Diagnostic meta-analysis

Paolo Eusebi

16/09/2021

Recap

- Important points from previous sessions

Perfect Reference Test

DTA-MA: perfect reference test

- There is an increasing interest in meta-analyzing data from diagnostic accuracy studies
- The data from the primary studies are summarized in a 2-by-2 cross-tabulation of the dichotomized test result against the true disease status (assuming we have a perfect reference test)

	D+	D-
T+	TP	FP
T-	FN	TN

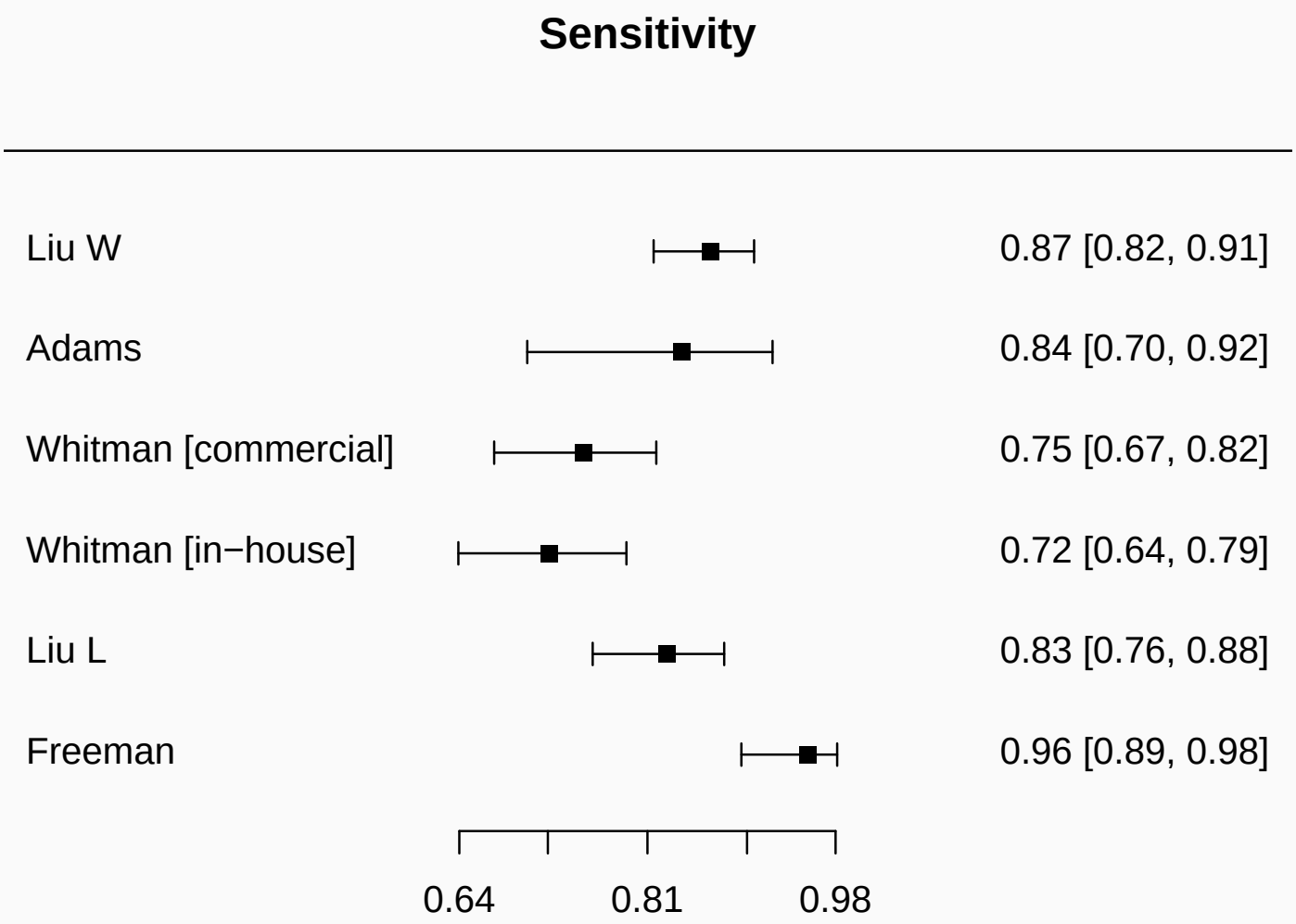
DTA-MA: perfect reference test

- Serological tests for covid-19 from 5 studies (but 6 observations) on evaluation of ELISA assay for covid-19 (Bastos et al 2020).

Study	TP	FN	TN	FP
Liu W	186	28	100	0
Adams	34	6	50	0
Whitman [commercial]	98	32	140	20
Whitman [in-house]	94	36	152	8
Liu L	127	26	116	4
Freeman	95	4	515	4

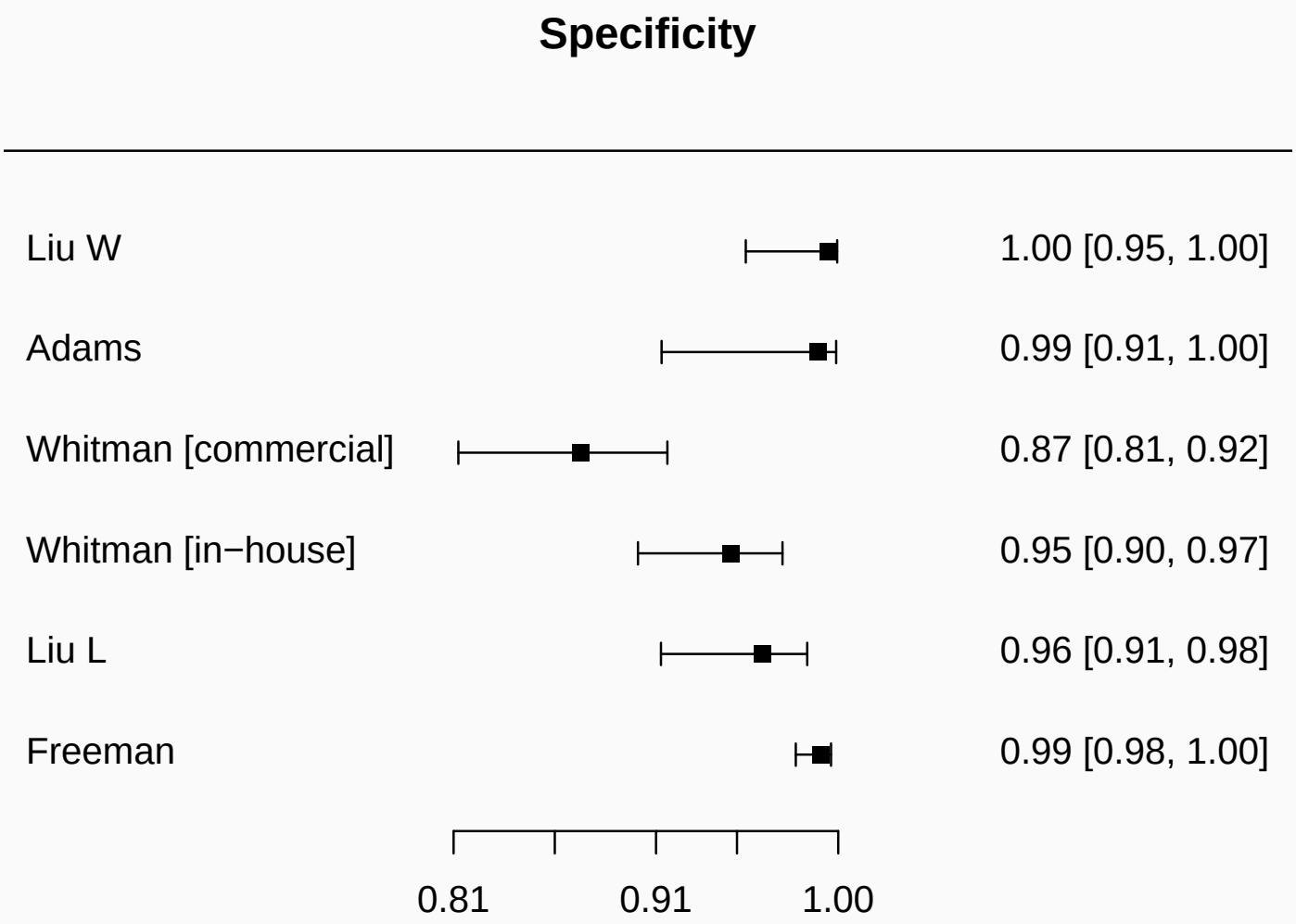
DTA-MA: perfect reference test

- Forest plot of sensitivity



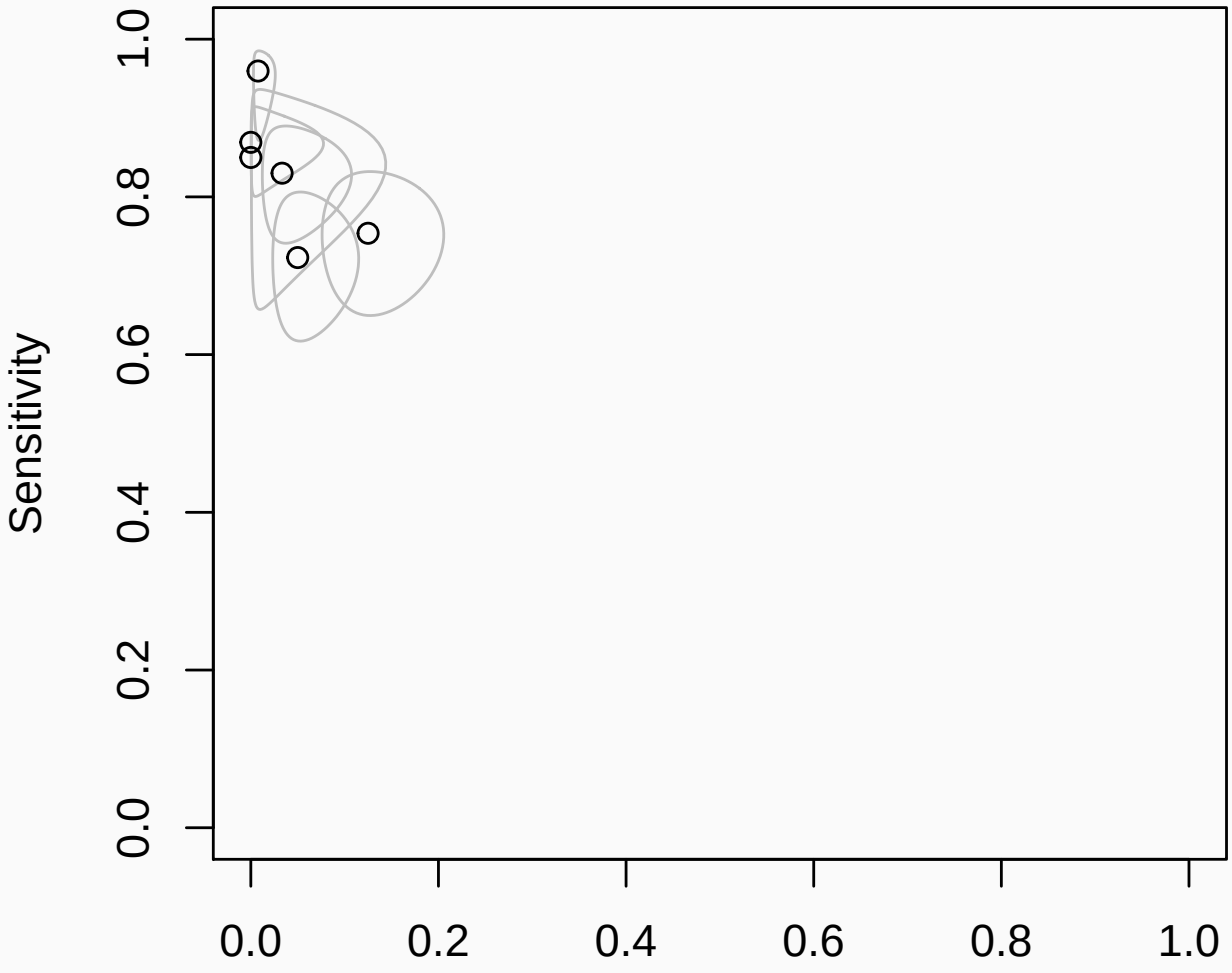
DTA-MA: perfect reference test

- Forest plot of specificity



DTA-MA: perfect reference test

- Data points with confidence ellipses on a ROC space

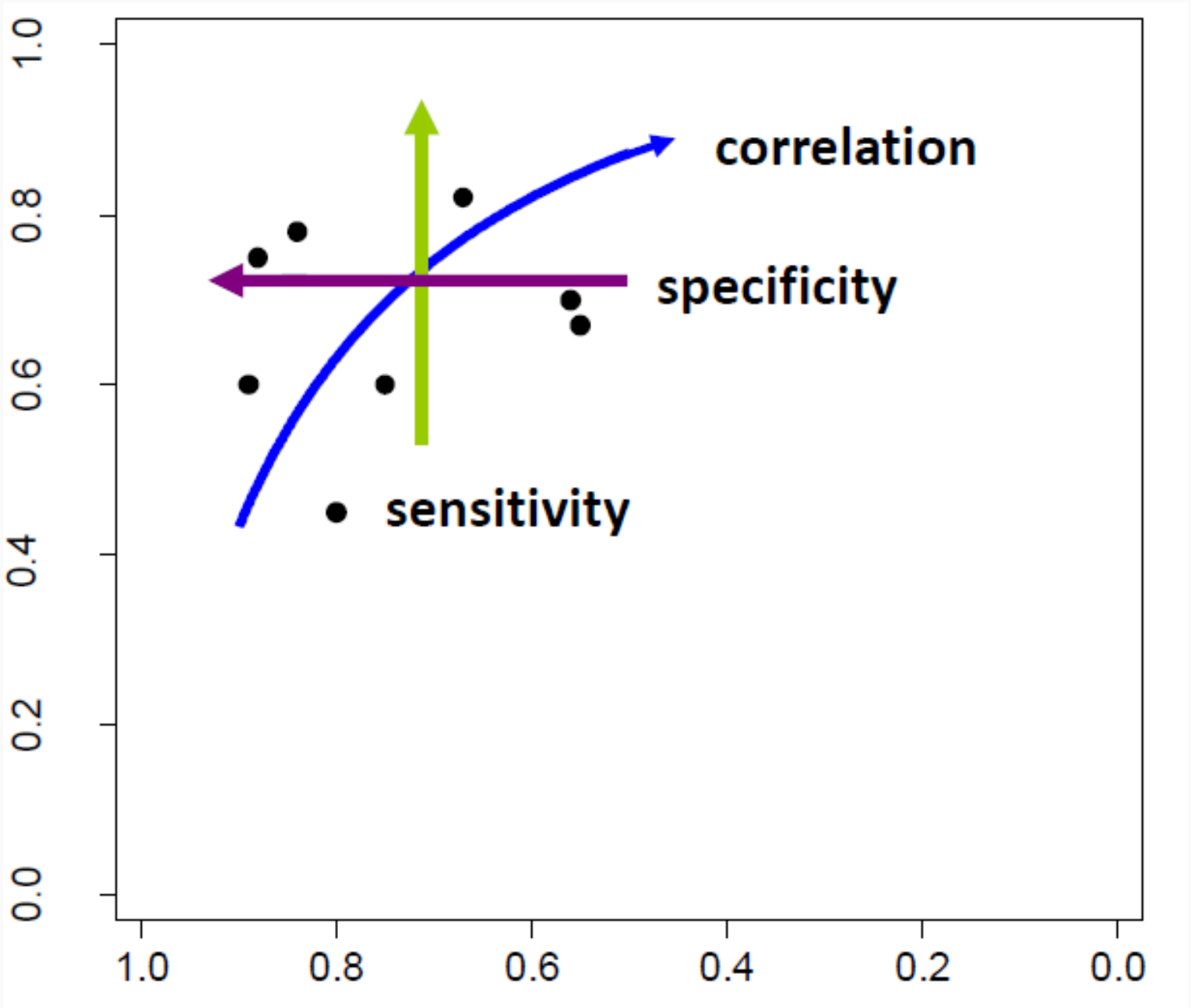


DTA-MA: perfect reference test

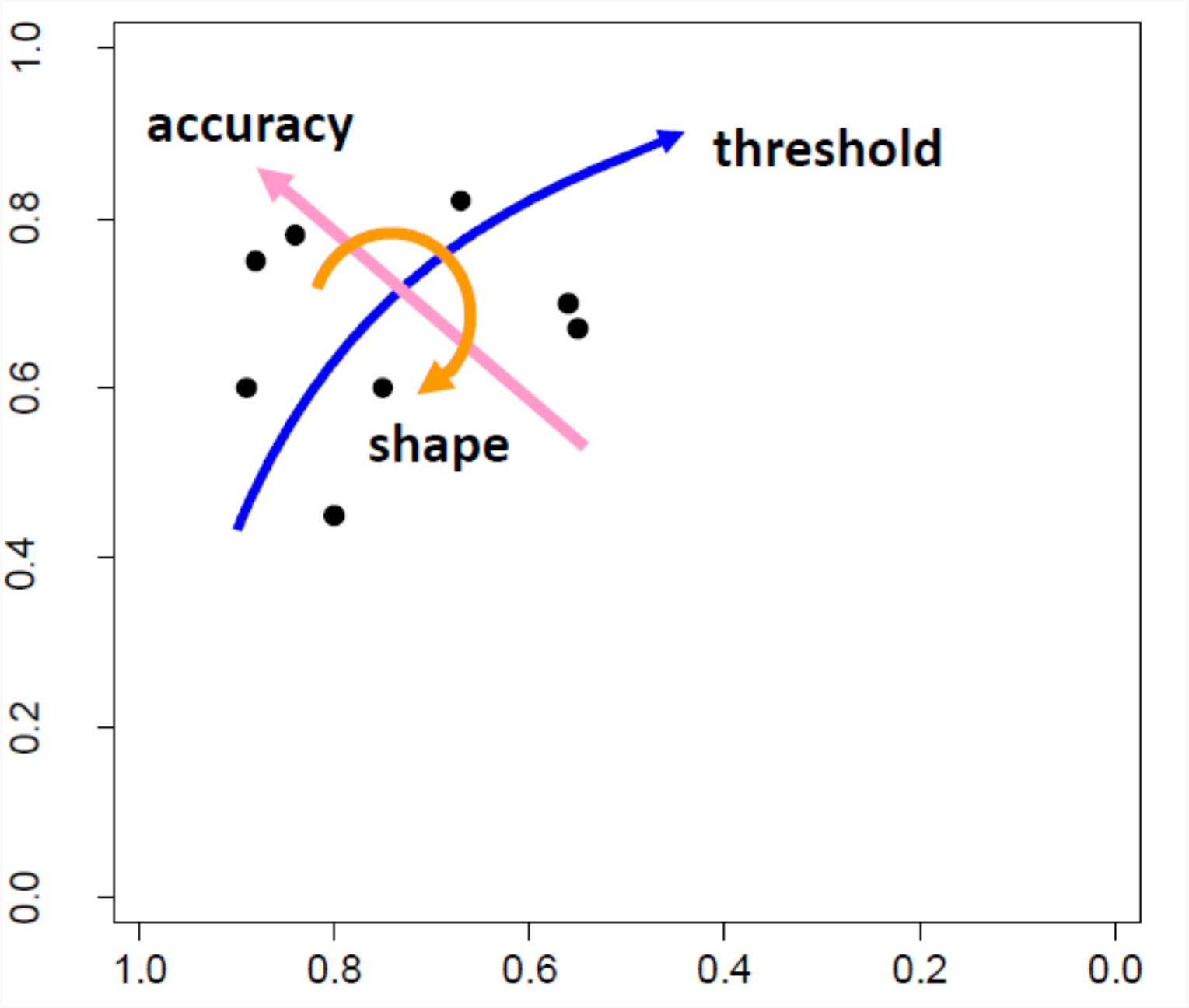
Two main frameworks:

- Hierarchical Summary ROC (Rutter and Gatsonis 2001)
- Bivariate analysis of sensitivity and specificity (Reitsma et al. 2005)

DTA-MA: bivariate analysis of sensitivity and specificity



DTA-MA: hierarchical summary ROC (HSROC)



DTA-MA: bivariate analysis of sensitivity and specificity

Some notation/definitions (no covariates)

$$(\mu_{A_i}\mu_{B_i}) \sim N((\mu_A\mu_B), \Sigma_{AB})$$

with

$$\Sigma_{AB} = \begin{pmatrix} \sigma_A^2 & \sigma_{AB}^2 \\ \sigma_{AB}^2 & \sigma_B^2 \end{pmatrix}$$

μ_{A_i} is the logit-transformed sensitivity in study i μ_{B_i} is the logit-transformed specificity in study i

DTA-MA: hierarchical summary ROC (HSROC)

Some notation/definitions (no covariates)

- The model is defined in terms of the probability π_{ij} that a patient in study i with a disease status j has a positive test result.
- $j = 0$ for a patient without the disease
- $j = 1$ for a patient with the disease
- $\pi_{i0} = 0$ is the false-positive rate (1-specificity)
- $\pi_{i1} = 0$ is the true-positive rate (sensitivity)

DTA-MA: hierarchical summary ROC (HSROC)

- level I (within study)

$$logit(\pi_{ij}) = (\theta_i + \alpha_i D_{ij}) \cdot exp(-\beta \cdot D_{ij})$$

- level II (between studies)

$$\theta_i \sim N(\Theta, \sigma_\theta^2)$$

$$\alpha_i \sim N(\Lambda, \sigma_\alpha^2)$$

θ_i are cutpoint parameters (or positivity criteria)

α_i are accuracy parameters, modeling the difference between true-positive and false-positive fractions

β is a shape parameter, allowing true-positive and false-positive fractions to increase at different rates as θ_i increases

DTA-MA: bivariate analysis of sensitivity and specificity

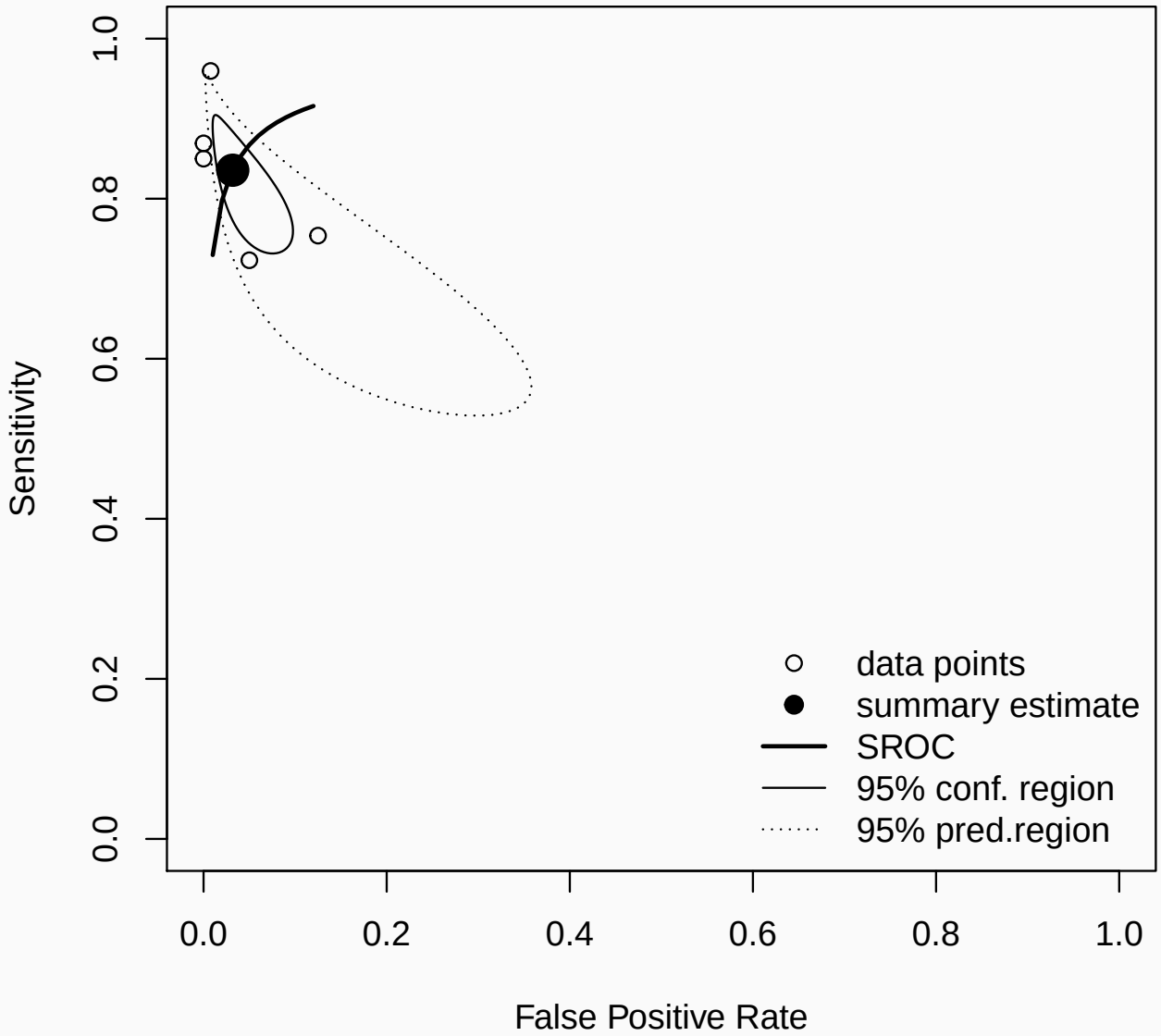
Let's run the model with reitsma function (mada R package)

```
fit.reitsma <- reitsma(d)
```

DTA-MA: bivariate analysis of sensitivity and specificity

```
##                      Estimate 95%ci.lb 95%ci.ub
## tsens.(Intercept)        1.63     1.13     2.13
## tfpr.(Intercept)        -3.41    -4.37    -2.46
## sensitivity              0.84     0.76     0.89
## false pos. rate          0.03     0.01     0.08
```

DTA-MA: bivariate analysis of sensitivity and specificity



DTA-MA: bivariate analysis of sensitivity and specificity

- Where is the summary measure of heterogeneity?
- There is I^2 for DTA-MA?
- Interpretation of confidence region/prediction region

DTA-MA: bivariate analysis of sensitivity and specificity

The function returns also HSROC parameters

```
## $coef_hsroc
## $coef_hsroc$Theta
## [1] -0.14
##
## $coef_hsroc$Lambda
## [1] 4.7
##
## $coef_hsroc$beta
## [1] 0.62
##
## $coef_hsroc$sigma2theta
## [1] 0.02
##
## $coef_hsroc$sigma2alpha
## [1] 2.3
```

DTA-MA: bivariate analysis of sensitivity and specificity

This is because Bivariate and HSROC approaches are equivalent when covariates are not included (Harbord et al. 2007)

- Parameter estimates from either model can be used to produce a summary operating point, an SROC curve, confidence regions, or prediction regions.
- The choice between these parameterizations depends partly on the degrees of and reasons for between-study heterogeneity and the treshold effect.

Imperfect Reference Test

DTA-MA: imperfect reference test(s)

Why?

- Ignoring the imperfect nature of the reference may result in biased estimates of pooled sensitivity and specificity of the test under evaluation

DTA-MA: imperfect reference test(s)

How?

- Multivariate generalized linear mixed model (MGLMM)
- Hierarchical summary receiver operating characteristic (HSROC)
- Exact relations between the parameters of these models can be provided.
- But some submodels of the MGLMM do not have corresponding equivalent submodels of the HSROC model, and vice versa.

DTA-MA: HSROC for imperfect reference test(s)

Dendukuri et al. Biometrics. 2012

- The data from the primary studies are summarized in a 2-by-2 cross-tabulation of the index test (T_1) result against the imperfect refernce (T_2)

	T2+	T2-
T1+	p*se*s2+(1-p)*(1-sp)*(1-c2)	p*(1-se)*s2+(1-p)*sp*(1-c2)
T1-	p*se*(1-s2)+(1-p)*(1-sp)*c2	p*(1-se)*(1-s2)+(1-p)*sp*c2

The sensitivity and the specificity of the reference test are defined as:

- $S_2 = P(T_2 = +|D+)$
- $C_2 = P(T_2 = -|D-)$

DTA-MA: discussion

- Comments?
- Questions?
- Ideas?

DTA-MA: hierarchical summary ROC (HSROC)

Let's do it with rjags

DTA-MA with JAGS: Likelihood

```
## model {
##
##  for(i in 1:l) {
##      # Likelihood
##      # se, sp are accuracy of CI
##      # s2, c2 are accuracy of LU
##      # pi is the prevalence
##
##      cell[i,1:4] ~ dmulti(prob[i,1:4],n[i])
##
##      prob[i,1] <- pi[i]*se[i]*s2+(1-pi[i])*(1-sp[i])*(1-c2)
##      prob[i,2] <- pi[i]*se[i]*(1-s2)+(1-pi[i])*(1-sp[i])*c2
##      prob[i,3] <- pi[i]*(1-se[i])*s2+(1-pi[i])*sp[i]*(1-c2)
##      prob[i,4] <- pi[i]*(1-se[i])*(1-s2)+(1-pi[i])*sp[i]*c2
```

DTA-MA with JAGS: model specifications

```
##
##
##      # Expressing accuracy in terms of HSROC parameters
##
##      b[i] <- exp((beta)/2)
##      logit(se[i]) <- (theta[i] + 0.5*alpha[i])/b[i]
##      logit(sp[i]) <- -(theta[i] - 0.5*alpha[i])*b[i]
##
##      # Priors for CI accuracy
##      theta[i] ~ dnorm(THETA,prec[1])
##      alpha[i] ~ dnorm(LAMBDA,prec[2])
```

DTA-MA with JAGS: model specifications

```
##
##      # Priors for prevalence parameters
##      pi[i] ~ dbeta(1,1)
##  }
##
```

DTA-MA with JAGS: model specifications

```
##
##
##  # CI accuracy
##  Se_overall <- 1/(1+exp((-THETA-0.5*LAMBDA)/exp(beta/2)))
##  Sp_overall <- 1/(1+exp((THETA-0.5*LAMBDA)*exp(beta/2)))
##
##  theta_new ~ dnorm(THETA,prec[1])
##  alpha_new ~ dnorm(LAMBDA,prec[2])
##
```

DTA-MA with JAGS: model specifications

```
##
##
##  # Predicted values for CI in a new study
##  Se_new <- 1/(1+exp(-(theta_new+0.5*alpha_new)/exp(beta/2)))
##  Sp_new <- 1/(1+exp((theta_new-0.5*alpha_new)*exp(beta/2)))
```

DTA-MA with JAGS: model specifications

```
##
##
##  # Priors over the accuracy parameters of CI
##  THETA ~ dunif(-2.6,2.6)
##  LAMBDA ~ dunif(-5.2,5.2)
##  beta ~ dunif(-1.3,1.3)
##
##  for(j in 1:2) {
##
##          prec[j] <- pow(sigma[j],-2)
##          sigma[j] ~ dgamma(4,2)
##  }
```

Exercise

Use Timsit paper data (Prev Vet Med 2016)

StudyID	TP	FP	FN	TN
Gardner	49	53	38	64
Buhman	37	1	90	18
Thompson	265	196	606	969
Schneider	121	42	910	592
Leach	195	60	1395	373
Tennant	157	29	1344	806
Rezac	127	157	4591	8316

1. Fit a bivariate model assuming perfect reference with reitsma() in mada
2. Fit a HSROC model assuming imperfect reference with HSROC() in HSROC
3. Fit a HSROC model assuming imperfect reference with model definitions in rjags