Welcome and Introductions - Introduction to Bayesian Analysis

CA18208 HARMONY Serbia Training School - https://harmony-net.eu/

Julio Alvarez Eleftherios Meletis

2022-08-31

Welcome and Introductions

Learning aims

By attending this training, participants will:

- Perceive the logic of latent class models and their applicability in diagnostic accuracy studies in veterinary medicine
- Get acquainted with Bayesian Latent Class Models (**BLCMs**) basic principles & challenges
- Perform hands-on training on Sensitivity (Se) and Specificity (Sp) estimation with BLCMs
- Understand the importance of standards for reporting of diagnostic accuracy studies that use BLCMs (STARD-BLCMs)

Schedule for today (31-08-2022)

Schedule tomorrow (01-09-2022)

Schedule (02-09-2022)

Some Housekeeping

- Please sign the attendance sheet every day to be eligible for reimbursement
- Make sure that you have filled in the circulated Google Doc for the reimbursement

Historical Sketch on (B)LCMs

- A brief introduction to the logic of diagnostic test evaluation
- An introduction to (Bayesian) Latent Class Models

Evaluation of diagnostic test accuracies

JOURNAL OF
MEDICAL VIROLOGY

LETTER TO THE EDITOR | Free Access

Should RT-PCR be considered a gold standard in the diagnosis of COVID-19?

Moustapha Dramé MD, PhD, Maturin Tabue Teguo MD, PhD, Emeline Proye MD, Fanny Hequet MD, Maxime Hentzien MD, PhD, Lukshe Kanagaratnam MD, PhD, Lidvine Godaert MD, PhD

First published: 08 May 2020 | https://doi.org/10.1002/jmv.25996 | Citations: 31

Evaluation of diagnostic test accuracies

JOURNAL OF
MEDICAL VIROLOGY

LETTER TO THE EDITOR | Free Access

Performance of VivaDiag COVID-19 IgM/IgG Rapid Test is inadequate for diagnosis of COVID-19 in acute patients referring to emergency room department

Irene Cassaniti, Federica Novazzi, Federica Giardina, Francesco Salinaro, Michele Sachs, Stefano Perlini, Raffaele Bruno, Francesco Mojoli, Fausto Baldanti ... See all authors

First published: 30 March 2020 | https://doi.org/10.1002/jmv.25800 | Citations: 61

Evaluation of diagnostic test accuracies

- «. . . indeed, when an existing test is considered as a reference, this suggests that the test in question is always correct, and that all misclassifications (false negatives, false positives) are due to the new test. . . »
- «Consequently, the new test will **never** be able to achieve sensitivity of 100%, since it is considered responsible for all misclassifications.»

JOURNAL OF
MEDICAL VIROLOGY

LETTER TO THE EDITOR | Free Access

Should RT-PCR be considered a gold standard in the diagnosis of COVID-19?

Moustapha Dramé MD, PhD, Maturin Tabue Teguo MD, PhD, Emeline Proye MD, Fanny Hequet MD, Maxime Hentzien MD, PhD, Lukshe Kanagaratnam MD, PhD, Lidvine Godaert MD, PhD

First published: 08 May 2020 | https://doi.org/10.1002/jmv.25996 | Citations: 31

Recap on diagnostic test accuracies

Sensitivity & Specificity

- Sensitivity is the ability of a diagnostic test, to correctly classify infected individuals
- Specificity is the ability of a diagnostic test, to correctly classify healthy individuals

	Infected	Healthy	
Test (+)	80	5	85
Test (-)	20	95	115
	100	100	200

- Se 80% and Sp of 95%

Recap on diagnostic test accuracies

Sensitivity & Specificity

- Sensitivity is the ability of a diagnostic test, to correctly classify infected individuals
- Specificity is the ability of a diagnostic test, to correctly classify healthy individuals

	Infected	Healthy	
Test (+)	80	5	85
Test (-)	20	95	115
	100	100	200

- Se and Sp?

Recap on diagnostic test accuracies

Sensitivity & Specificity

- Sensitivity is the ability of a diagnostic test, to correctly classify infected individuals
- Specificity is the ability of a diagnostic test, to correctly classify healthy individuals

	Test 2 (+)	Test 2 (-)	
Test 1 (+)	80	5	85
Test 1 (-)	20	95	115
	100	100	200

- Se_1, Se_2 and Sp_1, Sp_2?

Evaluation of diagnostic test accuracies in the absence of a true gold standard

JOURNAL OF
MEDICAL VIROLOGY

LETTER TO THE EDITOR | Open Access

Bayesian latent class models to estimate diagnostic test accuracies of COVID-19 tests

Sonja Hartnack ✉, Paolo Eusebi, Polychronis Kostoulas

First published: 08 August 2020 | https://doi.org/10.1002/jmv.26405

University of Zurich

Historical sketch on (Bayesian) Latent Class Models

- Hui-Walter paradigm (1980)
  - A particular model formulation that was originally designed for evaluating diagnostic tests in the absence of a gold standard
  - Not originally/necessarily Bayesian - implemented using Maximum Likelihood
  - But evaluating an imperfect test against another imperfect test is a bit like pulling a rabbit out of a hat
- If we don't know the true disease status, how can we estimate sensitivity or specificity for either test?

Hui-Walter paradigm (1980)

Hui-Walter models implementation to be further discussed in the next session.

Population 1

		T2+	T2-
D+	T1+	P1*Se1*Se2	P1*Se1*(1-Se2)
	T1-	P1*(1-Se1)*Se2	P1*(1-Se1)*(1-Se2)

		T2+	T2-
D-	T1+	(1-P1)*(1-Sp1)*(1-Sp2)	(1-P1)*(1-Sp1)*Sp2
	T1-	(1-P1)*Sp1*(1-Sp2)	(1-P1)*Sp1*Sp2

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)

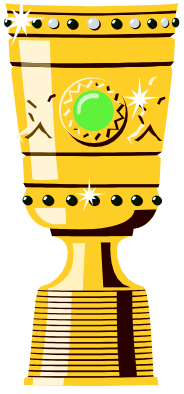
Population 1	
T1+T2+:	P1*Se1*Se2+(1-P1)*(1-Sp1)*(1-Sp2)
T1+T2-:	P1*Se1*(1-Se2)+(1-P1)*(1-Sp1)*Sp2
T1-T2+:	P1*(1-Se1)*Se2+(1-P1)*Sp1*(1-Sp2)
T1-T2-:	P1*(1-Se1)*(1-Se2)+(1-P1)*Sp1*Sp2

- 5 parameter and 3 degrees of freedom
- Non identifiable model

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)

Population 1	
T1+T2+:	P1*Se1*Se2+(1-P1)*(1-Sp1)*(1-Sp2)
T1+T2-:	P1*Se1*(1-Se2)+(1-P1)*(1-Sp1)*Sp2
T1-T2+:	P1*(1-Se1)*Se2+(1-P1)*Sp1*(1-Sp2)
T1-T2-:	P1*(1-Se1)*(1-Se2)+(1-P1)*Sp1*Sp2
Population 2	
T1+T2+:	P2*Se1*Se2+(1-P2)*(1-Sp1)*(1-Sp2)
T1+T2-:	P2*Se1*(1-Se2)+(1-P2)*(1-Sp1)*Sp2
T1-T2+:	P2*(1-Se1)*Se2+(1-P2)*Sp1*(1-Sp2)
T1-T2-:	P2*(1-Se1)*(1-Se2)+(1-P2)*Sp1*Sp2



Identifiable model!

6=6

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests

Conditionl dependence to be further discussed today in the last session.

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests

TABLE 2. Maximum Number of Estimable Parameters and Number of Parameters to Be Estimated in the Absence of Conditional Independence and Under Conditional Independence as a Function of the Number of Tests per Subject

Number of Tests	Maximum Number of Estimable Parameters	Parameters to be Estimated Under Conditional Dependence	Parameters to Be Estimated Under Conditional Independence
1	1	3	3
2	3	7	5
3	7	15	7
4	15	31	9
5	31	63	11
h	$2^h - 1$	$2^{h+1} - 1$	$2h + 1$

Berkvens D et al. (2006) Estimating Disease Prevalence in a Bayesian Framework Using Probabilistic Constraints. doi: 10.1097/01.ede.0000198422.64801.8d

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard

Historical sketch on (B)LCMs

Joseph et al. (1995) **Bayesian** estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard

- Remember Bayes' theorem? $P(\theta|Y) = \frac{P(\theta) \times P(Y|\theta)}{P(Y)}$
  - prevalence $\pi = P(D+)$
  - Sensitivity $Se_i = P(T_i + |D+)$
  - Specificity $Sp_i = P(T_i - |D-)$
  - Prior beta distributions for parameters of interest
    - $\pi \sim Beta(a_\pi, b_\pi)$
    - $Se_i \sim Beta(a_{Se_i}, b_{Se_i})$
    - $Sp_i \sim Beta(a_{Sp_i}, b_{Sp_i})$

$Posterior \propto Likelihood \times Prior$

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard
- Bayesian Using Gibbs Sampling (BUGS) (1997) - Software

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard
- Bayesian Using Gibbs Sampling (BUGS) (1997) - Software
- Enoe et al. (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease is unknown

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard
- Bayesian Using Gibbs Sampling (BUGS) (1997) - Software
- Enoe et al. (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease is unknown
- Just another Gibbs sampler (JAGS) (2007) - Software

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard
- Bayesian Using Gibbs Sampling (BUGS) (1997) - Software
- Enoe et al. (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease is unknown
- Just another Gibbs sampler (JAGS) (2007) - Software
- Toft et al. (2007) Assessing the convergence of Markov Chain Monte Carlo methods: an example from evaluation of diagnostic tests in absence of a gold standard

Historical sketch on (B)LCMs

- Hui-Walter paradigm (1980)
- Vacek (1985) - captures the conditional dependence between diagnostic tests
- Joseph et al. (1995) - Bayesian estimation of disease prevalence and diagnostic test evaluation in the absence of a gold standard
- Bayesian Using Gibbs Sampling (BUGS) (1997) - Software
- Enoe et al. (2000) Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease is unknown
- Just another Gibbs sampler (JAGS) (2007) - Software
- Toft et al. (2007) Assessing the convergence of Markov Chain Monte Carlo methods: an example from evaluation of diagnostic tests in absence of a gold standard
- Kostoulas et al. (2017) STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use BLCMs

BLCMs are endorsed by OIE

2.2.2. Samples from animals of unknown status

When the so-called reference standard is imperfect, which is the rule with any diagnostic tests, estimates of DSe and DSp for the candidate assay based on this standard will be flawed. A way to overcome this problem is to perform a latent class analysis of the joint results of the two tests assuming neither test is perfect.

Latent-class models do not rely on the assumption of a perfect reference test but rather estimate the accuracy of the candidate test and the reference standard with the joint test results (Branscum et al., 2005; Enøe et al., 2000; Georgiadis et al., 2003; Hui & Walter, 1980). If a Bayesian latent class analysis is used, prior knowledge about the performance of the reference test and the candidate test can be incorporated into the analysis.

Because these statistical models are complex and require critical assumptions, statistical assistance should be sought to help guide the analysis and describe the sampling from the target population(s), the characteristics of other tests included in the analysis, the appropriate choice of model and the estimation methods based on peer-reviewed literature (see *Terrestrial Manual* Chapter 3.6.5 [footnote [14]] for details).

Terminology *latent*

Yes, what does *latent* mean?

- The true infection status of an individual is unobserved-hidden/unknown, hence *latent*
- Instead of individuals being explicitly classified as *infected* or *uninfected*, each individual is assumed to have a probability of infection, given the combination of an observed diagnostic test outcome, knowledge on Se and Sp and prior knowledge of disease prevalence in the population of interest (Cheung et al. 2021).
- In a BLCM *latent* does mean soemthing different as in a "latent herpes infection".

Summary

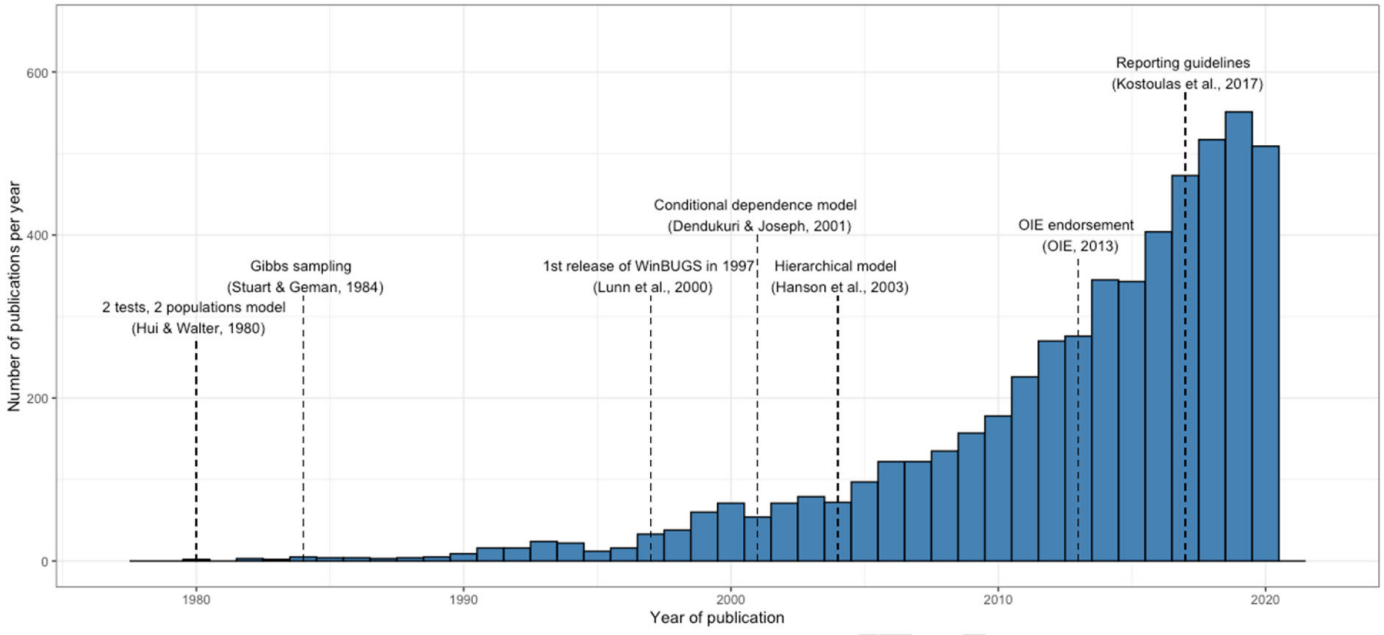


Fig. 2

Frequency histogram of the number of peer-reviewed articles published on latent class analysis when there is an imperfect reference test

References (I)

- Hui, S. L., & Walter, S. D. (1980). Estimating the error rates of diagnostic tests. Biometrics, 167-171.
- Vacek, P. M. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics, 959-968.The effect of conditional dependence on the evaluation of diagnostic tests
- Joseph, L., Gyorkos, T. W., & Coupal, L. (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. American journal of epidemiology, 141(3), 263-272.
- Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. J. (2002). WinBUGS: Bayesian Inference Using Gibbs Sampling Manual, Version 1.4. London: Imperial College; Cambridge, UK: MRC Biostatistics Unit.

References (II)

- Enøe, C., Georgiadis, M. P., & Johnson, W. O. (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Preventive veterinary medicine, 45(1-2), 61-81.
- Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, No. 125.10, pp. 1-10).
- Toft, N., Innocent, G. T., Gettinby, G., & Reid, S. W. (2007). Assessing the convergence of Markov Chain Monte Carlo methods: an example from evaluation of diagnostic tests in absence of a gold standard. Preventive veterinary medicine, 79(2-4), 244-256.
- Kostoulas, P., Nielsen, S. S., Branscum, A. J., Johnson, W. O., Dendukuri, N., Dhand, N. K., ... & Gardner, I. A. (2017). STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. Preventive veterinary medicine, 138, 37-47.

End of intro

Any questions so far?