Meta-Analytic Depiction Of Ordered Categorical Diagnostic Test Accuracy In ROC Space
No Thresholds Left Behind

Ben A. Dwamena, MD

The University of Michigan Radiology & VAMC Nuclear Medicine, Ann Arbor, Michigan

2009 Stata Conference, Washington, DC - July 31, 2009
Outline

1. Objectives
2. Diagnostic Test Evaluation
3. Example Data
5. Proposed Algorithm for Meta-analysis of Ordinal Data
6. Worked Examples
7. Concluding Remarks
Objectives

1. Review underlying concepts of medical diagnostic test evaluation

2. Provide illustrated overview of current methods for meta-analysis of diagnostic test accuracy studies with discrete outcomes

3. Describe a robust and flexible parametric algorithm for meta-analysis of ordered categorical data

4. Demonstrate implementation with Stata using two data sets, one with studies reporting same set of categories and the other with disparately categorized outcomes
Medical Diagnostic Test

Any measurement aiming to identify individuals who could potentially benefit from preventative or therapeutic intervention

This includes:

1. Elements of medical history e.g. Retrosternal chest pain
2. Physical examination e.g. Systolic blood pressure
3. Imaging procedures e.g. Chest xray
4. Laboratory investigations. e.g. Fasting blood sugar
5. Clinical prediction rules e.g. Geneva Score for Venous Thromboembolism
Diagnostic Test Types/Scales

1. **Dichotomous** using single implicit or explicit threshold
   
   eg. Presence or absence of a specific DNA sequence in blood serum
   
   eg. Fasting blood glucose $\geq 126$ mg/ml diagnostic of diabetes mellitus

2. **Ordered Categorical** with multiple implicit or explicit thresholds
   
   eg. the BIRADS scale for mammograms: 1 ‘Benign’; 2 ‘Possibly benign’; 3 ‘Unclear’; 4 ‘Possibly malignant’; 5 ‘Malignant’
   
   eg. Clinical symptoms classified as 1 ‘not present’, 2 ‘mild’, 3 ‘moderate’, or 4 ‘severe’

3. **Continuous**
   
   eg. biochemical tests such as serum levels of creatinine, bilirubin or calcium
Diagnostic Accuracy Studies

Figure: Basic Study Design

SERIES OF PATIENTS

INDEX TEST

REFERENCE TEST

CROSS-CLASSIFICATION
Figure: Distributions of test result for diseased and non-diseased populations defined by threshold (DT)
Binary Test Accuracy

Data Structure

Data often reported as $2 \times 2$ matrix

<table>
<thead>
<tr>
<th></th>
<th>Reference Test (Diseased)</th>
<th>Reference Test (Healthy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test Positive</td>
<td>True Positive (a)</td>
<td>False Positive (b)</td>
</tr>
<tr>
<td>Test Negative</td>
<td>False Negative (c)</td>
<td>True Negative (d)</td>
</tr>
</tbody>
</table>

1. The chosen threshold may vary between studies of the same test due to inter-laboratory or inter-observer variation

2. The higher the cut-off value, the higher the specificity and the lower the sensitivity
Binary Test Accuracy
Measures of Test Performance

**Sensitivity (true positive rate)** The proportion of subjects with disease who are correctly identified as such by test \((a/a+c)\)

**Specificity (true negative rate)** The proportion of subjects without disease who are correctly identified as such by test \((d/b+d)\)

**Positive predictive value** The proportion of test positive subjects who truly have disease \((a/a+b)\)

**Negative predictive value** The proportion of test negative subjects who truly do not have disease \((d/c+d)\)
Binary Test Accuracy
Measures of Test Performance

**Likelihood ratios (LR)** The ratio of the probability of a positive (or negative) test result in the patients with disease to the probability of the same test result in the patients without the disease (sensitivity/1-specificity) or (1-Sensitivity/specificity)

**Diagnostic odds ratio** The ratio of the odds of a positive test result in patients with disease compared to the odds of the same test result in patients without disease (LRP/LRN)
Non-binary Test Accuracy

ROC Curve Analysis

The accuracy of continuously or ordinally scaled tests is best summarized by ROC curve, a plot of all pairs of \((1-\text{specificity, sensitivity})\) as positivity threshold varies:

1. Provides complete description of potential performance

2. Facilitates comparison and combination of information across studies of the same test

3. Guides the choice of thresholds in applications

4. Provides a mechanism for relevant comparisons between different non-binary tests
Non-binary Test Accuracy

ROC Curve Analysis

Figure: ROC curve derived from changing test threshold

TP rate, $Se$

FP rate, $(1-Sp)$

Lower threshold

Starting point

Higher threshold
## Non-binary Test Accuracy

### ROC Curve Analysis

<table>
<thead>
<tr>
<th>Index Name</th>
<th>Notation</th>
<th>Definition</th>
<th>Interpretations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area under Curve</td>
<td>AUC</td>
<td>Integrate ROC over range(0-1)</td>
<td>Average TPF across all possible FPF</td>
</tr>
<tr>
<td>Specific ROC point</td>
<td>ROC($t_0$)</td>
<td>ROC($t_0$)</td>
<td>P[$Y_D &gt; q$]</td>
</tr>
<tr>
<td>Partial Area under curve</td>
<td>pAUC($t_0$)</td>
<td>Integrate ROC over range(0-$t_0$)</td>
<td>Average TPF across FPF ∈ (0-$t_0$)</td>
</tr>
<tr>
<td>Symmetry Point</td>
<td>Sym</td>
<td>ROC(Sym)=$\text{Sym}$</td>
<td>Sensitivity=Specificity</td>
</tr>
</tbody>
</table>

$Y_D$: Test result for diseased

$q=1-t_0$ quantile for $Y_D$
Test result for each individual $Y$ falls into one of $J$ categories ("ratings")

These categories are ordered in terms of increasing likelihood of disease

Data often reported as $2 \times j$ matrix

<table>
<thead>
<tr>
<th>Category</th>
<th>Diseased</th>
<th>Healthy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_1$</td>
<td>$n_{d1}$</td>
<td>$n_{h1}$</td>
<td>$n_1$</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>$C_j$</td>
<td>$n_{dj}$</td>
<td>$n_{hj}$</td>
<td>$n_j$</td>
</tr>
<tr>
<td>Total</td>
<td>$n_d$</td>
<td>$n_h$</td>
<td>$N$</td>
</tr>
</tbody>
</table>
Ordinal Data Analysis

Example Data

117 consecutive patients older than age 50 admitted to a Veterans Affairs (VA) nursing home (NH).

Screened for alcohol dependence using CAGE questionnaire as index test.

DSM-III-R criteria were used as Reference standard.

Forty-nine percent of study participants had lifetime alcohol abuse or dependence.
CAGE Scores for Alcoholism Screening

CAGE is an acronym for each of four questions:

1. Have you ever felt you should cut down on your drinking?
2. Have people annoyed you by criticizing your drinking?
3. Have you ever felt bad or guilty about your drinking?
4. Have you ever had a drink in the morning to get rid of a hangover?

Each question is scored 1 or 0 for YES or NO answers respectively.
### Ordinal Data Analysis

Example Data

<table>
<thead>
<tr>
<th>Score</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>45</td>
<td>9</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Abnormal</td>
<td>1</td>
<td>9</td>
<td>17</td>
<td>7</td>
<td>23</td>
</tr>
</tbody>
</table>

**Table:** Single Study CAGE Scores
Ordinal Data Analysis

Approaches

1. Dichotomization at single threshold and analysis as binary data
2. Empirical ROC plot of sensitivity and 1-specificity at different thresholds
3. Binormal ROC analysis
4. ROC analysis via Ordinal regression
Ordinal Data Analysis

Dichotomized Data

Recommended Positivity Threshold: Cage Score $\geq 2$

<table>
<thead>
<tr>
<th></th>
<th>DSM-III-R (Abnormal)</th>
<th>DSM-III-R (Normal)</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGE $\geq 2$</td>
<td>47</td>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>CAGE $&lt; 2$</td>
<td>10</td>
<td>54</td>
<td>64</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>60</td>
<td>117</td>
</tr>
</tbody>
</table>

Sensitivity (percent): $(47/57)*100 = 82$
Specificity (percent): $(54/60)*100 = 90$
Positive Predictive Value (percent): $(47/53)*100 = 90$
Negative Predictive Value (percent): $(54/64)*100 = 84$
Ordinal Data Analysis
Empirical ROC Analysis

Based on sensitivity/specificity pairs at multiple thresholds: The higher the cut-off value, the higher the specificity and the lower the sensitivity

**Sensitivity (TPR) at each threshold**
Number of diseased diagnosed positive/Number of diseased

**Specificity (TNR) at each threshold**
Number of non-diseased diagnosed negative/Number of non-diseased

<table>
<thead>
<tr>
<th>cutpoint</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;= 0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>&gt;= 1</td>
<td>98</td>
<td>75</td>
</tr>
<tr>
<td>&gt;= 2</td>
<td>82</td>
<td>90</td>
</tr>
<tr>
<td>&gt;= 3</td>
<td>52</td>
<td>96</td>
</tr>
<tr>
<td>&gt;= 4</td>
<td>40</td>
<td>100</td>
</tr>
<tr>
<td>&gt; 4</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
Ordinal Data Analysis

Empirical ROC Analysis with `roctab`

```
. roctab dtruth score [fw=dis], graph aspect(1)
```

Area under ROC curve = 0.9395
Ordinal Data Analysis
Binormal ROC Analysis

Test results of diseased and healthy subjects follow normal distributions with respective means $\mu_1$, $\mu_0$ and standard deviations $\sigma_1$ and $\sigma_0$

1. Scaled mean difference, $a = (\mu_1 - \mu_0)/\sigma_1$

2. Scale coefficient, $b = \sigma_0 / \sigma_1$

3. The binormal ROC curve: $TPR = a + b\Phi(FPR) \ (0 \leq FPR \leq 1)$

4. The area under curve, $AUROC = \Phi \left( \frac{a}{\sqrt{1+b^2}} \right)$

5. The symmetry point index, $Sym = \Phi \left( \frac{a}{1+b} \right)$
## Ordinal Data Analysis

Binormal ROC Analysis using `rocfit`

```bash
.rocfit dtruth score [fw=dis]
```

Binormal model of `dtruth` on `score`

|              | Coef.    | Std. Err. | z     | P>|z| | [95% Conf. Interval] |
|--------------|----------|-----------|-------|-----|----------------------|
| intercept    | 2.919589 | 0.648216  | 4.50  | 0.000 | 1.649110 4.190068    |
| slope (*)     | 1.443559 | 0.444514  | 1.00  | 0.318 | 0.572328 2.314789    |
| /cut1        | 0.663779 | 0.175958  | 3.77  | 0.000 | 0.318907 1.008651    |
| /cut2        | 1.355698 | 0.207460  | 6.53  | 0.000 | 0.949083 1.762312    |
| /cut3        | 1.950510 | 0.296180  | 6.59  | 0.000 | 1.370009 2.531011    |
| /cut4        | 2.207473 | 0.358501  | 6.16  | 0.000 | 1.504824 2.910122    |

Indices from binormal fit

<table>
<thead>
<tr>
<th>Index</th>
<th>Estimate</th>
<th>Std. Err.</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC area</td>
<td>0.951799</td>
<td>0.018976</td>
<td>0.914606 0.988991</td>
</tr>
<tr>
<td>delta(m)</td>
<td>2.022495</td>
<td>0.328053</td>
<td>1.379522 2.665467</td>
</tr>
<tr>
<td>d(e)</td>
<td>2.389621</td>
<td>0.270268</td>
<td>1.859906 2.919336</td>
</tr>
<tr>
<td>d(a)</td>
<td>2.351199</td>
<td>0.267932</td>
<td>1.826062 2.876336</td>
</tr>
</tbody>
</table>
Ordinal Data Analysis

Binormal ROC Curve using \texttt{rocplot}

\texttt{. rocplot, norefline aspect(1)}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{roc_curve}
\caption{Binormal ROC Curve with area under curve = 0.9491 and standard error (se(area)) of 0.0274.}
\end{figure}
Ordinal Data Analysis

Suppose, the test result $Y$ falls into one of $J$ categories ("ratings")
The probability of $Y$ falling in a given category $j$ or lower may be modeled
as a non-linear function using the ordinal regression equation:

$$g[Pr(Y \leq j \mid D)] = \frac{\theta_j - \alpha D}{\exp(\beta D)}$$

$g$: Cumulative link function

$D$ is a variable indicative of disease status

$\theta_j,...,\theta_{j-1}$: Cut-off values on an underlying latent scale

$\alpha$: Location parameter (measure of diagnostic accuracy)

$\beta$: Scale parameter (spread of responses across subjects)
Ordinal Data Analysis
Choice of Link Functions for Ordinal Regression

1. **Probit** This is the inverse standard normal cumulative distribution function. More suitable when a dependent variable is normally distributed.

2. **Logit** $f(x) = \log(x/(1 - x))$. This is usually used when the dependent ordinal variable has equal category.

3. **Log-log** $f(x) = -\log(-\log(x))$. Recommended when the probability of the lower category is high.

4. **Complementary log-log** $f(x) = \log(-\log(1 - x))$. Recommended when the probability of higher category is high.

5. **Cauchit** $f(x) = \tan(p(x - 0.5))$. This is used when extreme values are present in the data.
## Ordinal Data Analysis

Ordinal Probit ROC Analysis with oglm

```
. oglm score dtruth [fw=dis], link(probit) ls het(dtruth)
```

Heteroskedastic Ordered Probit Regression

|               | Coef.   | Std. Err. | z    | P>|z|   | [95% Conf. Interval] |
|---------------|---------|-----------|------|-------|----------------------|
| location      |         |           |      |       |                      |
| dtruth        | 2.022494| .3280529  | 6.17 | 0.000 | 1.379522  2.665466   |
| scale         |         |           |      |       |                      |
| dtruth        | -.3671114| .307929  | -1.19 | 0.233 | -.9706412  .2364184  |
| /cut1         | .6637788 | .1759585  | 3.77 | 0.000 | .3189066  1.008651   |
| /cut2         | 1.355697 | .2074599  | 6.53 | 0.000 | .9490835  1.762311   |
| /cut3         | 1.95051  | .2961795  | 6.59 | 0.000 | 1.370009  2.531011   |
| /cut4         | 2.207472 | .3585009  | 6.16 | 0.000 | 1.504824  2.910121   |

Number of obs = 117
LR chi2(2) = 93.77
Prob > chi2 = 0.0000
Pseudo R2 = 0.2707
Ordinal Data Analysis

Ordinal Probit ROC Curve with **roccat**

```plaintext
.roccat, avar('avar') avarlo('avarlo') avarhi('avarhi') bvar('bvar') ///
bvarlo('bvarlo') bvarhi('bvarhi') np(5000)
```

![ROC Curve Diagram](image)

AUROC = 0.95 [0.92 – 0.95]

Location = 2.92, Scale = 1.444
Diagnostic Meta-analysis
Critical review and statistical combination of previous research

Rationale

1. Too few patients in a single study
2. Too selected a population in a single study
3. No consensus regarding accuracy, impact, reproducibility of test(s)
4. Data often scattered across several journals
5. Explanation of variability in test accuracy
6. etc.
Diagnostic Meta-analysis

Scope

1. Identification of the number, quality and scope of primary studies

2. Quantification of overall classification performance (sensitivity and specificity), discriminatory power (diagnostic odds ratios) and informational value (diagnostic likelihood ratios)

3. Assessment of the impact of technological evolution (by cumulative meta-analysis based on publication year), technical characteristics of test, methodological quality of primary studies and publication selection bias on estimates of diagnostic accuracy

4. Highlighting of potential issues that require further research
Meta-analysis of diagnostic accuracy studies may be performed to provide summary estimates of test performance based on a collection of studies and their reported empirical or estimated smooth ROC curves.

Statistical methodology for meta-analysis of diagnostic accuracy studies focused on studies reporting estimates of test sensitivity and specificity or two by two data.

Both fixed and random-effects meta-analytic models have been developed to combine information from such studies.
1. To meta-analyze studies with results in more than two categories, results are often dichotomized in order to employ one of the binary methods.

2. It is more efficient and informative to take all thresholds into account.

3. Existing methods require the same number and set of thresholds, are computationally intensive adaptations of the binary methods or are based on hierarchical ordinal probit regression implementable using Bayesian inference.
Example Dataset 1
10 studies on CAGE for alcohol dependence screening (5 listed in table) using Similar Thresholds

Table: Observed Data

<table>
<thead>
<tr>
<th>Idnum</th>
<th>Author</th>
<th>Setting</th>
<th>Score</th>
<th>dis0</th>
<th>dis1</th>
<th>tdis0</th>
<th>tdis1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Saitz</td>
<td>PC</td>
<td>0</td>
<td>99</td>
<td>6</td>
<td>134</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>26</td>
<td>9</td>
<td>134</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>6</td>
<td>19</td>
<td>134</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td>21</td>
<td>134</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>21</td>
<td>134</td>
<td>76</td>
</tr>
<tr>
<td>2</td>
<td>McQuade</td>
<td>PC</td>
<td>0</td>
<td>197</td>
<td>7</td>
<td>247</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>31</td>
<td>11</td>
<td>247</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>17</td>
<td>12</td>
<td>247</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td>13</td>
<td>247</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
<td>10</td>
<td>247</td>
<td>53</td>
</tr>
<tr>
<td>4</td>
<td>Chan</td>
<td>PC</td>
<td>0</td>
<td>38</td>
<td>2</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>9</td>
<td>4</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>7</td>
<td>15</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>0</td>
<td>17</td>
<td>56</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>Bradley</td>
<td>AMP</td>
<td>0</td>
<td>69</td>
<td>32</td>
<td>117</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>33</td>
<td>20</td>
<td>117</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>13</td>
<td>28</td>
<td>117</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>1</td>
<td>20</td>
<td>117</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>10</td>
<td>117</td>
<td>110</td>
</tr>
<tr>
<td>10</td>
<td>Indran</td>
<td>AMP</td>
<td>0</td>
<td>179</td>
<td>0</td>
<td>483</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>120</td>
<td>4</td>
<td>483</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>126</td>
<td>24</td>
<td>483</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>53</td>
<td>19</td>
<td>483</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>483</td>
<td>52</td>
</tr>
</tbody>
</table>
Example Dataset 2
19 studies evaluating EBCT for diagnosis of coronary artery disease (15 listed in table)

Table: Disparate Thresholds

<table>
<thead>
<tr>
<th>Author</th>
<th>Abnormal</th>
<th>Normal</th>
<th>Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Budoff</td>
<td>427</td>
<td>283</td>
<td>2</td>
</tr>
<tr>
<td>Seese</td>
<td>87</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>Yao</td>
<td>45</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>Chen</td>
<td>74</td>
<td>42</td>
<td>4</td>
</tr>
<tr>
<td>Hosoi</td>
<td>202</td>
<td>80</td>
<td>5</td>
</tr>
<tr>
<td>Budoff</td>
<td>983</td>
<td>868</td>
<td>5</td>
</tr>
<tr>
<td>Almeda</td>
<td>160</td>
<td>86</td>
<td>4</td>
</tr>
<tr>
<td>Knez</td>
<td>1255</td>
<td>860</td>
<td>4</td>
</tr>
<tr>
<td>Wong</td>
<td>28</td>
<td>900</td>
<td>5</td>
</tr>
<tr>
<td>Shaw</td>
<td>249</td>
<td>10128</td>
<td>5</td>
</tr>
<tr>
<td>Greenland</td>
<td>84</td>
<td>945</td>
<td>4</td>
</tr>
<tr>
<td>Arad</td>
<td>129</td>
<td>4484</td>
<td>4</td>
</tr>
<tr>
<td>Taylor</td>
<td>14</td>
<td>1611</td>
<td>4</td>
</tr>
<tr>
<td>Vliengenhart</td>
<td>50</td>
<td>1745</td>
<td>4</td>
</tr>
<tr>
<td>La Monte</td>
<td>287</td>
<td>10459</td>
<td>4</td>
</tr>
</tbody>
</table>
Methodological Overview

1. Dichotomization At Single Threshold And Meta-Analysis As Binary Data

2. Proportional Odds Ordinal Regression Modeling

3. Bivariate Random-Effects Meta-Analysis of Slope And Intercept from Study-Specific Logit-Threshold Linear Regression

4. Bayesian Hierarchical Location-Scale Ordinal Regression Modeling
Methods for Dichotomized Data

Examples

1. Meta-analysis of sensitivity and specificity separately by direct pooling or modeling using fixed-effects or random-effects approaches.

2. Meta-analysis of positive and negative likelihood ratios separately using fixed-effects or random-effects approaches as applied to risk ratios in meta-analysis of therapeutic trials.

3. Meta-analysis of diagnostic odds ratios using fixed-effects or random-effects approaches as applied to meta-analysis of odds ratios in clinical treatment trials.

4. Summary ROC Meta-analysis using fixed-effects or random-effects approaches.
Methods for Dichotomized Data

Example Dataset: CAGE

<table>
<thead>
<tr>
<th>Author</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saitz</td>
<td>60</td>
<td>9</td>
<td>15</td>
<td>125</td>
</tr>
<tr>
<td>McQuade</td>
<td>35</td>
<td>20</td>
<td>18</td>
<td>227</td>
</tr>
<tr>
<td>Brown</td>
<td>44</td>
<td>9</td>
<td>19</td>
<td>52</td>
</tr>
<tr>
<td>Chan</td>
<td>42</td>
<td>9</td>
<td>6</td>
<td>47</td>
</tr>
<tr>
<td>Aertgeerts</td>
<td>80</td>
<td>90</td>
<td>95</td>
<td>1705</td>
</tr>
<tr>
<td>Buchsbaum</td>
<td>215</td>
<td>47</td>
<td>79</td>
<td>480</td>
</tr>
<tr>
<td>Joseph</td>
<td>48</td>
<td>6</td>
<td>10</td>
<td>54</td>
</tr>
<tr>
<td>Bradley</td>
<td>58</td>
<td>15</td>
<td>52</td>
<td>102</td>
</tr>
<tr>
<td>Jones</td>
<td>12</td>
<td>1</td>
<td>13</td>
<td>128</td>
</tr>
<tr>
<td>Indran</td>
<td>48</td>
<td>184</td>
<td>4</td>
<td>299</td>
</tr>
</tbody>
</table>
Methods for Dichotomized Data

Summary ROC Meta-analysis

The most commonly used and easy to implement method
It is a fixed-effects model

1. Linear regression analysis of the relationship
   \[ D = a + bS \] where:
   \[ D = (\text{logit TPR}) - (\text{logit FPR}) = \ln \text{DOR} \]
   \[ S = (\text{logit TPR}) + (\text{logit FPR}) = \text{proxy for the threshold} \]

2. \(a\) and \(b\) may be estimated by weighted or un-weighted least squares
   or robust regression, back-transformed and plotted in ROC space

3. Differences between tests or subgroups may examined by adding
   covariates to model
Methods for Dichotomized Data
Summary ROC Meta-analysis

\[ . \text{sroc } \text{tp fn fp tn} \]

Weighted Regression of D on S:
Slope = 0.088, Intercept = 3.152, n = 10
\[ t = 0.63, \text{ prob } >|t| = 0.545 \]

Homogeneous: thus \( \ln(\text{OR}) = 3.152 \) and OR = 23.380

AUC and Q*:
AUC = 0.898, se(AUC) = 0.020, 95% CI = (0.858, 0.937) (homogenous)
AUC = 0.896, se(AUC) = 0.019, 95% CI = (0.858, 0.934) (heterogenous)

Q* = (0.829, 0.171), se(Q*) = 0.021, 95% CI = ({0.787, 0.870}, {0.130, 0.213})

Correlation Test:
Spearman correlation (rho) = 0.709, p(rho=0) = 0.022

High correlation: use the summary ROC curve; do not use the summary TPR and FPR.
Methods for Dichotomized Data

Summary ROC Meta-analysis

\[ sroc \ tp \ fn \ fp \ tn \]

Weighted Regression of D on S:
Slope = 0.088, Intercept = 3.152, n = 10
\( t = 0.63, \text{ prob } > |t| = 0.545 \)

D as a function of S
(Circles are proportional to 1/var)

\[ D = \text{logit(tpr)} - \text{logit(fpr)} \]
\[ S = \text{logit(tpr)} + \text{logit(fpr)} \]
Slope = 0.088, Intercept = 3.152; \( t = 0.63, \text{ prob } > |t| = 0.545 \)
Methods for Dichotomized Data

Summary ROC Meta-analysis

```
sroc tp fn fp tn
```

sROC Curve

(Circles are proportional to 1/var)

AUC = 0.896, $Q^* = (0.829, 0.171)$
Mathematically equivalent models for estimating underlying SROC and average operating point and/or exploring heterogeneity

Hierarchical Summary ROC (HSROC) Model

1. Focused on inferences about the SROC curve, or comparing SROC curves but summary operating point(s) can be derived from the model parameters

Bivariate Mixed Effects Models

1. Focused on inferences about sensitivity and specificity but SROC curve(s) can be derived from the model parameters
2. Generalization of the commonly used DerSimonian and Laird random effects model
Level 1: Within-study variability

\[ y_{ij} \sim Bin(n_{ij}, \pi_{ij}) \]

\[ \text{logit}(\pi_{ij}) = (\theta_i + \alpha_i X_{ij}) \exp(-\beta X_{ij}) \]

\( \theta_i \) and \( \alpha_i \) Study-specific threshold and accuracy parameters

\( y_{ij} \) Number testing positive assumed to be binomially distributed

\( \pi_{ij} \) Probability that a patient in study \( i \) with disease status \( j \) has a positive test result

\( X_{ij} \) True disease status (coded -0.5 for those without disease and 0.5 for those with the disease)
Methods for Dichotomized Data
Hierarchical Summary ROC Regression

Level 2: Between-study variability

\[
\theta_i \sim N(\Theta, \sigma^2_\theta)
\]

\[
\alpha_i \sim N(A, \sigma^2_\alpha)
\]

\(\Theta\) and \(A\) Means of the normally distributed threshold and accuracy parameters

\(\sigma^2_\theta\) and \(\sigma^2_\alpha\) Variances of mean threshold and accuracy

\(\beta\) Shape parameter which models any asymmetry in the SROC curve
Methods for Dichotomized Data
Hierarchical Summary ROC Regression of CAGE Data

```plaintext
.metandi tp fp fn tn
```

Log likelihood = -74.385097  Number of studies = 10

|                      | Coef.  | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|----------------------|--------|-----------|-------|------|---------------------|
| HSRHC                |        |           |       |      |                     |
| Lambda               | 2.99829| .25168    | 2.505006 | 3.491573|
| Theta                | -.6057828| .2987906| -1.191402 | -.0201641|
| beta                 | .0472602| .2613612| 0.18 | 0.857| -.464984 | .5595187|
| s2alpha              | .1874206| .1608058| .0348731| 1.007265|
| s2theta              | .536654 | .264404  | .2043224| 1.409525|
| Summary pt.         |        |           |       |      |                     |
| Se                   | .7052654| .0527992| .5925838| .7974353|
| Sp                   | .8961592| .0246019| .8371398| .9354399|
| DOR                  | 20.65089| 3.888347| 14.27797| 29.86834|
| LR+                  | 6.791796| 1.281433| 4.692295| 9.83069|
| LR-                  | .3288864| .0526198| .2403582| .4500212|
| 1/LR-                | 3.040564| .4864715| 2.222118| 4.160458|
```
Methods for Dichotomized Data
Hierarchical Summary ROC Meta-analysis of CAGE Data

```
.metandi tp fp fn tn, plot
```
Methods for Dichotomized Data

Bivariate Mixed Model

**Level 1: Within-study variability: Approximate Normal Approach**

\[
\begin{pmatrix}
\text{logit}(p_{Ai}) \\
\text{logit}(p_{Bi})
\end{pmatrix}
\sim N\left(\begin{pmatrix}
\mu_{Ai} \\
\mu_{Bi}
\end{pmatrix}, C_i\right)
\]

\[
C_i = \begin{pmatrix}
    s_{Ai}^2 & 0 \\
    0 & s_{Bi}^2
\end{pmatrix}
\]

- \(p_{Ai}\) and \(p_{Bi}\): Sensitivity and specificity of the \(i\)th study
- \(\mu_{Ai}\) and \(\mu_{Bi}\): Logit-transforms of sensitivity and specificity of the \(i\)th study
- \(C_i\): Within-study variance matrix
- \(s_{Ai}^2\) and \(s_{Bi}^2\): Variances of logit-transforms of sensitivity and specificity
Methods for Dichotomized Data

Bivariate Mixed Model

Level 1: Within-study variability: Exact Binomial Approach

\[ y_{Ai} \sim Bin(n_{Ai}, p_{Ai}) \]

\[ y_{Bi} \sim Bin(n_{Bi}, p_{Bi}) \]

- \( n_{Ai} \) and \( n_{Bi} \) Number of diseased and non-diseased
- \( y_{Ai} \) and \( y_{Bi} \) Number of diseased and non-diseased with true test results
- \( p_{Ai} \) and \( p_{Bi} \) Sensitivity and specificity of the \( i \)th study
Methods for Dichotomized Data
Bivariate Mixed Model

Level 2: Between-study variability

\[
\begin{pmatrix}
\mu_{Ai} \\
\mu_{Bi}
\end{pmatrix}
\sim
\mathcal{N}
\left(
\begin{pmatrix}
M_A \\
M_B
\end{pmatrix},
\Sigma_{AB}
\right)
\]

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

\(\mu_{Ai}\) and \(\mu_{Bi}\) Logit-transforms of sensitivity and specificity of the \(i^{th}\) study

\(M_A\) and \(M_B\) Means of the normally distributed logit-transforms

\(\Sigma_{AB}\) Between-study variances and covariance matrix
Methods for Dichotomized Data

Bivariate Mixed Binary Regression of CAGE Data

. midas tp fp fn tn, res(all)

SUMMARY DATA AND PERFORMANCE ESTIMATES

Number of studies = 10
Reference-positive Units = 953
Reference-negative Units = 3609
Pretest Prob of Disease = 0.21

Correlation (Mixed Model)= -0.84
Proportion of heterogeneity likely due to threshold effect = 0.71
Interstudy variation in Sensitivity: ICC_SEN = 0.17, 95% CI = [0.02-0.32]
Interstudy variation in Specificity: ICC_SPE = 0.17, 95% CI = [0.03-0.31]
Heterogeneity (Chi-square): LRT_Q = 178.971, df =2.00, LRT_p =0.000
Inconsistency (I-square): LRT_I2 = 99, 95% CI = [98-99]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.72</td>
<td>[0.60, 0.81]</td>
</tr>
<tr>
<td>Specificity</td>
<td>0.90</td>
<td>[0.84, 0.94]</td>
</tr>
<tr>
<td>Positive Likelihood Ratio</td>
<td>7.3</td>
<td>[4.9, 10.7]</td>
</tr>
<tr>
<td>Negative Likelihood Ratio</td>
<td>0.31</td>
<td>[0.22, 0.44]</td>
</tr>
<tr>
<td>Diagnostic Odds Ratio</td>
<td>23</td>
<td>[16, 34]</td>
</tr>
</tbody>
</table>
Methods for Dichotomized Data

Bivariate Summary ROC Meta-analysis of CAGE data

```
.midas tp fp fn tn, sroc(curve mean data conf pred) level(95)
```

![ROC curve graph with observed data, summary operating point, 95% confidence region, and 95% prediction region]
Suppose, the test result $Y$ falls into one of $J$ categories ("ratings")

The probability of $Y$ falling in a given category $j$ or lower may be modeled using the ordinal regression equation:

$$\text{logit}[\Pr(Y \leq j \mid D)] = \theta_j - \alpha D$$

$D$ is a variable indicative of disease status

$\theta_j, \ldots, \theta_{j-1}$: Cut-off values on an underlying latent scale

$\alpha$: Location parameter (measure of diagnostic accuracy = log-odds ratio)
Proportional Odds Regression (POR) Framework

Alternative Fixed- or Random-effects Approaches

1. Single POR and log-odds ratio of pooled data

2. Single POR and log-odds ratio with adjustment for study using dummy variables

3. Study-specific POR and log-odds ratios

All ROC curves are symmetric because of the assumption of a constant odds ratio for test accuracy
Proportional Odds Regression Model

Fixed-effects POR of Pooled Data (FEPOR)

. oglm score resp [fw=dis], link(logit)

Ordered Logistic Regression

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>Std. Err.</th>
<th>z</th>
<th>P&gt;\mid z\mid</th>
<th>[95% Conf. Interval]</th>
</tr>
</thead>
<tbody>
<tr>
<td>resp</td>
<td>2.88352</td>
<td>0.0794469</td>
<td>36.29</td>
<td>0.000</td>
<td>2.727807 3.039233</td>
</tr>
<tr>
<td>/cut1</td>
<td>1.20843</td>
<td>0.0393556</td>
<td>30.71</td>
<td>0.000</td>
<td>1.131293 1.285564</td>
</tr>
<tr>
<td>/cut2</td>
<td>2.0914</td>
<td>0.0496541</td>
<td>42.12</td>
<td>0.000</td>
<td>1.99408  2.18872</td>
</tr>
<tr>
<td>/cut3</td>
<td>3.28055</td>
<td>0.0676223</td>
<td>48.51</td>
<td>0.000</td>
<td>3.148014 3.413088</td>
</tr>
<tr>
<td>/cut4</td>
<td>4.48068</td>
<td>0.0913089</td>
<td>49.07</td>
<td>0.000</td>
<td>4.30172  4.659645</td>
</tr>
</tbody>
</table>
Proportional Odds Regression Model

Random-effects POR of Pooled Data (REPOR)

.gllamm score resp, i(study) weight(wgt) link(ologit) eq(resp) adapt

number of level 1 units = 4562
number of level 2 units = 10
Condition Number = 9.5321335
log likelihood = -4296.7662

| score    | Coef. | Std. Err. | z     | P>|z|    | [95% Conf. Interval] |
|----------|-------|-----------|-------|--------|-----------------------|
| resp     | 3.046648 | .2144375 | 14.21 | 0.000  | 2.626358 3.466938     |
| _cut11   | 1.204772 | .0393759 | 30.60 | 0.000  | 1.127596 1.281947     |
| _cut12   | 2.106683 | .0501227 | 42.03 | 0.000  | 2.008444 2.204922     |
| _cut13   | 3.349023 | .0699056 | 47.91 | 0.000  | 3.212011 3.486036     |
| _cut14   | 4.601222 | .0950536 | 48.41 | 0.000  | 4.41492 4.787523      |

Variances and covariances of random effects

***level 2 (study) var(1): .37787108 (.18985077)
Proportional Odds Regression Model
Fixed-effects POR with Studies as Dummy Variables (FEPORD)

. oglm score resp std2-std10 [fw=dis], link(logit)

Ordered Logistic Regression
Number of obs = 4562
LR chi2(10) = 2042.14
Prob > chi2 = 0.0000
Pseudo R2 = 0.2009
Log likelihood = -4061.6084

| score | Coef.  | Std. Err. |  z   | P>|z|  | [95% Conf. Interval] |
|-------|--------|-----------|------|------|---------------------|
| resp  |  2.948556 |  0.0861223 | 34.24 | 0.000 | 2.779759 - 3.117352 |
| std2  |  -0.4815434 |  0.1887587 |  -2.55 | 0.011 |  -0.8515037 -  -0.1115832 |
| std3  |  -0.1028179 |  0.2302515 |  -0.45 | 0.655 |  -0.5541025  0.3484667 |
| std4  |   0.3582957 |  0.2305584 |   1.55 | 0.120 |  -0.0935903   0.8101818 |
| std5  |  -1.132683 |   0.1516207 |  -7.47 | 0.000 |  -1.429854 -  -0.8355117 |
| std6  |  -0.3640983 |  0.1543584 |  -2.36 | 0.018 |  -0.666352    0.0165615 |
| std7  |   0.2108051 |   0.2254389 |   0.94 | 0.350 |   -0.231047    0.6526571 |
| std8  |  -0.4197682 |  0.1926105 |  -2.18 | 0.029 |  -0.7972779  -0.0422585 |
| std9  |  -1.088437 |   0.2458275 |  -4.43 | 0.000 |  -1.57025    -0.6066241 |
| std10 |   1.158538 |   0.1589892 |   7.29 | 0.000 |   0.8469249   1.4701515 |

| /cut1 |   0.8164335 |   0.1402567 |   5.82 | 0.000 |  0.5415355  1.091332 |
| /cut2 |   1.809484 |   0.1437148 |  12.59 | 0.000 |   1.527808   2.09116 |
| /cut3 |   3.08217 |   0.1503633 |  20.50 | 0.000 |   2.787464   3.376877 |
| /cut4 |   4.319102 |   0.1617342 |  26.70 | 0.000 |   4.002108   4.636095 |
Proportional Odds Regression Model
Random-effects POR with Studies as Dummy Variables (REPORD)

```
.gllamm score resp std2-std10, i(study) weight(wgt) link(ologit) eq(resp) adapt
log likelihood = -4036.4392
```

| score | Coef. | Std. Err. | z     | P>|z|   | [95% Conf. Interval] |
|-------|-------|-----------|-------|-------|----------------------|
| resp  | 3.026 | .237      | 12.75 | 0.000 | 2.561                |
| std2  | -.318 | .239      | -1.33 | 0.185 | -.788                |
| std3  | -.034 | .333      | -0.10 | 0.918 | -.687                |
| std4  | .385  | .318      | 1.21  | 0.226 | -.238                |
| std5  | -.838 | .199      | -4.22 | 0.000 | -1.227               |
| std6  | -.387 | .215      | -1.80 | 0.072 | -.808                |
| std7  | .071  | .325      | 0.22  | 0.827 | -.566                |
| std8  | .479  | .261      | 1.84  | 0.066 | -.031                |
| std9  | -.922 | .310      | -2.97 | 0.003 | -1.529               |
| std10 | 1.527 | .204      | 7.48  | 0.000 | 1.127                |

| _cut11 | 1.057 | .186      | 5.68  | 0.000 | .692                |
| _cut12 | 2.069 | .189      | 10.93 | 0.000 | 1.699               |
| _cut13 | 3.363 | .196      | 17.18 | 0.000 | 2.979               |
| _cut14 | 4.619 | .206      | 22.42 | 0.000 | 4.216               |

Variances and covariances of random effects

***level 2 (study) var(1): .454 (.236)
Proportional Odds Regression Model
Study-specific POR

```stata
levelsof author, local(levels)
postutil clear
nois postfile porfile str30 Study ldor ldorse ldorlo ldorhi ///
using porresults, replace
foreach l of local levels{
    local study "'l'"
    nois oglm score dtruth [fw=dis] if author == "'l'", link(logit)
    nlcom (avar: _b[dtruth]), post
    local ldor= _b[avar]
    local ldorse=_se[avar]
    local ldorlo=_b[avar]-invnorm(1-$alph)*_se[avar]
    local ldorhi=_b[avar]+invnorm(1-$alph)*_se[avar]
    nois post porfile ("'study'") ("ldor") ("ldorse") ("ldorlo") ("ldorhi")
}
nois postclose porfile
```

Proportional Odds Regression Model
Study-specific Log-odds Ratios

. use porresults, clear
. nois list Study  ldor  ldorse  ldorlo  ldorhi, ///
sep(0) div ab(32) abs noo compress

<table>
<thead>
<tr>
<th>Study</th>
<th>ldor</th>
<th>ldorse</th>
<th>ldorlo</th>
<th>ldorhi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aertgeerts</td>
<td>2.543996</td>
<td>.1625985</td>
<td>2.225309</td>
<td>2.862683</td>
</tr>
<tr>
<td>Bradley</td>
<td>1.627137</td>
<td>.263943</td>
<td>1.109819</td>
<td>2.144456</td>
</tr>
<tr>
<td>Brown</td>
<td>2.655694</td>
<td>.4095413</td>
<td>1.853008</td>
<td>3.45838</td>
</tr>
<tr>
<td>Buchsbaum</td>
<td>3.492826</td>
<td>.1832168</td>
<td>1.853008</td>
<td>3.45838</td>
</tr>
<tr>
<td>Chan</td>
<td>3.679464</td>
<td>.5191509</td>
<td>2.661947</td>
<td>4.696981</td>
</tr>
<tr>
<td>Indran</td>
<td>2.17833</td>
<td>.2748064</td>
<td>1.63972</td>
<td>2.716941</td>
</tr>
<tr>
<td>Jones</td>
<td>4.223248</td>
<td>.6260694</td>
<td>2.996175</td>
<td>5.450322</td>
</tr>
<tr>
<td>Joseph</td>
<td>4.090933</td>
<td>.52996</td>
<td>3.05223</td>
<td>5.129636</td>
</tr>
<tr>
<td>McQuade</td>
<td>3.407526</td>
<td>.3518678</td>
<td>2.717878</td>
<td>4.097174</td>
</tr>
<tr>
<td>Saitz</td>
<td>3.832313</td>
<td>.3883162</td>
<td>3.071227</td>
<td>4.593399</td>
</tr>
</tbody>
</table>
Proportional Odds Regression Model

Meta-analysis of Study-specific Log-odds Ratios

```
. mvmeta ldor ldorvar, (fixed|ml|mm|reml) vars(ldor1)
```

| Model      | ldor      | Std. Err. | z    | P>|z|  | [95% Conf. Interval] |
|------------|-----------|-----------|------|------|---------------------|
| FESSPOR    | 2.865804  | 0.0891923 | 32.13| 0.000| 2.69099 - 3.040618  |
| RESSPOR ML | 3.087069  | 0.2629351 | 11.74| 0.000| 2.571725 - 3.602412 |
| RESSPOR MM | 3.0881    | 0.2624814 | 11.77| 0.000| 2.573646 - 3.602554 |
| RESSPOR REML| 3.094911 | 0.2771367 | 11.17| 0.000| 2.551733 - 3.638089 |
### Proportional Odds Regression Model

#### Summary AUROCs

<table>
<thead>
<tr>
<th>Approach</th>
<th>AUROC</th>
<th>CI</th>
<th>CIW</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEPOR</td>
<td>0.88</td>
<td>0.87-0.89</td>
<td>0.02</td>
</tr>
<tr>
<td>REPOR</td>
<td>0.89</td>
<td>0.86-0.92</td>
<td>0.06</td>
</tr>
<tr>
<td>FEPORD</td>
<td>0.88</td>
<td>0.87-0.90</td>
<td>0.03</td>
</tr>
<tr>
<td>REPORD</td>
<td>0.89</td>
<td>0.85-0.92</td>
<td>0.07</td>
</tr>
<tr>
<td>FESSPOR</td>
<td>0.88</td>
<td>0.86-0.89</td>
<td>0.03</td>
</tr>
<tr>
<td>RESSPOR</td>
<td>0.89</td>
<td>0.85-0.93</td>
<td>0.08</td>
</tr>
</tbody>
</table>

CI: Confidence Interval  
CIW: Confidence Interval width
### Proportional Odds Regression Model

#### Summary Log-odds Ratios

<table>
<thead>
<tr>
<th>Approach</th>
<th>Logor</th>
<th>CI</th>
<th>CIW</th>
</tr>
</thead>
<tbody>
<tr>
<td>FEPOR</td>
<td>2.844</td>
<td>2.728-3.039</td>
<td>0.311</td>
</tr>
<tr>
<td>REPOR</td>
<td>3.047</td>
<td>2.626-3.467</td>
<td>0.841</td>
</tr>
<tr>
<td>FEPORD</td>
<td>2.949</td>
<td>2.780-3.117</td>
<td>0.337</td>
</tr>
<tr>
<td>REPORD</td>
<td>3.026</td>
<td>2.561-3.491</td>
<td>0.930</td>
</tr>
<tr>
<td>FESSPOR</td>
<td>2.866</td>
<td>2.691-3.041</td>
<td>0.350</td>
</tr>
<tr>
<td>RESSPOR</td>
<td>3.095</td>
<td>2.552-3.638</td>
<td>1.086</td>
</tr>
</tbody>
</table>

CI: Confidence Interval
CIW: Confidence Interval width
Logit-Threshold/Bivariate Meta-Regression Model

This consists of:

1. Study-specific Logit-Threshold Linear Regression (Moses-Shapiro-Littenberg)
2. Bivariate Mixed Modeling Of Study-Specific Intercepts And Slopes
3. Parametric Estimation Of Summary ROC And Indices Using Mean Intercept And Slope Estimates
Logit-Threshold/Bivariate Meta-Regression
Study-specific Logit-Threshold Linear Regression

For the $jth$ threshold of the $ith$ study,

\[ D_{ij} = \alpha_i + \beta_i S_{ij} \text{ where:} \]

\[ D_{ij} = \logit(TPR_{ij}) - \logit(FPR_{ij}) \]

\[ S_{ij} = \logit(TPR_{ij}) + \logit(FPR_{ij}) \]

TPR = True Positive Rate; FPR = False Positive Rate

\[ \alpha_i = \text{Study-specific Intercept} \]

\[ \beta_i = \text{Study-specific Slope} \]

\[ \alpha_i \text{ and } \beta_i \text{ estimated by maximum likelihood} \]
Logit-Threshold/Bivariate Meta-Regression

Bivariate Meta-Regression: Within-study Variability

\[
\begin{pmatrix}
\alpha_i \\
\beta_i
\end{pmatrix}
\sim \mathcal{N}
\left(
\begin{pmatrix}
\mu_{\alpha i} \\
\mu_{\beta i}
\end{pmatrix}, \Sigma_W
\right)
\]

\[
\Sigma_W =
\begin{pmatrix}
\sigma^2_{\alpha i} & \rho_i \sigma_{\alpha i} \sigma_{\beta i} \\
\rho_i \sigma_{\alpha i} \sigma_{\beta i} & \sigma^2_{\beta i}
\end{pmatrix}
\]

- \(\alpha_i\) and \(\beta_i\): Estimated intercept and slope estimates of the \(i\)th study
- \(\mu_{\alpha i}\) and \(\mu_{\beta i}\): True intercept and slope estimates of the \(i\)th study
- \(\Sigma_W\): Within-study correlation \((\rho_i)\) variances \((\sigma^2_{\alpha i}\) and \(\sigma^2_{\beta i}\)) and covariance \((\rho_i \sigma_{\alpha i} \sigma_{\beta i})\) matrix
Logit-Threshold/Bivariate Meta-Regression

Bivariate Meta-Regression: Between-study Variability

\[
\begin{pmatrix}
\mu_{\alpha i} \\
\mu_{\beta i}
\end{pmatrix}
\sim
\mathcal{N}
\begin{pmatrix}
\begin{pmatrix}
\mu_{\alpha} \\
\mu_{\beta}
\end{pmatrix}
, \Sigma_B
\end{pmatrix}
\]

\[
\Sigma_B = 
\begin{pmatrix}
\tau_{\alpha}^2 & \kappa \tau_{\alpha} \tau_{\beta} \\
\kappa \tau_{\alpha} \tau_{\beta} & \tau_{\beta}^2
\end{pmatrix}
\]

\(\mu_{ai}\) and \(\mu_{bi}\) True intercept and slope estimates of the \(i\)th study

\(\mu_{a}\) and \(\mu_{b}\) Overall intercept and slope estimates

\(\Sigma_B\) Between-study correlation (\(\kappa\)) variances (\(\tau_{\alpha}^2\) and \(\tau_{\beta}^2\)) and covariance (\(\kappa \tau_{\alpha} \tau_{\beta}\)) matrix
### Logit-Threshold/Bivariate Meta-Regression

**Example data: CAGE**

<table>
<thead>
<tr>
<th>Author</th>
<th>Thresh</th>
<th>TPR</th>
<th>FPR</th>
<th>Author</th>
<th>Thresh</th>
<th>TPR</th>
<th>FPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saitz</td>
<td>1</td>
<td>0.92</td>
<td>0.27</td>
<td>Buchsbaum</td>
<td>1</td>
<td>0.89</td>
<td>0.19</td>
</tr>
<tr>
<td>Saitz</td>
<td>2</td>
<td>0.80</td>
<td>0.07</td>
<td>Buchsbaum</td>
<td>2</td>
<td>0.73</td>
<td>0.09</td>
</tr>
<tr>
<td>Saitz</td>
<td>3</td>
<td>0.55</td>
<td>0.02</td>
<td>Buchsbaum</td>
<td>3</td>
<td>0.44</td>
<td>0.02</td>
</tr>
<tr>
<td>Saitz</td>
<td>4</td>
<td>0.27</td>
<td>0.01</td>
<td>Buchsbaum</td>
<td>4</td>
<td>0.19</td>
<td>0.01</td>
</tr>
<tr>
<td>McQuade</td>
<td>1</td>
<td>0.87</td>
<td>0.20</td>
<td>Joseph</td>
<td>1</td>
<td>0.98</td>
<td>0.25</td>
</tr>
<tr>
<td>McQuade</td>
<td>2</td>
<td>0.66</td>
<td>0.08</td>
<td>Joseph</td>
<td>2</td>
<td>0.82</td>
<td>0.10</td>
</tr>
<tr>
<td>McQuade</td>
<td>3</td>
<td>0.43</td>
<td>0.01</td>
<td>Joseph</td>
<td>3</td>
<td>0.53</td>
<td>0.03</td>
</tr>
<tr>
<td>McQuade</td>
<td>4</td>
<td>0.19</td>
<td>0.01</td>
<td>Joseph</td>
<td>4</td>
<td>0.40</td>
<td>0.01</td>
</tr>
<tr>
<td>Brown</td>
<td>1</td>
<td>0.79</td>
<td>0.23</td>
<td>Bradley</td>
<td>1</td>
<td>0.71</td>
<td>0.41</td>
</tr>
<tr>
<td>Brown</td>
<td>2</td>
<td>0.70</td>
<td>0.15</td>
<td>Bradley</td>
<td>2</td>
<td>0.53</td>
<td>0.13</td>
</tr>
<tr>
<td>Brown</td>
<td>3</td>
<td>0.52</td>
<td>0.05</td>
<td>Bradley</td>
<td>3</td>
<td>0.27</td>
<td>0.02</td>
</tr>
<tr>
<td>Brown</td>
<td>4</td>
<td>0.27</td>
<td>0.02</td>
<td>Bradley</td>
<td>4</td>
<td>0.09</td>
<td>0.01</td>
</tr>
<tr>
<td>Chan</td>
<td>1</td>
<td>0.96</td>
<td>0.32</td>
<td>Jones</td>
<td>1</td>
<td>0.88</td>
<td>0.12</td>
</tr>
<tr>
<td>Chan</td>
<td>2</td>
<td>0.87</td>
<td>0.16</td>
<td>Jones</td>
<td>2</td>
<td>0.48</td>
<td>0.01</td>
</tr>
<tr>
<td>Chan</td>
<td>3</td>
<td>0.56</td>
<td>0.04</td>
<td>Jones</td>
<td>3</td>
<td>0.24</td>
<td>0.01</td>
</tr>
<tr>
<td>Chan</td>
<td>4</td>
<td>0.34</td>
<td>0.01</td>
<td>Jones</td>
<td>4</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>Aertgeerts</td>
<td>1</td>
<td>0.61</td>
<td>0.13</td>
<td>Indran</td>
<td>1</td>
<td>0.99</td>
<td>0.63</td>
</tr>
<tr>
<td>Aertgeerts</td>
<td>2</td>
<td>0.46</td>
<td>0.05</td>
<td>Indran</td>
<td>2</td>
<td>0.92</td>
<td>0.38</td>
</tr>
<tr>
<td>Aertgeerts</td>
<td>3</td>
<td>0.24</td>
<td>0.02</td>
<td>Indran</td>
<td>3</td>
<td>0.46</td>
<td>0.12</td>
</tr>
<tr>
<td>Aertgeerts</td>
<td>4</td>
<td>0.11</td>
<td>0.01</td>
<td>Indran</td>
<td>4</td>
<td>0.10</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Logit-Threshold/Bivariate Meta-Regression
Study-specific Linear Regression Intercepts and Slopes

<table>
<thead>
<tr>
<th>Author</th>
<th>$\alpha$</th>
<th>SE($\alpha$)</th>
<th>$\beta$</th>
<th>SE($\beta$)</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aertgeerts</td>
<td>2.498</td>
<td>0.277</td>
<td>-0.024</td>
<td>0.061</td>
<td>0.900</td>
</tr>
<tr>
<td>Bradley</td>
<td>1.587</td>
<td>0.391</td>
<td>-0.162</td>
<td>0.090</td>
<td>0.753</td>
</tr>
<tr>
<td>Brown</td>
<td>2.571</td>
<td>0.126</td>
<td>-0.088</td>
<td>0.044</td>
<td>0.743</td>
</tr>
<tr>
<td>Buchsbaum</td>
<td>3.498</td>
<td>0.185</td>
<td>0.032</td>
<td>0.049</td>
<td>0.727</td>
</tr>
<tr>
<td>Chan</td>
<td>3.718</td>
<td>0.177</td>
<td>0.006</td>
<td>0.054</td>
<td>0.425</td>
</tr>
<tr>
<td>Indran</td>
<td>2.874</td>
<td>0.363</td>
<td>0.144</td>
<td>0.081</td>
<td>0.104</td>
</tr>
<tr>
<td>Jones</td>
<td>4.372</td>
<td>0.966</td>
<td>0.194</td>
<td>0.189</td>
<td>0.854</td>
</tr>
<tr>
<td>Joseph</td>
<td>4.308</td>
<td>0.337</td>
<td>0.119</td>
<td>0.101</td>
<td>0.468</td>
</tr>
<tr>
<td>McQuade</td>
<td>3.270</td>
<td>0.512</td>
<td>-0.063</td>
<td>0.128</td>
<td>0.763</td>
</tr>
<tr>
<td>Saitz</td>
<td>3.702</td>
<td>0.235</td>
<td>-0.033</td>
<td>0.067</td>
<td>0.649</td>
</tr>
</tbody>
</table>
Logit-Threshold/Bivariate Meta-Regression
Mean Intercepts and Slopes by Bivariate Mixed Modeling

<table>
<thead>
<tr>
<th>Method</th>
<th>$\alpha$</th>
<th>$\text{Se}(\alpha)$</th>
<th>$\beta$</th>
<th>$\text{Se}(\beta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>fixed</td>
<td>3.098</td>
<td>0.072</td>
<td>0.019</td>
<td>0.020</td>
</tr>
<tr>
<td>reml</td>
<td>3.199</td>
<td>0.252</td>
<td>-0.006</td>
<td>0.027</td>
</tr>
<tr>
<td>ml</td>
<td>3.198</td>
<td>0.239</td>
<td>-0.006</td>
<td>0.026</td>
</tr>
<tr>
<td>mm</td>
<td>3.199</td>
<td>0.237</td>
<td>-0.005</td>
<td>0.027</td>
</tr>
</tbody>
</table>

**REML**: Restricted maximum likelihood

**ML**: Full maximum likelihood

**MM**: Method of moments

**Intercept($\alpha$)**: Average accuracy/discriminatory power of test

**Slope($\beta$)**: Measures symmetry of ROC Curve
Logit-Threshold/Bivariate Meta-Regression
Summary ROC Curve
Bayesian Hierarchical Ordinal Regression Model

Conceptual Framework

1. Random-effects formulation of meta-analysis of studies with an unequal number of nonnested categories

2. Employs a hierarchical ordinal regression model, accounting for heterogeneity of studies within-study correlation

3. Assumes that each study estimates a study-specific ROC curve that can be viewed as a random sample from a population of all ROC curves of such studies

4. Accounts for different sources of variation in the data, through study-specific location and scale parameters

5. There are several ways to construct summary ROC curves and their credible bands
Bayesian Hierarchical Ordinal Regression Model

Model Specification

**Level I (Within study variability)**

\[
M_{ik} \mid D_{ik}, \alpha_k, \beta_k \sim \begin{cases} 
\mathcal{N}(0,1), & \text{if } D_{ik} = 0 \\
\mathcal{N}(\beta_k, \exp(2\alpha_k)), & \text{if } D_{ik} = 1 
\end{cases}
\]

\[
Y_{ik} = j \text{ when } \theta_{j-1,k} \leq M_{ik} < \theta_{j,k}
\]

**Level II (Between study variability)**

\[
\alpha_k \sim \mathcal{N}(\Gamma'V_k, \sigma^2_{\alpha})
\]

\[
\beta_k \sim \mathcal{N}(\Lambda'W_k, \sigma^2_{\beta})
\]

\[
\theta_{0,k} \sim \mathcal{N}(0,100), \quad \theta_{j,k} = \sum_{i=0}^{j-1} \theta_{i,k} + \text{Exp}(1), \quad \text{for } j > 0:
\]

**Level III (Hyperpriors)**

\[
\Gamma_{l_1}, \Lambda_{l_2} \sim \mathcal{N}(0,10^6), \quad \sigma^2_{\alpha}, \sigma^2_{\beta} \sim \mathcal{IG}(0.001, 0.001)
\]
Bayesian Hierarchical Ordinal Regression Model

Specification

1. The model explicitly uses latent variables $M$ that give rise to the data $Y$ via a discretization process depending on thresholds $\theta$.

2. $D_{ik}$ indicate the true disease status of the patient $i$ in study $k$ with $D_{ik} = 1$ if disease is present and $D_{ik} = 0$ if not.

3. $\beta_k$ is the location parameter and $\alpha_k$ the scale parameter for the ROC curve of study $k$.

4. $V_k$ and $W_k$ are study-level covariate vectors of dimensions $v1$ and $v2$, respectively.
Bayesian Hierarchical Ordinal Regression Model
Parameter Estimation

1. Markov Chain Monte Carlo Simulation using Gibbs Sampling

2. Estimation via posterior means and medians

3. Every simulated pair \((\beta_k, \alpha_k)\) defines an ROC curve

4. The sensitivity of the posterior estimates to choice of priors may be examined using several different priors for the variances of study location and scale parameters
Bayesian Hierarchical Ordinal Regression Model
Summary ROCs, Functionals and Variability

1. Summary ROC Curves
   1. Mean SROC
   2. Pointwise SROC
   3. Loess SROC
   4. Mean Qstar and AUROC

2. Variability
   1. Envelope Bands for ROC Curves
   2. Pointwise Bands for ROC Curves
   3. Credible intervals for TPR at fixed FPR
Bayesian Hierarchical Ordinal Regression Model
Methodology and Application


1. They meta-analyzed 20 out of 27 eligible studies, published from 1980 to 1996.

2. Among the selected studies, seven had 2 categories, four had 4, eight had 5, and one had 7.

3. Thirteen of the studies were prospective and 7 retrospective.
Proposed Algorithm for Meta-analysis of Ordinal Data

Multi-stage SROC Modeling Algorithm

This consists of:

1. Estimation Of Study-Specific ROC Parameters From Observed 2 By J Data By Heteroskedastic Ordinal Regression

2. Estimation Of Mean Location And Scale From Study-Specific Estimates By Bivariate Linear Mixed Modeling

3. Estimation Of Summary ROC And Indices Using Mean Location And Scale Estimates
Suppose, the test result $Y_{ik}$ for $ith$ patient from $kth$ study falls into one of $J$ categories ("ratings"). The probability of $Y_{ik}$ falling in a given category $j$ or lower may be modeled as a non-linear function using the ordinal regression equation:

$$g[Pr(Y_{ik} \leq j \mid D_{ik})] = \frac{\theta_{jk} - \alpha D_{ik}}{\exp(\beta D_{ik})}$$

$g$: Cumulative link function

$D_{ik}$: a variable indicative of disease status

$\theta_j, \ldots, \theta_{j-1}$: Cut-off values on an underlying latent scale

$\alpha$: Location parameter (measure of diagnostic accuracy)

$\beta$: Scale parameter (spread of responses across subjects)
Bivariate Random-effects Estimation of Mean parameters

Within-study Variability (Level 1) model

\[
\begin{pmatrix}
  y_{1i} \\
  y_{2i}
\end{pmatrix}
\sim
\mathcal{N}
\left(
\begin{pmatrix}
  \mu_{1i} \\
  \mu_{2i}
\end{pmatrix},
\Sigma_W
\right)
\]

\[
\Sigma_W = 
\begin{pmatrix}
  \sigma_{1i}^2 & \rho_i \sigma_{1i} \sigma_{2i} \\
  \rho_i \sigma_{1i} \sigma_{2i} & \sigma_{2i}^2
\end{pmatrix}
\]

- \( y_{1i} \) and \( y_{2i} \): Estimated location and scale effects of the \( i \)th study
- \( \mu_{1i} \) and \( \mu_{2i} \): True location and scale effect of the \( i \)th study
- \( \Sigma_W \): Within-study correlation (\( \rho_i \)) variances (\( \sigma_{1i}^2 \) and \( \sigma_{2i}^2 \)) and covariance (\( \rho_i \sigma_{1i} \sigma_{2i} \)) matrix
Bivariate Random-effects Estimation of Mean parameters

Between-study Variability (Level 2) model

\[
\begin{pmatrix}
\mu_{1i} \\
\mu_{2i}
\end{pmatrix}
\sim \mathcal{N}
\left(
\begin{pmatrix}
\mu_1 \\
\mu_2
\end{pmatrix},
\Sigma_B
\right)
\]

\[
\Sigma_B =
\begin{pmatrix}
\tau_1^2 & \kappa \tau_1 \tau_2 \\
\kappa \tau_1 \tau_2 & \tau_2^2
\end{pmatrix}
\]

\(\mu_{1i}\) and \(\mu_{2i}\) True location and scale effects of the \(i\)th study

\(\mu_1\) and \(\mu_2\) Overall location and scale effects

\(\Sigma_B\) Between-study correlation (\(\kappa\)) variances (\(\tau_1^2\) and \(\tau_2^2\)) and covariance (\(\kappa \tau_1 \tau_2\)) matrix
Proposed Algorithm for Meta-analysis of Ordinal Data

Bivariate Random-effects Estimation of Mean parameters

Estimation Methods

1. Maximum Likelihood (ML)
2. Restricted Maximum Likelihood (REML)
3. DerSimonian and Laird Method Of Moments (MM)
Proposed Algorithm for Meta-analysis of Ordinal Data

Estimation of Summary ROC and Functionals
Binormal ROC Analysis

1. TPR = a + bΦ(FPR) (0 ≤ FPR ≤ 1)

2. a = meta-analytic location parameter

3. b = meta-analytic scale parameter

4. AUROC = Area under curve = Φ\left(\frac{a}{\sqrt{1+b^2}}\right)

5. Sym = Symmetry point index = Φ\left(\frac{a}{1+b}\right)
Proposed Algorithm for Meta-analysis of Ordinal Data

Estimation of Summary ROC and Functionals

Bilogistic ROC Analysis

1. \( TPR = \text{invlogit}(a + b \times \text{logit}(FPR)) \) (\( 0 \leq FPR \leq 1 \))

2. \( a \) = meta-analytic location parameter

3. \( b \) = meta-analytic scale parameter

4. Area under curve (AUROC) and Symmetry point index (Sym) derived from integration of \( TPR = \text{invlogit}(a + b \times \text{logit}(FPR)) \)
### Table: Study-specific Estimates by Ordinal Probit

<table>
<thead>
<tr>
<th>Study</th>
<th>Location (Se)</th>
<th>Scale (Se)</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aertgeerts</td>
<td>1.37 (0.19)</td>
<td>0.96 (0.11)</td>
<td>-0.14</td>
</tr>
<tr>
<td>Bradley</td>
<td>0.74 (0.16)</td>
<td>0.67 (0.11)</td>
<td>0.22</td>
</tr>
<tr>
<td>Brown</td>
<td>1.53 (0.38)</td>
<td>0.95 (0.28)</td>
<td>0.54</td>
</tr>
<tr>
<td>Buchsbaum</td>
<td>2.21 (0.20)</td>
<td>1.14 (0.13)</td>
<td>0.63</td>
</tr>
<tr>
<td>Chan</td>
<td>2.17 (0.44)</td>
<td>1.04 (0.31)</td>
<td>0.73</td>
</tr>
<tr>
<td>Indran</td>
<td>1.79 (0.28)</td>
<td>1.47 (0.20)</td>
<td>0.20</td>
</tr>
<tr>
<td>Jones</td>
<td>2.22 (0.64)</td>
<td>0.92 (0.34)</td>
<td>0.66</td>
</tr>
<tr>
<td>Joseph</td>
<td>2.92 (0.65)</td>
<td>1.44 (0.44)</td>
<td>0.72</td>
</tr>
<tr>
<td>McQuade</td>
<td>1.73 (0.33)</td>
<td>0.83 (0.19)</td>
<td>0.55</td>
</tr>
<tr>
<td>Saitz</td>
<td>2.16 (0.34)</td>
<td>0.99 (0.21)</td>
<td>0.68</td>
</tr>
</tbody>
</table>
Example Dataset 1
Similar Thresholds
Example Dataset 1
Similar Thresholds

Table: Summary performance indices by estimation method

<table>
<thead>
<tr>
<th>Method</th>
<th>Location</th>
<th>Scale</th>
<th>Area</th>
<th>Sympoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>reml</td>
<td>1.82 (1.43-2.20)</td>
<td>1.00 (0.85-1.16)</td>
<td>0.90 (0.86-0.94)</td>
<td>0.82 (0.78-0.86)</td>
</tr>
<tr>
<td>ml</td>
<td>1.81 (1.44-2.17)</td>
<td>1.00 (0.85-1.15)</td>
<td>0.90 (0.86-0.94)</td>
<td>0.82 (0.78-0.85)</td>
</tr>
<tr>
<td>mm</td>
<td>1.83 (1.41-2.25)</td>
<td>1.01 (0.85-1.16)</td>
<td>0.90 (0.86-0.95)</td>
<td>0.82 (0.78-0.86)</td>
</tr>
</tbody>
</table>

REML: Restricted maximum likelihood
ML: Full maximum likelihood
MM: Method of moments
Location: Measure of accuracy/discriminatory power of test
Scale: Measures symmetry of ROC curve
Sympoint: Symmetry point (sensitivity = specificity)
Example Dataset 1
Similar Thresholds

Table: Estimated between-studies SDs and correlation

<table>
<thead>
<tr>
<th>Method</th>
<th>SD(Location)</th>
<th>SD(Scale)</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>REML</td>
<td>0.510</td>
<td>0.151</td>
<td>1.00</td>
</tr>
<tr>
<td>ML</td>
<td>0.473</td>
<td>0.140</td>
<td>1.00</td>
</tr>
<tr>
<td>MM</td>
<td>0.583</td>
<td>0.166</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Example Dataset 1
Similar Thresholds: Using summary data from REML

![Summary ROC Curve](image)
# Example Dataset 2

Disparate Thresholds

<table>
<thead>
<tr>
<th>Study</th>
<th>Cutpoints</th>
<th>Location (Se)</th>
<th>Scale (Se)</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Almeda</td>
<td>4</td>
<td>1.32 (0.18)</td>
<td>1.09 (0.19)</td>
<td>0.72</td>
</tr>
<tr>
<td>Arad</td>
<td>4</td>
<td>1.32 (0.14)</td>
<td>0.96 (0.11)</td>
<td>0.42</td>
</tr>
<tr>
<td>Bielak</td>
<td>6</td>
<td>1.86 (0.23)</td>
<td>1.02 (0.18)</td>
<td>0.70</td>
</tr>
<tr>
<td>Budoff</td>
<td>7</td>
<td>1.24 (0.06)</td>
<td>1.44 (0.08)</td>
<td>0.54</td>
</tr>
<tr>
<td>Chen</td>
<td>4</td>
<td>2.17 (0.38)</td>
<td>1.15 (0.31)</td>
<td>0.77</td>
</tr>
<tr>
<td>Greenland</td>
<td>4</td>
<td>0.54 (0.13)</td>
<td>0.92 (0.13)</td>
<td>0.27</td>
</tr>
<tr>
<td>Hosoi</td>
<td>5</td>
<td>1.34 (0.16)</td>
<td>0.93 (0.14)</td>
<td>0.67</td>
</tr>
<tr>
<td>Knez</td>
<td>4</td>
<td>1.90 (0.09)</td>
<td>1.35 (0.11)</td>
<td>0.82</td>
</tr>
<tr>
<td>LaMonte</td>
<td>4</td>
<td>1.75 (0.13)</td>
<td>1.27 (0.12)</td>
<td>0.68</td>
</tr>
<tr>
<td>Nixdorff</td>
<td>2</td>
<td>0.72 (4.01)</td>
<td>0.20 (3.99)</td>
<td>1.00</td>
</tr>
<tr>
<td>Raggi</td>
<td>4</td>
<td>1.61 (0.37)</td>
<td>1.62 (0.32)</td>
<td>0.08</td>
</tr>
<tr>
<td>Schepis</td>
<td>5</td>
<td>1.54 (0.36)</td>
<td>1.19 (0.31)</td>
<td>0.57</td>
</tr>
<tr>
<td>Seese</td>
<td>2</td>
<td>5.61 (300.45)</td>
<td>3.06 (234.45)</td>
<td>1.00</td>
</tr>
<tr>
<td>Shaw</td>
<td>5</td>
<td>0.87 (0.09)</td>
<td>0.97 (0.07)</td>
<td>-0.06</td>
</tr>
<tr>
<td>Taylor</td>
<td>4</td>
<td>0.42 (0.50)</td>
<td>0.53 (0.35)</td>
<td>-0.33</td>
</tr>
<tr>
<td>Vliengenthart</td>
<td>4</td>
<td>1.10 (0.22)</td>
<td>1.21 (0.20)</td>
<td>0.20</td>
</tr>
<tr>
<td>Wong</td>
<td>5</td>
<td>1.00 (0.27)</td>
<td>1.12 (0.26)</td>
<td>0.28</td>
</tr>
<tr>
<td>Yao</td>
<td>2</td>
<td>3.22 (78.35)</td>
<td>3.14 (97.37)</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Example Dataset 2
Disparate Thresholds

Study-specific ROC Curves

B.A. Dwamena (UofM-VAMC)
## Example Dataset 2

Disparate Thresholds

**Table:** Summary performance indices by estimation method

<table>
<thead>
<tr>
<th>Method</th>
<th>Location</th>
<th>Scale</th>
<th>Area</th>
<th>Sympoint</th>
</tr>
</thead>
<tbody>
<tr>
<td>reml</td>
<td>1.36 (1.12-1.60)</td>
<td>1.11 (0.98-1.24)</td>
<td>0.83 (0.79-0.87)</td>
<td>0.74 (0.71-0.77)</td>
</tr>
<tr>
<td>ml</td>
<td>1.36 (1.13-1.60)</td>
<td>1.11 (0.98-1.23)</td>
<td>0.83 (0.79-0.86)</td>
<td>0.74 (0.71-0.77)</td>
</tr>
<tr>
<td>mm</td>
<td>1.36 (1.13-1.59)</td>
<td>1.11 (0.99-1.23)</td>
<td>0.83 (0.79-0.86)</td>
<td>0.74 (0.71-0.77)</td>
</tr>
</tbody>
</table>

**REML:** Restricted maximum likelihood  
**ML:** Full maximum likelihood  
**MM:** Method of moments  
**Location:** Measure of accuracy/discriminatory power of test  
**Scale:** Measures symmetry of ROC curve  
**Sympoint:** Symmetry point (sensitivity = specificity)
Example Dataset 2
Disparate Thresholds

Table: Estimated between-studies SDs and correlation

<table>
<thead>
<tr>
<th>Method</th>
<th>SD(Location)</th>
<th>SD(Scale)</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>REML</td>
<td>0.441</td>
<td>0.183</td>
<td>0.563</td>
</tr>
<tr>
<td>ML</td>
<td>0.423</td>
<td>0.74</td>
<td>0.563</td>
</tr>
<tr>
<td>MM</td>
<td>0.420</td>
<td>0.174</td>
<td>0.562</td>
</tr>
</tbody>
</table>
Example Dataset 2
Disparate Thresholds: Using summary results from REML
Concluding Remarks

Conclusions

1. Dichotomization of ordinal data is simple with abundance of meta-analytical methods and software programs but inefficient with loss of information.

2. The "no thresholds left behind" proposed algorithm is very robust, flexible, informative and efficient.

3. It is invariant to the number/set of thresholds, link function or estimation procedure.
Concluding Remarks

Conclusions

1. Easily extended for covariate meta-regression and covariate-adjusted SROC analysis
2. Easily implemented in Stata using Stata-native and User-written commands
3. `midacat` module for automated implementation will be available shortly
4. Datasets, do-files and unpublished ado-files available from author on request
Aertgeerts B., Buntinx F., and Kester A.
The value of the CAGE in screening for alcohol abuse and alcohol dependence in general clinical populations: a diagnostic meta-analysis.
J clin Epidemiol 2004;57:30-39

Bivariate Random Effects Meta-Analysis of ROC Curves.
Med Decis Making 2008;28:621-628

Begg C.B. and Mazumdar M.
Operating characteristics of a rank correlation test for publication bias.
Biometrics 1994;50:1088-1101

Chu H. and Cole S.R.
Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach.
J Clin Epidemiol 2006;59:1331-1332

Dendukuri N., Chui K. and Brophy J.M.
Validity of EBCT for coronary artery disease: a systematic review and meta-analysis.
BMC Medicine 2007;5:35
Dukic V. and Gatsonis C.
Meta-analysis of diagnostic test accuracy studies with varying number of thresholds.
Biometrics 2003;59:936-946

Dwamena, B.
midas: Module for Meta-Analytical Integration of Diagnostic Accuracy Studies
Boston College Department of Economics, Statistical Software Components 2007;
s456880: http://ideas.repec.org/c/boc/bocode/s456880.html.

Ewing J.A.
Detecting Alcoholism: The CAGE questionnaire.
JAMA 1984;252:1905-1907

Unification of models for meta-analysis of diagnostic accuracy studies.
Biostatistics 2007;8:239-251

An empirical comparison of methods for meta-analysis of diagnostic accuracy showed
hierarchical models are necessary
Journal of Clinical Epidemiology 2008;61;1095-1103
Harbord R.M., and Whitting P.  
metandi: Meta-analysis of diagnostic accuracy using hierarchical logistic regression  
Stata Journal 2009;2:211-229

Irwig L., Macaskill P., Glasziou P. and Fahey M.  
Meta-analytic methods for diagnostic test accuracy.  

Kester A.D.M., and Buntinx F.  
Meta-Analysis of ROC Curves.  

Littenberg B. and Moses L. E.  

Macaskill P.  
Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis.  
J Clin Epidemiol 2004;57:925-932
Moses L.E., Shapiro D. and Littenberg B.
Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations.

Pepe M.S.
Receiver Operating Characteristic Methodology.

Pepe M.S.
The Statistical Evaluation of Medical Tests for Classification and Prediction.
2003; Oxford: Oxford University Press

Reitsma J.B., Glas A.S., Rutjes A.W.S., Scholten R.J.P.M., Bossuyt P.M. and Zwinderman A.H.
Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews.
J Clin Epidemiol 2005;58:982-990

Rutter C.M., and Gatsonis C.A.
A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations
Stat Med 2001;20:2865-2884
Toledano A. and Gatsonis C.A.
Regression analysis of correlated receiver operating characteristic data.
Academic Radiology 1995;2:S30-S36

Tosteson A.A. and Begg C.B.
A general regression methodology for ROC curve estimation.
Medical Decision Making 1988;8:204-215

Williams R.
Using Heterogeneous Choice Models To Compare Logit and Probit Coefficients Across Groups
Sociological Methods and Research 2009;37: 531-559

White I.R.
Multivariate Random-effects Meta-analysis.
Stata Journal 2009;1:40-56