

# Introduction to BDM for health scientists

## Training Material

Eleftherios Meletis, Konstantinos Pateras, Julio Alvarez, Matt  
Denwood, Polychronis Kostoulas

25-27 October 2022

Sensitivity - Specificity estimation with and  
without a gold standard

# 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%

Diagnostic test evaluation: with gold standard =  
simple!/straightforward!

# 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

	<del>Infected</del>	<del>Healthy</del>	
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  $Sp_1$ ,  $Sp_2$ ?

## Hui-Walter paradigm/model (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
- ▶ 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/model (1980)

[https://www.youtube.com/watch?v=z6devQmW2xE&ab\\_channel=PolychronisKostoulas](https://www.youtube.com/watch?v=z6devQmW2xE&ab_channel=PolychronisKostoulas)



## Hui-Walter paradigm (1980)

- ▶ Hui-Walter models implementation to be further discussed in the next session.
- ▶ But we will use the data/observations from the manuscript published back in 1980.

# Hui-Walter (1980) dataset

**Table 1**  
*Results of Mantoux and Tine tests for tuberculosis in two populations*

Mantoux test	Population 1			Population 2		
	Tine test			Tine test		
	Positive	Negative	Total	Positive	Negative	Total
Positive	14	4	18	887	31	918
Negative	9	528	537	37	367	404
Total	23	532	555	924	398	1322

## Encode the Table\_1 data in RStudio

```
pop_1 = matrix(nrow=3,ncol=3)
rownames(pop_1) = c("Mantoux_Test_Pos", "Mantoux_Test_Neg",
colnames(pop_1) = c("Tine_Test_Pos", "Tine_Test_Neg", "Total")

pop_1[1,1] = 14
pop_1[1,2] = 4
pop_1[2,1] = 9
pop_1[2,2] = 528
#Total rows and columns
pop_1[1,3] = pop_1[1,1] + pop_1[1,2]
pop_1[2,3] = pop_1[2,1] + pop_1[2,2]
pop_1[3,1] = pop_1[1,1] + pop_1[2,1]
pop_1[3,2] = pop_1[1,2] + pop_1[2,2]
N_1 = sum(pop_1[1,1] + pop_1[1,2] + pop_1[2,1] + pop_1[2,2])
pop_1[3,3] = N_1
pop_1
```

## Now let's do pop\_2

```
pop_2 = matrix(nrow=3,ncol=3)
rownames(pop_2) = c("Mantoux_Test_Pos", "Mantoux_Test_Neg",
colnames(pop_2) = c("Tine_Test_Pos", "Tine_Test_Neg", "Total"

pop_2[1,1] = 887
pop_2[1,2] = 31
pop_2[2,1] = 37
pop_2[2,2] = 367

#Total rows and columns

pop_2[1,3] = pop_2[1,1] + pop_2[1,2]
pop_2[2,3] = pop_2[2,1] + pop_2[2,2]
pop_2[3,1] = pop_2[1,1] + pop_2[2,1]
pop_2[3,2] = pop_2[1,2] + pop_2[2,2]
N_2 = sum(pop_2[1,1] + pop_2[1,2] + pop_2[2,1] + pop_2[2,2])
pop_2[3,3] = N_2
pop_2
```

## Exercise

Assuming Mantoux test as a gold standard, estimate and save the sensitivity and specificity of tine test in both populations?

## Solution

```
# 1st population  
(sensitivity_1 <- pop_1[1,1] / (pop_1[1,3]))
```

```
## [1] 0.7777778
```

```
(specificity_1 <- pop_1[2,2] / (pop_1[2,3]))
```

```
## [1] 0.9832402
```

```
#2nd population
```

```
(sensitivity_2 <- pop_2[1,1] / (pop_2[1,3]))
```

```
## [1] 0.9662309
```

```
(specificity_2 <- pop_2[2,2] / (pop_2[2,3]))
```

```
## [1] 0.9084158
```

# Hui-Walter paradigm (1980)

Population 1

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

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

# 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

# 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

$$T1+T2+: P2 * Se1 * Se2 + (1-P2) * (1-Sp1) * (1-Sp2)$$

$$T1+T2-: P2 * Se1 * (1-Se2) + (1-P2) * (1-Sp1) * Sp2$$

$$T1-T2+: P2 * (1-Se1) * Se2 + (1-P2) * Sp1 * (1-Sp2)$$

$$T1-T2-: P2 * (1-Se1) * (1-Se2) + (1-P2) * Sp1 * Sp2$$

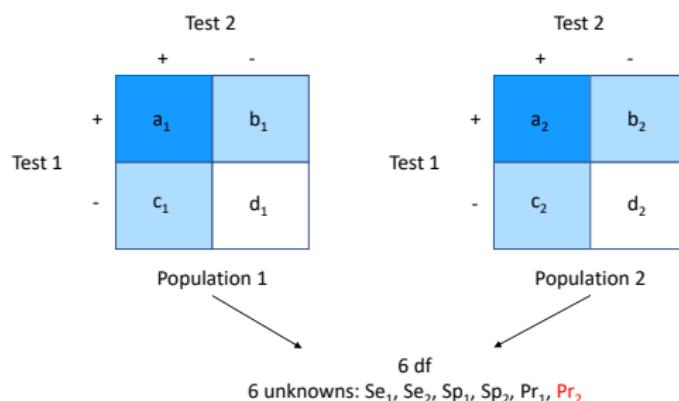


Identifiable model!



## Hui-Walter model

- ▶ A particular model formulation that was originally designed for evaluating diagnostic tests in the absence of a gold standard
- ▶ Also known as the two\_test - two\_population setting/paradigm



## Model Specification ('hw\_definition')

```
library('rungjags')

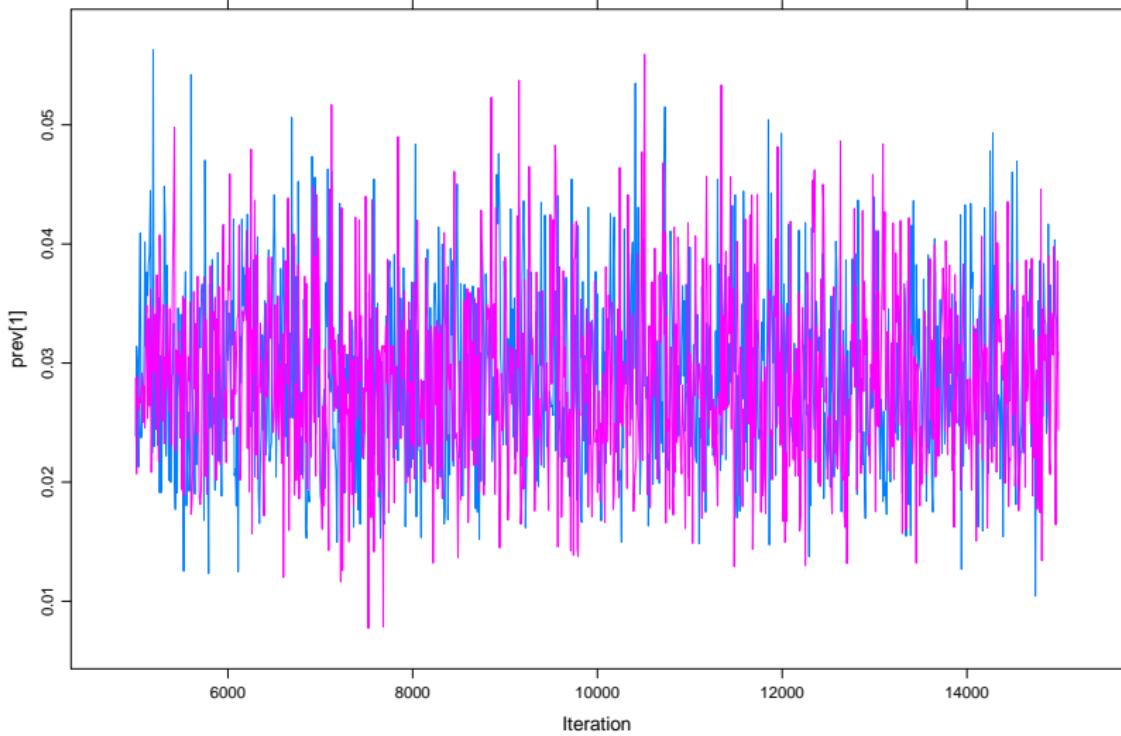
Population_1 <- as.numeric(pop_1[1:2,1:2])
Population_2 <- as.numeric(pop_2[1:2,1:2])

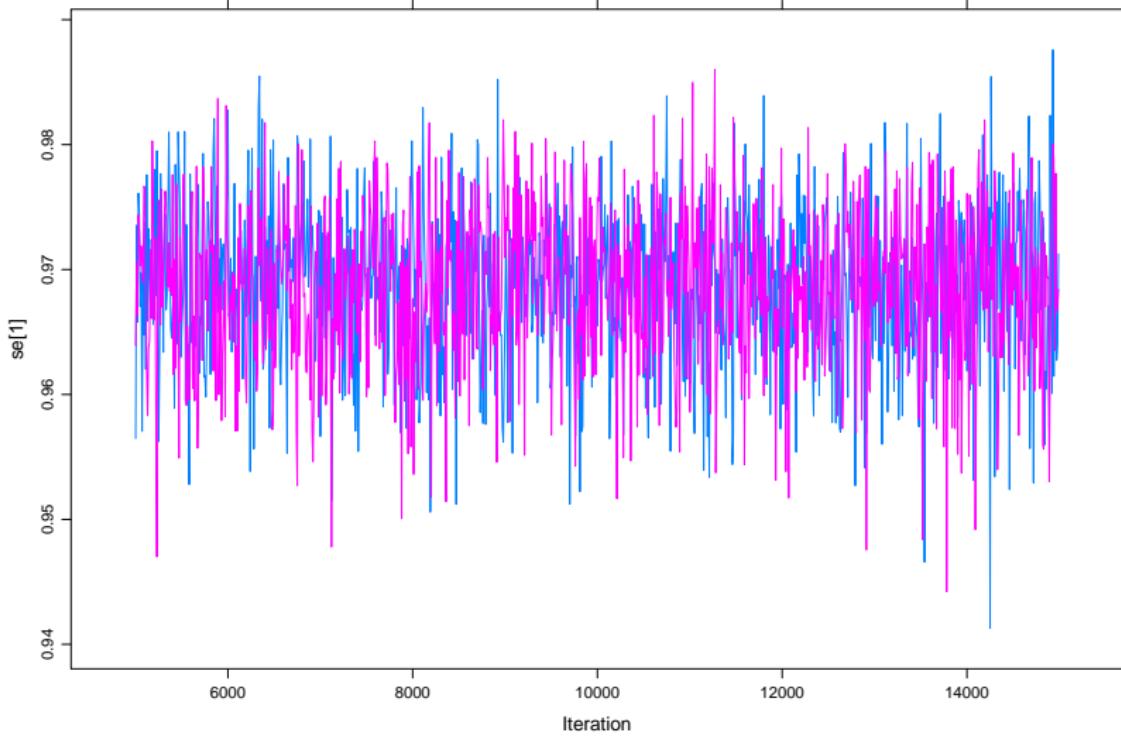
prev <- list(chain1=c(0.05,0.99), chain2=c(0.95,0.05))
se <- list(chain1=c(0.5,0.99), chain2=c(0.99,0.5))
sp <- list(chain1=c(0.5,0.99), chain2=c(0.99,0.5))

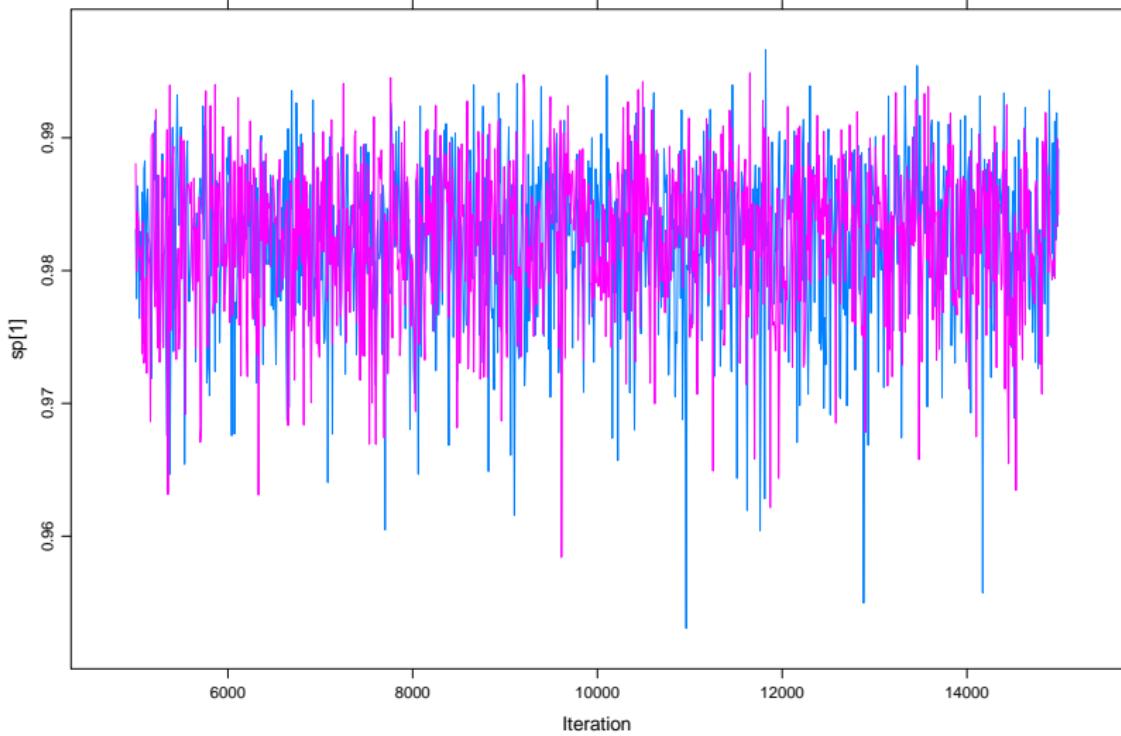
results <- run.jags(hw_definition, n.chains=2)
```

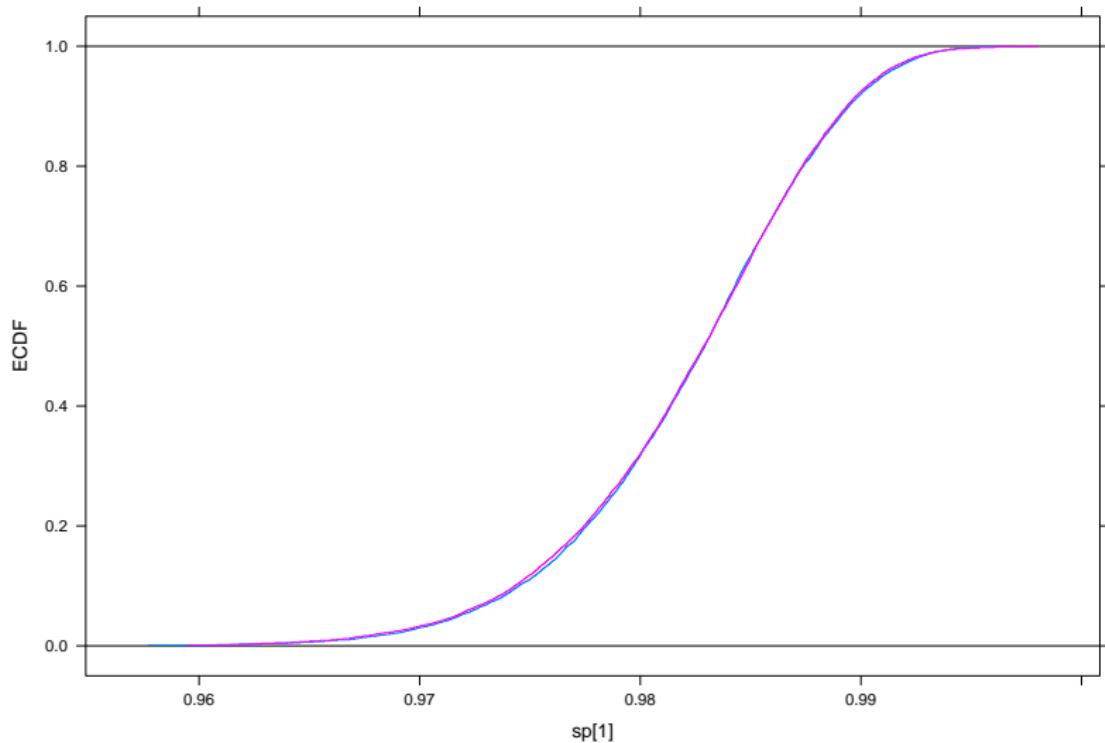
Remember to check convergence and effective sample size!

```
plot(results)
```









```
summary(results)
```

## Exercise 1

Run the `hw_definition` model under the following different scenarios and interpret the results in each case.

1. Change the priors for  $Se$  and  $Sp$  and try  $Beta(2,1)$ .
2. Remove the  $T(1-sp[1], )$  from the model and run it again. What happens now?
3. Try to run the model with different initial values and removing  $T(1-sp[1], )$ . For example try it with:

```
se <- list(chain1=c(0.01,0.99), chain2=c(0.99,0.01))  
sp <- list(chain1=c(0.01,0.99), chain2=c(0.99,0.01))
```

4. Run the model with only 1 population (either `pop_1` or `pop_2`). What happens then?

## Event Closure

Group photo



# HARMONY

Novel tools for test evaluation and  
disease prevalence estimation