

Test Performance Characteristics and Basic Statistics

G. Scott Gazelle, MD, MPH, PhD
Institute for Technology Assessment
Massachusetts General Hospital

Jeffrey H. Blume, PhD
Dept of Biostatistics
Vanderbilt University Medical Center

Diana Miglioretti, PhD
Group Health Cooperative
Center for Health Studies

Nancy Obuchowski, PhD
Dept of Biostatistics
Cleveland Clinic Foundation

Key Concepts

	Diagnostic Test Accuracy	Statistical Concepts
Basic -1	Sensitivity, specificity, PPV, NPV	Population, sampling, null and alternative hypotheses, type I and II errors
Basic -2	ROC curves and summary indices derived from them	p-values, confidence intervals, correlation, agreement
Intermediate -1	Study design, MRMC studies, blinding, randomization, biases	Study designs, RCTs, paired/unpaired designs, biases
Intermediate -2	Chi square test, McNemar's test, reader agreement	Chi square test, t-tests, paired t-tests, Wilcoxon test
Advanced - 1	Sample size calculations for proportions, ROC area	Sample size calculations for proportions, means
Advanced - 2	Screening studies – design, biases, outcome measures	ANOVA, linear and logistic regression

Suggested Readings:

	Diagnostic Test Accuracy	Statistical Concepts
Basic -1	<ol style="list-style-type: none"> 1. Weinstein S, Obuchowski NA, Lieber ML. Clinical evaluation of diagnostic tests. AJR 2005; 184: 14-19. 2. Jarvik JG. Fundamentals of Clinical Research for Radiologists: The research framework. AJR 2001; 176: 873-877. 3. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. Ann Intern Med 1981; 94: 553-592. 	<ol style="list-style-type: none"> 1. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet. 2002 Jan 5;359(9300):57-61. Review. 2: Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet. 2002 Jan 12;359(9301):145-9. Review. 3. Kevin Chu (1999) An introduction to statistics, significance testing and the P value. Emergency Medicine Australasia 11 (1), 28-34.
Basic -2	<ol style="list-style-type: none"> 1. Obuchowski NA. ROC analysis. AJR 2005; 184: 364-372. 2. Obuchowski NA. Special Topics I: ROC curves and their use in Radiology. Radiology 2003; 229:3-8. 3. Dwyer AJ. In pursuit of a piece of the ROC. Radiology 1997; 201: 621-625. 	<ol style="list-style-type: none"> 1. Blume JD and Peipert JF. "What your statistician never told you about P-values." Journal of the American Association of Gynecologic Laparoscopists, vol. 10(4):439-444, 2003. 2. Blume JD and Peipert JF. "Randomization in Controlled Clinical

		Trials: Why the flip of a coin is so important." Journal of the American Association of Gynecologic Laparoscopists, vol. 11(3):320-325, 2004.
Intermediate -1	<ol style="list-style-type: none"> 1. Begg CB, McNeil BJ. Assessment of radiologic tests, control of bias, and other design considerations. Radiology 1988; 167: 565-569. 2. Obuchowski NA. Special Topics III: Bias. Radiology 2003; 229: 617-621. 3. Black WC. How to evaluate the radiology literature. AJR 1990; 154: 17-22. 	<ol style="list-style-type: none"> 1. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet. 2002 Jan 26;359(9303):341-5. 2. Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet. 2002 Feb 2;359(9304):431-4.
Intermediate -2	<ol style="list-style-type: none"> 1. Dawson and Trapp, Chapter 6 2. Crewson PE. Reader agreement studies. AJR 2005; 184: 1391-1397. 3. Pagano, Chapter s 15 and 16. 	<ol style="list-style-type: none"> 1. Joseph L, Reinhold C. Statistical inference for proportions. AJR 2005; 184: 1057-1064 2. Joseph L, Reinhold C. Statistical inference for continuous variables. AJR 2005; 184: 1047-1056. 3. Dawson and Trapp, chapter 6
Advanced - 1	<ol style="list-style-type: none"> 1. Beam CA. Strategies for improving power in diagnostic radiology research. AJR 1992; 159: 631-637. 2. Beam CA. Statistically engineering the study for success. AJR 2002; 179: 47-52. 3. Obuchowski NA. Sample size tables for receiver operating characteristic studies. AJR 2000; 175: 603-608. 	<ol style="list-style-type: none"> 1. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet. 2002 Mar 2;359(9308):781-5. 2. Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet. 2005 Apr 9-15;365(9467):1348-53.
Advanced - 2	<ol style="list-style-type: none"> 1. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet. 2005 Apr 23-29;365(9469):1500-5. 2. Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet. 2002 Jan 19;359(9302):248-52. 3. Grimes DA, Schulz KF. Uses and abuses of screening tests. Lancet. 2002 Mar 9;359(9309):881-4. Review. 	<ol style="list-style-type: none"> 1. Dendukuri N, Reinhold C. Correlation and regression. AJR 2005; 184: 3-18. 2. Dawson and Trapp, Chapter s 7 and 8. 3. Pagano, Chapter 12,17,18

Dawson B and Trapp RG. *Basic and Clinical Biostatistics* (4th ed). Lange Medical Books, NY, NY 2004.

Principles of Biostatistics, 2nd edition. Pagano and Gauvreau, Duxbury, 2000.

Additional references for clinical research:

1: Grimes DA, Schulz KF.

An overview of clinical research: the lay of the land.
Lancet. 2002 Jan 5;359(9300):57-61. Review.
PMID: 11809203 [PubMed - indexed for MEDLINE]

2: Grimes DA, Schulz KF.

Descriptive studies: what they can and cannot do.
Lancet. 2002 Jan 12;359(9301):145-9. Review.
PMID: 11809274 [PubMed - indexed for MEDLINE]

3: Grimes DA, Schulz KF.

Bias and causal associations in observational research.
Lancet. 2002 Jan 19;359(9302):248-52.
PMID: 11812579 [PubMed - indexed for MEDLINE]

4: Grimes DA, Schulz KF.

Cohort studies: marching towards outcomes.
Lancet. 2002 Jan 26;359(9303):341-5.
PMID: 11830217 [PubMed - indexed for MEDLINE]

5: Schulz KF, Grimes DA.

Case-control studies: research in reverse.
Lancet. 2002 Feb 2;359(9304):431-4.
PMID: 11844534 [PubMed - indexed for MEDLINE]

6: Schulz KF, Grimes DA.

Generation of allocation sequences in randomised trials: chance, not choice.
Lancet. 2002 Feb 9;359(9305):515-9. Review.
PMID: 11853818 [PubMed - indexed for MEDLINE]

7: Schulz KF, Grimes DA.

Allocation concealment in randomised trials: defending against deciphering.
Lancet. 2002 Feb 16;359(9306):614-8. Review.
PMID: 11867132 [PubMed - indexed for MEDLINE]

8: Schulz KF, Grimes DA.

Blinding in randomised trials: hiding who got what.
Lancet. 2002 Feb 23;359(9307):696-700.
PMID: 11879884 [PubMed - indexed for MEDLINE]

9: Schulz KF, Grimes DA.

Sample size slippages in randomised trials: exclusions and the lost and wayward.
Lancet. 2002 Mar 2;359(9308):781-5.
PMID: 11888606 [PubMed - indexed for MEDLINE]

10: Grimes DA, Schulz KF.

Uses and abuses of screening tests.
Lancet. 2002 Mar 9;359(9309):881-4. Review.
PMID: 11897304 [PubMed - indexed for MEDLINE]

11: Schulz KF, Grimes DA.

Unequal group sizes in randomised trials: guarding against guessing.

Lancet. 2002 Mar 16;359(9310):966-70.
PMID: 11918933 [PubMed - indexed for MEDLINE]

12: Schulz KF, Grimes DA.
Sample size calculations in randomised trials: mandatory and mystical.
Lancet. 2005 Apr 9-15;365(9467):1348-53.
PMID: 15823387 [PubMed - indexed for MEDLINE]

13: Grimes DA, Schulz KF.
Compared to what? Finding controls for case-control studies.
Lancet. 2005 Apr 16-22;365(9468):1429-33.
PMID: 15836892 [PubMed - indexed for MEDLINE]

14: Grimes DA, Schulz KF.
Refining clinical diagnosis with likelihood ratios.
Lancet. 2005 Apr 23-29;365(9469):1500-5.
PMID: 15850636 [PubMed - indexed for MEDLINE]

15: Schulz KF, Grimes DA.
Multiplicity in randomised trials I: endpoints and treatments.
Lancet. 2005 Apr 30-May 6;365(9470):1591-5.
PMID: 15866314 [PubMed - indexed for MEDLINE]

16: Schulz KF, Grimes DA.
Multiplicity in randomised trials II: subgroup and interim analyses.
Lancet. 2005 May 7-13;365(9471):1657-61.
PMID: 15885299 [PubMed - indexed for MEDLINE]