Using the Java Web Services Architecture to Select Sample Size for Biomedical Studies

University of Denver
Denver, CO
April 27, 2012

Sarah M. Kreidler, DPT, MS
Uttara R. Sakhadeo
Vijay C. Akula
Keith E. Muller, MS, PhD, MS
Deborah H. Glueck, PhD

Outline

- Introduce our team and our research
- Define sample size and statistical power
- Motivate the need for sample size software in biomedical research
- Describe our software product and related technologies
- Outline future collaboration opportunities

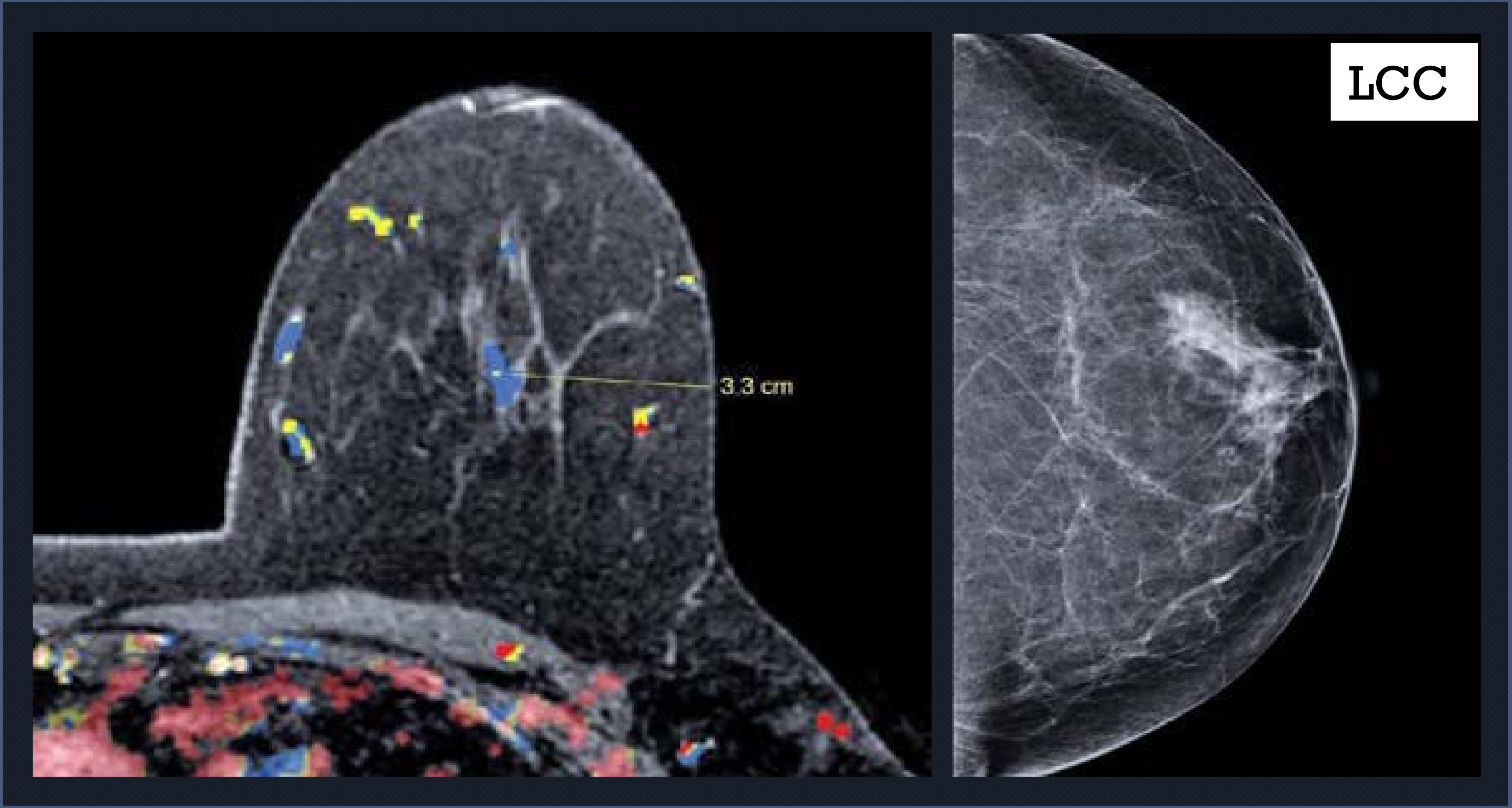
Funding information

- National Institute of Dental and Craniofacial Research
- NIDCR 1 R01 DE020832-01A1

Our lab group

- Statisticians
- Computer scientists
- Technical writers
- Graphic designers

What is the best way to screen for breast cancer?

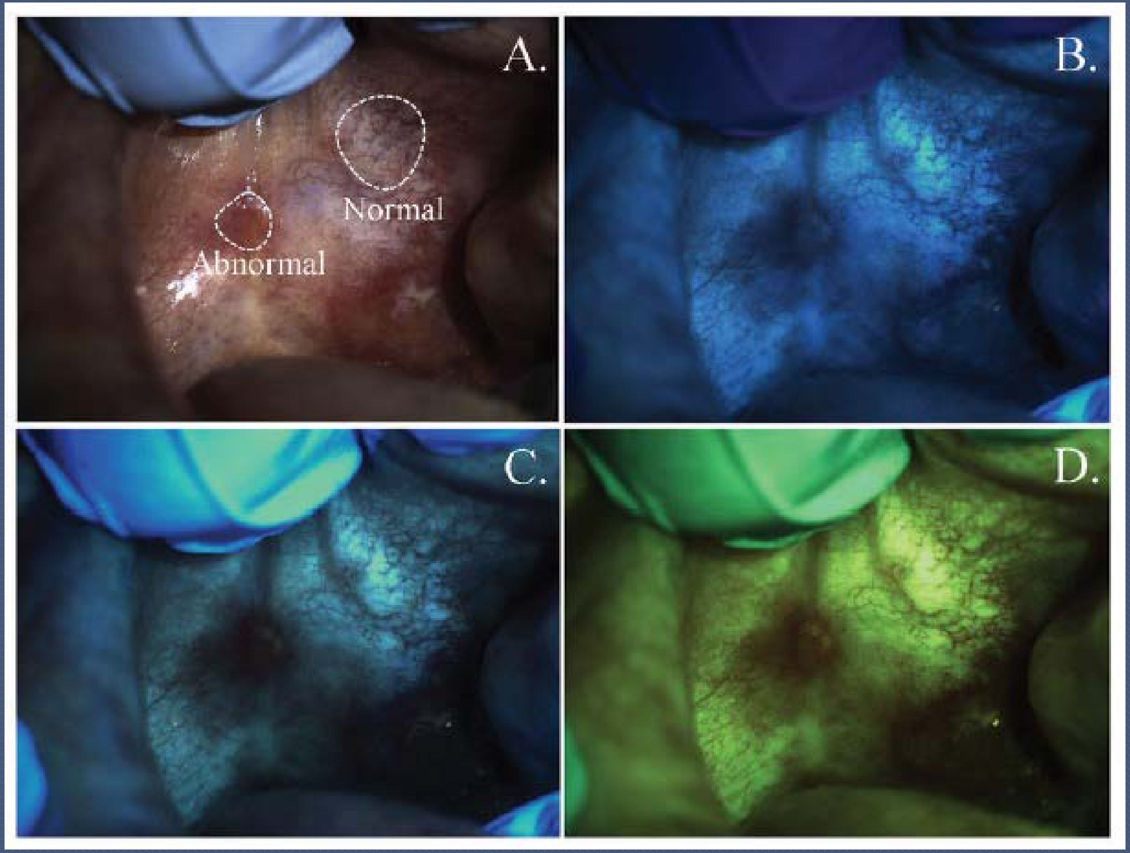


Radiology Imaging Associates, 2010

Do community gardens increase fruit and vegetable intake?

Denver Urban Gardens, http://dug.org/research

How should we screen for oral cancer?



Roblyer et al., 2009, *Cancer Prevention Research*

Outline

- Introduce our team and our research
- **Define sample size and statistical power**
- Motivate the need for sample size software in biomedical research
- Describe our software product and related technologies
- Outline future collaboration opportunities

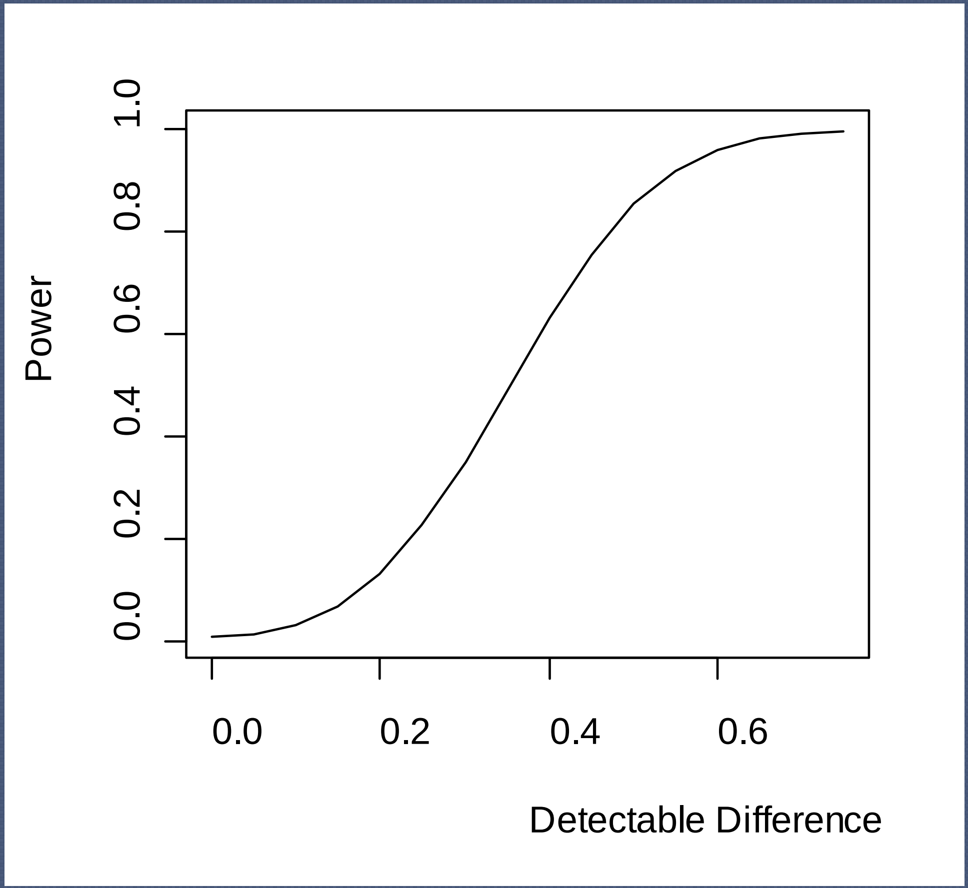
What is sample size?

- Total number of participants in a study
- Must be large enough to answer the scientific question
- Must limit risk to study participants

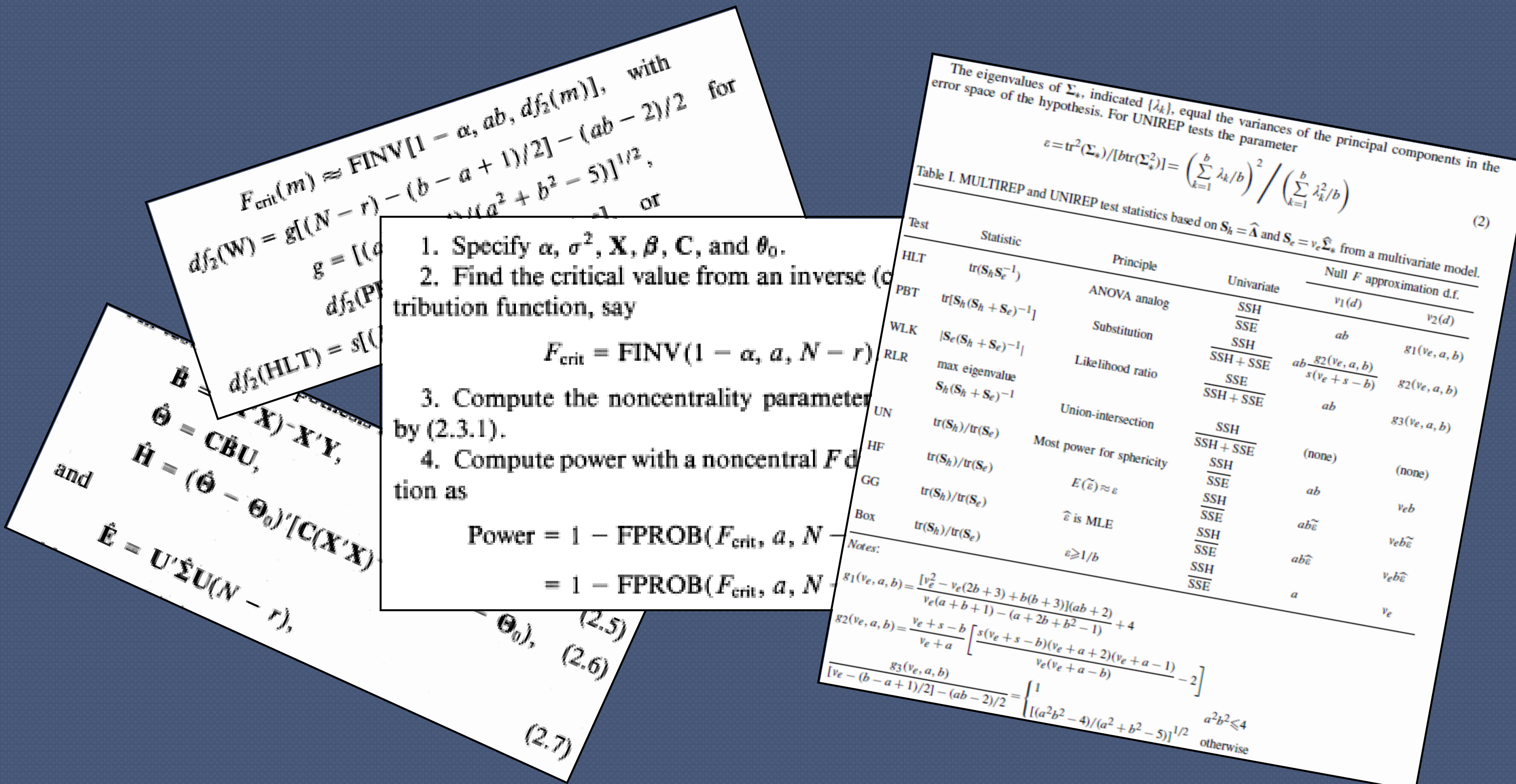
Why do we care?

"One DPP participant treated with troglitazone developed hepatic failure requiring liver transplantation and, in the presence of other complicating illness, subsequently died"

What is statistical power?



How do statisticians select sample size?



Scientists are busy. They don't have time to learn all of this!

Outline

- Introduce our team and our research
- Define sample size and statistical power
- **Motivate the need for sample size software in biomedical research**
- Describe our software product and related technologies
- Outline future collaboration opportunities

How do you make complex methods available to scientists?

Software bridges the gap between statistical theory and scientific practice

- Must be user-friendly
- Must be free
- Must be easily accessible

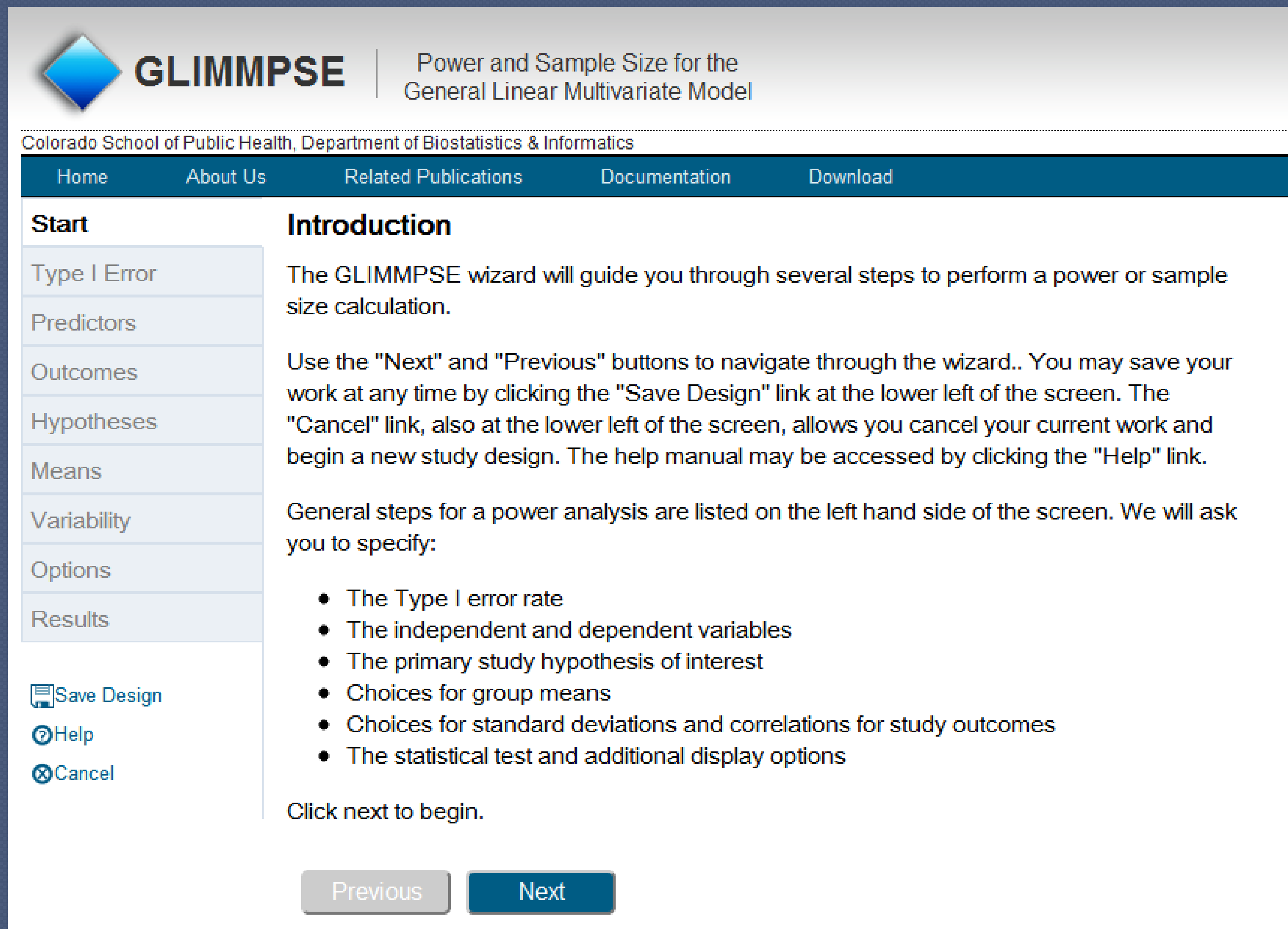
Software best practices

- Functional specification
- Software design specification
- Coding standards
- Revision control
- Code review
- Testing
- Bug tracking
- Deployment

Outline

- Introduce our team and our research
- Define sample size and statistical power
- Motivate the need for sample size software in biomedical research
- **Describe our software product and related technologies**
- Outline future collaboration opportunities

GLIMMPSE


GLIMMPSE
Power and Sample Size for the General Linear Multivariate Model
Colorado School of Public Health, Department of Biostatistics & Informatics
Home
About Us
Related Publications
Documentation
Download
Start
Type I Error
Predictors
Outcomes
Hypotheses
Means
Variability
Options
Results
Save Design
Help
Cancel
Introduction
The GLIMMPSE wizard will guide you through several steps to perform a power or sample size calculation.
Use the "Next" and "Previous" buttons to navigate through the wizard.. You may save your work at any time by clicking the "Save Design" link at the lower left of the screen. The "Cancel" link, also at the lower left of the screen, allows you cancel your current work and begin a new study design. The help manual may be accessed by clicking the "Help" link.
General steps for a power analysis are listed on the left hand side of the screen. We will ask you to specify:
The Type I error rate
The independent and dependent variables
The primary study hypothesis of interest
Choices for group means
Choices for standard deviations and correlations for study outcomes
The statistical test and additional display options
Click next to begin.
Previous
Next

Results

GLIMMPSE | Power and Sample Size for the General Linear Multivariate Model

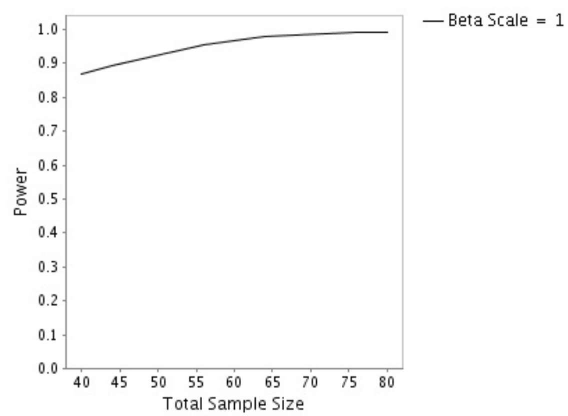
Colorado School of Public Health, Department of Biostatistics & Informatics

Home | About Us | Related Publications | Documentation | Download

- Start
- Type I Error
- Predictors
- Outcomes
- Hypotheses
- Means
- Variability
- Options
- **Results**

Save Design
Help
Cancel

Power Curve



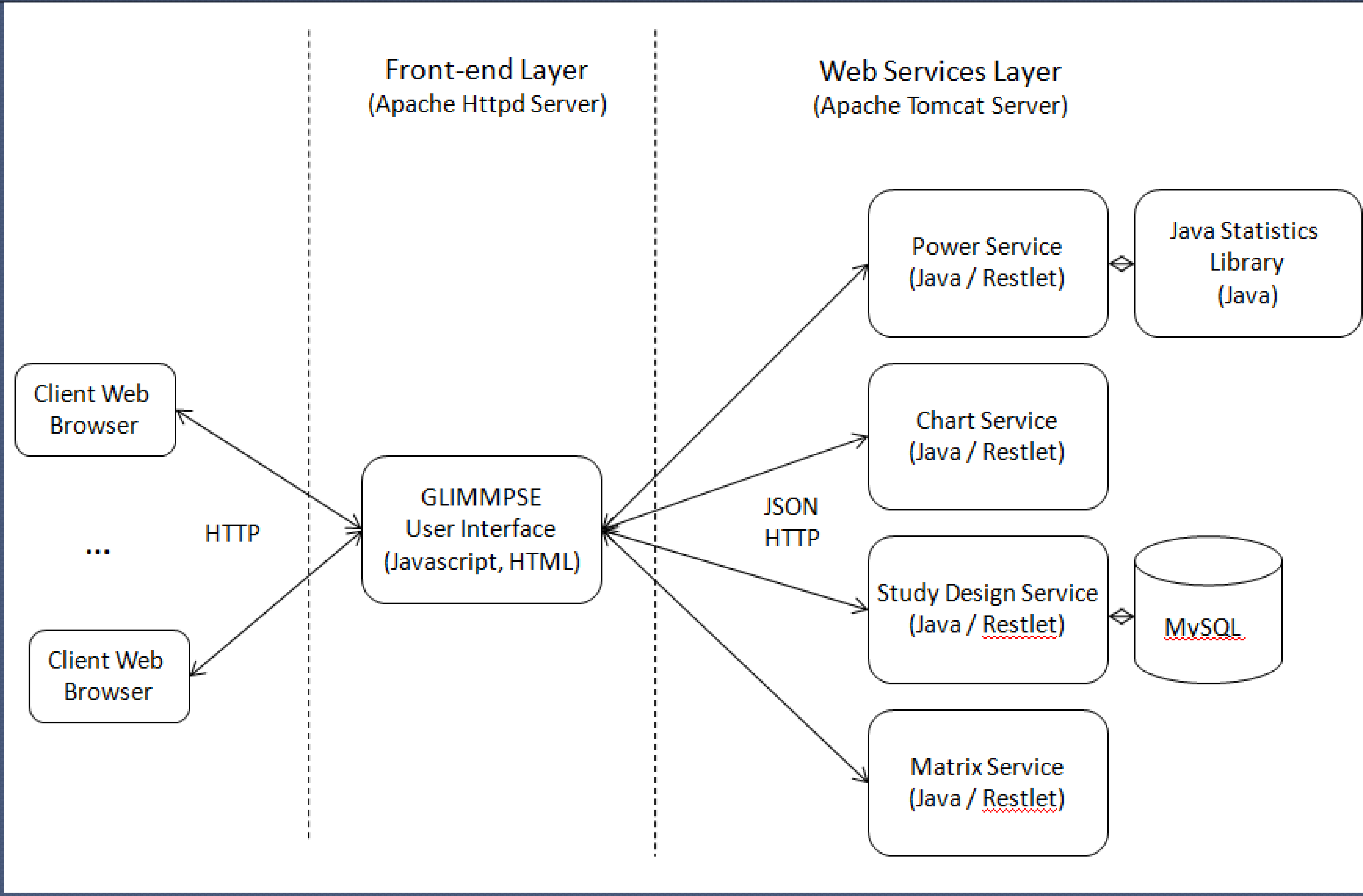
Power Results

Test	Actual Power	Total Sample Size	Beta Scale	Sigma Scale	Alpha	Nominal Power	Power Method
unirep	0.8690	40	1.0	1.0	0.05	0.8690	conditional
unirep	0.9677	60	1.0	1.0	0.05	0.9677	conditional
unirep	0.9930	80	1.0	1.0	0.05	0.9930	conditional

Save results | View matrices used for these results

Previous | Next

Java web services architecture



Architecture advantages

- Modularity
- Scalability
- Platform independence

Version 2.0.0 improvements

- Multilevel and longitudinal designs
- Simplified callback structure in GUI
- JSON Communication Layer
- New web services
  - Study Design Service
  - Matrix Service

GLIMMPSE user interface

- Web-based
- Written with the Google Web Toolkit
- Wizard-style input format

Google Web Toolkit

- Compiles pseudo-Java source code into highly optimized JavaScript
- Built-in AJAX support
- Built in JSON parser
- Piriti library auto-generates JSON parsing code

Pros and cons

- Pros
  - Coding style is similar to Java Swing
  - Integrated with Eclipse
- Cons
  - Compatibility issues across versions
  - Only “Mostly” browser independent
  - Slow compilation

Matrix service

- Matrices commonly used to represent statistical models
- Validation of user inputs
- Provides matrix algebra operations

Matrix service technologies

- Restlet 2.0.10
- Apache Commons Math 2.1
- JavaStatistics library 1.1.0
- Tomcat 6.0 / 7.0

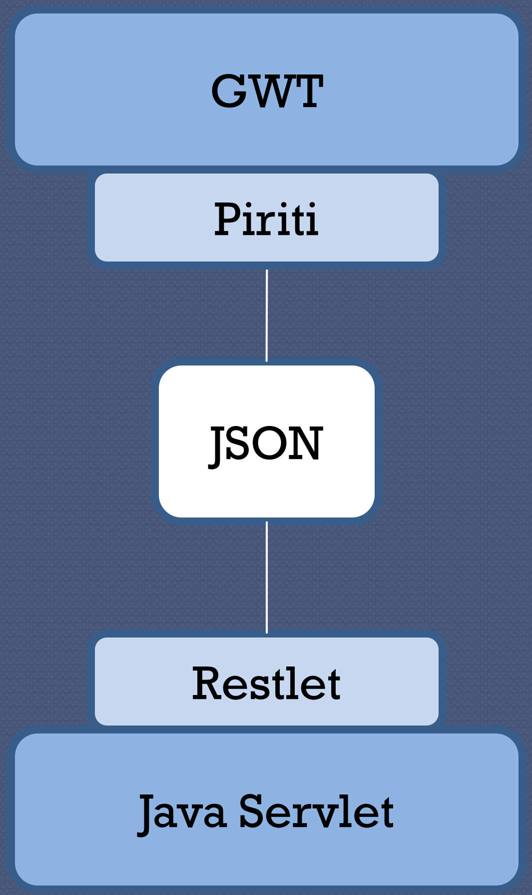
Study design service

- The study design object includes all required information for a sample size calculation
- Persistence to a MySQL database
- RESTful API providing CRUD functionality

Study design service technologies

- Restlet 2.0.10
- Hibernate 4.0.0
- MySQL 5.5
- Tomcat 6.0 / 7.0

Client-server communication



JSON

- String based syntax for transmitting Java objects over HTTP
- Supports objects, sub-objects, and arrays
- Encoding / decoding can be automated
- Parsing time is less since JSON is less verbose

Technical challenges

- Managing complex relationships between study sub-components
- Writing domain object code compatible with Hibernate, Jackson, and Google Web Toolkit

Future directions

- Mobile platforms
- Desktop application
- 3D power surfaces
- Integrate new statistical methods

Outline

- Introduce our team and our research
- Define sample size and statistical power
- Motivate the need for sample size software in biomedical research
- Describe our software product and related technologies
- **Outline future collaboration opportunities**

Some unsolved problems…

- Distributed statistical simulations
- Performance of Davies' algorithm
- Object modeling of statistical tests
- Algorithmic complexity of power analysis

Opportunities with the University of Colorado Denver

- Collaboration opportunities
- Employment opportunities
- https://www.jobsatcu.com/
- Contact us

  Deborah.glueck@ucdenver.edu

Thank you!

Colorado School of Public Health
University of Colorado Denver

Deborah.glueck@ucdenver.edu

References

AJAX Tutorial. Retrieved April 26, 2012, from http://www.w3schools.com/ajax/default.asp

Apache Software Foundation. *Apache Commons Mathematics Library. Version 2.1*. Retrieved from http://commons.apache.org/math/

Apache Software Foundation. Apache Tomcat. Version 6.0. Retrieved May 17, 2011, from http://tomcat.apache.org/

Davies, R. B. (1980). Algorithm AS 155: The Distribution of a Linear Combination of Chi-Square Random Variables. *Applied Statistics*, *29*(3), 323-333.

Denver Urban Gardens - Research. Retrieved April 26, 2012, from http://dug.org/research

Fielding, R. (2000). *Architectural Styles and the Design of Network-based Software Architectures*. University of California, Irvine. Retrieved from http://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm

Google Web Toolkit — Google Developers. Retrieved April 26, 2012, from https://developers.google.com/web-toolkit/

References

Hibernate - JBoss Community. Retrieved April 26, 2012, from http://www.hibernate.org/

JSON. Retrieved April 26, 2012, from http://json.org/

McGovern, J., Tyagi, S., Stevens, M., & Mathew, S. (2003). *Java Web Services Architecture*. Morgan Kaufmann.

Muller, K. E., Lavange, L. M., Ramey, S. L., & Ramey, C. T. (1992). Power Calculations for General Linear Multivariate Models Including Repeated Measures Applications. *Journal of the American Statistical Association*, *87*(420), 1209-1226.

MySQL :: The world's most popular open source database. Retrieved April 26, 2012, from http://www.mysql.com/

piriti - JSON and XML Mapper for GWT - Google Project Hosting. Retrieved April 26, 2012, from http://code.google.com/p/piriti/

Restlet - RESTful web framework for Java. Retrieved April 26, 2012, from http://www.restlet.org/

Roblyer, D., Kurachi, C., Stepanek, V., Williams, M. D., El-Naggar, A. K., Lee, J. J., Gillenwater, A. M., et al. (2009). Objective Detection and Delineation of Oral Neoplasia Using Autofluorescence Imaging. *Cancer Prevention Research*, *2*(5), 423-431. doi:10.1158/1940-6207.CAPR-08-0229