CMPS 119 Software for Society

Introduction to Student Projects
April 2, 2014
Your Instructors

Philip Strong MD
- Recovering software engineer
- Physician Champion for Specialty Clinical Analytics, PAMF; Medical Director for IT, El Camino Hospital
- Hospitalist at ECH Mountain View
- Has been participating in or leading short-term overseas medical trips since 2005.
- Founder of Health Records for Everyone (HR4E)
- pcstrong@comcast.net
- Mobile: (408) 242-6907

Clinton Lewis
- Refugee from corporate America
- Product manager at Toshiba MRI
- Software Engineer for orbit determination trajectory and force modeling package for Air Force spy satellites. (pre-drone)
- Teacher of astronomy at West Valley Community College in Saratoga
- Computer modeling of detection system in genetic sequencer
- webscreen@ScreenGenes.org
- Mobile: 650 906 9629
HR4E
(Phil Strong)

• Health Records for Everyone
• 501(c)(3) Aptos-based non-profit
• “... promoting the adoption of individual health records in developing countries”
• Worked w James Davis in two previous courses, including 2012 CMPS 119
• As a sideline, started working on a questionnaire for deployment on a tablet in mammography suites, to determine suitability for genetic testing (more on this later)
Why/How Questionnaires

- Initially a sideline for HR4E: developed to identify women at PAMF at higher than average risk for breast/ovarian cancer who might benefit from genetic counseling
  - By conventional means, we only identify about half (based on population estimates)
- Re-engineered a similar program created by Intermountain Health in Salt Lake City (dogged with intellectual property issues), instead based on <published> NCCN guidelines
- Committed from the start to make this available as “Open Source”
Why/How Questionnaires

• Along with participants from UCSC classes, we attempted student projects that involved:
  – Extending the HR4E code base for collecting health data for use in developing countries
  – Evaluating “standard packages” for implementing the HR4E mammography questionnaire (compared to de-novo code)

• The best student projects
  – Compared the HR4E health record software to other approaches (scanning, hand-entry by subject matter expert) to data collection in low-resource environments
  – Compared standard mobile questionnaire platform configuration to de-novo code
Why/How Questionnaires

Conclusions:
• For small health-assessment clinics, it can be less expensive to have a subject matter expert hand enter data into a spreadsheet than to configure and support an electronic health record system
  – <The HR4E approach to health records doesn’t make for tractable student projects>
• It was easier to write de novo code than configure and use “standard” questionnaire software for “non-trivial” questionnaires
Meanwhile, Requirements Change

• The US Preventive Services Task Force wants screening for genetic predisposition for breast and ovarian cancer to occur at **primary care visits**
  – mammography suite assessment, because it doesn’t start until age 40-50, still misses too many

• The NCCN guidelines have changed twice since our original software was written

• We needed a more flexible approach, but has to be easier to configure than existing packages

• We’re still committed to Open Source software

• <Phil still wants to keep working on software for use in developing countries>
Potential Student Projects

• Use the HR4E database structure to create & deploy a “non-trivial” questionnaire
  – Can use existing UI
  – Can directly write into the db

• Create a module to restructure and use HR4E questionnaire output (for 10 patients, we’ll supply) to invoke a risk assessment algorithm via web-service (we’ll supply the specifications) and get back the risk score

• Set up github for managing the HR4E code-base (we’ll supply), and show us how we can fork the code to create a new version.
NCCN Guidelines

NCCN Guidelines Version 1.2014
Hereditary Breast and/or Ovarian Cancer Syndrome

HEREDITARY BREAST AND/OR OVARIAN CANCER SYNDROME TESTING CRITERIA\textsuperscript{a,b,c}

- Individual from a family with a known deleterious \textit{BRCA1}/\textit{BRCA2} mutation
- Personal history of breast cancer\textsuperscript{b} + one or more of the following:
  \begin{itemize}
    \item Diagnosed \textless 45 y
    \item Diagnosed \textless 50 y with:
      \begin{itemize}
        \item An additional primary\textsuperscript{d}
        \item \geq 1 close blood relative\textsuperscript{e} with breast cancer at any age
        \item An unknown or limited family history\textsuperscript{a}
      \end{itemize}
    \item Diagnosed \textless 60 y with:
      \begin{itemize}
        \item Triple negative breast cancer
        \item Diagnosed at any age with:
          \begin{itemize}
            \item \geq 2 close blood relatives\textsuperscript{e} with breast cancer diagnosed \textless 50 y
            \item \geq 2 close blood relatives\textsuperscript{e} with breast cancer at any age
            \item \geq 1 close blood relative\textsuperscript{e} with epithelial ovarian\textsuperscript{f} cancer
            \item \geq 2 close blood relatives\textsuperscript{e} with pancreatic cancer and/or prostate cancer (Gleason score \geq 7) at any age
            \item A close male blood relative\textsuperscript{e} with breast cancer
          \end{itemize}
      \end{itemize}
  \end{itemize}
- Family history only (significant limitations of interpreting test results for an unaffected individual should be discussed):
  \begin{itemize}
    \item First- or second-degree blood relative meeting any of the above criteria
    \item Third-degree blood relative with breast cancer\textsuperscript{b} and/or ovarian\textsuperscript{f} cancer with \geq 2 close blood relatives\textsuperscript{e} with breast cancer (at least one with breast cancer \textless 50 y) and/or ovarian\textsuperscript{f} cancer
    \item Clinical judgment should be used to determine if the patient has reasonable likelihood of a mutation, considering the unaffected patient’s current age and the age of female unaffected relatives who link the patient with the affected relatives.
    \item Testing of unaffected individuals should only be considered when an appropriate affected family member is unavailable for testing.
\end{itemize}

- Personal history of epithelial ovarian\textsuperscript{f} cancer
- Personal history of male breast cancer
- Personal history of pancreatic cancer or prostate cancer (Gleason score \geq 7) at any age with \geq 2 close blood relatives\textsuperscript{e} with breast and/or ovarian\textsuperscript{f} and/or pancreatic or prostate cancer (Gleason score \geq 7) at any age
- For pancreatic cancer, if Ashkenazi Jewish ancestry, only one additional affected relative is needed

\textsuperscript{a}Meeting one or more of these criteria warrants further personalized risk assessment, genetic counseling, and often genetic testing and management. The probability of mutation detection associated with these criteria will vary based on family structure. Individuals with unknown or limited family history/structure, such as fewer than 2 first- or second-degree female relatives having lived beyond age 45 in either lineage, may have an underestimated probability of familial mutation detection. The likelihood of mutation detection may be very low in families with a large number of unaffected female relatives. Clinical judgment should be used to determine the appropriateness of genetic testing. The maternal and paternal sides should be considered independently.

\textsuperscript{b}For the purposes of these guidelines, invasive and ductal carcinoma in situ breast cancers should be included.

\textsuperscript{c}Patients who have received an allogeneic bone marrow transplant should not have molecular genetic testing via blood or buccal samples due to unreliable test results from contamination by donor DNA. If available, DNA should be extracted from a fibroblast culture. If this source of DNA is not possible, buccal samples can be considered, subject to the risk of donor DNA contamination.

\textsuperscript{d}Two breast primaries includes bilateral (contralateral) disease or two or more clearly separate ipsilateral primary tumors either synchronously or asynchronously.

\textsuperscript{e}Close blood relatives include first-, second-, and third-degree relatives on same side of family.

\textsuperscript{f}See BRCO-V3.

\textsuperscript{g}For the purposes of these guidelines, fallopian tube and primary peritoneal cancers are included. Ovarian/fallopian tube/primary peritoneal cancers are component tumors of Lynch syndrome/hereditary non-polyposis colorectal cancer; be attentive for clinical evidence of this syndrome. See NCCN Guidelines for Genetic/Familial High-Risk Assessment: Colorectal.

\textsuperscript{h}Testing for Ashkenazi Jewish founder-specific mutation(s) should be performed first. Full sequencing may be considered if ancestry also includes non-Ashkenazi Jewish relatives or other HBOC criteria are met. Founder mutations exist in other populations.
CRITERIA FOR FURTHER GENETIC RISK EVALUATION

An affected individual with one or more of the following:

- A known mutation in a breast cancer susceptibility gene within the family
- Early-age-onset breast cancer
- Triple negative (ER-, PR-, HER2-) breast cancer
- Two breast cancer primaries in a single individual
- Breast cancer at any age, and
  - ≥1 close blood relative with breast cancer ≤50 y, or
  - ≥1 close blood relative with epithelial ovarian cancer at any age, or
  - ≥2 close blood relatives with breast cancer and/or pancreatic cancer at any age
  - From a population at increased risk
- ≥1 family member on same side of family with a combination of breast cancer and ≥1 of the following (especially if early onset): pancreatic cancer, prostate cancer (Gleason score ≥7), sarcoma, adenocortical carcinoma, brain tumors, endometrial cancer, leukemia/lymphoma; thyroid cancer, dermatologic manifestations and/or macrocephaly, hamartomatous polyps of GI tract; diffuse gastric cancer
- Ovarian cancer
- Male breast cancer

An unaffected individual with a family history of one or more of the following:

- A known mutation in a breast cancer susceptibility gene within the family
- ≥2 breast primaries in single individual
- ≥2 individuals with breast primaries on the same side of family
- ≥1 ovarian cancer primary from the same side of family
- First- or second-degree relative with breast cancer ≤45 y
- ≥1 family member on same side of family with a combination of breast cancer and ≥1 of the following (especially if early onset): pancreatic cancer, prostate cancer (Gleason score ≥7), sarcoma, adenocortical carcinoma, brain tumors, endometrial cancer, leukemia/lymphoma; thyroid cancer, dermatologic manifestations and/or macrocephaly, hamartomatous polyps of GI tract; diffuse gastric cancer
- Male breast cancer

Note: All recommendations are category 2A unless otherwise indicated.

Clinical Trials: NCCN believes that the best management of any cancer patient is in a clinical trial. Participation in clinical trials is especially encouraged.

For the purposes of these guidelines, fallopian tube and primary peritoneal cancers are included. Ovarian/fallopian tube/primary peritoneal cancers are component tumors of Lynch syndrome/hereditary non-polyposis colorectal cancer; be attentive for clinical evidence of this syndrome. See NCCN Guidelines for Genetic/Familial High-Risk Assessment: Colorectal.
Software for Creating, and Deploying Questionnaires

- Subject of the questionnaires: risk assessment for healthcare
- Deployment
  - How: Web based, multi-platform
  - Portable server, self contained, minimum infrastructure
- Access
  - Waiting room: health history form
  - Patient’s home
  - at a remote clinic
Questionnaire examples

2. Is your healthcare provider, the provider you usually see if you need a check-up, want advice about a health problem, or get sick or hurt?
   - Yes
   - No

3. How long have you been going to your healthcare provider?
   - Less than 6 months
   - At least 6 months but less than 1 year
   - At least 1 year but less than 3 years
   - At least 3 years but less than 5 years
   - 5 years or more

Breast Health Screening

Please answer all of the following questions about your breast cancer:

Was your cancer “triple negative”, without receptors for estrogen, progesterone and HER2?
   - Yes
   - No

Have you had cancer in both breasts?
   - Yes
   - No

Was your cancer found in more than one location in the same breast?
   - Yes
   - No

What was your age when your breast cancer was first diagnosed?
   - Age 50 or under
   - Age 51 or older

[Back] [Next]
What do you want to do in this class?

• √ Python, general purpose language
  • How to set up and run
  • Nifty tools: lists and dictionaries
• √ Django, web frameworks
  • database interaction
  • location of the documentation
• Website for programmers:  http://stackoverflow.com
• User interface for a Questionnaire Editor
• √ Redesign of the UI logic.
• Rewrite the software
What to do for HR4E (non-profit)?

• Need software to create the CDA document (standard for information exchange)
• Clinical Document Architecture (CDA)
• Interact with an on-line risk calculator
  • send patient data from the questionnaire
    • CDA document
  • accept results
  • insert results into database
• Help with the current “industrial” UI. The interface logic needs streamlining and simplification.
• Place source code into GitHub (half hour?)
Assignment: Run a Questionnaire

• For next week (30 minutes), answer the questions (not truthfully) on a Questionnaire and send results to webscreen@ScreenGenes.org

• For the following weeks:
  • study risk assessment journal article assigned to your group
  • build a questionnaire
  • test and verify
  • ask another group to test your questionnaire
  • send verification to webscreen@ScreenGenes.org

• Other help needed by our group
  • place the source code into GitHub for collaborative software effort.
  • need further testing
  • software need updating (Python and Django)
Assignment
for Friday, April 18

– Each of your group create an account in the Questionnaire Editor

– Select a “Project” a.k.a. “Team” or “Group”
  – Rename the Project to whatever name your group pleases. Renaming a Project renames it for everyone in your group!

– Run one of the Sample Questionnaires
  • Answer the questions.
  • Dump the output to your computer.
  • Email the results
    • to yourself, then to webscreen@ScreenGenes.org.

– Play with the Questionnaire Editor
Respond to the Questionnaire

• Answer the questions.
• However, use bogus names and data.
  • health questionnaire
• Enter your correct email (why?)
  • Your Questionnaire responses will be mailed to your email address.
  • Another reason: in order to get a grade for this assignment.
• Dump the output to your computer.
• Email the results you receive via email to webscreen@ScreenGenes.org.
Finding the Questionnaire Editor

• Create a new account
  • http://50.79.41.35/scrntest/multiquest/registration/createNewAccount/

• User menu (after login)
  • on the Internet:
    • http://50.79.41.35/scrntest/multiquest/registration/userLanding/
  • on “ClassServer” network in the classroom:
    • http://169.254.165.219/scrntest/multiquest/registration/userLanding/
  • When the Server is “on”!

18
User Interface “Scope”

- Upon logging in, Who can see what?
  - Only Project members can see the Questionnaire being edited.
    - Be careful within a project! You can trample on each others edits!
  - Anyone can see and execute the Questionnaire.
  - Can change Project
Possible Problems on the website

• May be buggy. Report problems to webscreen@ScreenGenes.org, preferably with a screenshot.
  • The load may slow the website, e.g. healthcare.gov
• However, the website is self contained, and information on this website is private. Caveats and clarifications:
  • Django security measures are in place.
  • The NSA may have copied the server. However Edward Snowden does not work for us.
  • Hackers exist!
• Questionnaire results are stored in the local database.
  • however, the data doesn’t go anywhere except to those logged in.
Welcome to the Questionnaire Editor

Your Home Page for editing and executing a Questionnaire

Your username is: p1. Your login is associated with the Project: Amber

The default Questionnaire you have selected is: SampleQuestionnaire

Select your next action from the following list:

Select and run a Questionnaire

- Run questionnaires associated with any Project
- Select a default Questionnaire for the Project: Amber

Project management

- Select a different Project
- Edit your Project information: Amber
- Create a new Sample Questionnaire
- Logout of the Questionnaire Editor

Edit a Questionnaire

- Create page transition logic based upon question response

Get Questionnaire results

- Select and view responses to the Questionnaires
- Dump submission data to a spreadsheet file for Project Amber
How to create a Questionnaire

- Each questionnaire
  - has a “splash” page as the first page
  - has a “userID” page as the second page
  - has a “completion” page as the last page,
- Draw a diagram showing the page transitions.
The User!
The User!

Application software
The User!

Questionnaires “have” Pages and the Pages have Questions.

Questions are “in” Pages and Pages are in Questionnaires.
Questionnaire Data Model

- Object model “has a” relations:
  - User
  - Project
  - Questionnaire
  - Page
  - Question
Entity-Relationship Diagram
Database strategy

• Draw diagram with database and layers of software
• Database tables have information to define multiple objects.
• Relationships (ManyToMany) between objects are controlled using a “connector” table referenced by the application.
The name “tag” or “shortTag” refers to a easily remembered text mnemonic used for object retrieval from the database. This device is used for the database objects:

- Projects
- Questionnaires
- Pages

With added qualifications, these short tags point to a unique record in a table in the database, and even more immediately, a Django view function.
Questionnaire page transition logic

for your sample questionnaire: BRCA_S
Questionnaire page transition logic

splash (start screen) → P1 (ovarian pt history) → always → P2 (ovarian family history) → P3 (breast ca pt history) → P3a (br ca pt de...
Questionnaire page transition logic

for your sample questionnaire: BRCA_S

P2 (ovarian family history) → P3 (breast ca pt history) → P4 (brca & bc fm history) → P3a (br ca pt details) → P3b (br ca fm details) → P3c (f condil)

Always

BRCA mutation risk extremely low
BRCA mutation risk higher than average
BRCA mutation risk indeterminate

Q response:

YYY  YNY  NYY  NNY

YYN  YNN  NYN  NNN
Web pages are “stateless”

• The Server must begin by treating each Client query as unrelated to any previous queries.
  • The Server responds to many queries from many Client computers all over the world in quick order.
• A “Session” is created on the Server when there is a current “cookie” on the Client. This maintains consistency for the user interaction with the website.
  • Example: a “login” on a website creates a Session.
Login Process
March 30, 2014

http://50.79.41.35/scrntest/multiquest/registration/

login

try your new login!

createNewAccount

if no Project

selectProjectDefault

userLanding

editProjectRecord

http://50.79.41.35/scrntest/multiquest/registration/X/

http://50.79.41.35/

selectProjectsQuestionnairesToExecute

http://50.79.41.35/scrntest/multiquest/registration/X/

http://50.79.41.35/X

Page 404: Page Not Found