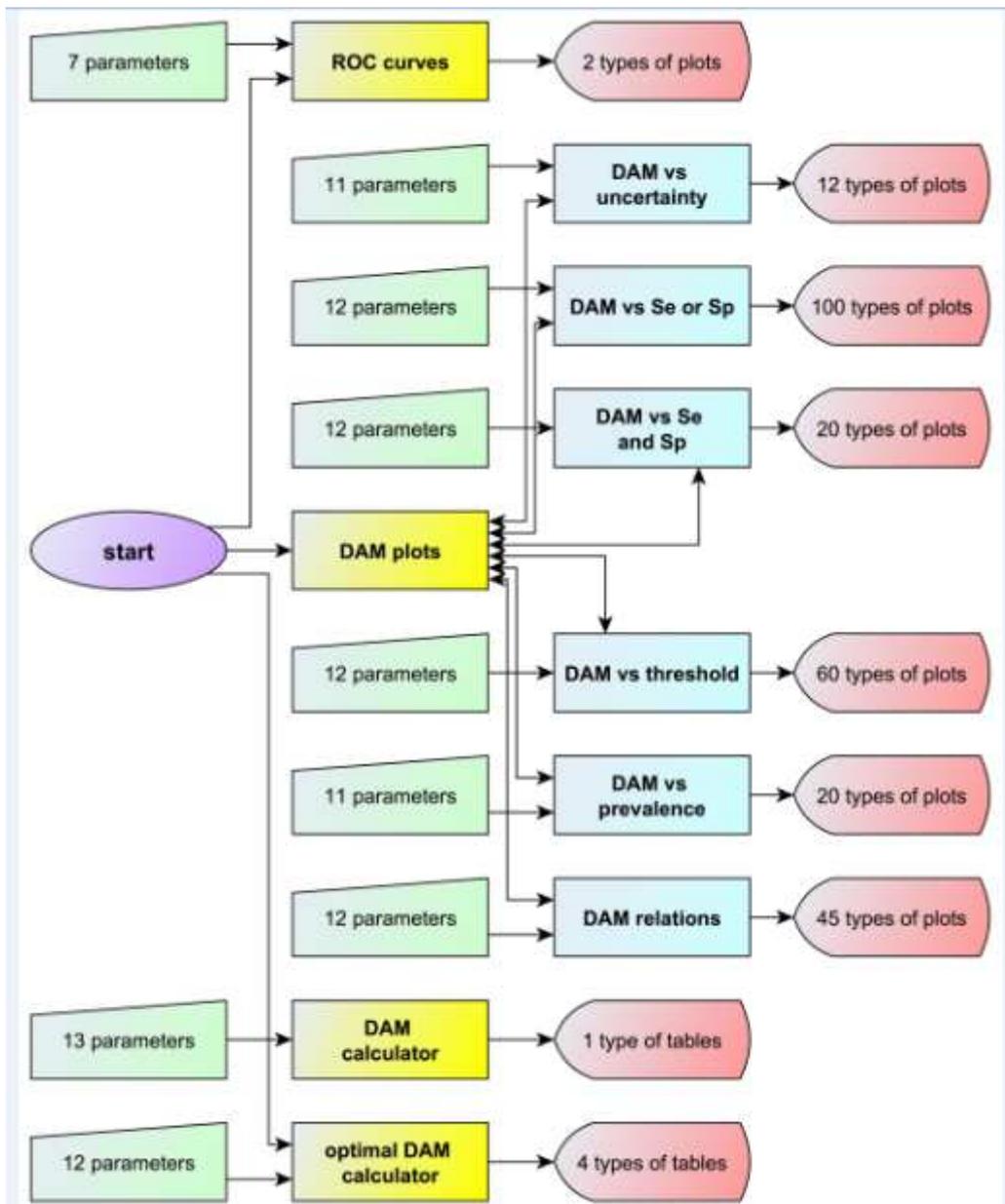
# Figures



— nondiseased population — diseased population — diagnostic threshold

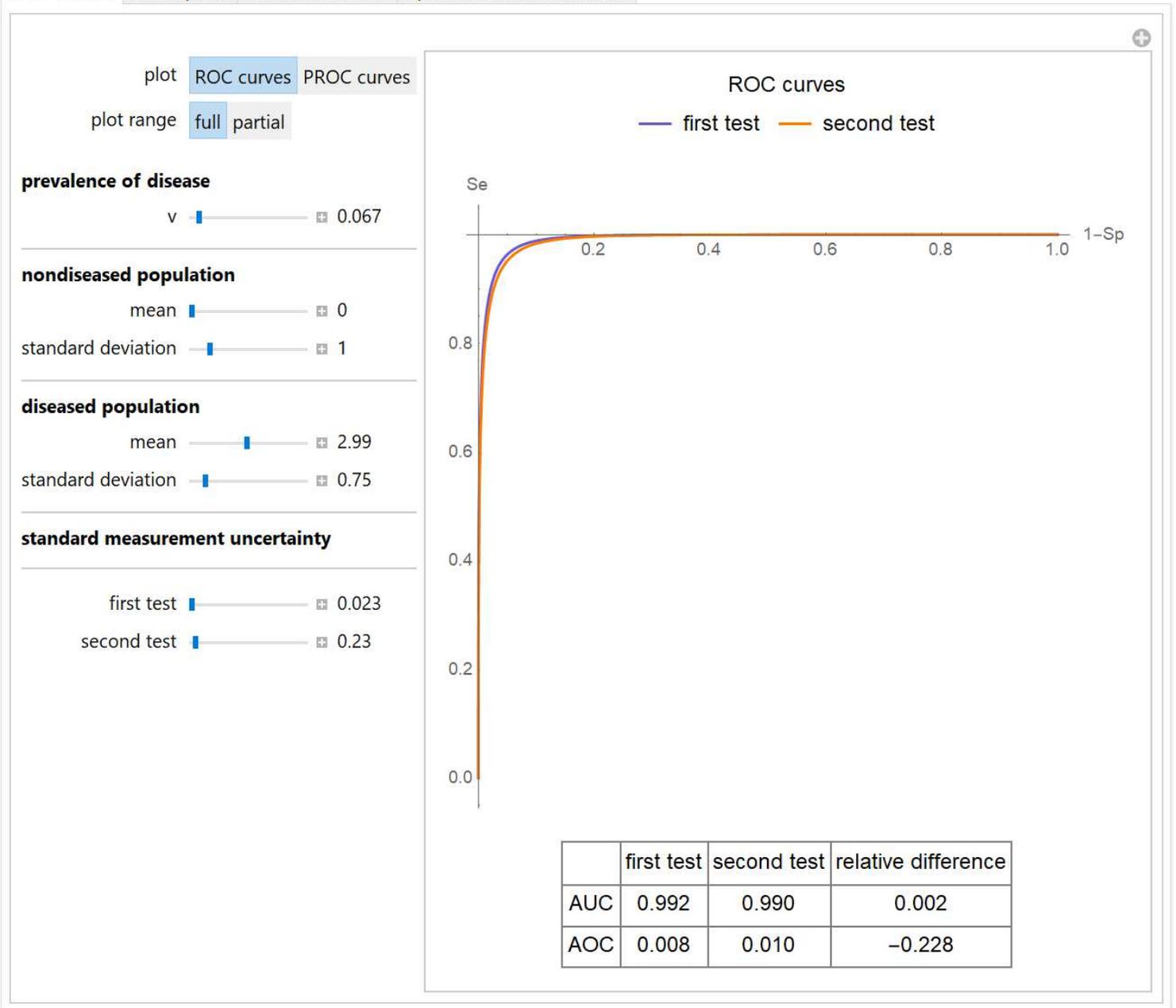
**Figure 1**

Probability density function plots. The probability density functions plots of a measurand in a non-diseased and diseased population.



**Figure 2**

Program flowchart. The flowchart of the program with the number of the input parameters and of the output types for each module or submodule.



**Figure 3**

ROC curves module screenshot. ROC curves plots with the settings shown at the left.

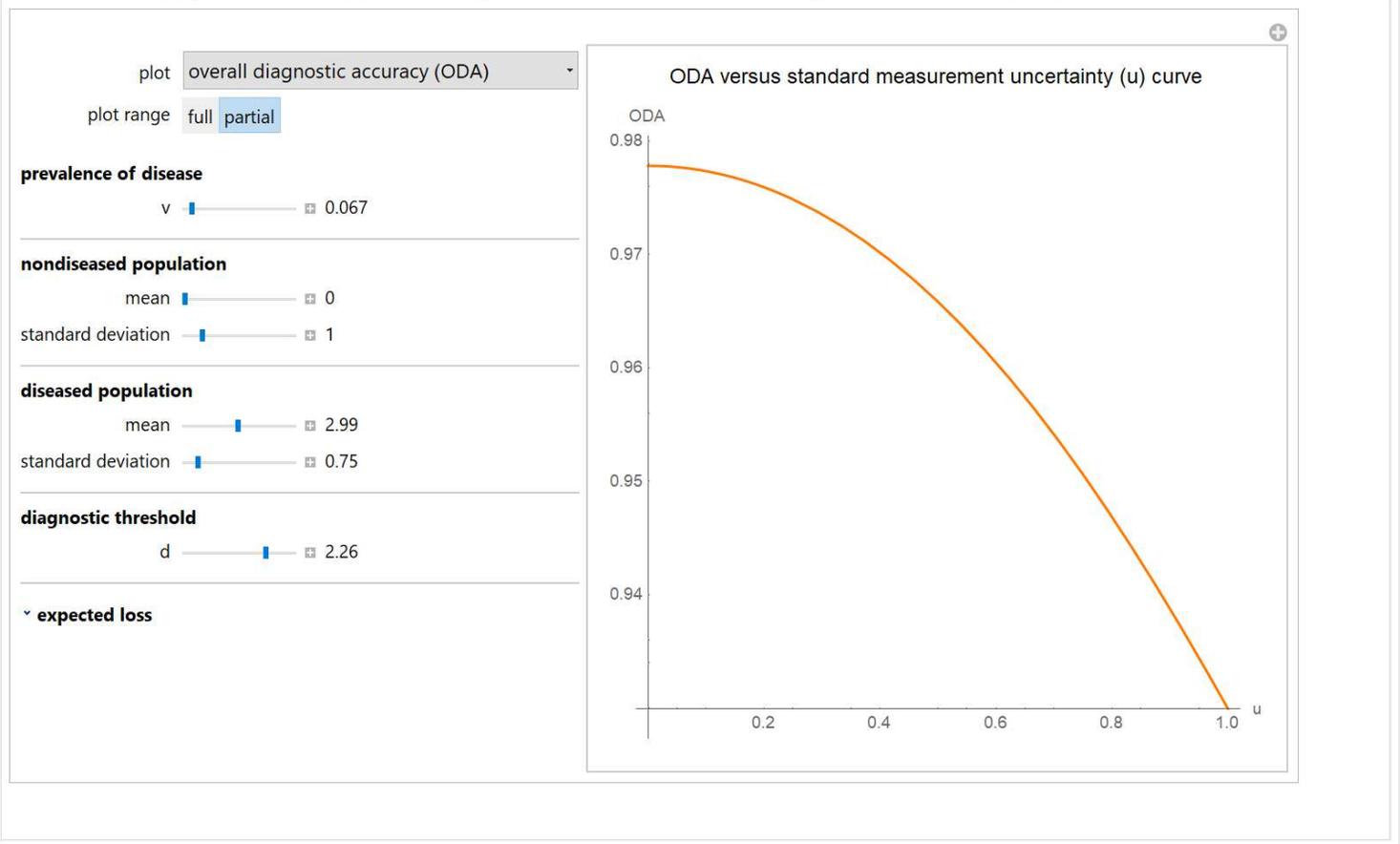
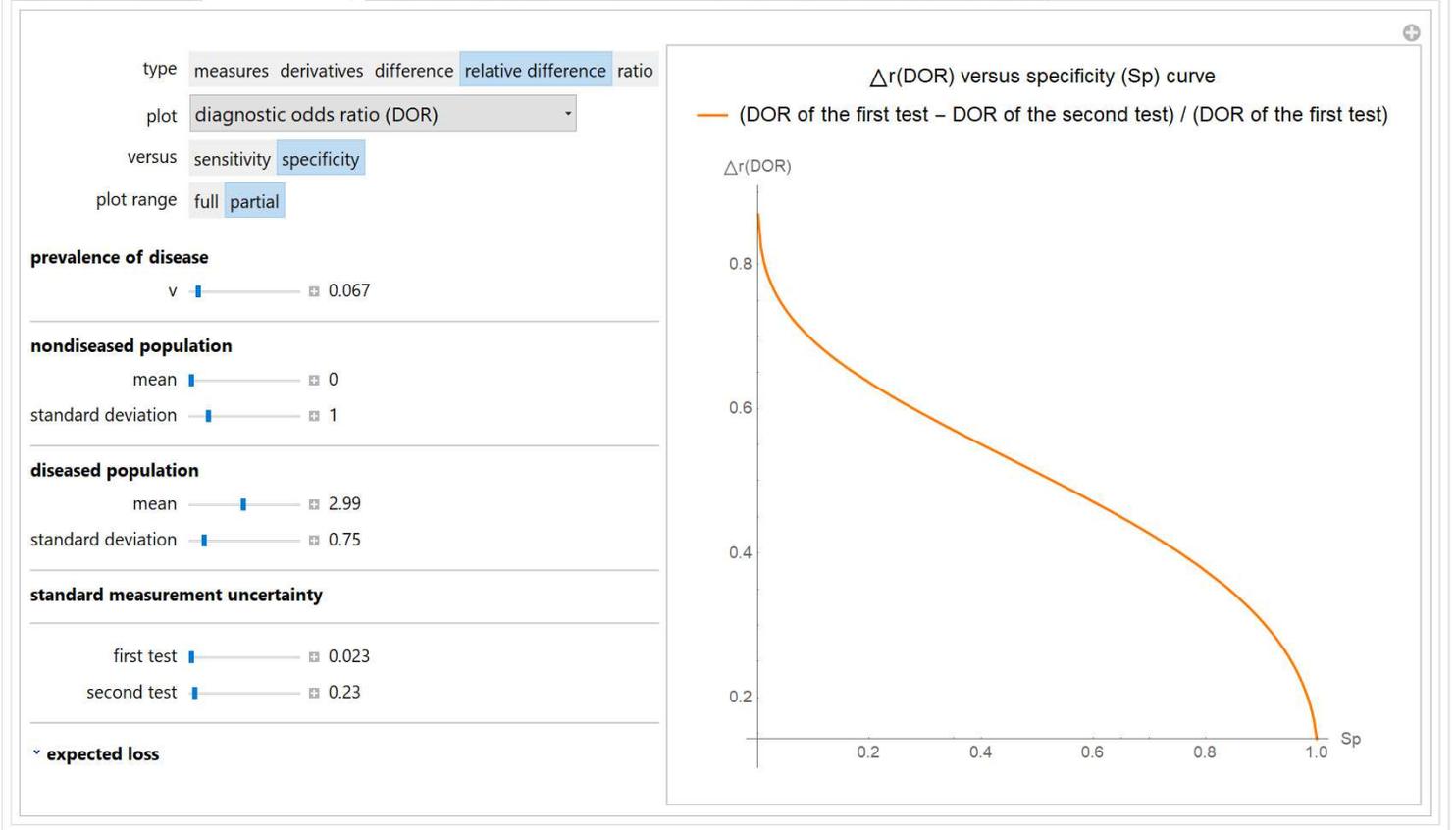


Figure 4

DAM plots module, DAM versus uncertainty submodule screenshot. Overall diagnostic accuracy (ODA ) vs standard measurement uncertainty ( u ) curve plot with the settings shown at the left.



**Figure 5**

DAM plots module, DAM versus sensitivity or specificity submodule screenshot. Relative difference of the diagnostic odds ratio (DOR ) of the two screening or diagnostic tests versus specificity (Sp ) curve plot, with the settings shown at the left.

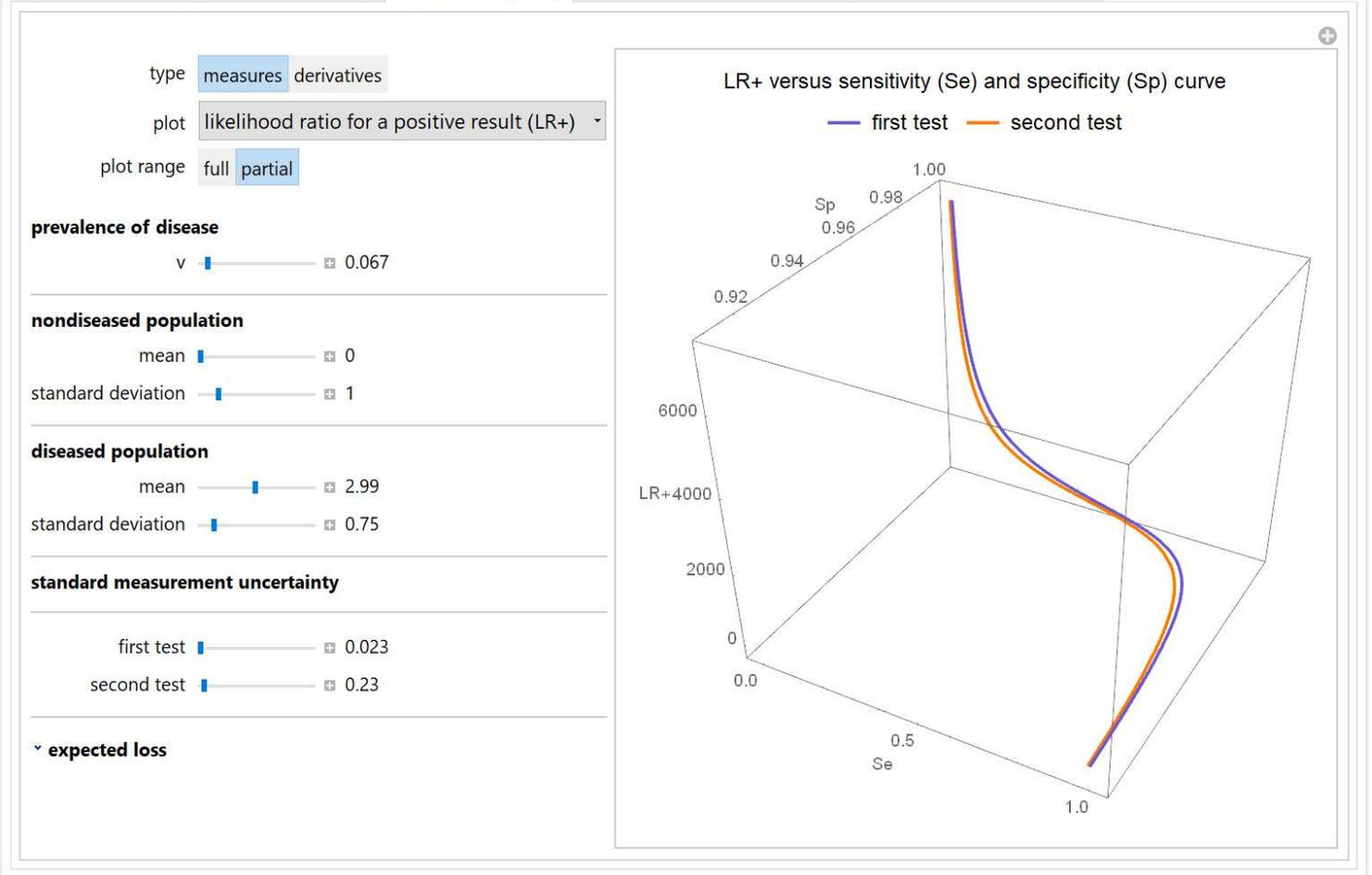
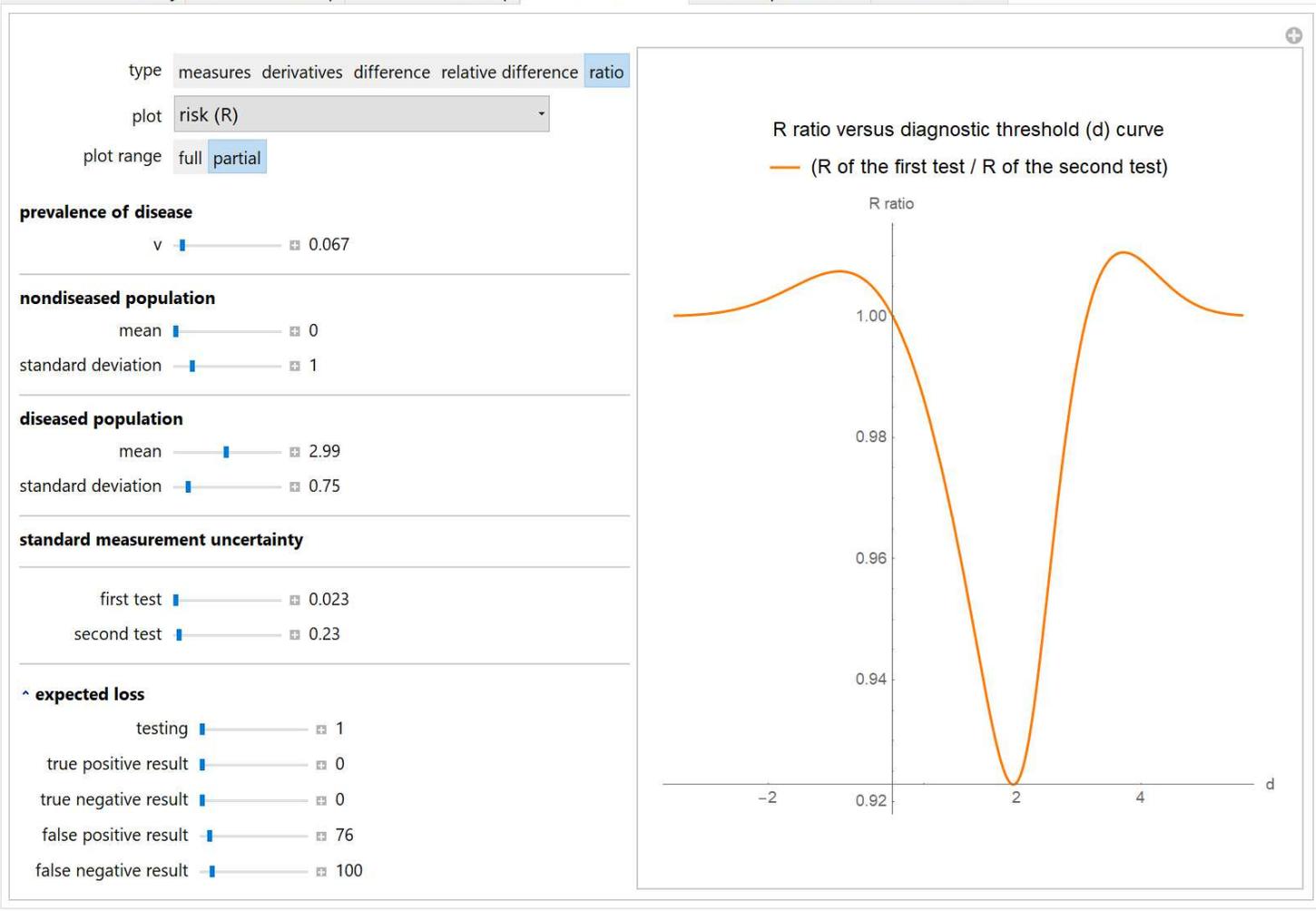


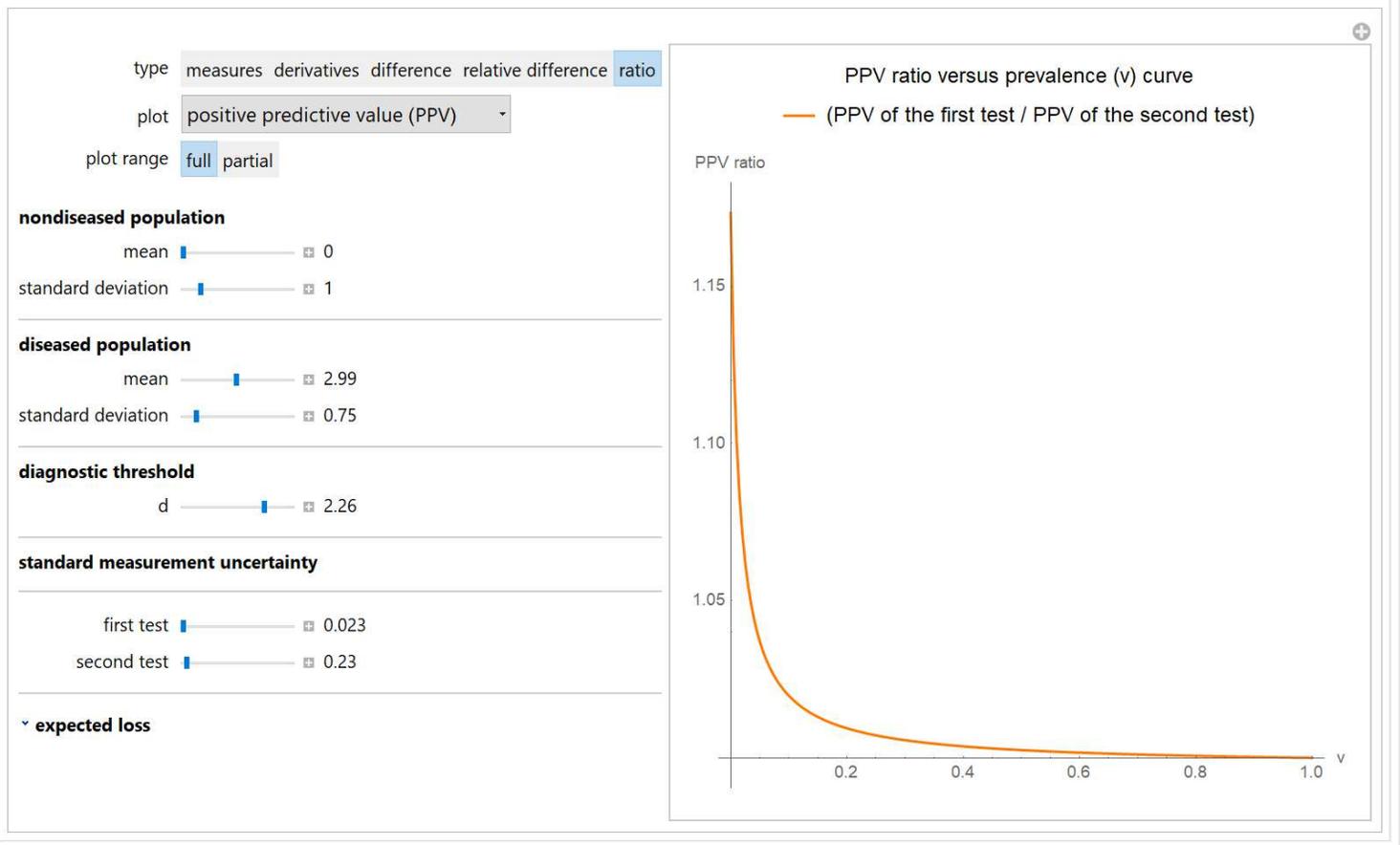
Figure 6

DAM plots module, DAM versus sensitivity and specificity submodule screenshot. Likelihood ratio for a positive test result (LR +) versus sensitivity (Se ) and specificity (Sp ) curves plot, with the settings shown at the left



**Figure 7**

DAM plots module, DAM vs threshold submodule screenshot. Relative difference of the risk (R) of the two screening or diagnostic tests vs diagnostic threshold (d) curve plot, with the settings at the left



**Figure 8**

DAM plots module, DAM vs prevalence submodule screenshot. Ratio of the positive predictive values (PPV) of the two screening or diagnostic tests versus prevalence (v) of the disease curve plot, with the settings at the left.

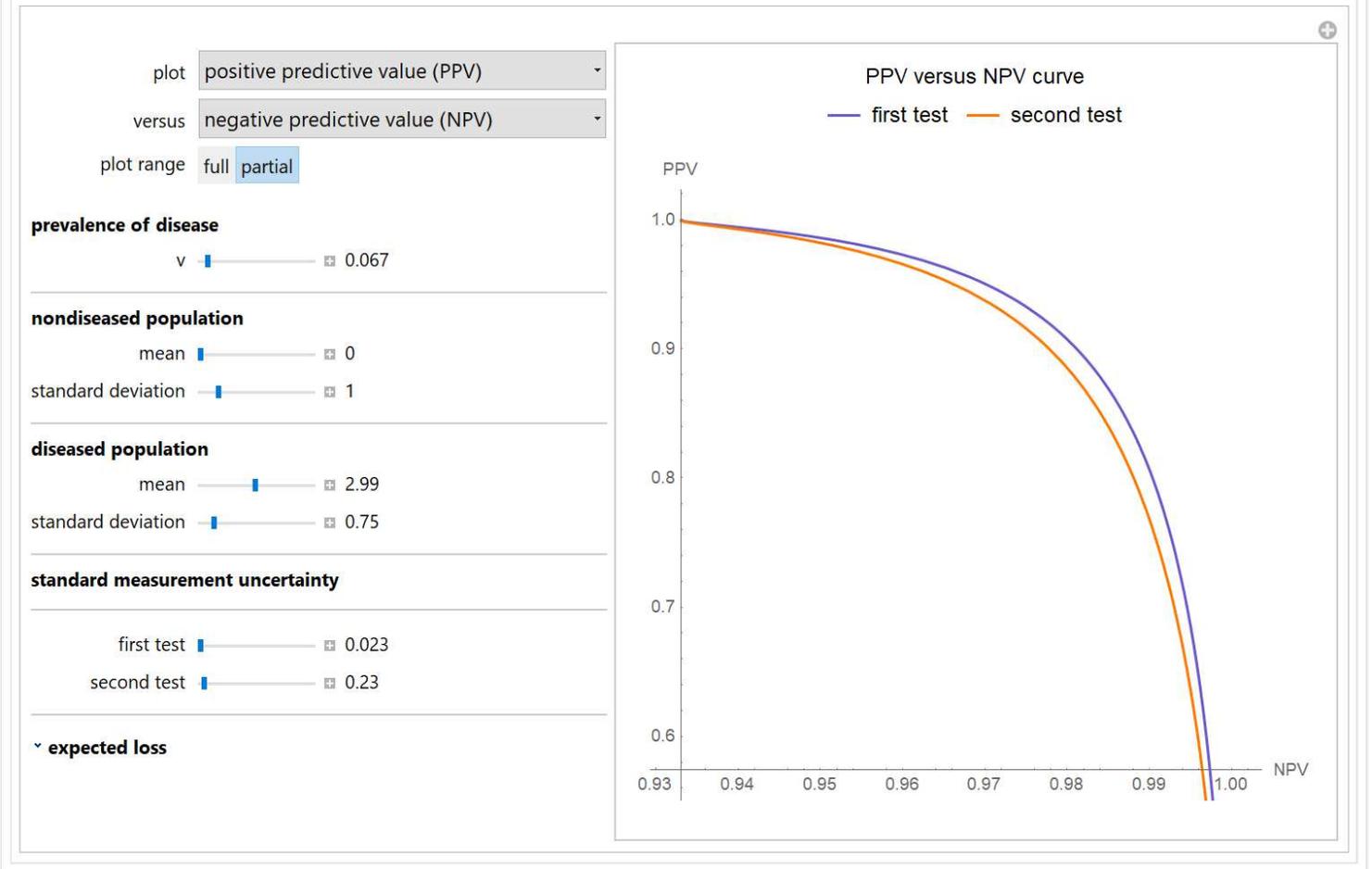


Figure 9

DAM plots module, DAM relations submodule screenshot. Positive predictive value (PPV ) versus negative predictive value (NPV ) curves plot, with the settings at the left.

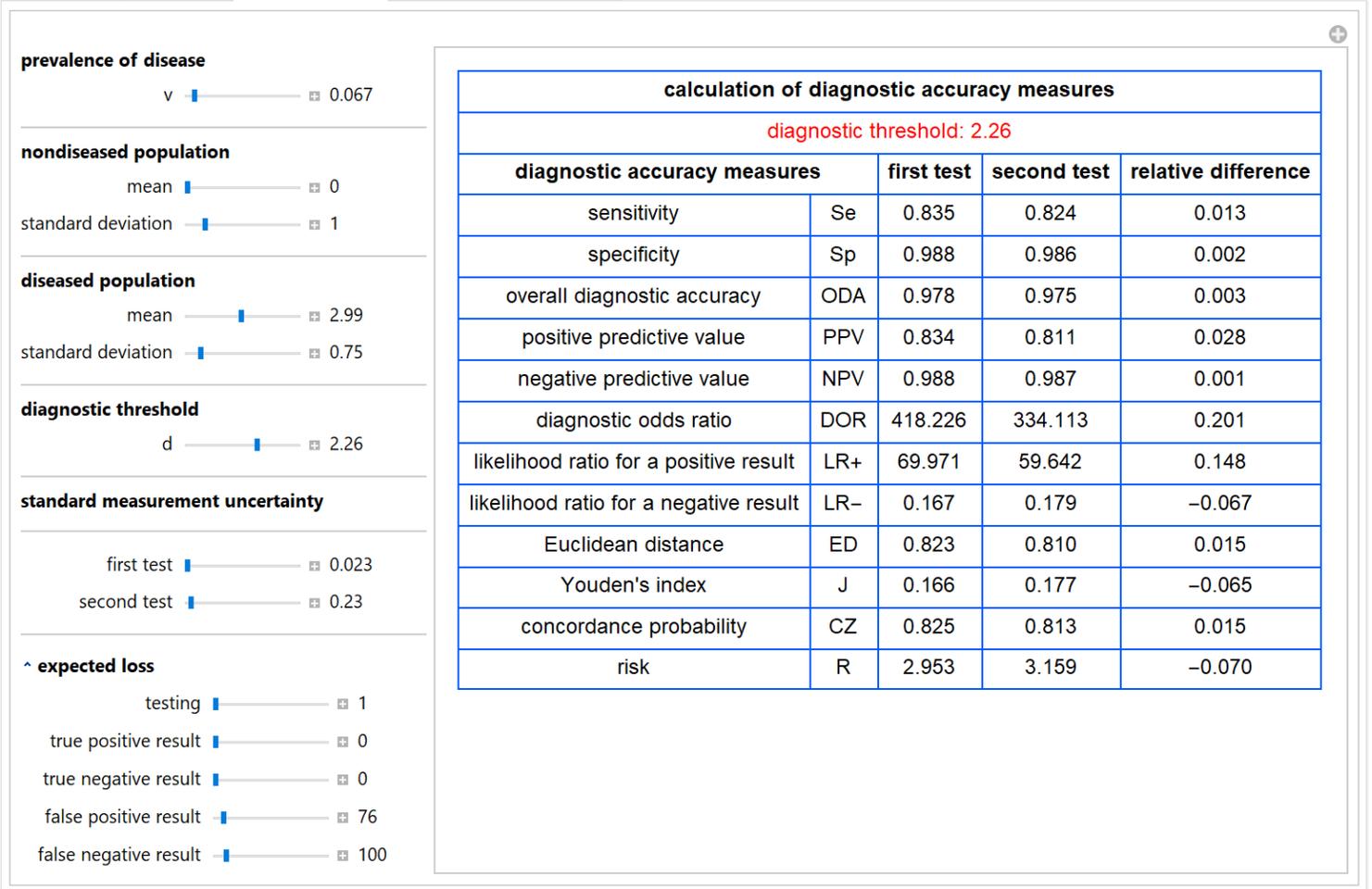


Figure 10

DAM calculator module screenshot. Calculated diagnostic accuracy measures and their relative differences, with the settings at the left

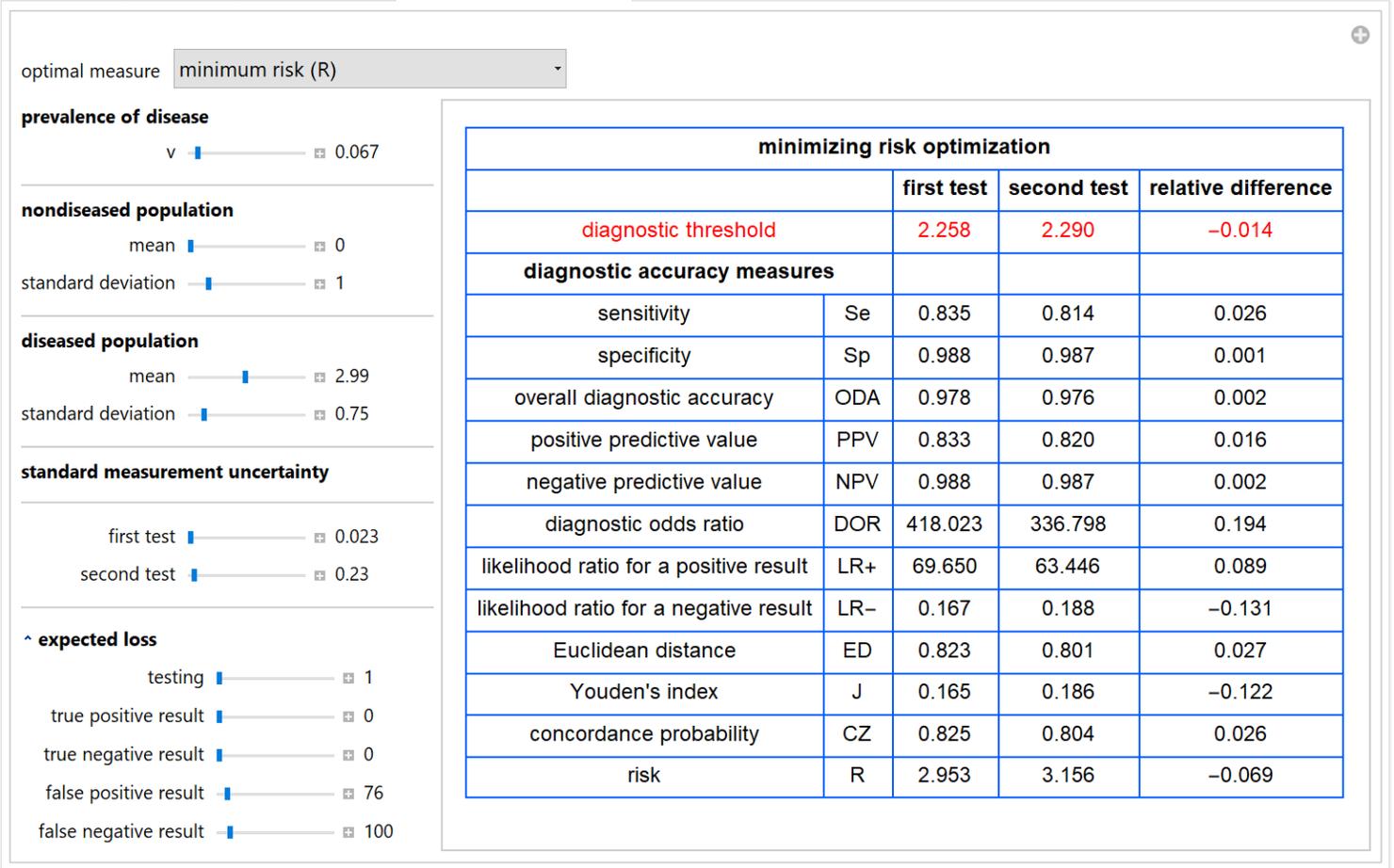
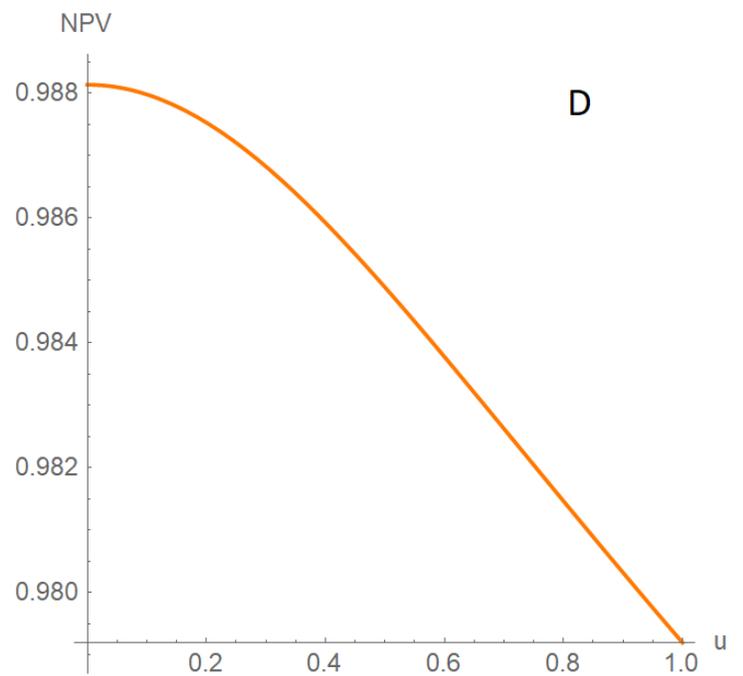
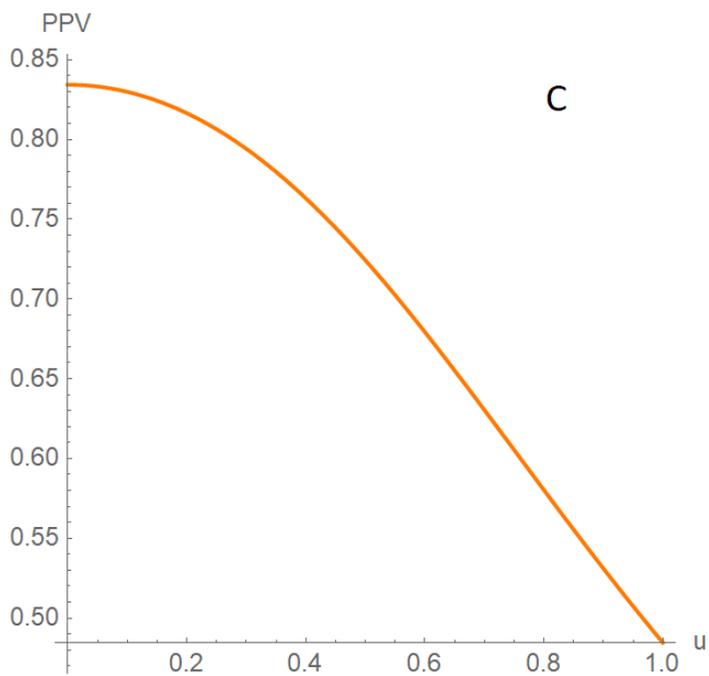
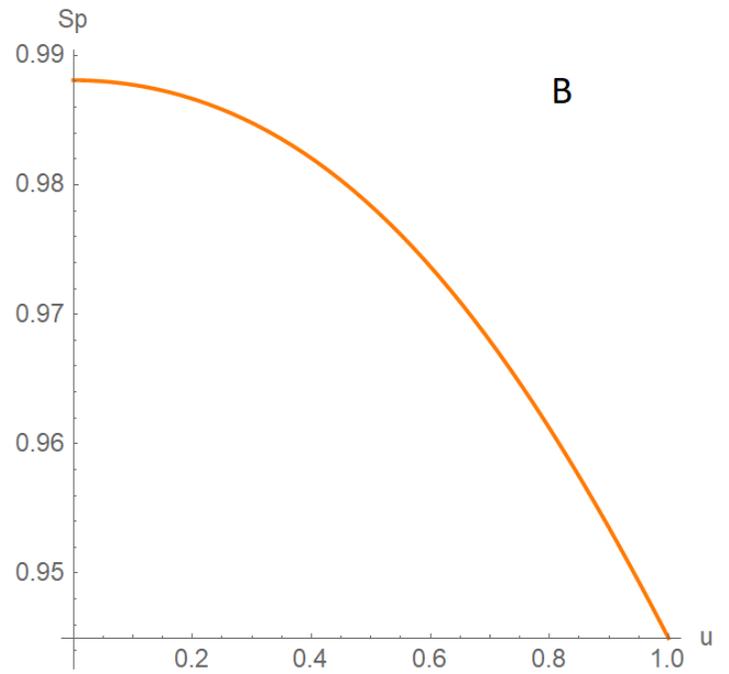
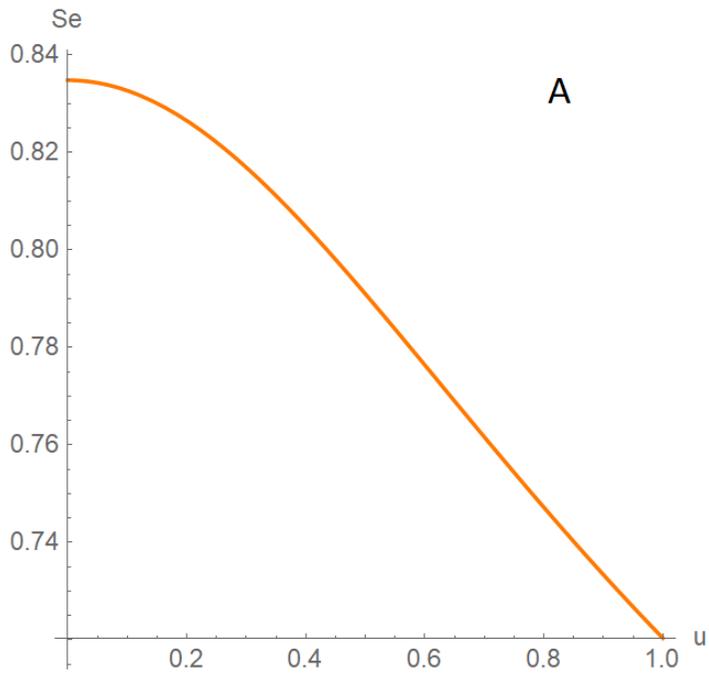


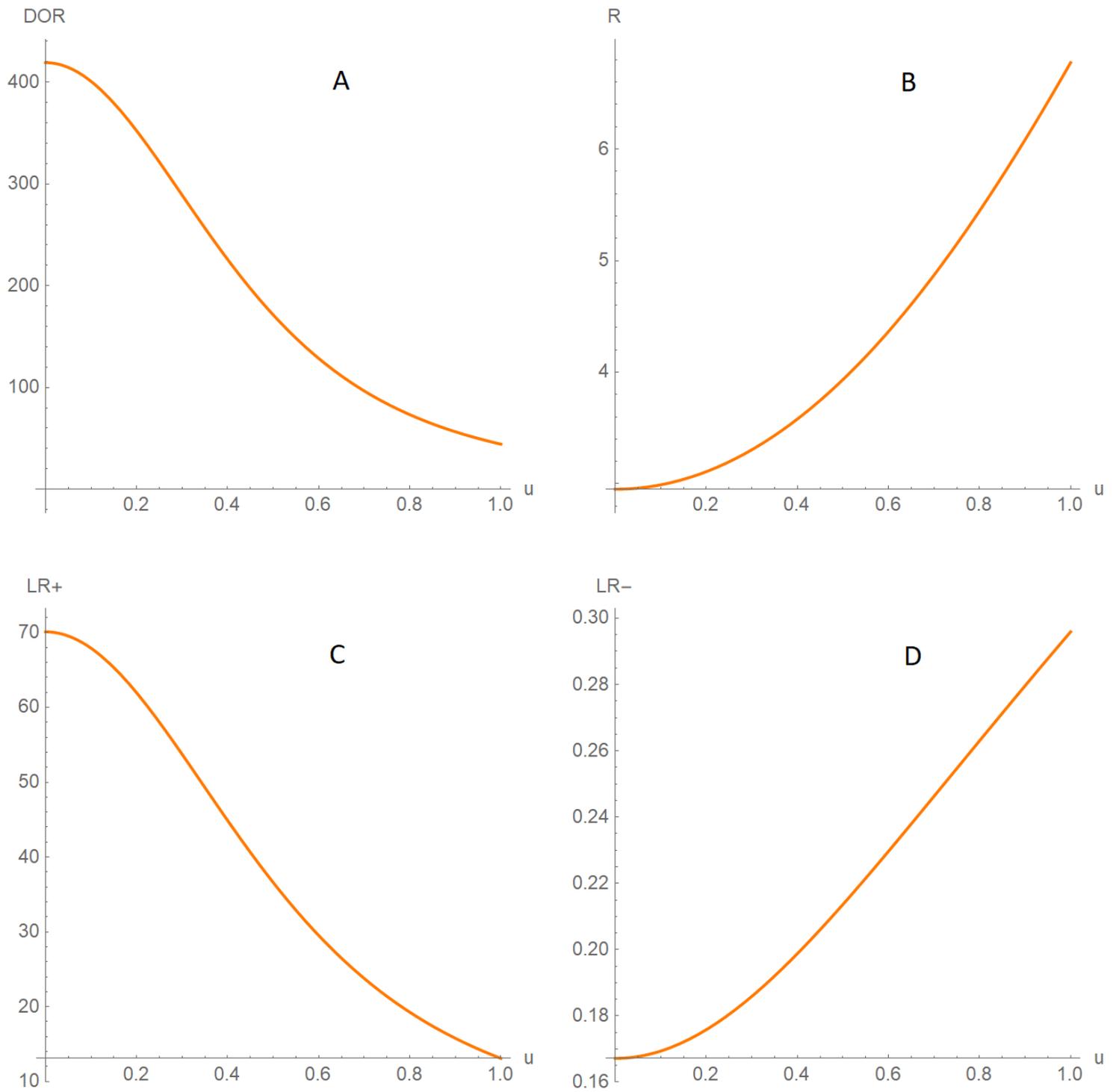
Figure 11

Optimal DAM calculator module screenshot. Calculated diagnostic accuracy measures minimizing risk (R) and their relative differences, with the settings at the left.



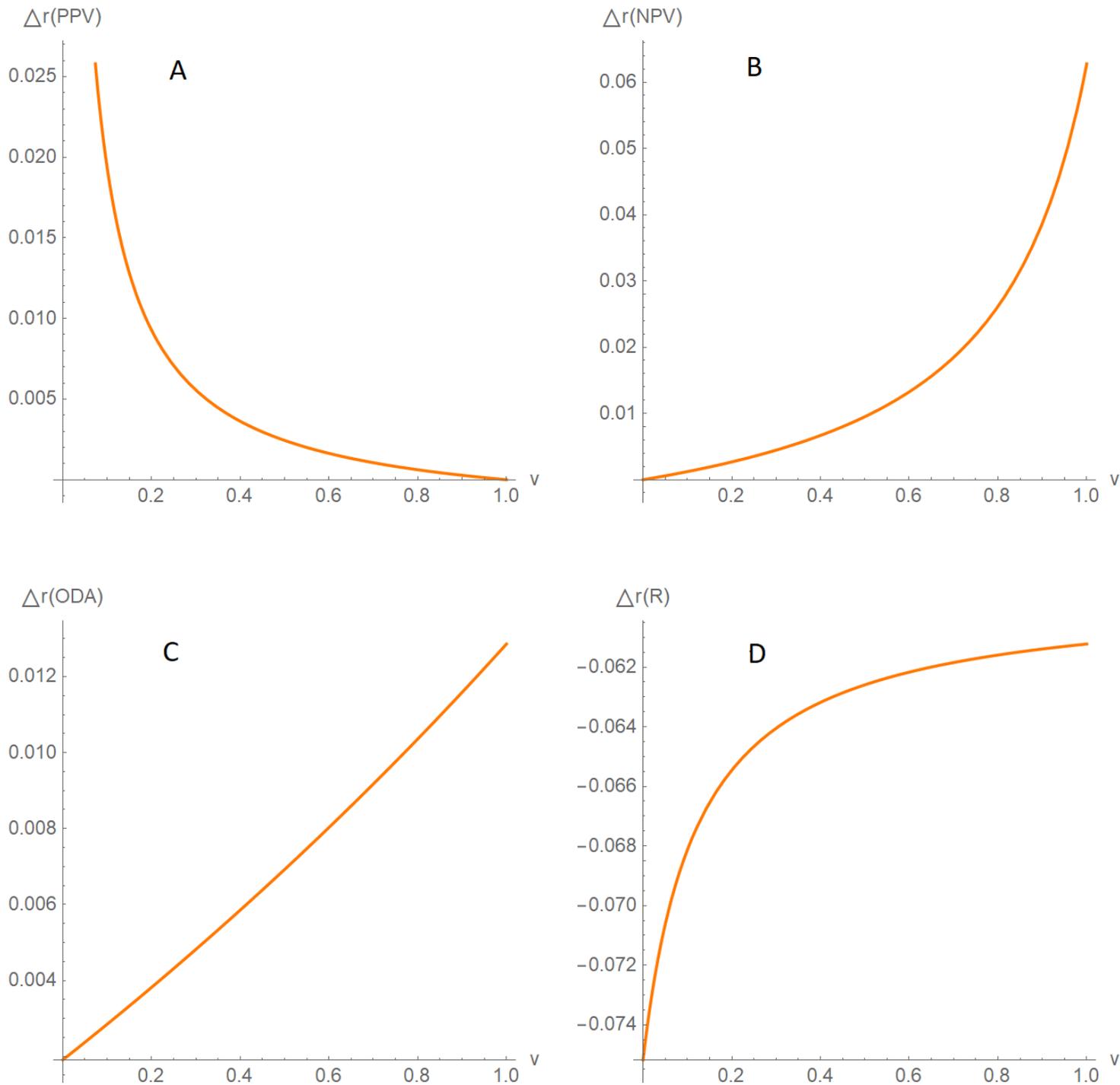
**Figure 12**

DAM versus uncertainty plots. Plots of (A) Sensitivity (Se ), (B) specificity (Sp ), (C) positive predictive value (PPV ) and (D) negative predictive value (NPV ) versus standard measurement uncertainty (u ) curves, with the respective parameters in Table 4.



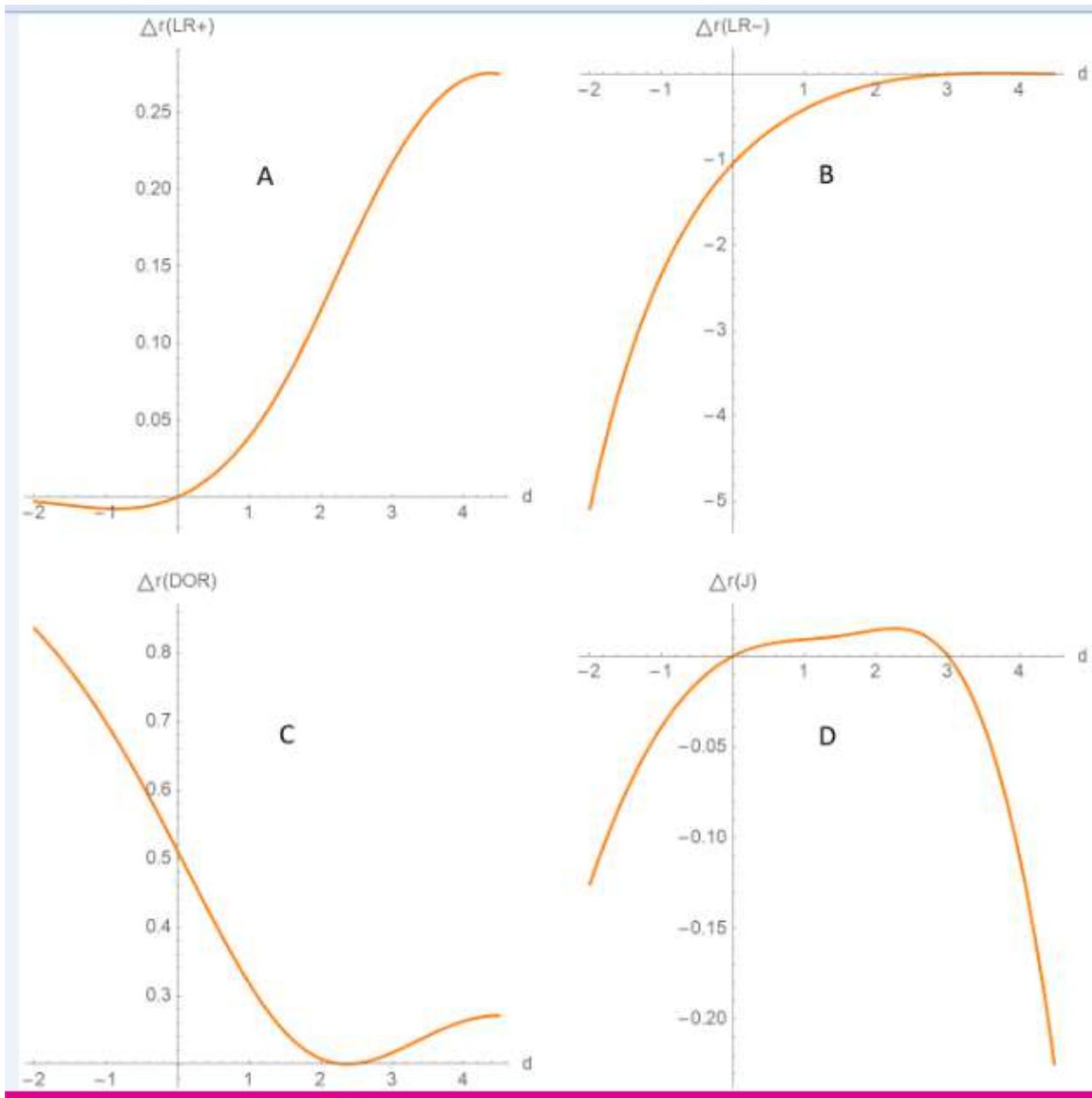
**Figure 13**

DAM versus uncertainty plots. Plots of (A) diagnostic odds ratio (DOR), (B) risk ( $R$ ), (C) likelihood ratio for a positive result ( $LR+$ ) and (D) likelihood ratio for a negative result ( $LR-$ ) versus standard measurement uncertainty ( $u$ ) curves, with the respective parameters in Table 4.



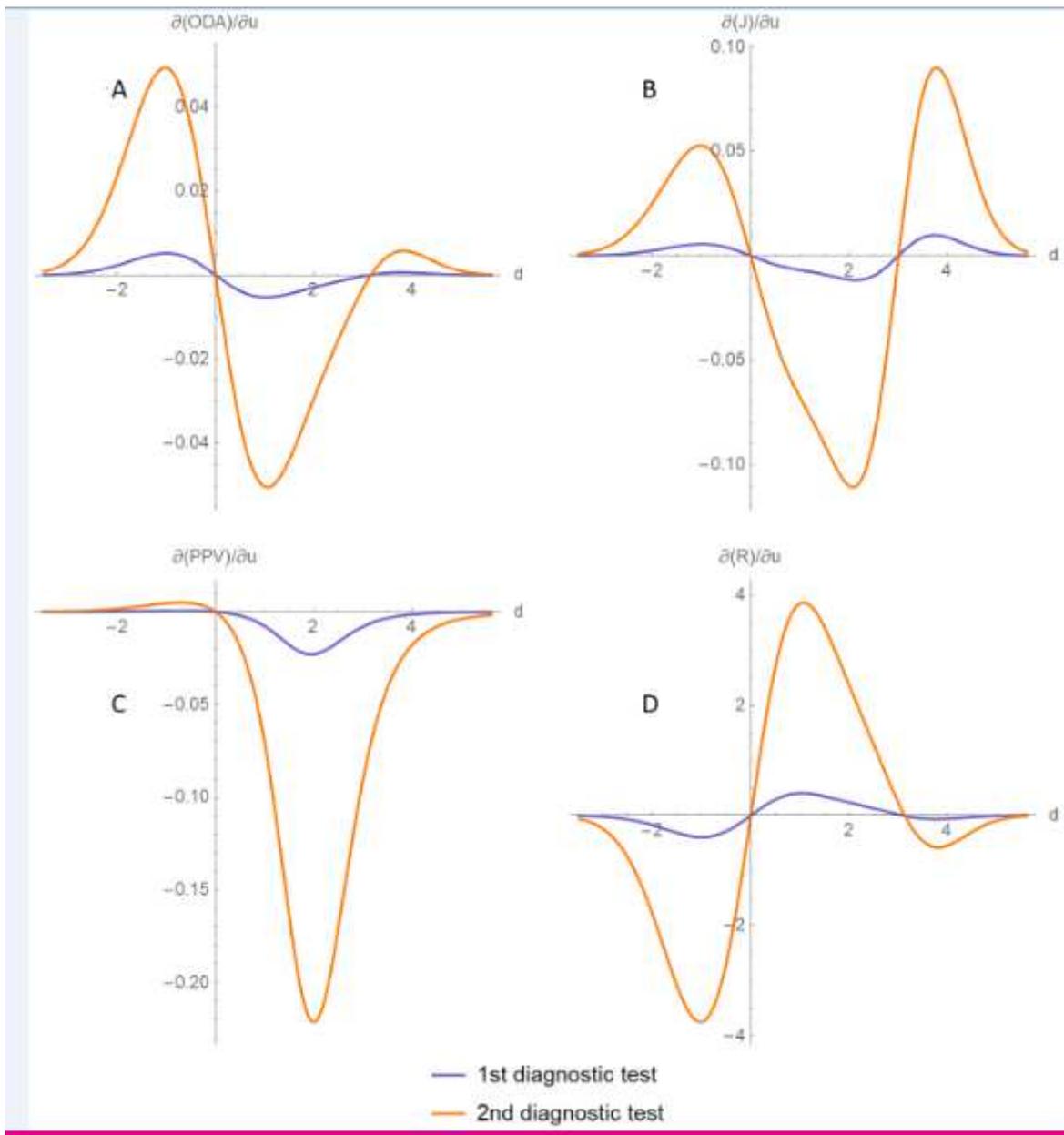
**Figure 14**

DAM relative differences versus prevalence plots. Plots of the relative difference of the (A) positive predictive value (PPV), (B) negative predictive value (NPV), (C) overall diagnostic accuracy (ODA) and (D) risk (R) of two diagnostic or screening tests, measuring the same measurand with different uncertainties, versus prevalence ( $v$ ) curves, with the respective parameters in Table 4.



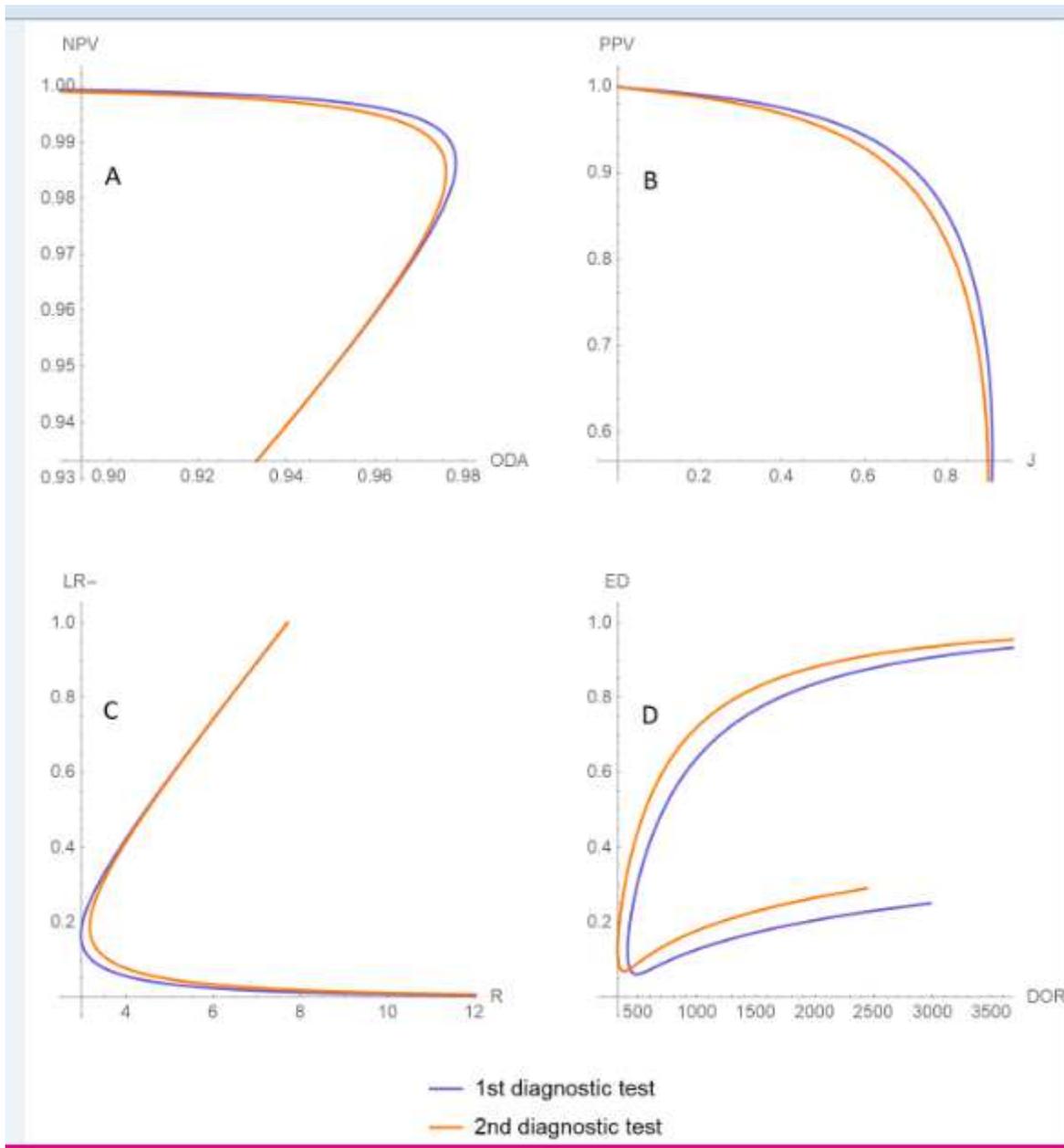
**Figure 15**

DAM relative differences versus diagnostic threshold plots Plots of the relative difference of the (A) likelihood ratio for a positive result (LR + ), (B) likelihood ratio for a negative result (LR - ), (C) diagnostic odds ratio (DOR ) and (D) Youden's index (J ) of two screening or diagnostic tests, measuring the same measurand with different uncertainties, versus diagnostic threshold ( $d$  ) curves, with the respective parameters in Table 4.



**Figure 16**

DAM partial derivatives versus diagnostic threshold plots. Plots of partial derivatives of (A) overall diagnostic accuracy (ODA), (B) Youden's index (J), (C) positive predictive value (PPV) and (D) risk (R), with respect to measurement uncertainty, of two tests, measuring the same measurand with different uncertainties, versus diagnostic threshold ( $d$ ) curves, with the parameters in Table 4.



**Figure 17**

DAM relations plots. Plots of the relations between (A) negative predictive value (NPV ) and overall diagnostic accuracy (ODA ), (B) positive predictive value (PPV ) and Youden's index (J ), (C) likelihood ratio for a negative result (LR - ) and risk (R ) and (D) Euclidean distance (ED ) and diagnostic odds ratio (DOR ), of two tests, measuring the same measurand with different uncertainties, with the respective parameters in Table 4

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Relation.cdf](#)