Web-based Supplementary Materials for "A composite likelihood method for bivariate meta-analysis in diagnostic systematic reviews"

Yong Chen, Yulun Liu, Jing Ning, Lei Nie, Hongjian Zhu and Haitao Chu

Section 1: Additional tables and figures

We first numerically demonstrate the advantage of the proposed method over the univariate analysis in estimating functions of sensitivity and specificity. Specifically, we consider the positive and negative likelihood ratios, defined as

$$\text{LR+} = \Pr(+|D)/\Pr(+|\bar{D}) = \text{Sensitivity}/(1 - \text{Specificity}),$$

$$\text{LR}- = \Pr(-|D)/\Pr(-|\bar{D}) = \text{1-Sensitivity/Specificity}.$$

The inference based on univariate analyses for LR+ and LR− is likely to overestimate the standard errors due to the ignored negative correlation between estimated sensitivity and specificity. Thus, the coverage of the confidence intervals for LR+ and LR− based on the univariate analysis cannot reach the nominal level, even when the sample size is large. On the other hand, the inference based on the proposed composite likelihood (CL) method can properly account for the correlation between estimated sensitivity and specificity, and thus leads to coverage close to the nominal level.

Figure S1 presents the empirical coverage probability (CP) of the 95% confidence intervals for LR+ (Upper panels) and LR− (Lower panels) based on univariate method and the composite likelihood (CL) method under different settings of between-study correlation ρ and number of studies (m). It is clear that the CP of the univariate method is biased and is too conservative (which can be explained by the ignored negative correlation between the estimated sensitivity and specificity). On the other hand, the CP of the CL method is close to the nominal level, especially when the number of studies is sufficiently large.

Table S1: Summary of 5,000 simulations with data generated from $\mathcal{M}_1$: bias (BIAS), standard errors (SE), model-based standard errors (MBSE) and coverage probabilities (CP) of estimates. True values of model parameters are: sensitivity = 70.0, specificity = 80.0, LR+ = 3.50, and LR− = 0.38. All entries for sensitivity and specificity are multiplied by 100.

			SL† method				CL* method			
m	ρ		BIAS	SE	MBSE	CP(%)	BIAS	SE	MBSE	CP(%)
10	0.0	Se	-0.4	7.2	6.6	89.8	-0.00	0.07	0.07	89.5
		Sp	-0.6	5.5	5.1	89.3	-0.00	0.06	0.05	89.7
		LR+	0.15	1.15	1.04	89.7	0.18	1.15	1.05	89.4
		LR−	0.02	0.10	0.09	90.9	0.01	0.09	0.09	90.4
	-0.3	Se	-0.4	7.2	6.6	88.4	-0.5	7.1	6.6	89.9
		Sp	-0.5	5.5	5.1	87.5	-0.5	5.5	5.1	89.4
		LR+	0.13	1.05	0.94	87.5	0.13	1.04	0.93	88.1
		LR−	0.01	0.09	0.08	88.8	0.01	0.09	0.08	89.7
	-0.6	Se	-0.4	7.2	6.6	85.2	-0.5	7.3	6.6	88.2
		Sp	-0.5	5.4	5.1	85.0	-0.5	5.6	5.1	88.7
		LR+	0.10	0.92	0.84	84.2	0.10	0.92	0.83	89.2
		LR−	0.01	0.08	0.07	85.5	0.01	0.08	0.07	88.7
25	0.0	Se	-0.3	4.6	4.4	92.5	-0.3	4.6	4.4	93.0
		Sp	-0.3	3.5	3.4	93.0	-0.2	3.5	3.4	93.2
		LR+	0.05	0.68	0.65	93.4	0.06	0.68	0.65	92.8
		LR−	0.01	0.06	0.06	92.8	0.01	0.06	0.06	92.8
	-0.3	Se	-0.3	4.6	4.4	92.7	-0.1	4.5	4.4	93.1
		Sp	-0.2	3.5	3.4	92.9	-0.2	3.4	3.4	93.0
		LR+	0.04	0.62	0.59	92.8	0.06	0.61	0.59	92.9
		LR−	0.00	0.06	0.05	92.7	0.00	0.06	0.05	93.0
	-0.6	Se	-0.3	4.6	4.4	92.7	-0.2	4.5	4.4	93.3
		Sp	-0.2	3.5	3.4	93.1	-0.3	3.5	3.4	93.1
		LR+	0.03	0.55	0.53	92.4	0.03	0.55	0.52	92.4
		LR−	0.00	0.05	0.05	92.5	0.00	0.05	0.05	93.2
50	0.0	Se	-0.2	3.3	3.2	93.5	-0.1	3.3	3.2	93.5
		Sp	-0.2	2.4	2.4	94.4	-0.1	2.5	2.4	94.1
		LR+	0.02	0.47	0.46	93.9	0.03	0.47	0.46	93.9
		LR−	0.00	0.04	0.04	93.4	0.00	0.04	0.04	93.6
	-0.3	Se	-0.2	3.3	3.2	93.5	-0.1	3.2	3.2	93.8
		Sp	-0.1	2.4	2.4	94.5	-0.1	2.5	2.4	94.1
		LR+	0.02	0.42	0.42	93.9	0.02	0.43	0.42	93.4
		LR−	0.00	0.04	0.04	93.5	0.00	0.04	0.04	93.7
	-0.6	Se	-0.2	3.3	3.2	93.6	-0.1	3.3	3.2	94.0
		Sp	-0.1	2.4	2.4	94.1	-0.1	2.5	2.4	93.8
		LR+	0.01	0.38	0.37	94.3	0.01	0.38	0.37	93.7
		LR−	0.00	0.04	0.04	93.6	0.00	0.04	0.04	93.8

SL method†: standard maximum likelihood method based on the BGLMM
CL method*: proposed composite likelihood method

Table S2: Summary of 5,000 simulations with data generated from $\mathcal{M}_2$: bias (BIAS), standard errors (SE), model- based standard errors (MBSE) and coverage probabilities (CP) of estimates. True values of model parameters are: sensitivity = 70.0, specificity = 80.0, LR + = 3.50, and LR − = 0.38. All entries for sensitivity and specificity are multiplied by 100.

			SL† method				CL* method			
m	ρ		BIAS	SE	MBSE	CP(%)	BIAS	SE	MBSE	CP(%)
10	0.0	Se	-0.8	8.7	8.2	89.9	-0.8	8.8	8.2	89.4
		Sp	-0.8	7.0	6.4	90.3	-1.1	7.0	6.5	89.3
		LR+	0.24	1.56	1.35	90.1	0.20	1.54	1.32	87.7
		LR−	0.02	0.12	0.11	90.8	0.02	0.12	0.11	90.1
	-0.3	Se	-0.9	9.0	8.3	88.3	-0.9	8.9	8.3	89.7
		Sp	-1.0	7.1	6.5	89.0	-0.9	7.0	6.5	89.3
		LR+	0.16	1.36	1.20	87.5	0.16	1.33	1.17	88.5
		LR−	0.02	0.11	0.10	88.6	0.02	0.11	0.10	90.4
	-0.6	Se	-0.9	9.0	8.3	86.4	-1.0	9.1	8.4	88.6
		Sp	-1.0	7.1	6.5	86.9	-0.9	7.2	6.5	88.9
		LR+	0.12	1.17	1.05	85.8	0.13	1.19	1.02	88.3
		LR−	0.01	0.10	0.09	86.8	0.01	0.10	0.09	89.0
25	0.0	Se	-0.3	5.6	5.5	92.4	-0.6	5.7	5.5	92.9
		Sp	-0.3	4.3	4.2	92.6	-0.4	4.3	4.3	93.3
		LR+	0.09	0.85	0.82	93.0	0.06	0.84	0.82	92.5
		LR−	0.01	0.07	0.07	93.3	0.01	0.08	0.07	93.7
	-0.3	Se	-0.4	5.7	5.5	93.0	-0.6	5.7	5.5	92.9
		Sp	-0.4	4.4	4.2	92.9	-0.5	4.4	4.3	93.0
		LR+	0.05	0.76	0.74	92.7	0.04	0.78	0.74	92.2
		LR−	0.01	0.07	0.07	93.0	0.01	0.07	0.07	93.6
	-0.6	Se	-0.3	5.6	5.5	93.1	-0.5	5.6	5.5	92.6
		Sp	-0.5	4.4	4.3	93.2	-0.4	4.4	4.3	93.2
		LR+	0.04	0.68	0.65	92.5	0.04	0.68	0.65	92.5
		LR−	0.01	0.06	0.06	93.3	0.01	0.06	0.06	92.9

SL method†: standard maximum likelihood method based on the BGLMM
CL method*: proposed composite likelihood method

Table S2: Summary of 5,000 simulations with data generated from $\mathcal{M}_2$: bias (BIAS), standard errors (SE), model-based standard errors (MBSE) and coverage probabilities (CP) of estimates. True values of model parameters are: sensitivity = 90.0, specificity = 95.0, LR + = 18.00, and LR − = 0.11. All entries for sensitivity and specificity are multiplied by 100. **(continued)**

			SL† method				CL* method			
m	ρ		BIAS	SE	MBSE	CP(%)	BIAS	SE	MBSE	CP(%)
10	0.0	Se	-0.8	4.4	4.1	89.7	-0.7	4.4	4.1	90.3
		Sp	-0.4	2.5	2.3	89.8	-0.5	2.5	2.3	90.8
		LR+	2.81	13.74	10.87	88.9	2.42	12.99	8.90	87.5
		LR−	0.01	0.05	0.04	89.7	0.01	0.05	0.04	90.9
	-0.3	Se	-0.8	4.4	4.1	84.7	-0.7	4.4	4.1	90.3
		Sp	-0.5	2.5	2.3	84.6	- 0.5	2.5	2.3	90.0
		LR+	2.40	12.54	10.30	83.1	2.19	14.10	8.90	87.7
		LR−	0.01	0.05	0.04	84.9	0.01	0.05	0.04	91.2
	-0.6	Se	-0.6	4.5	4.1	77.0	-0.8	4.3	4.1	90.8
		Sp	-0.6	2.5	2.4	80.2	-0.5	2.5	2.3	90.9
		LR+	2.19	18.69	10.54	76.0	2.03	11.74	8.40	87.6
		LR−	0.01	0.05	0.04	77.3	0.01	0.05	0.04	91.3
25	0.0	Se	-0.3	2.7	2.6	93.7	-0.3	2.6	2.7	94.0
		Sp	-0.1	1.4	1.5	93.5	-0.2	1.5	1.5	93.7
		LR+	1.04	5.86	5.75	93.2	0.89	5.87	5.46	93.1
		LR−	0.00	0.03	0.03	93.7	0.00	0.03	0.03	94.4
	-0.3	Se	-0.2	2.6	2.6	93.7	-0.3	2.7	2.6	93.5
		Sp	-0.2	1.5	1.5	93.9	-0.2	1.5	1.5	93.5
		LR+	0.88	5.67	5.58	92.9	0.78	5.63	5.32	92.2
		LR−	0.00	0.03	0.03	93.7	0.00	0.03	0.03	93.6
	-0.6	Se	-0.3	2.6	2.6	92.6	-0.3	2.6	2.6	93.5
		Sp	-0.1	1.5	1.4	91.8	-0.2	1.5	1.5	94.2
		LR+	1.14	5.65	5.51	91.3	0.62	5.34	5.10	92.6
		LR−	0.00	0.03	0.03	92.7	0.00	0.03	0.03	93.7
50	0.0	Se	-0.1	1.8	1.9	94.7	-0.1	1.9	1.9	94.4
		Sp	0.0	1.0	1.0	94.6	0.0	1.0	1.0	95.3
		LR+	0.55	3.91	3.94	94.9	0.53	3.79	3.86	95.0
		LR−	0.00	0.02	0.02	94.8	0.00	0.02	0.02	94.7
	-0.3	Se	-0.1	1.9	1.9	94.9	-0.2	1.9	1.9	94.7
		Sp	0.0	1.0	1.0	94.1	0.0	1.0	1.0	94.6
		LR+	0.71	3.89	3.92	95.5	0.52	3.78	3.77	94.7
		LR−	0.00	0.02	0.02	95.0	0.00	0.02	0.02	94.8
	-0.6	Se	-0.1	1.8	1.9	94.7	-0.1	1.9	1.9	94.6
		Sp	0.0	1.0	1.0	94.3	0.0	1.0	1.0	94.0
		LR+	0.65	3.70	3.80	95.0	0.56	3.78	3.69	94.3
		LR−	0.00	0.02	0.02	94.6	0.00	0.02	0.02	94.8

SL method†: standard maximum likelihood method based on the BGLMM

CL method*: proposed composite likelihood method

Table S3: Summary of 5,000 simulations with data generated from $\mathcal{M}_1$: bias (BIAS), standard errors (SE), model-based standard errors (MBSE) and coverage probabilities (CP) of estimates. True values of model parameters are: sensitivity = 90.0, specificity = 95.0, LR+ = 18.00, LR− = 0.11, dOR = 171.00.

			Univariate method				CL* method			
m	ρ		BIAS	SE	MBSE	CP(%)	BIAS	SE	MBSE	CP(%)
10	0.0	LR+	1.63	9.84	7.99	88.1	1.63	9.84	7.92	88.3
		LR−	0.01	0.04	0.03	89.8	0.01	0.04	0.04	90.5
		dOR	26.95	117.16	107.98	88.1	26.95	117.16	105.47	87.1
	-0.3	LR+	1.46	9.06	7.67	88.2	1.46	9.06	7.41	87.8
		LR−	0.01	0.04	0.03	90.4	0.01	0.04	0.03	90.7
		dOR	22.15	103.93	106.29	90.5	22.15	103.93	92.40	87.0
	-0.6	LR+	1.51	8.96	7.69	89.2	1.51	8.96	7.15	88.4
		LR−	0.01	0.04	0.03	90.4	0.01	0.04	0.03	90.5
		dOR	19.53	100.14	107.67	93.2	19.53	100.14	77.62	85.8
25	0.0	LR+	0.64	4.91	4.53	92.0	0.64	4.91	4.51	91.9
		LR−	0.00	0.02	0.02	93.5	0.00	0.02	0.02	93.6
		dOR	13.16	69.80	64.94	92.0	13.16	69.80	64.35	91.5
	-0.3	LR+	0.51	4.66	4.46	92.6	0.51	4.66	4.35	92.2
		LR−	0.00	0.02	0.02	93.1	0.00	0.02	0.02	93.0
		dOR	10.10	60.86	63.61	94.2	10.10	60.86	55.76	91.4
	-0.6	LR+	0.47	4.36	4.46	93.2	0.47	4.36	4.24	92.5
		LR−	0.00	0.02	0.02	93.6	0.00	0.02	0.02	92.7
		dOR	6.83	50.27	62.08	96.3	6.83	50.27	45.89	90.9
50	0.0	LR+	0.30	3.21	3.11	93.9	0.30	3.21	3.11	93.9
		LR−	0.00	0.02	0.02	94.1	0.00	0.02	0.02	94.0
		dOR	5.70	44.96	43.57	93.8	5.70	44.96	43.44	93.5
	-0.3	LR+	0.30	3.11	3.11	94.3	0.30	3.11	3.04	93.8
		LR−	0.00	0.02	0.02	93.9	0.00	0.02	0.02	93.5
		dOR	4.72	39.92	43.37	95.7	4.72	39.92	38.02	93.0
	-0.6	LR+	0.23	2.99	3.10	94.4	0.23	2.99	2.95	93.5
		LR−	0.00	0.02	0.02	95.4	0.00	0.02	0.02	94.7
		dOR	3.24	33.67	42.93	97.9	3.24	33.67	31.82	92.8

CL method*: proposed composite likelihood method

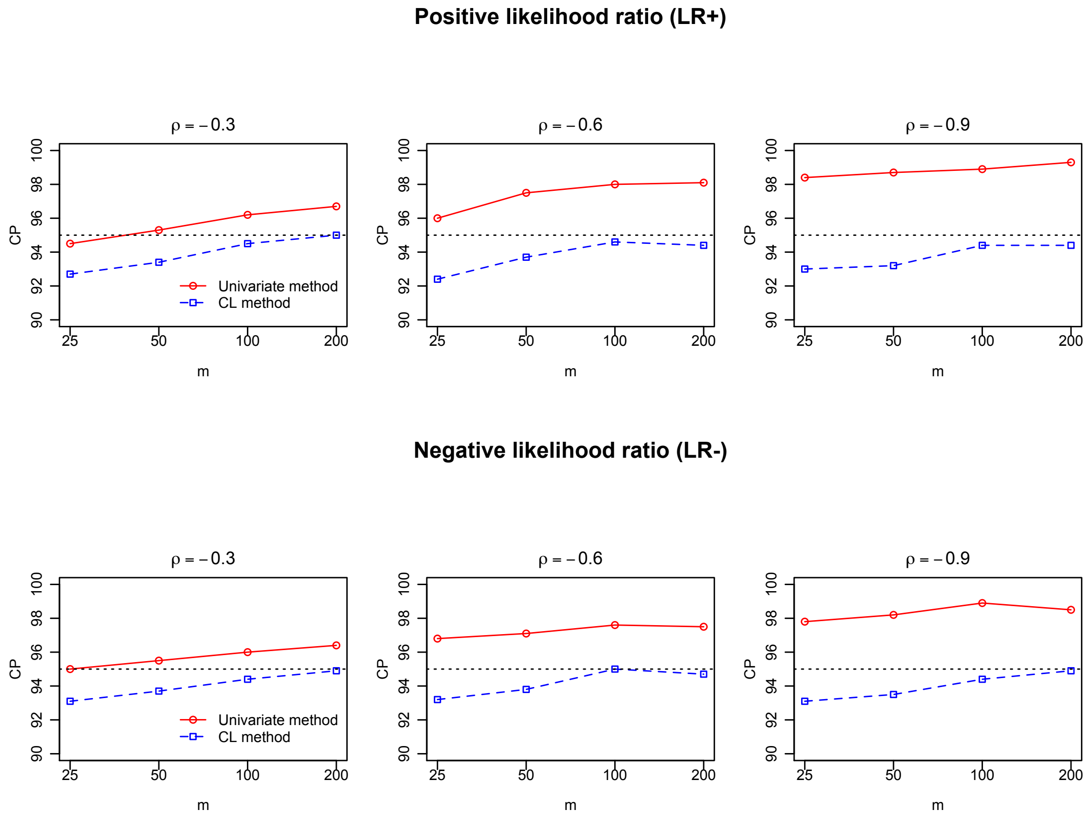

Figure S1: Coverage probability (CP) of the 95% confidence intervals for LR+ (Upper panels) and LR− (Lower panels) based on univariate method and the composite likelihood (CL) method under different settings of between-study correlation ρ and number of studies (m)

Section 2: R program for the composite likelihood method and SAS program for the standard maximum likelihood method and a working example

```
####################
## R program
####################

### functions to compute the score functions w.r.t theta1 and theta2
## density function of logit-normal
dlogitnorm = function(p, mu, tau2){
(2*pi*tau2)^(-1/2)*exp(-(qlogis(p)-mu)^2/(2*tau2))/(p*(1-p))
}

integrand1 = function(n,y, mypar2, p.grid){
dbinom(y, size=n, prob=p.grid)*dlogitnorm(p.grid, mypar2[1], mypar2[2])
}

integrand2 = function(n,y,mypar2, p.grid){
dbinom(y, size=n, prob=p.grid)*dlogitnorm(p.grid, mypar2[1], mypar2[2])*2*(qlogis(p.grid)-mypar2[1])/(2*mypar2[2])
}

integrand3 = function(n,y,mypar2, p.grid){
dbinom(y, size=n, prob=p.grid)*dlogitnorm(p.grid, mypar2[1], mypar2[2])
*(-1/(2*mypar2[2])+(qlogis(p.grid)-mypar2[1])^2/(2*mypar2[2]^2))
}

### R library for use of "integrate" and "glmmML" functions
library(stats)
library(glmmML)

### Example: dataset are collected from the distant metastais with the
### computed tomography (CT) diagnsotic test (Xing et al., 2011) in Section 4
estim = estim.orig = mbse.orig = matrix(NA,nrow=4,ncol=1)
```

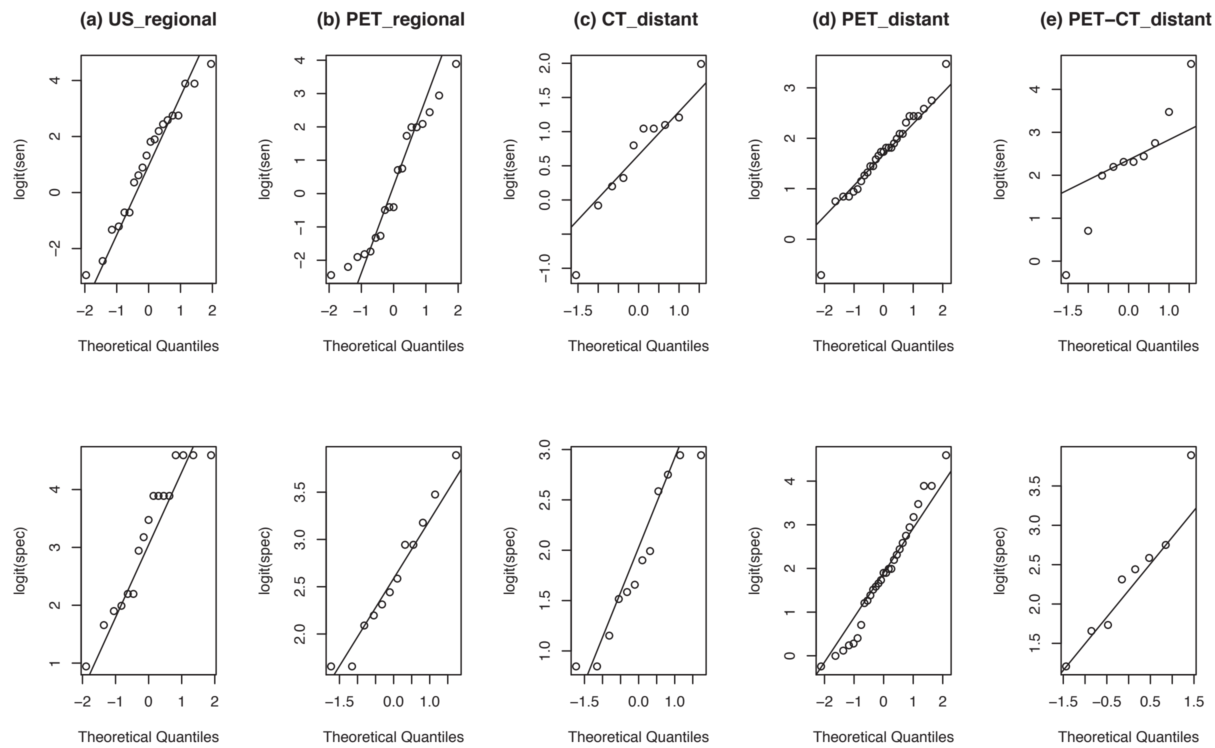

Figure S2: Upper panel: normal Q-Q plots with respect to sensitivity (logit scale) in each stratum; Lower panel: normal Q-Q plots with respect to specificity (logit scale) in each stratum. Here two regional subgroups, including PET-CT and CT, are removed from this analysis.

```
mySandwich = matrix(NA,nrow=4,ncol=4)

id = c(23,24,26,27,28,29,30,31,32,36,37,38)
Nd = c(0,4,37,87,9,5,44,46,59,13,89,15)
SeY = c(0,4,20,50,7,4,21,35,41,9,66,4)
Nn = c(151,85,39,17,9,29,59,18,47,102,161,59)
SpY = c(124,65,33,12,8,28,56,13,44,85,141,55)
mydat = data.frame(cbind(id,Nd,SeY,Nn,SpY))
names(mydat) = c("id", "Nd", "SeY", "Nn", "SpY")
m=nrow(mydat)

### The method of the composite likelihood in Section 2
## 1.1 point estimate: fit GLMM model(using Gauss-Hermite quadrature method with points 8)
fit.SeY = glmmML(cbind(SeY, Nd-SeY)~1, data=mydat, cluster=id, method="ghq", n.points=8,
control = list(epsilon = 1e-08, maxit = 50000, trace = FALSE))
fit.SpY = glmmML(cbind(SpY, Nn-SpY)~1, data=mydat, cluster=id, method="ghq", n.points=8,
control = list(epsilon = 1e-08, maxit = 50000, trace = FALSE))
```

Table S3: Numbers of studies collected by Xing et al., (2011) stratified by the type of metastasis (i.e., regional versus distant metastasis) and the type of imaging modalities.

Types	**# of study**
Regional metastasis	
Ultrasonography (US)	20
Computed Tomography (CT)	3
Positron Emission Tomography (PET)	21
Combination of both (PET-CT)	5
Distant metastasis	
Computed Tomography (CT)	12
Positron Emission Tomography (PET)	29
Combination of both (PET-CT)	8

```
## Logit scale for sensitivity and specificity
estim[c(1:2)] = c(as.numeric(fit.SeY$coefficients), (fit.SeY$sigma)^2)
estim[c(3:4)] = c(as.numeric(fit.SpY$coefficients), (fit.SpY$sigma)^2)

## translate to the original scale for sensitivity and specificity
estim.orig[c(1:2)] = c(plogis(estim[1]), estim[2])
estim.orig[c(3:4)] = c(plogis(estim[3]), estim[4])

## 1.2 The variance estimate: using the Sandwich method
## calculate I11.hat and I22.hat
I11.hat = solve(m*fit.SeY$variance)
I22.hat = solve(m*fit.SpY$variance)

## calculation of the score function (study-specific) w.r.t theta1
int1 = int2 = int3 = rep(NA, length=m)
est1 = estim[c(1: 2)] ##plug in the point estimate based on sensitivity
for(i in 1: m){

int1[i] = integrate(integrand1, lower=0, upper=1, n=mydat$Nd[i], mypar2=est1, y=mydat$SeY[i])$value
int2[i] = integrate(integrand2, lower=0, upper=1, n=mydat$Nd[i], mypar2=est1, y=mydat$SeY[i])$value
int3[i] = integrate(integrand3, lower=0, upper=1, n=mydat$Nd[i], mypar2=est1, y=mydat$SeY[i])$value
}

## derivation for SeY w.r.t beta1 and tau1^2 based on the sensitivity data
deriveSeY.mu = int2/int1
deriveSeY.tau2 = int3/int1
B.SeY = cbind(deriveSeY.mu, deriveSeY.tau2)

## calculation of the score function (study-specific) w.r.t theta2
int1 = int2 = int3 = rep(NA, length=m)
est2 = estim[c(3:4)] ##plug in the point estimate based on specificity
for(i in 1: m){

int1[i] = integrate(integrand1, lower=0, upper=1, n=mydat$Nn[i], mypar2=est2, y=mydat$SpY[i])$value
int2[i] = integrate(integrand2, lower=0, upper=1, n=mydat$Nn[i], mypar2=est2, y=mydat$SpY[i])$value
int3[i] = integrate(integrand3, lower=0, upper=1, n=mydat$Nn[i], mypar2=est2, y=mydat$SpY[i])$value
}

## derivation for SpY w.r.t beta2 and tau2^2 based on the specificity data
deriveSpY.mu = int2/int1
deriveSpY.tau2 = int3/int1
B.SpY = cbind(deriveSpY.mu, deriveSpY.tau2)

I12.hat = t(B.SeY)%*%B.SpY/m

myoff.diag = solve(I11.hat)%*%I12.hat%*%solve(I22.hat)/m
myupper = cbind(fit.SeY$variance, (myoff.diag))
mylower = cbind(t(myoff.diag), fit.SpY$variance)
myV = rbind(myupper, mylower)

## calculate the model-based standard error for sensitivity and specificity using delta method
se.mbse = (estim.orig[1]*(1-estim.orig[1]))*sqrt(diag(myV)[1])
sp.mbse = (estim.orig[3]*(1-estim.orig[3]))*sqrt(diag(myV)[3])
s1.2.mbse = sqrt(diag(myV)[2]); s2.2.mbse = sqrt(diag(myV)[4])
mbse.orig = rbind(se.mbse, s1.2.mbse, sp.mbse, s2.2.mbse)

####################
## SAS program
####################

/*The method of the full likelihood function in Section 2*/
/*Input the data*/
data CTdistant;
input id Nd SeY Nn SpY;
datalines;
23  0   0 151 124
24  4   4  85  65
26 37  20  39  33
27 87  50  17  12
28  9   7   9   8
29  5   4  29  28
30 44  21  59  56
31 46  35  18  13
32 59  41  47  44
36 13   9 102  85
37 89  66 161 141
38 15   4  59  55
run;

data bvmeta;
set CTdistant;
keep id Y ntp nfp ntn nfn;
ntp=SeY; nfn=Nd-ntp;
ntn=SpY; nfp=Nn-ntn;
Y=1;
run;
```

```
/*Apply the NLMIXED procedure to fit the equation(3)in Section 2 */
proc nlmixed data=bvmeta fd df=1000 gtol=1e-10;
parms alpha0=2.2 beta0=2.9 sigse=0 sigsp=0 fz3=0.7;
lsei = alpha0 + muse ;
lspi = beta0  + musp ;
Sei = 1/(1+exp(-lsei));
Spi = 1/(1+exp(-lspi));
RhoSeSp = (exp(2*fZ3)-1)/(exp(2*fZ3)+1);
logL =  ntp*log(Sei)+nfn*(log(1-Sei))+ntn*log(Spi)+nfp*log(1-Spi);
model Y~ general(logL);
random muse musp ~normal([0, 0],[exp(2*sigse),
                                 RhoSeSp*exp(sigse+sigsp), exp(2*sigsp)])
                          subject = id;
estimate "RhoSeSp" (exp(2*fZ3)-1)/(exp(2*fZ3)+1);
estimate "sigse" exp(sigse);
estimate "sigsp" exp(sigsp);
estimate "Se" 1/(1+exp(-alpha0));
estimate "Sp" 1/(1+exp(-beta0));
estimate "LRp" exp(alpha0)*(1+exp(beta0))/(1+exp(alpha0));
estimate "LRn"  (1+exp(beta0))/(exp(beta0)*(1+exp(alpha0)));
estimate "logitSe" alpha0;
estimate "logitSp" beta0;
ods output ParameterEstimates=IaParaEst0 AdditionalEstimates=IaAddEst0 FitStatistics=IaFit0;
run;
}
```