Statistical Competencies for Clinical & Translational Science: Evolution in Action

Felicity Enders, PhD

Association of Clinical and Translational Statisticians Annual Meeting
Boston, MA, August 2, 2014
Introduction

• CTS is a relatively new discipline
  • Competencies in flux
• Competencies serve to define the discipline
  • Program evaluation
  • Learner evaluation
    • Self-assessment
    • Validated instruments
• Provide guidance regarding training
  • Statistics is too big to teach everything
Overview

• Refining the competencies
• Assessing the competencies
• Assessing CTS learners
• Consensus on the competencies
Refining the Competencies

Who Are CTS Learners?

- CTS learners are typically not statisticians
  - Similar to Public Health students
  - Graduate level
- Variety of learner goals
  - Read the literature
  - Co-I; co-author
  - PI; first or last author
- Variety of learning modalities
  - Traditional graduate level coursework
  - CME or other on-demand access
CTS Statistics Competencies (2009)

Describe the basic principles and practical importance of random variation, systematic error, sampling error, measurement error, hypothesis testing, type I and type II errors, and confidence limits.

Compute sample size, power, and precision for comparisons of two independent samples with respect to continuous and binary outcomes.

Explain the uses, importance, and limitations of early stopping rules in clinical trials.

Scrutinize the assumptions behind different statistical methods and their corresponding limitations.

Generate simple descriptive and inferential statistics that fit the study design chosen and answer research question.

Describe the uses of meta-analytic methods.

Statistics-Related Competencies

Sources of Error

• Describe the concepts and implications of reliability and validity of study measurements
• Evaluate the reliability and validity of measures
  - Assess sources of bias and variation in published studies.
  - Assess threats to study validity (bias) including problems with sampling, recruitment, randomization, and comparability of study groups

Scientific Communication

• Communicate clinical and translational research findings to difference groups of individuals, including colleagues, students, the lay public, and the media
• Write summaries of scientific information for use in the development of clinical health care policy

Study Design

• Propose study designs for addressing a clinical or translational research question

Other Sources of Competencies


Why Public Health?

• Public Health graduate students are nonstatisticians in a medical field
• They will use statistics to read and publish in the medical research literature
• Competencies were compared, condensed, and extended in Enders (2011)
## Competency Comparison

<table>
<thead>
<tr>
<th>Clinical and Translational Science</th>
<th>Public Health</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assess sources of <strong>bias</strong> and variation in published studies</td>
<td></td>
<td>CONSORT TREND STROBE</td>
</tr>
<tr>
<td>Assess threats to study validity (bias) including problems with sampling, recruitment, randomization, and comparability of study groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propose <strong>study designs</strong> for addressing a clinical or translational research question</td>
<td></td>
<td>STROBE</td>
</tr>
<tr>
<td>Describe the <strong>basic principles</strong> and practical importance of random variation, systematic error, sampling error, measurement error, hypothesis testing, type I and type II errors, and confidence limits</td>
<td>Describe the <strong>basic concepts</strong> of probability, random variation, and commonly used statistical probability distributions</td>
<td>CONSORT TREND STROBE</td>
</tr>
<tr>
<td>Compute <strong>sample size, power, and precision</strong> for comparisons of two independent samples with respect to continuous and binary outcomes</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Competency Comparison

<table>
<thead>
<tr>
<th>Clinical and Translational Science</th>
<th>Public Health</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explain the uses, importance, and limitations of early stopping rules in clinical trials</td>
<td></td>
<td>Consort TREND</td>
</tr>
<tr>
<td>Describe the concepts and implications of reliability and validity of study measurements</td>
<td>Calculate basic epidemiologic measures</td>
<td>TREND</td>
</tr>
<tr>
<td>Evaluate the reliability and validity of measures</td>
<td>Draw appropriate inferences from epidemiologic data</td>
<td></td>
</tr>
<tr>
<td>Scrutinize the assumptions behind different statistical methods and their corresponding limitations</td>
<td>Describe preferred methodologic alternatives to commonly used statistical methods when assumptions are not met</td>
<td></td>
</tr>
</tbody>
</table>
## Competency Comparison

<table>
<thead>
<tr>
<th>Clinical and Translational Science</th>
<th>Public Health</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distinguish among the different measurement scales and the implications for selection of statistical methods to be used on the basis of these distinctions</td>
<td>Generate simple descriptive and inferential statistics that fit the study design chosen and answer research question</td>
<td>Apply descriptive techniques commonly used to summarize public health data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apply common statistical methods for inference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apply descriptive and inferential methodologies according to the type of study design for answering a particular research question</td>
</tr>
<tr>
<td></td>
<td>Describe the uses of meta-analytic methods</td>
<td>CONSORT TRENDS STROBE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Competency Comparison

<table>
<thead>
<tr>
<th>Clinical and Translational Research</th>
<th>Public Health</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communicate clinical and translational research findings to different groups of individuals, including colleagues, students, the lay public, and the media</td>
<td>Interpret results of statistical analyses found in public health studies</td>
<td>CONSORT TREND STROBE</td>
</tr>
<tr>
<td>Write summaries of scientific information for use in the development of clinical health care policy</td>
<td>Develop written and oral presentations on the basis of statistical analyses for both public health professionals and educated lay audiences</td>
<td></td>
</tr>
</tbody>
</table>
## Competencies Added from Guidelines

<table>
<thead>
<tr>
<th>Competency</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Describe size of the effect</strong> with a measure of precision</td>
<td>CONSORT TRENDS STROBE</td>
</tr>
<tr>
<td><strong>Describe the study sample</strong>, including sampling methods, the amount and type of missing data, and the implications for generalizability</td>
<td>TRENDS STROBE</td>
</tr>
<tr>
<td><strong>Interpret results in light of multiple comparisons</strong></td>
<td>CONSORT TRENDS STROBE</td>
</tr>
<tr>
<td><strong>Identify inferential methods appropriate for</strong> clustered, matched, paired, or longitudinal studies</td>
<td>TRENDS STROBE</td>
</tr>
<tr>
<td><strong>Describe</strong> <strong>adjusted</strong> inferential methods appropriate for the study design, including examination of interaction</td>
<td>CONSORT TRENDS STROBE</td>
</tr>
<tr>
<td>Describe statistical methods appropriate to address loss to followup</td>
<td>STROBE</td>
</tr>
</tbody>
</table>
Assessing the Competencies

BERD Education Working Group

• Goal: assess need for training in each competency

• Methods: 18 BERD members surveyed for training needed (high/some/none) by trainee type

• Trainees categorized by background (no research experience/reader/research experience) and intended role (reader/co-I/PI)

• Asked for missing competencies
## Competencies Added by Reviewers

<table>
<thead>
<tr>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Describe statistical methods appropriate to address loss to follow-up</td>
</tr>
<tr>
<td>Understand the reasons for performing research that is reproducible from data collection through publication of results</td>
</tr>
<tr>
<td>Understand appropriate methods for data presentation, especially effective statistical graphs and tables</td>
</tr>
<tr>
<td>Characterization of diagnostic testing, including sensitivity, specificity, and ROC curves</td>
</tr>
<tr>
<td>Describe appropriate data quality and data management procedures</td>
</tr>
</tbody>
</table>
Division of the Competencies

- Based on predicted **high need for training for future PIs**
  - **Fundamental** $\geq 70\%$ high need for training
  - **Intermediate** 60%-69% high need for training
  - **Specialized** <60% high need for training

- Cutoffs assigned post hoc
Fundamental Competencies

Goal: PI

- Assess sources of bias and variation
- Propose study designs
- Communicate research findings
- Understand the reasons for performing reproducible research
- Describe basic statistical principles and their practical importance
- Describe concepts and implications of reliability and validity
- Describe the study sample, including sampling methods
- Describe size of the effect with a measure of precision
- Understand appropriate methods for data presentation
Fundamental Competencies
Goal: Informed Reader

- Assess sources of bias and variation
- Propose study designs
- Communicate research findings
- Understand the reasons for performing reproducible research
- Describe basic statistical principles and their practical importance
- Describe concepts and implications of reliability and validity
- Describe the study sample, including sampling methods
- Describe size of the effect with a measure of precision
- Understand appropriate methods for data presentation
Intermediate Competencies
Goal: PI

- Distinguish among different measurement scales
- Generate simple descriptive and inferential statistics
- Identify adjusted inferential methods appropriate for the study design
- Scrutinize assumptions and limitations behind different statistical methods
- Interpret results in light of multiple comparisons
- Identify inferential methods appropriate for clustered, matched, paired or longitudinal studies
Intermediate Competencies

Goal: Informed Reader

- Distinguish among different measurement scales
- Generate simple descriptive and inferential statistics
- Identify adjusted inferential methods appropriate for the study design
- Scrutinize assumptions and limitations behind different statistical methods
- Interpret results in light of multiple comparisons
- Identify inferential methods appropriate for clustered, matched, paired or longitudinal studies
Specialized Competencies

Goal: PI

1. Describe statistical methods appropriate to address loss to follow-up
2. Describe appropriate data quality and data management procedures
3. Compute sample size, power, and precision for comparisons
4. Explain the uses, importance, and limitations of early stopping rules in clinical trials
5. Characterization of diagnostic testing, including sensitivity, specificity, and ROC curves
6. Describe the uses of meta-analytic methods
Specialized Competencies

Goal: Informed Reader

- Describe statistical methods appropriate to address loss to follow-up
- Describe appropriate data quality and data management procedures
- Compute sample size, power, and precision for comparisons
- Explain the uses, importance, and limitations of early stopping rules in clinical trials
- Characterization of diagnostic testing, including sensitivity, specificity, and ROC curves
- Describe the uses of meta-analytic methods
Suggested Modifications

• Comments on the competencies suggested wording changes
  • Critical verb often proposed to change level of competency

• Example [proposed modification]
  • Propose [Assess] study designs for addressing a clinical or translational research question

• Modifications have not yet been assessed
Summary of Competency Assessment

• Strengths
  • First assessment of competencies
  • Added competencies
  • Considered learner’s background and role

• Limitations
  • N of 18
  • Cutoffs for competency types determined post hoc
  • Survey did not use terms “specialized” and “fundamental”
Assessing CTS Learners


Related Work: Assessment Tools for CTS

• In order to evaluate learner competency, validated measures will be needed

• Such instruments need to be:
  • Linked to CTS statistical competencies
  • Normed for CTS learners
  • Relatively quick to complete
Assessment Instruments: Existing tools

- Windish (2007)
  - Well validated
  - For physicians reading the literature
- delMas (2007) CAOS test
  - Well validated
  - For undergraduates after 1st stat course

- For others, see Enders (2011)
Assessment Instruments: New Tools

• Biostatistics Mastery Assessment of Proficiency (Biostatistics MAP)
  • Aligned with first course in biostatistics
  • Univariate and bivariate methods

• Global Regression Expectations in StatisticS (REGRESS)
  • Aligned with training on linear regression
  • Simple & multiple regression models
Biostatistics MAP

- Students Pre-Course
- Students Post-Course
## Biostatistics MAP

<table>
<thead>
<tr>
<th></th>
<th>Student Pre-course N=68</th>
<th>Student Post-course N=68</th>
<th>Post-course-Pre-course N=68</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Score (of 30)</strong></td>
<td>12 (1-25)</td>
<td>18 (0-29)</td>
<td>4 (-10 to 15)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Domains</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definitions (of 3)</td>
<td>1 (0-3)</td>
<td>2 (0-3)</td>
<td>0 (-3 to 2)</td>
<td>0.043</td>
</tr>
<tr>
<td>Study Design (of 3)</td>
<td>1 (0-3)</td>
<td>2 (0-3)</td>
<td>0 (-1 to 3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Choice of Method (of 7)</td>
<td>3 (0-7)</td>
<td>5 (0-7)</td>
<td>1 (-5 to 7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Application (of 6)</td>
<td>3 (0-6)</td>
<td>4 (0-6)</td>
<td>1 (-2 to 6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interpretation (of 8)</td>
<td>3 (0-7)</td>
<td>5 (0-8)</td>
<td>2 (-4 to 5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Assumptions (of 3)</td>
<td>1 (0-3)</td>
<td>1 (0-3)</td>
<td>0 (-2 to 2)</td>
<td>0.019</td>
</tr>
<tr>
<td><strong>Topics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General (of 8)</td>
<td>4 (0-8)</td>
<td>5 (0-8)</td>
<td>1 (-6 to 5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Categorical Outcome (of 8)</td>
<td>4 (0-8)</td>
<td>5 (0-8)</td>
<td>1 (-3 to 5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Continuous Outcome (of 8)</td>
<td>3 (0-7)</td>
<td>5 (0-8)</td>
<td>1.5 (-3 to 6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time to Event Outcome (of 2)</td>
<td>0 (0-2)</td>
<td>0 (0-2)</td>
<td>0 (-1 to 2)</td>
<td>0.123</td>
</tr>
<tr>
<td>Diagnostic Testing &amp; Agreement (of 4)</td>
<td>2 (0-4)</td>
<td>3 (0-4)</td>
<td>1 (-3 to 3)</td>
<td>0.010</td>
</tr>
</tbody>
</table>

©2014 MFMER | slide-32
REGRESS: Internal Validation

Total REGRESS Score

- Students Pre-Course
- Students Post-Course
- Practicing Statisticians
## REGRESS

<table>
<thead>
<tr>
<th></th>
<th>Mayo pre-course N=52</th>
<th>Mayo post-course N=59</th>
<th>p-value</th>
<th>Practicing Statisticians N=22</th>
<th>p-value vs. pre</th>
<th>p-value vs. post</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REGRESS Score (of 27)</td>
<td>9.3 (4.3)</td>
<td>19.0 (3.5)</td>
<td>&lt;0.001</td>
<td>20.1 (3.5)</td>
<td>&lt;0.001</td>
<td>0.21</td>
</tr>
<tr>
<td>SLR Score (of 11)</td>
<td>6.3 (2.5)</td>
<td>8.3 (1.6)</td>
<td>&lt;0.001</td>
<td>9.8 (1.2)</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Domains</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpreting &amp; Using SLR Equation (of 8)</td>
<td>4 (0-8)</td>
<td>6 (3-8)</td>
<td>&lt;0.001</td>
<td>7 (5-8)</td>
<td>&lt;0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Interpreting &amp; Using MLR Equation (of 4)</td>
<td>1 (0-4)</td>
<td>3 (0-4)</td>
<td>&lt;0.001</td>
<td>3 (2-4)</td>
<td>&lt;0.001</td>
<td>0.37</td>
</tr>
<tr>
<td>Modeling &amp; Statistical Significance (of 4)</td>
<td>2 (0-4)</td>
<td>3 (2-4)</td>
<td>&lt;0.001</td>
<td>3 (2-4)</td>
<td>&lt;0.001</td>
<td>0.86</td>
</tr>
<tr>
<td>Assessing Assumptions (of 4)</td>
<td>1 (0-4)</td>
<td>2 (1-4)</td>
<td>&lt;0.001</td>
<td>2 (0-4)</td>
<td>&lt;0.001</td>
<td>0.84</td>
</tr>
<tr>
<td>Confounding &amp; Colinearity (of 4)</td>
<td>0 (0-3)</td>
<td>3 (0-4)</td>
<td>&lt;0.001</td>
<td>3 (0-4)</td>
<td>&lt;0.001</td>
<td>0.53</td>
</tr>
<tr>
<td>Interaction (of 3)</td>
<td>0 (0-3)</td>
<td>2 (0-3)</td>
<td>&lt;0.001</td>
<td>2 (0-3)</td>
<td>&lt;0.001</td>
<td>0.82</td>
</tr>
</tbody>
</table>
REGRESS: Initial External Validation

<table>
<thead>
<tr>
<th>REGRESS Scores (out of 27)</th>
<th>N</th>
<th>Mean (SD)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pre-Course</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Michigan Students</td>
<td>52</td>
<td>11.3 (2.8)</td>
<td></td>
</tr>
<tr>
<td>Mayo Students</td>
<td>52</td>
<td>9.3 (4.3)</td>
<td>0.018</td>
</tr>
<tr>
<td><strong>Post-Course</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>University of Michigan Students</td>
<td>52</td>
<td>15.2 (3.1)</td>
<td></td>
</tr>
<tr>
<td>Mayo Students</td>
<td>59</td>
<td>19.0 (3.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Statisticians</td>
<td>22</td>
<td>20.1 (3.5)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
REGRESS: Initial External Validation

![Histogram showing Pre-Course and Post-Course Total REGRESS Score distributions. The graph indicates a shift in the distribution of scores post-course, with Post-Course scores being more concentrated in the higher score range.]
### REGRESS: Initial External Validation

<table>
<thead>
<tr>
<th></th>
<th>Pre-course</th>
<th>Post-Course</th>
<th>Pre vs. Post</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Summary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REGRESS (out of 27)</td>
<td>11.29 (2.84)</td>
<td>15.17 (3.05)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>SLR Score (out of 11)</td>
<td>7.37 (1.46)</td>
<td>7.92 (1.61)</td>
<td>0.0145</td>
</tr>
<tr>
<td><strong>Domain</strong></td>
<td>Median (range)</td>
<td>Median (range)</td>
<td></td>
</tr>
<tr>
<td>Interpretating and Using SLR Equation (of 8)</td>
<td>5.0 (2-7)</td>
<td>5.5 (2-8)</td>
<td>0.00554</td>
</tr>
<tr>
<td>Interpreting and using MLR Equation (of 4)</td>
<td>1.5 (0-4)</td>
<td>2.0 (0-4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Modeling and Statistical Significance (of 4)</td>
<td>2.0 (0-4)</td>
<td>3.0 (1-4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Assessing Assumptions (of 4)</td>
<td>1.0 (0-3)</td>
<td>2.0 (0-4)</td>
<td>0.0052</td>
</tr>
<tr>
<td>Confounding and Colinearity (of 4)</td>
<td>1.0 (0-3)</td>
<td>2.0 (0-4)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Interaction (of 3)</td>
<td>1.0 (0-2)</td>
<td>1.0 (0-3)</td>
<td>0.0007</td>
</tr>
</tbody>
</table>
Assessment Instruments: Current Status

• Biostatistics Mastery Assessment of Proficiency (Biostatistics MAP)
  • Internal validation done

• Global Regression Expectations in StatisticS (REGRESS)
  • Internal validation done (Enders, 2013)
  • 1st external validation done (Kidwell, 2014)
  • Currently creating next version
  • Planning a larger external validation for 2014-2015
Consensus on the Competencies
CTS Statistics Competencies (2011)

Four Competencies Added by CTSA Consortium

Describe the role that biostatistics serves in biomedical and public health research.

Defend the significance of data and safety monitoring plans.

Collaborate with biostatisticians in the design, conduct, and analyses of clinical and translational research.

Evaluate computer output containing the results of statistical procedures and graphics.

Next Steps

• Update CTS competencies for statistics
  • Modify further?
  • Get buy-in from CTSA PIs

• Assess which competencies are taught to whom
  • For which competencies would Consortium-wide training materials be helpful?

• Update instruments/items to better reflect CTS competencies
Survey of ACTStat Attendees

• Consensus among statisticians
  • Preferred wording
  • Additions
  • Exclusions
• Competency level
  • Fundamental – Specialized
ACTS Survey

• Competencies
  • Which wording do you prefer?
  • Should this competency be excluded?
  • Specialized = 1
    • Only advanced learners in some areas need to achieve these competencies
  • Fundamental = 9
    • Every CTS learner needs to achieve these competencies

• What’s missing?
  • Propose additional statistical competencies for CTS learners
Questions & Discussion