OAI Online and Access to Data, Images, and Biospecimens

Susan M. Rubin
OAI Coordinating Center
Dept. of Epidemiology and Biostatistics
University of California, San Francisco
OAI data is publicly available

- Accessible via the OAI Online website
  - www.oai.ucsf.edu
- Developed by data team at OAI Coordinating Center, UCSF
- OAI data and documentation available to anyone
  - Register
  - Agree to terms of data use agreements
OAI data users (as of 10-24-08)

- Registered OAI Online users 1,063
  - US 747
  - International 316
- Number of countries 51
  (from Argentina to New Zealand)
- Datasets downloaded 2,380
- Image sets distributed 126
Public data releases

- Data released in 2 groups:
  - 1\textsuperscript{st} group = 1\textsuperscript{st} half (56% enrolled, n=2,686)
  - 2\textsuperscript{nd} group = Entire cohort (n=4,796)

- Regular release schedule
  - Two releases per visit cycle

- Concurrent release of images and clinical data
Data currently available

- Baseline * (1st group; n=2,686)
- 12-month (1st group; n=2,686)
- 18-month (subset of progression cohort; n=287)
  - plus corresponding baseline and 12-month MRI images for this subset
- 24-month (1st group; n=2,686)
- Paired X-ray readings (JSW) for 160 longitudinal image set

* Clinical data for entire cohort available; however, images available for 1st group only.
Upcoming data releases

- November 2008
  - Baseline images (entire cohort)
  - 12-month questionnaire, exams, images (entire cohort)

- Winter 2008/2009
  - Data from ongoing centralized image assessments

- Spring 2009
  - 30-month questionnaire, MRIs (subset of progression cohort)
  - 24-month questionnaire, exams, images (2nd group) *

- Summer/Fall 2009
  - 36-month questionnaire, exams, images * (1st group)

* Release schedule may be modified.
Access to OAI data

- Register OAI Online (www.oai.ucsf.edu)
- Agree to terms in Data Use Agreement
- Download questionnaire, exam data and documentation
- For images:
  - Fax 'Image Use Agreement' and request form to UCSF
  - Smaller releases (<250 GB):
    UCSF sends images on OAI 'lego'-drive; user returns drive
  - Larger releases:
    User provides 2-TB hard-drive; UCSF returns it with images
Create an account

- Required to access 'Data and Documentation'
- Contact information
  - Notified of new releases, features, errata, etc.

Logon Help

- If you are a new user, please Create My Account first. The process only takes a few minutes. You need an account to be able to access the Data & Documentation section of the website.
- If you intend to download data or analyze data with the DataExplorer, you have to read and accept the terms of the Data Use Agreement (you must be logged on to view it).
- Remember that passwords are case-sensitive.
The Osteoarthritis Initiative
a multi-center, four-year observational study of men and women

The OAI is a nationwide research study, sponsored by the National Institutes of Health (part of the Department of Health & Human Services), that will help us better understand how to prevent and treat knee osteoarthritis, one of the most common causes of disability in adults.

OAIONline is the name for this website, which provides access to the Osteoarthritis Initiative (OAI) study documentation and data.

Data News

1-2-3 Data. Three quick steps to download the clinical data.

What's available as of 4/15/2008

Clinical (download data from OAIONline) -
- Baseline questionnaire and exam data for entire cohort (4,796 participants)
- Baseline biomarkers dataset (includes baseline results from paired baseline/12-month x-ray readings on a sample of 150 progression cohort participants)
- 12-month follow-up questionnaire and exam data for first half of cohort (2,688 participants)

My Cart: Create a personal cart of variables of interest that can be used 1) to build a custom codebook able to be read offline, and 2) to filter downloadable datasets.
# Summary of Data and Images Currently Available

1-2-3 Data. [Three quick steps](#) to download the clinical data.

What's available as of 4/15/2008

| Clinical (download data from OAIONline) | Baseline questionnaire and exam data for entire cohort (4,796 participants)  
| --- | ---  
|  | Baseline biomarkers dataset (includes baseline results from paired baseline/12-month x-ray readings on a sample of 160 progression cohort participants)  
|  | 12-month follow-up questionnaire and exam data for first half of cohort (2,686 participants)  
|  | 12-month biomarkers dataset (includes 12-month results from paired baseline/12-month x-ray readings on a sample of 160 progression cohort participants)  
|  | 18-month questionnaire data for 288 progression cohort participants  
|  | 24-month follow-up questionnaire and exam data for first half of cohort (2,686 participants)  

| Imaging (request images from OAI Coordinating Center) | Baseline X-ray and MRI images for a sample of 200 progression and incidence cohort participants  
| Baseline and 12-month X-ray and MRI images for a sample of 160 participants |
Where to start

About OAI

The Osteoarthritis Initiative (OAI)

- Study overview and objectives
- Study design protocol (0.85 mb) Appendices (2.5 mb)
- Collaboration worldwide
- Funding
- UCSF IRB approval letter for OAI

Visit Schedule and Assessments

- Examination Measures
- Questionnaire Measures

OAI Related Links

- OAI News & Events at the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS)
- National Institute on Aging (NIA)
How to find data

- Refer to list of examination and questionnaire measures, by visit
# Visit schedule and measurements

## OAI Protocol

**Osteoarthritis Initiative: A Knee Health Study**

<table>
<thead>
<tr>
<th>Questionnaire Measures and Frequency</th>
<th>Initial Eligibility</th>
<th>Screening Visit</th>
<th>Enrollment Visit</th>
<th>12-mo</th>
<th>18-mo</th>
<th>24-mo</th>
<th>30-mo</th>
<th>36-mo</th>
<th>48-mo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contact Information</strong></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Demographics</strong> (such as age, gender, ethnicity, education, marital status, residency, income)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Employment, current and past</strong></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Health care and health insurance</strong></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MRI contraindications</strong></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Knee Symptoms</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Frequency of knee symptoms &amp; medication use for knee symptoms, past 12 mos, 30 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- Knee pain 0-10 rating scale, past 7, 30 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- WOMAC knee pain and stiffness, past 7 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- KOOS knee pain and symptoms, past 7 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td><strong>Knee-related function and QOL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- WOMAC physical function, past 7 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- KOOS sport, recreation, past 7 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- KOOS Quality of life, past 7 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- Participant global assessment of knee impact</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>- Limitation of activity due to knee Sx, past 30 days</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Tools for data searching

Data & Documentation

OAI has collected both clinical data from forms (questionnaires and examinations) and images from X-RAY and MRI images of joints.

FAQ - frequently asked questions

Summary of Data and Images Currently Available

Clinical

- **Clinical data** - "Getting Started with OAI Data: Overview of Structure, Use, and Conventions", version logic, visit schedule and assessments, variable guides, formats, documentation (includes variable distributions), datasets, and dataset descriptions.

- **Search/browse** a database of categorized variable documentation displayed longitudinally with links to statistics, comments, and annotated data collection forms.

- **Data collection forms** - annotated pdfs of the questionnaires used for collecting data, including the visit schedule and assessments.

- **Operations manuals** - pdfs of the detailed procedures for collecting data and performing examinations.

- Create a personal cart of variables of interest that can be used 1) to build a custom codebook able to be read offline, and 2) to filter downloadable datasets.

- **DataExplorer** - explore the clinical data online. Generate descriptive statistics: cross tabs, subsetting, frequency distributions and more. This tool requires acceptance of the terms of a Data Use Agreement.

- **Tutorial** - walks you through the variety of tools available for touring and comprehending the clinical data. Variables for Sample Research Question.
Search and browse

- Powerful tool to become familiar with data
- Search any variable or category
- Longitudinal display
- Access data collection form
- Look at frequencies
- Review notes and SAS code
Search keyword and/or category

1. Keyword or phrase: buttocks
   - in variable label, root, or category
   - Clear Text

2. Visit: P02 Initial Eligibility Intvw, P01 Screening, V00 Enrollment, V01 12 month Annual

3. Search/Browse, Clear All, tutorial
### Search results

#### Back pain

<table>
<thead>
<tr>
<th>Label</th>
<th>Root</th>
<th>P02</th>
<th>P01</th>
<th>V00</th>
<th>V01</th>
<th>V02</th>
<th>V03</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any back pain, past 30 days</td>
<td>BP30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often bothered by back pain, past 30 days</td>
<td>BP300FT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When had back pain how bad was it on average, past 30 days</td>
<td>BPBAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain location, upper back</td>
<td>BPUB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain location, middle back</td>
<td>BPMB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain location, lower back</td>
<td>BPLB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain location, buttocks</td>
<td>BPB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain location, don't know</td>
<td>BPDK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limit activities due to back pain, past 30 days (calc)</td>
<td>BPACTCV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total days in bed and/or limited activity due to back pain, past 30 days (calc)</td>
<td>BPTOT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many days stay in bed due to back pain, past 30 days (calc)</td>
<td>BPBEDC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many days limit activities due to back pain, past 30 days (calc)</td>
<td>BPBDAYCV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Add Checked Items to Cart*
Annotated data collection form

BACK PAIN and FUNCTION

The next questions are about back pain.

18 During the past 30 days, have you had any back pain? **BP30**
   1 O Yes  0 O No  .D O Don’t know  .R O Refused

 **BP30OFT**
   (Examiner Note: Read response options. OPTIONAL - Show Card #3.)
   3 O All of the time  2 O Most of the time  1 O Some of the time  0 O Rarely  .D O Don’t know

b. When you had back pain, how bad was it on average? **BPBAD**
   (Examiner Note: Read response options.)
   1 O Mild  2 O Moderate  3 O Severe  .D O Don’t know

c. In what part or parts of your back is the pain usually located?
   (Examiner Note: REQUIRED - Show Card #4. Mark all that apply.)

1 O Upper Back **BPUB**
   1 O Middle Back **BPMB**
   1 O Lower Back **BPLB**
   1 O Buttocks **BPB**  .D O Don’t know  **BPKD**

d. During the past 30 days, have you limited your activities because of back pain? **BPACT**
   1 O Yes  0 O No  .D O Don’t know  .R (Refused)

   **BPACTCV**

i. How many days did you stay in bed because of your back? If you are unsure, please make your best guess. **BPBEBED**
   1 O Days  .D O Don’t know

   **BPBEBEDCV**

ii. How many days did you limit your activities because of your back? Do not include days you stayed in bed. If you are unsure, please make your best guess.
   **BPDAY**
   1 O Days  .D O Don’t know

   **BPDAYCV**
### Descriptive statistics

#### Dataset: Back pain location, buttocks

#### Category: Joint symptoms/function

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Screening</th>
<th>12 month Annual</th>
<th>24 month Annual</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01BPB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V01BPB</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V03BPB</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### SAS Format Name

<table>
<thead>
<tr>
<th>Name</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>YNOK</td>
<td>155</td>
<td>100</td>
</tr>
<tr>
<td>YNOK</td>
<td>254</td>
<td>100</td>
</tr>
<tr>
<td>YNOK</td>
<td>156</td>
<td>100</td>
</tr>
</tbody>
</table>

#### Related Labels:

- Back pain location: buttocks
- How often bothered by back pain, past 30 days
- Back pain location, upper back
- When back pain was most severe, past 30 days
- Back pain location, middle back
- Back pain location, lower back

**Osteoarthritis Initiative**
Data explorer

- Another powerful tool available on OAI Online to become familiar with data
- Generate descriptive statistics
  - Frequency distributions
  - Cross tabulations
Baseline:

- Knee pain status (KSX) vs. Radiographic OA (XRKOA)
- Only in elderly people who aren’t overweight
### Data explorer cross tabulation

**Table of P01KSX by P01XROA**

<table>
<thead>
<tr>
<th></th>
<th>D01KSX</th>
<th>D01XROA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frequency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Percent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Row Pct</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Col Pct</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table Data**

- **0**: Neither | 1: Right | 2: Left | 3: Both | **Total**

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th><strong>Total</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Missing</strong></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td><strong>.M:</strong></td>
<td>0.51</td>
<td>0.34</td>
<td>0.33</td>
<td>0.17</td>
<td>1.00</td>
</tr>
<tr>
<td><strong>0: No pain either knee on either knee</strong></td>
<td>124</td>
<td>77</td>
<td>16</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>21.23</td>
<td>21.16</td>
<td>12.10</td>
<td>17.74</td>
<td>124</td>
</tr>
<tr>
<td><strong>1: Infreq pain knee, no pain other knee</strong></td>
<td>122</td>
<td>8.90</td>
<td>13.91</td>
<td>7.38</td>
<td>122</td>
</tr>
<tr>
<td></td>
<td>20.89</td>
<td>20.89</td>
<td>18.99</td>
<td>18.99</td>
<td>122</td>
</tr>
<tr>
<td><strong>2: Infreq pain other knee</strong></td>
<td>125</td>
<td>11.69</td>
<td>2.05</td>
<td>1.54</td>
<td>36</td>
</tr>
</tbody>
</table>
Shop 'til you drop

- Personal shopping cart ('My Cart')
- Store variables that you select
- Access later
- Create custom codebook and save for offline work
  - Variable names
  - Statistics
  - Comments
- Download dataset(s) containing your variables
Imaging Data

- Requesting images: procedures to request and obtain X-RAY and MRI images.
- Imaging Data Use Agreement: complete and send to the UCSF Coordinating Center when requesting images.
- Operations Manuals: This link is to all the manuals ordered by visit; you will need to click on a visit to view the operations manuals for MRI and XRAY images.
- Example Images: Examples showing types of images typically available from participants with varying degrees of OA.
- Summary QA reports on MR scanner performance are available. Please contact us by email at OAIOnlineFeedback@psf.ucsf.edu if you would like a copy.

Visit Schedule and Assessments

- Examination Measures
- Questionnaire Measures

Image Releases

Image releases are listed below. Please refer to the reff# when making a request for images.

See OAI data release version logic for an explanation of both the clinical and imaging data version conventions.

Images will be distributed upon request via hard-drives and will be accompanied by a Dicom Image Release Notes (DIRN) file for that release. DIRNs are available for download via the links below. They contain the following sections:

- Participants and Images included - selection criteria for participants and types of images available
- Disk Contents file hierarchy on the disk

Osteoarthritis Initiative
OAI biospecimens available soon

- Annual collection of biospecimens
  - Serum
  - Plasma
    - EDTA
    - Citrated
  - DNA
  - Urine
- Biospecimens (including DNA) from baseline, 12-mo, and 24-mo available for request
Application process for biospecimens

- Interested investigators encouraged to apply

- Complete application form and proposal
  - Specify and justify amount of specimen requested
  - Include data sharing plan
    - All data generated from biospecimen use will be made available to public via OAI Online website
  - Submit to OAIbiospecimens@kai-research.com

- Review process
  - Biospecimen Review & Allocation Committee (BRAC)
  - Proposals reviewed for feasibility, scientific merit and consistency with OAI goals
  - Full review process takes 6-10 weeks
  - Specimens released once funding confirmed
## Application submission deadlines

<table>
<thead>
<tr>
<th>OAI application submitted to KAI</th>
<th>BRAC review</th>
<th>NIH application receipt dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 15</td>
<td>Feb/March</td>
<td>June/July</td>
</tr>
<tr>
<td>May 15</td>
<td>June/July</td>
<td>Oct/Nov</td>
</tr>
<tr>
<td>Sept 15</td>
<td>Oct/Nov</td>
<td>Feb/March</td>
</tr>
</tbody>
</table>

Osteoarthritis Initiative
Publication guidelines for public data users

- Encourage use of public datasets
- Ensure appropriate citation and acknowledgement of OAI
- Inform us about OAI publications
  - Publications tracked on OAI Online
  - Electronic archive of OAI publications
    - PDFs posted if available and permitted by copyright holder
Another option: collaborate with OAI investigators

- Participate in OAI publications review process
  - E-mail: OAIPublications@psg.ucsf.edu

- Submit analysis proposal to OAI Publications Committee for feedback and approval
  - Constructive critiques
  - Analytic guidance
  - Avoid unintended duplication of effort
  - Foster collaboration among OAI investigators
  - High quality publications based on valid use and interpretation

- Good housekeeping seal:
  - “This manuscript has received the approval of OAI Publications Committee based on review of scientific content and data interpretation”
Summary

- Wealth of information (n=4,796)
  - Imaging (MRI, x-rays, image assessments)
  - Examination and questionnaire measures
  - Biospecimens

- Online tools available to facilitate data access
  - Tutorial
  - Search and browse
  - Data explorer
  - Descriptive statistics, cross tabs
  - Variable guide

- Publications Committee
  - E-mail: OATPublications@psg.ucsf.edu
  - Submit analysis proposals
    - Constructive critiques
    - Analytic guidance