

Methods for Algorithm Comparison Assessment

by
Nancy A. Obuchowski, PhD
Cleveland Clinic Foundation

Outline

- TAVR study description
- Study objectives
- Analysis when ground truth is known
- Absence of ground truth
 - test for interchangeability

TAVR Study

- 110 patients undergoing transcatheter aortic valve replacement (TAVR)
- Pre-surgical 3D CT used to match prosthesis size to annulus size
- Radiologist manually measures annulus from CT image = **clinical standard**

TAVR Study

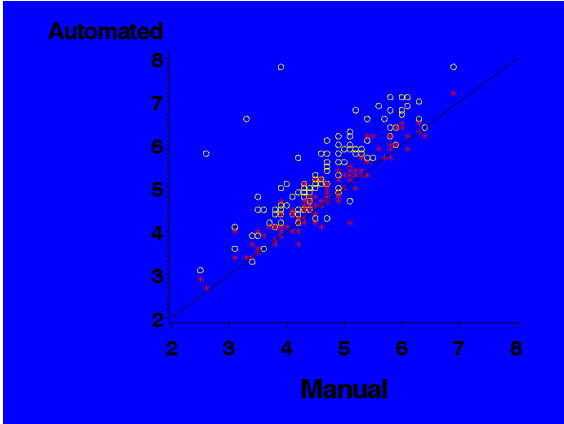
- Automated analysis – faster, potentially more reproducible than manual
- Two Automated Methods:
 - Fully-automated method
 - Semi-automated (insertion points of measurements reviewed and corrected, as needed)

Study Design for TAVR Study

	Expert Radiologist	Less experienced Radiologist
Manual measurements	2x	1x
Fully-Automated	2x	2x
Semi-automated	2x	2x

Study Goals:

1. Compare agreement in area measurements of automated methods with manual
 - Bias
 - Precision
 - Summary measure
2. Assess effect of automated methods on patient management:
 - Test for interchangeability

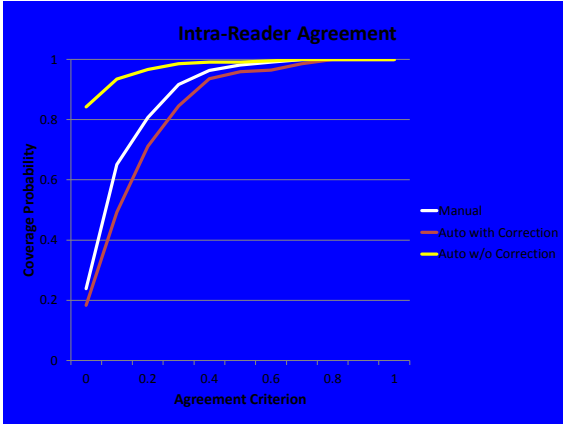


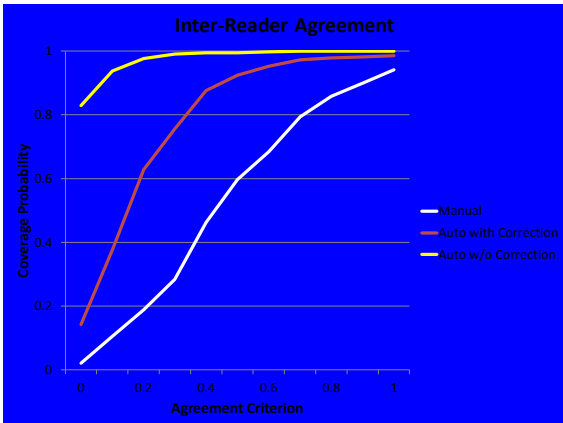
Bias

(mean difference betw Automated method and Expert's Manual measurement)

	Reader 1	Reader 2	overall
Fully-automated	0.715	0.677	0.696
Semi-automated	0.158	0.267	0.212
p-value	<0.001	<0.001	<0.001

- ### Precision
- Intra-reader variability
 - Inter-reader variability



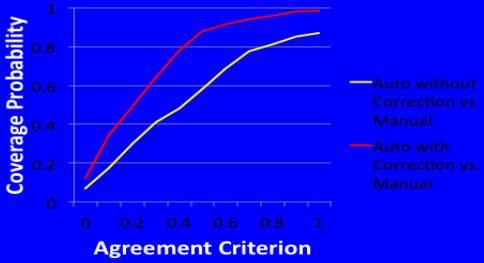


Results from Disaggregate Analysis

	Least Bias	Most Precision
Fully Automated		✓
Semi-Automated	✓	

Summary Measures (aggregate analysis)

- **ICC (intraclass correlation coefficient):** its magnitude depends on between-subject variability
- **CP (Coverage Probability):** simple proportion of cases with agreement as the definition of agreement becomes more lax



Goal #2: Effect of Automated Method on Patient Management

Three Valve Sizes:

23 valve: Annulus area 3.3-4.0

26 valve: Annulus area 4.2-5.0

29 valve: Annulus area 5.3-6.3

<3.3 or >6.3 → no replacement performed

Can clinicians switch between manual and semi-automated methods without affecting patient management?

Test for Interchangeability

Proportion of cases where Expert Reader agrees about valve size

2 Manual measurements by Expert	1 Manual and 1 Semi-Automated by Expert
0.945	0.876

Proportion of cases where Different Readers agree about valve size

2 Manual measurements by different readers	Manual measurement by one reader and Semi-Automated by another
0.594	0.824

Are Manual and Automated methods Interchangeable?

Scenario	Expected Outcome of Switching between Manual and Automated:
Expert Reader	1.7-12.2% increase in frequency of disagreement in valve size
Different Readers	14.8-31.3% decrease in frequency of disagreement in valve size

Conclusion

- Semi-automated less bias, but also less precise than fully-automated method.
- Semi-automated method agrees with manual more often.
- Semi-automated method can be used interchangeably with manual in typical clinical scenario.
