Hellenic Complex Systems Laboratory

Latent class probabilities and uncertainty estimation without a gold standard

Technical Report XXXII

Theodora Chatzimichail, Rallou A. Chatzimichail and
Aristides T. Hatjimihail
2026

Latent class probabilities and uncertainty estimation without a gold standard

Theodora Chatzimichail[a], MRCS, Rallou A. Chatzimichail[b], MEng, MSc, PhD, Aristides T. Hatjimihail[c], MD, PhD

[a] tc@hcsl.com, [b] rc@hcsl.com, [c] ath@hcsl.com

[a, b, c] *Hellenic Complex Systems Laboratory*

DOI: 10.5281/zenodo.24455278

Abstract

Background: When neither assay provides a gold standard, an individual probability estimate requires a defined outcome. We developed the BDPU-*LatentClass* software, estimating the probability that an individual belongs to the latent class with the higher mean log concentration on assay 1. It separates uncertainty in the fitted parameters from uncertainty due to analytical error in a new assay result.

Methods: We fitted a two-class Gaussian mixture to paired fasting plasma glucose (FPG) and two-hour oral glucose tolerance test (OGTT) glucose measurements from 1,361 NHANES participants. We modeled both measurements on the log scale and allowed the assays to covary within each class. We predicted class membership from a new FPG result by weighting posterior draws by the predictive density of that result.

Results: At an FPG of 126 mg/dL, the estimated probability of class 1 membership was 0.441 (95% parameter credible interval 0.204–0.684). The parameter, measurement and total standard uncertainties were 0.122, 0.082 and 0.147 in probability units. The estimate rose to 0.666 under an exploratory model with diagonal covariance.

Conclusions: BDPU-*LatentClass* reports both class probabilities and their credible intervals, together with the parameter and measurement contributions to uncertainty. The estimate depended on the covariance model. Identifying the higher-FPG class as a disease state requires independent clinical evidence.

Keywords

Bayesian latent class, no gold standard, diagnostic probability, analytical uncertainty, credible interval, NHANES, fasting plasma glucose, oral glucose tolerance test, Wolfram Language

1. Background

An imperfect assay can contribute evidence about an unobserved state only through a model linking that state to the measured quantity [1]. The classical two-test Hui–Walter design uses populations with different prevalences and assumptions about shared accuracy and conditional independence [2]. Dependence can change inferred test performance [3], and surveillance applications require additional assumptions [4]. Parameter counting alone cannot establish identifiability [5].

Here both assays are quantitative. The target is the class with the higher mean log value of the first assay. Separating uncertainty in that class probability from variation in a future binary outcome makes it possible to ask how much analytical error in a new result contributes under a specified model.

Earlier diagnostic software addressed Bayesian diagnostic probabilities and their uncertainty [18–20]. The companion analysis in Technical Report XXXI defines its outcome by a threshold on the continuous reference assay. Here the two continuous assays define latent classes, and the individual prediction concerns class membership.

2. Implementation

2.1.Study population and assay measurements

The demonstration uses paired fasting plasma glucose (LBXGLU) and two-hour OGTT glucose (LBXGLT) from NHANES 2013–2014 [6–9]. We retained participants aged 20–79 years with DIQ010 = 2, two positive assay values, no known pregnancy (RIDEXPRG ≠ 1 and RHD143 ≠ 1), a completed full challenge (GTDCODE = 0 and GTXDRANK = 1), sampling at 105–135 minutes (GTDDR2MN) and a positive OGTT weight (WTSOG2YR). Unknown pregnancy status remained eligible. The protocol and these criteria defined fasting eligibility.

Of 10,175 demographic records, 5,417 met the age criterion, 4,604 met the diagnosis criterion and 1,709 had paired positive measurements. The pregnancy and completion criteria retained 1,706. Timing retained 1,702, all with a positive weight. The original split sampled approximately 20% within each measured-OGTT stratum (cutoff 200 mg/dL, seed 31014). The 1,361 training and 341 held-out participants included 54 and 14 above or at the cutoff, respectively. The likelihood used only the paired measurements. The analysis is unweighted.

2.2.Latent classes and analytical error

For participant i, D_i is a latent class, Z_i is the pair of log concentrations before analytical error, and Y_i is the observed log pair. Each concentration is divided by its stated unit before taking logarithms. Class d has weight π_d, mean μ_d and latent covariance Σ_d. With independent zero-mean Gaussian analytical error of covariance R, the observed covariance is $T_d = \Sigma_d + R$.

$$D_i \sim \text{Bernoulli}(\pi_1), \quad Z_i \mid D_i = d \sim N_2(\mu_d, \Sigma_d)$$
$$Y_i = Z_i + \varepsilon_i, \quad \varepsilon_i \sim N_2(0, R), \quad T_d = \Sigma_d + R$$

The diagonal entries of R are $\log(1 + CV^2)$ for the respective assays. Ordering the first component of the mean vectors orients class 1 toward higher mean log FPG. The two class distributions can overlap. Table 1 gives the primary model settings.

Table 1. Data and primary fit

Setting	Value
Data and split	NHANES 2013–2014, paired FPG and OGTT. Training 1,361, holdout 341.
Class covariance	Two full observed covariance matrices. Both latent matrices are positive definite.
Analytical CVs	FPG 0.0133209788, OGTT 0.0102304629 (assumed proxy).
Prior centers	(90, 95) and (130, 180) mg/dL, transformed componentwise to log scale.
Prior parameters	Dirichlet(1, 1) weights. NIW with $\kappa = 0.1$, $\nu = 4$, $\Psi = \text{diag}(0.04, 0.25)$.
Primary sampler	Four chains, seeds 72101–72104. Each ran for 6,500 iterations. We discarded the first 2,000 and retained every third subsequent draw.
Retained posterior	1,500 draws per chain, 6,000 draws pooled.

The sampler alternates latent allocation, weight and covariance–mean updates. A joint proposal must preserve mean ordering and leave every latent covariance eigenvalue above 10^{-10}. The joint prior has one global normalizing constant. Prior centers specify log locations and supply no independent clinical class anchor.

2.3.Probability from one observed assay result

Before a new result is observed, the fitted class-1 probability is the mean class-1 mixing weight across posterior draws from the training data. For a new log assay-1 value y, the probability is the ratio of the integrated class-1 predictive density to the total predictive density (Appendix A2). Draws that assign a higher predictive density to y receive greater weight.

Let Q be the probability of class 1 given the underlying assay-1 log concentration before analytical error and the model parameters. Conditional on the observed result, the law of total variance separates variance in Q into parameter and measurement components. Their square roots are the corresponding standard uncertainties. The parameter interval uses weighted quantiles of the class probabilities for individual posterior draws. The total-Q interval also accounts for uncertainty in the underlying concentration.

2.4.Notebook workflow

In the main notebook, select ‘Start analysis’. Check the active dataset, enter the observed assay-1 value in the displayed units, choose an interval level and select ‘Calculate probabilities’. The calculator shows the class-1 probability before and after the observation. The ‘Tables’ tab gives both class probabilities on the 0–1 scale, their parameter and total-Q intervals, and the parameter, measurement and total standard uncertainties.

Under ‘Plots’, select ‘Generate curves’ for the chosen grid. Under ‘Tables’, ‘Export result’ saves the current point calculation as a JSON file. ‘Data and model’ offers ‘Fit a custom dataset’ and ‘Restore reference’. ‘Restore reference’ reloads the embedded fit. ‘Open Help and Methods’ opens the companion notebook.

3. Results

3.1.Individual prediction at FPG 126 mg/dL

Before observing the new FPG result, the fitted probability of the higher-mean-log-FPG class was 0.0654. At 126 mg/dL, its posterior probability was 0.4411 and that of the complementary class was 0.5589 (Table 2). Parameter uncertainty was larger than the contribution from analytical error in this new result. Total standard uncertainty was 0.1466. The class probabilities sum to one, and both classes have the same standard uncertainties.

Table 2. Conditional probabilities and uncertainty at FPG 126 mg/dL

Quantity	Point estimate	95% lower	95% upper
Model-defined class 1	0.4411	0.2041	0.6843
Complementary class 0	0.5589	0.3157	0.7959
Class 1 total-Q interval	—	0.1794	0.7371
Class 0 total-Q interval	—	0.2629	0.8206
Parameter standard uncertainty	0.1218	—	—
Measurement standard uncertainty	0.0815	—	—
Total-Q standard uncertainty	0.1466	—	—

The first two rows give the posterior class probabilities and their 95% parameter credible limits. The total-Q limits come from a saved simulation of 24,000 draws. A new simulation can give slightly different endpoints. Standard uncertainties are expressed in probability units.

3.2.Probability and uncertainty across the observed range

Figure 1 extends the point calculation across the observed FPG range. Mean ordering identifies the classes, but the fitted distributions have different variances. The class-1 probability can therefore rise at both ends of the observed range. A reading in the lower tail does not by itself imply higher clinical risk.

The curves connect 103 FPG values from 67 to 333 mg/dL, including 110 and 126 mg/dL. The limits at each FPG value are pointwise 95% credible limits. The total-Q limits use 24,000 draws per value. The reference line marks 126 mg/dL. The squared standard uncertainties add to give the total variance.

At a fixed FPG, the parameter interval summarizes variation in the class-1 probability across posterior draws weighted by the predictive density of that FPG. The total-Q interval also reflects uncertainty in the underlying concentration before analytical error. Each interval applies to one FPG value.

1.1.Sensitivity to model assumptions

At 126 mg/dL, the class-1 probability increased by 0.0021 under the stronger joint prior, by 0.0256 with zero analytical error and by 0.2248 under the diagonal-covariance model (Table 3). Each alternative used one exploratory chain. The calculator embeds the primary fit. The diagonal model also uses a different prior family, so its estimated change reflects both assumptions.

Table 3. Assumption sensitivity at FPG 126 mg/dL

Scenario	Class-1 probability	Probability difference
Primary full covariance	0.4411	0.0000
Stronger joint prior	0.4432	+0.0021
Zero analytical CVs	0.4667	+0.0256
Diagonal covariance	0.6659	+0.2248

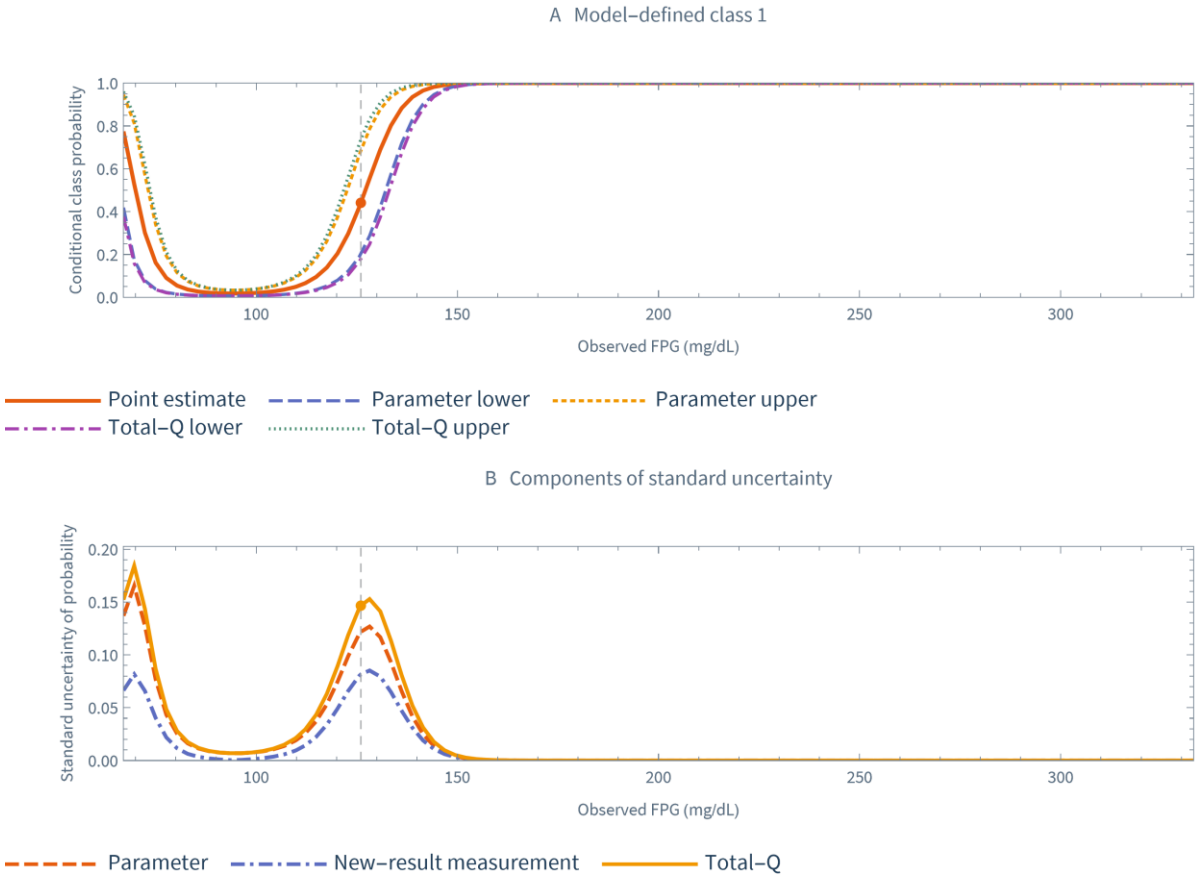


Figure 1. Class probability and conditional standard uncertainty across observed FPG

1.2.Finite-chain precision and measured-data prediction

The probability ratio Monte Carlo standard error (MCSE) was 0.0042. Across the mixing weight, four means, six covariance entries and probability at 126 mg/dL, the maximum rank-normalized split/folded R-hat was 1.00294 and minimum rank-normalized bulk effective sample size was 815 [10]. Predictive-weight effective sample size describes concentration of the new-result weights and does not adjust for chain autocorrelation.

Accuracy summaries use a threshold of 126 mg/dL and 95% training-posterior credible intervals. The holdout score uses the joint density on the concentration scale. Neither measure establishes clinical accuracy.

In 500 posterior predictive replicates of the selected training cohort (seed 20261021), the observed means and 95th percentiles for both assays and the paired log-assay correlation all lay within their respective 95% predictive intervals. Across 20 conditional simulations of 24,000 draws (seeds 8281–8300), lower and upper total-Q endpoints had Monte Carlo SDs of 0.00172 and 0.00268. These SDs describe simulation variability conditional on the saved posterior.

Table 4. Primary fit diagnostics and threshold estimates

Quantity	Result	Interpretation
Probability ratio MCSE at 126	0.004223	Precision of the posterior ratio estimate
R-hat for probability at 126	1.002631	Rank-normalized split/folded check
R-hat for class-1 weight	1.002943	Rank-normalized split/folded check
Predictive-weight ESS	5,933.1	Weight concentration among 6,000 draws
Model-implied sensitivity	0.4363 [0.3107, 0.5971]	Relative to provisional class 1
Model-implied specificity	0.9958 [0.9932, 0.9978]	Relative to provisional class 0
Held-out joint log score	−8.6080 per participant	341 observed assay pairs

2. Discussion

2.1.Interpreting the latent class

Ordering the class means keeps the class labels consistent across posterior draws. Establishing whether class 1 corresponds to clinical disease requires independent evidence. Table 3 shows how much the estimated probability can change across models even when the interval within a given model appears precise.

Models for categorical tests [11, 12] and mixed-membership diagnostic-error models [13] address related questions. A probability of belonging to a binary class does not measure an individual's degree of disease. Precision limits identified for particular binary-test designs [14] do not yield a universal sample-size rule for this continuous model. Repeated measurements, other assays, clinically adjudicated subsets or multiple populations may add information when the joint model supports their use.

2.2.Comparison and scope

Table 5. Scientific tasks addressed by related software

Software	**Model and scope**
BDPU-LatentClass 1.0.0	Two quantitative assays and fixed analytical CVs. Predictive weighting gives individual class probabilities, parameter and total-Q intervals, and the parameter/measurement variance decomposition.
meddecide legacy module [11]	Categorical tests. The documented procedures summarize prevalence, sensitivity, specificity and agreement for the supported latent-class and reference models.
mclust [15]	Finite Gaussian mixtures for clustering, classification and density estimation, with covariance-model selection and resampling.

The entries compare documented modeling tasks. There was no head-to-head refit, speed comparison or assessment of clinical superiority.

The selected cohort was analyzed without survey weights, so its fitted class weights do not estimate national prevalence [17]. The primary model assumes two lognormal classes, fixed analytical CVs and stable measurement processes. Its conditional total excludes uncertainty about model structure, clinical interpretation, transport, assay bias and the CVs. The holdout log score and five predictive checks concern measured data. Neither establishes interval coverage or clinical decision benefit.

3. Conclusions

The notebook estimates class membership from an observed assay result and reports the parameter and measurement contributions to uncertainty. In this example, the exploratory diagonal-covariance model changed the estimate more than setting analytical error to zero. Clinical interpretation requires external evidence.

4. Supplemental material

The following supplemental files are available for download as a ZIP archive at: https://www.hcsl.com/Supplements/SBDPU-LC.zip (accessed on November 23, 2026):

a) Supplemental File I:
BDPU-LatentClass.nb. The main Wolfram notebook fits the model and calculates class probabilities and uncertainties.
b) Supplemental File II:
BDPU-LatentClassHelper.nb. The helper notebook explains the model, probability estimands, uncertainty decomposition and input checks, and gives examples of the calculation functions.

5. List of abbreviations

CV, coefficient of variation. ESS, effective sample size. FPG, fasting plasma glucose. JSON, JavaScript Object Notation. MCSE, Monte Carlo standard error. NHANES, National Health and Nutrition Examination Survey. NIW, normal–inverse-Wishart. OGTT, oral glucose tolerance test. SD, standard deviation.

6. Declarations

6.1.Ethics approval and consent to participate

This secondary analysis used public, deidentified NHANES observations. The NCHS Research Ethics Review Board approved the original 2011–2016 survey under protocol #2011-17 [16]. The authors report that data collection followed the Declaration of Helsinki and participants provided written informed consent. No new recruitment or intervention occurred, and no additional institutional determination was obtained for this analysis.

6.2.Consent for publication

No identifiable individual information or patient images are reported.

6.3.Availability of data and materials

The two notebooks accompany this report. The primary notebook embeds the 1,361 paired training observations and 6,000 retained posterior draws used for the default predictions. The holdout score, alternative fits and repeated-stream diagnostics summarize the original study analyses. The source component documentation identifies the NHANES variables and eligibility fields [6–9].

6.4.Competing interests and funding

The authors declare no competing interests. This research received no external funding.

6.5.Authors' contributions

For the contribution statement, TC denotes Theodora Chatzimichail, RCA denotes Rallou A. Chatzimichail and ATH denotes Aristides T. Hatjimihail. Conceptualization, TC and RCA; methodology, TC, RCA and ATH; software, TC, RCA and ATH; validation, TC; formal analysis, RCA and ATH; investigation, TC; resources, ATH; data curation, TC; writing—original draft, RCA; writing—review and editing, ATH; visualization, TC and RCA; supervision, ATH; project administration, RCA. All authors have read and agreed to the published version of the manuscript.

6.6.Originality and submission approval

The manuscript is original, is not under consideration elsewhere, and has no undisclosed publication overlap. All authors approve submission.

6.7.Acknowledgements and use of generative artificial intelligence

The authors acknowledge assistance from OpenAI Codex with data retrieval and verification scripts, calculator implementation, preparation of supplementary materials, and manuscript drafting and editing. The authors retain responsibility for the scientific claims, interpretation and final submission.

7. References

1. Sun A, Zhou XH. Estimation of diagnostic test accuracy without gold standards. Stat Med. 2025;44(3-4):e10315. doi:10.1002/sim.10315. PMID:39854081.

2. Hui SL, Walter SD. Estimating the error rates of diagnostic tests. Biometrics. 1980;36(1):167–71. PMID:7370371.

3. Keddie SH, Baerenbold O, Keogh RH, Bradley J. Estimating sensitivity and specificity of diagnostic tests using latent class models that account for conditional dependence between tests: a simulation study. BMC Med Res Methodol. 2023;23:58. doi:10.1186/s12874-023-01873-0.

4. Bermingham ML, Handel IG, Glass EJ, Woolliams JA, de Clare Bronsvoort BM, McBride SH, et al. Hui and Walter's latent-class model extended to estimate diagnostic test properties from surveillance data: a latent model for latent data. Sci Rep. 2015;5:11861. doi:10.1038/srep11861.

5. Okkaoglu Y. Estimating the accuracy of a diagnostic test without a gold standard [dissertation]. University of Bristol. 2026. Sections 1.4.3–1.4.4 and 6.2. Source

6. National Center for Health Statistics. NHANES 2013–2014 oral glucose tolerance test data and documentation, OGTT_H. Accessed 9 October 2026. Source

7. National Center for Health Statistics. NHANES 2013–2014 demographic variables and sample weights, DEMO_H. Accessed 9 October 2026. Source

8. National Center for Health Statistics. NHANES 2013–2014 diabetes questionnaire, DIQ_H. Accessed 9 October 2026. Source

9. National Center for Health Statistics. NHANES 2013–2014 laboratory procedure manual. Fasting glucose in plasma. Accessed 9 October 2026. Source

10. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: an improved R-hat for assessing convergence of MCMC. Bayesian Anal. 2021;16:667–718. doi:10.1214/20-BA1221.

11. ClinicoPath Development Team. Analyzing diagnostic tests without a gold standard. meddecide. Accessed 21 October 2026. Source

12. Li T, Liu P. Comparison of two Bayesian methods in evaluation of the absence of the gold standard diagnostic tests. Biomed Res Int. 2019;2019:1374748. doi:10.1155/2019/1374748. PMCID:PMC6720053.

13. Erosheva EA, Joutard C. Estimating diagnostic error without a gold standard: a mixed membership approach. In: Airoldi EM, Blei DM, Erosheva EA, Fienberg SE, editors. Handbook of mixed membership models and their applications. Chapman & Hall/CRC; 2014. p. 141–57. Source

14. Dendukuri N, Bélisle P, Joseph L. Bayesian sample size for diagnostic test studies in the absence of a gold standard: comparing identifiable with non-identifiable models. Stat Med. 2010;29:2688–97. doi:10.1002/sim.4037.

15. mclust authors. Gaussian mixture modelling for model-based clustering, classification, and density estimation. Accessed 21 October 2026. Source

16. National Center for Health Statistics. NHANES NCHS Ethics Review Board approval. Accessed 19 October 2026. Source

17. National Center for Health Statistics. NHANES tutorials. Weighting module. Accessed 9 October 2026. Source

18. Chatzimichail T, Hatjimihail AT. A software tool for estimating uncertainty of Bayesian posterior probability for disease. Diagnostics (Basel). 2024;14(4):402. doi:10.3390/diagnostics14040402.

19. Chatzimichail T, Hatjimihail AT. A software tool for applying Bayes' theorem in medical diagnostics. BMC Med Inform Decis Mak. 2024;24:399. doi:10.1186/s12911-024-02721-x.

20. Chatzimichail T, Hatjimihail AT. A Bayesian inference based computational tool for parametric and nonparametric medical diagnosis. Diagnostics (Basel). 2023;13:3135. doi:10.3390/diagnostics13193135.

8. Appendix A Mathematical details

8.1.A1 Target and restricted joint prior

Let D denote the latent class and $\mathcal{D}$ the training measurements. For dimension p = 2 and class d ∈ {0, 1}, θ comprises class weights, log locations and observed covariances. The joint prior is proportional to the product below. In the normal–inverse-Wishart density, the conditional covariance of the mean is T_d/κ.

$$p(\theta) \propto \mathrm{Dir}(\pi;\, \alpha) \prod_{d=0}^{1} \mathrm{NIW}(\mu_d, T_d;\, m_d, \kappa, \nu, \Psi)\; I_C(\theta)$$

The indicator requires both T_d − R matrices to be positive definite and the class-1 mean log assay-1 value to exceed that of class 0. In computation, the minimum latent eigenvalue must exceed 10^{-10}. Table 1 gives the full-model hyperparameters. The diagonal scenario instead uses independent normal–inverse-gamma priors, with inverse-gamma shape (ν − 1)/2 and scale $\Psi_{jj}/2$, and retains the mean-ordering and positive latent-variance restrictions.

8.2.A2 Individual probability from an observed result

Let y be the new log assay-1 observation and φ the univariate normal density. The observed class mean is $\mu_{d,1}$ and its variance is $T_{d,11}$. Write $a_\theta(y)$ for the weighted class-1 predictive density and $b_\theta(y)$ for the sum of weighted class densities. Their ratio for one θ is $m_\theta(y)$.

$$f_\theta{}^d(y) = \varphi(y;\, \mu_{d,1}, T_{d,11}), \quad m_\theta(y) = \frac{\pi_1 f_\theta{}^1(y)}{\pi_0 f_\theta{}^0(y) + \pi_1 f_\theta{}^1(y)}$$

$$\Pr(D = 1 \mid y, \mathcal{D}) = \frac{\int a_\theta(y)\, p(\theta \mid \mathcal{D})\, d\theta}{\int b_\theta(y)\, p(\theta \mid \mathcal{D})\, d\theta} = \mathrm{E}_{\theta \mid y,\mathcal{D}}[m_\theta(y)]$$

For a possible analytical-error-free log value f, $q_\theta(f)$ uses the latent variances $\Sigma_{d,11} = T_{d,11} - R_{11}$. Conditional on y and θ, F has two normal components weighted by their class-specific contributions to $b_\theta(y)$. Their conditional means and variances follow below.

$$q_\theta(f) = \frac{\pi_1\, \varphi(f;\, \mu_{1,1}, \Sigma_{1,11})}{\pi_0\, \varphi(f;\, \mu_{0,1}, \Sigma_{0,11}) + \pi_1\, \varphi(f;\, \mu_{1,1}, \Sigma_{1,11})}$$

$$\mathrm{E}_{F \mid y,d,\theta}[F] = \mu_{d,1} + \frac{\Sigma_{d,11}}{T_{d,11}}(y - \mu_{d,1}), \quad \mathrm{Var}_{F \mid y,d,\theta}(F) = \frac{\Sigma_{d,11} R_{11}}{T_{d,11}}$$

8.3.A3 Parameter and measurement uncertainty

Define Q = $q_\theta(F)$. Integrating over F conditional on θ and y gives $m_\theta(y)$. The predictive-weighted parameter distribution and the conditional distribution of F then give the variance identity and bounds below. Here p denotes the new-result class probability.

$$u_{par}{}^2 = \mathrm{Var}_{\theta \mid y,\mathcal{D}}[m_\theta(y)], \quad u_{meas}{}^2 = \mathrm{E}_{\theta \mid y,\mathcal{D}}[\mathrm{Var}_{F \mid y,\theta}(q_\theta(F))]$$

$$u_{total}{}^2 = u_{par}{}^2 + u_{meas}{}^2 = \mathrm{Var}(Q \mid y, \mathcal{D}), \quad 0 \le u_{total}{}^2 \le p(1-p) \le \frac{1}{4}$$

Parameter limits are weighted quantiles of $m_\theta(y)$. Total-Q limits are quantiles of jointly sampled θ and F after conditioning on y. The complementary interval is obtained by subtracting the reversed endpoints from one. Future binary-outcome variance includes the additional conditional expectation of Q(1 − Q).

8.4.A4 Model-implied accuracy

For a positive observed assay-1 threshold c, the following expressions give sensitivity and specificity relative to the provisional classes. Φ denotes the standard normal distribution function. Summaries use the retained training-posterior draws without an individual-result update.

$$Se_\theta(c) = 1 - \Phi\left(\frac{\log c - \mu_{1,1}}{T_{1,11}{}^{1/2}}\right), \quad Sp_\theta(c) = \Phi\left(\frac{\log c - \mu_{0,1}}{T_{0,11}{}^{1/2}}\right)$$

9. Appendix B Programming and reproducibility details

9.1.B1 Computation and numerical interpretation

In the helper, select 'Load functions for examples' and press Shift+Enter in an example cell. `lCPriorProbability[lcDefaultFit]` gives the fitted class-1 probability before a new result. lCProbability[lcDefaultFit, 126.] conditions on FPG 126 mg/dL. lCAccuracy gives class-relative sensitivity and specificity. Specify interval levels as fractions, such as 0.95. lCRecord and lCExportRecord prepare and export the result.

The calculator evaluates Gaussian class odds in batches for total-Q samples and normal tail probabilities across draws for threshold summaries. Checks cover class complements, interval ordering, the variance identity and bounds (Appendix A3), and zero measurement variance at zero new-result CV.

An independent 48-node Gauss–Hermite calculation used all 6,000 draws at the 103 saved FPG values. The class-1 probabilities and three standard uncertainties agreed to within 2.0×10^{-14}. The check excluded simulated endpoints.

9.2.B2 Saved and recalculated results

Figure 1 and the saved display use 101 regular FPG nodes plus 110 and 126 mg/dL. The records at 90%, 95% and 99% share one four-chain posterior. The 24,000 total-Q draws per node give marginal limits. Lines between nodes do not depict joint posterior function draws.

At 126 mg/dL, the saved class-1 total-Q interval is 0.1794–0.7371. At the same 95% level, 6,000 posterior draws and 24,000 conditional samples with seed 8271 gave 0.1752–0.7396 in Wolfram. The simulation realization affects the endpoints. Probabilities, parameter limits and standard uncertainties agree to the reported precision.

9.3.B3 Inputs for a new fit

A headerless CSV must contain at least 20 positive numeric pairs, assay 1 then assay 2. In 'Data and model', set assay names, units, CVs, priors and sampler settings, then select 'Import CSV and fit one exploratory chain'. The helper demonstrates lCFit. For inference, run independent chains, examine sampler diagnostics and predictive checks, and assess sensitivity to model assumptions.

10. Software availability and requirements

Software name: BDPU-LatentClass

Version: 1.0.0

DOI: 10.5281/zenodo.24455076

Project home page:

https://www.hcsl.com/Tools/BayesianDiagnosticUncertaintyLatentClass/index.html

Computational source notebook: BDPU-LatentClass.nb. Available at:
https://www.hcsl.com/Tools/BayesianDiagnosticUncertaintyLatentClass/BDPU-LatentClass.nb

Programming language: Wolfram Language

Operating systems: Microsoft Windows 10+, Linux 4.18+, Apple macOS 13+

Other software requirements: Wolfram Player (freely available) or Wolfram Mathematica.

Hardware requirements: x86-64-v2 and higher or Apple Silicon microarchitecture CPU, and 4 GB+ RAM, recommended.

Software citation: Chatzimichail T, Chatzimichail RA, Hatjimihail AT. *BDPU-LatentClass: Latent-Class Probabilities and Uncertainty from Paired Quantitative Assays*. Ver. 1.0.0. Hellenic Complex Systems Laboratory; 2026. https://doi.org/10.5281/zenodo.24455076

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

11. Citation

Chatzimichail T, Chatzimichail RA, Hatjimihail AT. *Latent class probabilities and uncertainty estimation without a gold standard*. Technical Report XXXII. Hellenic Complex Systems Laboratory; 2026. https://www.doi.org/10.5281/zenodo.24455278

12. License

Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

First Published: November 23, 2026