# BLCMs with Covariates

Sonja Hartnack  Valerie Hungerbühler  Eleftherios Meletis

2022-07-15

## Example from Veterinary Epidemiologic research, 2nd ed., Dohoo, p. 111-113

- Test are considered conditionally independent if the probability of getting a given test results on one test does not depend on the result from the other test, given the disease status of the individual.
- If tests are conditionally independent, the formulae for $Se$ and $Sp$ under parallel $(Se_p, Sp_p)$ and series $(Se_s, Sp_s)$ interpretation are:
- $Se_p = Se_1 + Se_2 - (Se_1 * Se_2)$
- $Sp_p = Sp_1 * Sp_2$
- $Se_s = Se_1 * Se_2$
- $Sp_s = Sp_1 + Sp_2 - (Sp_1 * Sp_2)$

# Example from Veterinary Epidemiologic research, 2nd ed., Dohoo, p. 111-113

**Example 5.8 Multiple tests—series versus parallel interpretation**
data = ISA_test

The data in this example are from the ISA_test dataset. The tests we are using are the indirect fluorescent antibody test (IFAT) and the polymerase chain reaction (PCR) test, with clinical disease status (see dataset description Chapter 31) as the gold standard. The observed joint distributions of test results and virus presence are shown below along with the 4 possible test interpretation criteria.

| | Number of fish by test-result category | | | | Totals |
|---|---|---|---|---|---|
| **IFAT result** | + | + | 0 | 0 | |
| **PCR result** | + | 0 | + | 0 | |
| **Diseased fish** | 134 | 4 | 29 | 9 | 176 |
| **Non-diseased fish** | 0 | 28 | 12 | 534 | 574 |
| **Series interpretation** | + | 0 | 0 | 0 | |
| **Parallel interpretation** | + | + | + | 0 | |

*Se* of IFAT only = 138/176 = 0.784

*Sp* of IFAT only = 546/574 = 0.951

*Se* of PCR only = 163/176 = 0.926

*Sp* of PCR only = 562/574 = 0.979

*Se* of series interpretation = 134/176 = 0.761

*Se* of parallel interpretation = (134+4+29)/176 = 0.949

*Sp* of series interpretation = (28+12+534)/574 = 1.000

*Sp* of parallel interpretation = 534/574 = 0.930

# Example from Veterinary Epidemiologic research, 2nd ed., Dohoo, p. 111-113

| | Number of fish by test-result category | | | | Totals |
|---|---|---|---|---|---|
| **IFAT result** | + | + | 0 | 0 | |
| **PCR result** | + | 0 | + | 0 | |
| **Diseased fish** | 134 | 4 | 29 | 9 | 176 |
| **Non-diseased fish** | 0 | 28 | 12 | 534 | 574 |
| **Series interpretation** | + | 0 | 0 | 0 | |
| **Parallel interpretation** | + | + | + | 0 | |

*Se* of IFAT only = 138/176 = 0.784

*Sp* of IFAT only = 546/574 = 0.951

*Se* of PCR only = 163/176 = 0.926

*Sp* of PCR only = 562/574 = 0.979

*Se* of series interpretation = 134/176 = 0.761

*Se* of parallel interpretation = (134+4+29)/176 = 0.949

*Sp* of series interpretation = (28+12+534)/574 = 1.000

*Sp* of parallel interpretation = 534/574 = 0.930

**Table 5.3 Expected Se and Sp levels with combined tests for ISA assuming conditional independence (data from Example 5.8)**

| | Sensitivity | | Specificity | |
|---|---|---|---|---|
| **Interpretation** | **Expected** | **Observed** | **Expected** | **Observed** |
| **Parallel** | 0.784+0.926 - 0.784*0.926=0.984 | 0.949 | 0.951*0.979=0.931 | 0.930 |
| **Series** | 0.784*0.926=0.726 | 0.761 | 0.951+0.979 - 0.979*0.951=0.999 | 1.000 |

## Estimating covariances between test results

- Using the $Se$ and $Sp$ estimates from the ISA example, the covariances in the $D+$ and the $D-$ groups are:
  - $covar(+) = Se_s(obs) - (Se_1 * Se_2) = 0.761 - 0.726 = 0.035$
  - $covar(-) = Sp_p(obs) - (Sp_1 * Sp_2) = 0.930 - 0.931 = -0.001$

## Conditional dependencies

- Conditional independence implies that given an animal is diseased (or not) the probability $P$ of positive (or negative) outcomes for $T_1$, the test results of the first test, is the same - regardless of the known outcome for the second test, $T_2$.

# Example of a COVID-19 data set

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

## Example of a COVID-19 data set

1. how to prepare the data set in the correct format
   - create_data_cassaniti.R
2. how to describe the model
   - model_final.bug
3. how to run the model in JAGS with runjags
   - runjags.version.R
4. how to check convergence and how to analyse the data

## How to prepare the data set

data from Cassaniti et al. 2020 DOI: 10.1002/jmv.25800
https://github.com/shartn/BLCM-COVID19

## Exercises

- Ex. 1
  - Can you re-run the exercises?
  - Assess what happens if you add other covariances?
  - How many could you add and still have “meaningful results”?
  - Try different priors
  - Looking at the runjags reference manual, could ou customize the plots (just showing trace plots and histograms)?
- Ex. 2 (Bonus)
  - Could you expand the model with a fourth test with simulated data?

# Exercise 3 Covariates

- Explore the data set 'echinococcus.csv' PCR for either E. multilocularis or E. granulosus, ELISA for both, eggs found by arecoline purgation, Tawnia co-infection, age and sex
- Run classical 'risk factor analysis': is sex, Taenia co-infection or age a risk factor for echinococcus (PCR-prevalence, seroprevalence or purges)? Obtain p-values and ORs with confidence intervals.

# Exercise 4 Covariates

- Prepare the data set in the correct format (dump, add ones) for BLCM
- name it : m.short <- as.matrix(dat)
- Run a model for three tests (assume a very high sensitivity for arecoline purgation)
- Try different priors
- Evidence of conditional dependencies
- Is there evidence for a covariate effect on the prevalence? -compare your finding with Ex.3