Central Reading of Knee X-rays for Kellgren & Lawrence Grade and Individual Radiographic Features of Tibiofemoral Knee OA

1. Overview

1.1 SAS dataset

Name: **kXR_SQ_BUxx** (xx identifies the time point)
Display label: **kXR SQ reading (BU)**

1.2 Contents of dataset

This dataset contains central longitudinal readings of serial knee X-rays for tibiofemoral radiographic OA done by Dr Piran Aliabadi, MD and Dr Burt Sack, MD under the direction of Dr David Felson, MD of the Boston University Clinical Epidemiology Research and Training Unit, for baseline through 48-month visits (Project 15) and by the same readers in Boston, under the direction of Michael Nevitt, PhD of the OAI Coordinating Center for visits over the entire 96-month follow-up period in OAI (Project 37 (or 42)). **Please note that although some participants are coded READPRJ=42, they are in fact participants in Project 37. Users should recode these participants from READPRJ=42 to READPRJ=37.**

There are data for these two separate reading projects in the dataset:

1. Project 15: readings of Kellgren and Lawrence Grades (KLG)[1] and OARSI Joint Space Narrowing (JSN) Grades[2] in **all participants** who had at least one follow-up X-ray, up to the 48-month follow-up visit.
2. Project 37 (or 42): readings of KLG and OARSI JSN grades for 72-month/96-month follow-up in:

a. **Any knees eligible to develop incident radiographic knee OA between the 48-month and 96-month visits** i.e.: knees had 72-month and/or 96-month visit knee X-rays and were KLG 0 or 1 at baseline and did not have KLG≥2 with JSN≥1 on the 48-month X-ray (or the last time point with Project 15 data) to determine incidence by 96-months.

b. **Any knees that had developed incident radiographic knee OAI between baseline and 48-month visits** to assess worsening of disease over the next follow-up period (knees which were end stage prior to the 72-month visit were excluded).

c. **A random selection of 246 knees which were KLG 2 or 3 at baseline** to determine worsening of disease over the next follow-up period (knees which were end stage prior to the 72-month visit were excluded).

Project 15 provides semi-quantitative scores for assessing incidence and progression of radiographic knee OA over the first 48 months of the OAI. Project 37 (or 42) provides scores for assessing incidence of radiographic knee OA over the entire 96 months of the study, and scores for a sample of knees with prevalent radiographic knee OA to look at progression over 96 months.

IMPORTANT NOTE regarding incidence of OA: Project 15 provides complete data for incidence of radiographic OA over the first 48 months of OAI, and when combined with Project 37 (or 42) provides complete data for incidence of radiographic OAI over the entire 96-month follow-up period.

IMPORTANT NOTE regarding progression of OA: Project 15 provides complete data for progression of radiographic OA over the first 48 months of OAI, but even when combined with Project 37 (or 42) data for progression of radiographic OA over the entire 96-month follow-up are incomplete. Readings of progression of radiographic OAI from the 48-month to 96-month follow-up period are only available in a subset of knees comprising (a) any knees that developed radiographic OA in the first 48 months of the study, and (b) a random selection of knees which were KLG 2 or 3 at the start of the study and remained eligible for JSN after the 48-month visit.

The kXR_QJSW_DuryeaXX datasets provide complete quantitative JSW measurements which may be used for examining progression of radiographic OA for the entire 96-month follow-up of OAI for all knees which (a) developed radiographic OA at any time point in the study, (b) were KLG 2-3 at baseline, and (c) a random selection of knees that remained KLG 0 or 1 at all time points throughout the study.

In both Projects 15 and 37 (or 42) knees that were scored KLG≥2 at baseline or at a follow-up period, were also scored for other individual radiographic features (IRFs) such as osteophytes and chondrocalcinosis.

For Project 15, all available knee radiographs from baseline thru 48-month visits (up to 5 images) were scored, with the baseline visit known, and the other visits blinded to chronological order.

For Project 37 (or 42), it is important to note that because the knees already had readings for Project 15, readers were presented with the baseline X-ray, and the last X-ray scored for Project 15 (usually 48 month) along with the unscored images from 72-month and/or 96-month visits (blinded to order). Scores available from Project 15 were provided to the readers, and the knees were scored for all visits provided, usually 4 visits: baseline, 48-month, 72-month and 96-month visits. If the 48-month X-ray was not acquired the previous visit with a knee radiograph was used, most often the 36-month visit.

Each dataset contains the reading data for a single visit, or time point. Within each dataset, a knee may have readings from project 15 only or from both Projects 15 and 37 (or 42) (indicated by the variable READPRJ). The corresponding longitudinal data from reading projects by this vendor for other time points are in separate datasets. For example, the data from this vendor's reading of baseline X-rays are in a dataset ending in "00" (i.e., kXR_SQ_BU00), while the corresponding data from this vendor for the 12-month visit X-ray readings are in a dataset with the same name but ending in "01" (i.e., kXR_SQ_BU01),

and so on. Please note that no X-rays were acquired at 18-month (V02) and 30-month (V04) clinic visits, which were only performed on a subset of participants, nor were X-rays acquired at the 60-month (V07) and the 84-month (V09) phone contacts. Therefore, the 24-month visit is V03, the 36-month visit is V05, the 48-month visit is V06, the 72-month visit is V08, and the 96-month visit is V10. For a guide to visit numbering in OAI, see "*VisitPrefixDefinitions.pdf*".

Important similarities with previously released datasets

- kXR_SQ_BUxx datasets include KLG scores for each knee at each time point, with possible values 0, 1, 2, 3, 4.
- kXR_SQ_BUxx datasets include JSN scores for each knee at each time point, with possible values of 0, 1, 1.x, 2, 2.x and 3. Values of 1.x and 2.x indicate partial (or within) grade changes compared to previous time points (see 2.3 and 3.3 for further information).
- The OUTCOMES99 dataset includes variables that summarize JSN progression across all time points for each knee (see 1.3.2 and Appendix A).
- The OUTCOMES99 dataset includes variables that summarize KLG≥2 incidence across all time points for each knee (see 1.3.2 and Appendix A), using two different definitions of incidence: "incident KLG≥2 with or without JSN" and "incident KLG≥2 with JSN". (See 1.3.2 and Appendix A.)

Important differences vs. previously released datasets

- Knees with follow-up X-rays at 72- or 96-months were read for Project 37 (or 42) only if the knees were still eligible to develop incident KLG≥2 OA with JSN after 48 months, or if the knee was in a random selection of knees that were KLG 2 or 3 at baseline.
- The OUTCOMES99 dataset variables that summarized whether medial or lateral JSN progression had occurred have now been replaced by variables that summarize whether whole-grade JSN progression or partial grade JSN progression have occurred, at which visit narrowing occurred, and in which tibio-femoral compartment (see 1.3.2 and Appendix A).

Additional details about the specifics of defining radiographic OA, incidence and progression, and reading procedures for Projects 15 and 37 (or 42) are given in separate Appendices (A-E) to this document.

1.3 Merging datasets to assess longitudinal changes and incident OA

IMPORTANT: All users are strongly encouraged to read the *Overview and Description of Central Image Assessments* document carefully before attempting to use this dataset!

1.3.1 Assessing longitudinal changes in variables

To compare values of a variable across serial time points from the same project by this vendor, or to calculate change scores, it is necessary that the kXR_SQ_BUxx datasets for the various time points are merged taking into account the reading project number in the variable READPRJ.

Project 15 data can be used to examine serial KLG and JSN scores over the first 48 months of the study.

To examine serial KLG and JSN scores over the entire 96 months of OAI, kXR_SQ_BUxx datasets with Projects 15 and 37 (or 42) must be combined. **However, we encourage users to first determine whether the summary JSN progression and summary incident KLG≥2 variables in the OUTCOMES99 data are appropriate for an analysis as this can save time and effort (see Appendix A for more information about this).**

We understand that there are many other situations where it may be necessary to merge together data from Project 15 and Project 37 (or 42). Appendix B should be read carefully before attempting this, since it provides examples of the advantages and disadvantages of various different methods of merging and analyzing the data.

IMPORTANT: Please read APPENDIX A carefully for information about pre-calculated outcome variables for development and progression of radiographic knee OA in OAI, and APPENDIX B for information about strategies for merging together the data from Project 15 and Project 37 (or 42).

1.3.2 Variables for incidence and progression of OA in OUTCOMES99 dataset

The OAI SAS dataset OUTCOMES99 contains all the variables needed to analyze the status of radiographic knee OA (ROA) in OAI (i.e.: does a knee have prevalent ROA? Is there incident ROA and at which visit? Does a knee with ROA progress due to a change in JSN?). Please see Appendix A of this document and the documentation provided with the OUTCOMES99 dataset for more information.

OUTCOMES99 has variables that distinguish between two different definitions of incident radiographic OA:

(1) Development of KLG≥2 from KLG 0 or 1 regardless of whether either medial or lateral compartments exhibited any JSN at the time point when incidence occurs;

(2) Development of KLG≥2 from KLG 0 or 1 where the knee has both osteophytes AND an OARSI JSN score ≥1 in either medial of lateral compartments at the time point when incidence occurs.

Definition #1 can be considered the traditional definition of incident OA (requiring only development of a definite osteophyte), and Definition #2 can be considered a modified definition which requires both presence of a definite osteophyte and definite JSN at the visit of incidence. Further information and examples of these definitions of incident radiographic OA are described by Felson *et al*[3].

Notice that all knees that meet Definition #2 at a specific visit, also meet definition #1 at that visit, although they may have met Definition #1 at an earlier visit if the presence of a definite osteophyte preceded development of definite JSN.

The OUTCOMES99 dataset includes summary variables that indicate whether each type of incident OA has occurred in the right and left knees, and at which time point (visit) each definition is first met. (Table 1).

Table 1. Variables in the OUTCOMES99 dataset which are used for defining whether incident radiographic OA has occurred during the study, and at which visit

Definition #1 (Traditional)	**RIGHT knee variable**	**LEFT knee variable**
First visit for which KLG≥2 regardless of JSN	V99ERXIOA	V99ELXIOA
Last visit for which KLG≥2 regardless of JSN	V99ERKLOA	V99ELKLOA
Last visit for which KLG<2	V99ERXNOA	V99ELXNOA
Definition #2 (Modified)	**RIGHT knee variable**	**LEFT knee variable**
First visit for which KLG≥2 and JSN≥1	V99ERXIOAN	V99ELXIOAN
Last visit for which KLG≥2 and JSN≥1	V99ERKLOAN	V99ELKLOAN
Last visit for which KLG<2 or at which KLG=2 and JSN=0	V99ERXNOAN	V99ELXNOAN
Duration of follow-up available		
Last visit for which KLG and JSN scores are available	V99EXLVSQD	

There are also variables to indicate whether JSN has worsened: the first visit at which it has worsened, and in which tibio-femoral compartment the worsening initially occurred (Table 2). There are separate sets of variables for (a) worsening defined as a JSN score increase of a whole grade or more from baseline, and (b) worsening defined as JSN score increase of partial grade or more from baseline. In a similar manner to variables for incident radiographic OA, there are also variables for the last visit at which JSN has worsened from baseline, and also for the last visit at which JSN has not worsened from baseline. (See Appendix A for more information.)

Table 2. Variables in the OUTCOMES99 dataset which are used for defining whether worsening of JSN has occurred during the study, and at which visit

Definition #1 (Whole grade)	**RIGHT knee variable**	**LEFT knee variable**
First visit for which JSN is (whole grade or more) larger than at baseline	V99ERJSFW	V99ELJSFW
Last visit for which JSN is (whole grade or more) larger than at baseline	V99ERJSLW	V99ELJSLW
Last visit for which JSN is the same as at baseline (excluding partial grade narrowing)	V99ERNJSLW	V99ELNJSLW
Tibio-Femoral compartment in which whole grade or more narrowing initially occurred	V99ERJSTFW	V99ELJSTFW
Definition #2 (Partial grade)	**RIGHT knee variable**	**LEFT knee variable**
First visit for which JSN is (partial grade or more) larger than at baseline	V99ERJSFP	V99ELJSFP
Last visit for which JSN is (partial grade or more) larger than at baseline	V99ERJSLP	V99ELJSLP
Last visit for which JSN is the same as at baseline	V99ERNJSLP	V99ELNJSLP
Tibio-Femoral compartment in which whole grade or more narrowing initially occurred	V99ERJSTFP	V99ELJSTFP

Effect of reading blinded to chronological order. It is important to note that because the readers were blinded to the time point of follow-up X-rays, KLG at a later visit may be less than the KLG at an earlier visit. Therefore, a variable has been included that indicates the last time point at which a knee did not meet the definition for each type of incidence (Table 1). Sometimes the visit for the first time point at which the definition was met will be earlier than the last visit at which the definition was not met. For example, a knee may have been scored for the first time as KLG=2 at 24 months, but as a result of being blinded to order the readers scored the 48-month X-ray as KLG=1.

Such knees are ones which “wobble” from KLG≥2 and KLG<2 during the study, and users need to consider how to analyze data from such knees carefully. For example, sensitivity analyses can test whether associations with incident knee ROA are robust to whether or not knees that "wobble" are included in the analysis as incident OA.

The variable V99EXLVSQD is provided to give information on the length of follow-up data available for a participant. For some participants, there are gaps of substantial duration between the last visit for which KLG and JSN scores are available and the visit prior to that with reading data.

Further information and examples of how to use these variables are given in Appendix A of this document and in the documentation of the OUTCOMES99 dataset.

Appendix B provides information about ways in which data from Project 37 (or 42) and Project 15 can be combined. In addition, to find comprehensive information about the structure and contents of the central image assessment datasets, general guidelines for defining change between time points and duration of follow-up, and strategies for merging datasets for analysis please see "*Overview and Description of Central Image Assessments*".

1.4 Condition

- Known data errors: problems/cautions for use are listed by variable in the "Release Comments", which can be found in the various *kXR_SQ_BUxx_Comments.pdf* files.

- Since the last release of these datasets, a small number of data entry errors were found in some of the previously released scores.
- Dataset strengths/weaknesses:
    - Data are expected for all participants included in a project sample. If expected data do not exist for a knee, SAS special missing values are assigned to denote why the data were not obtained.
    - For some knees, this dataset contains two rows of data (2 records) for the knee at a given time point, and other OAI datasets may also contain multiple records per knee. This needs to be taken into account when merging with other datasets. Please see "*Overview and Description of Central Image Assessments*" for more information on merging.

    - To examine longitudinal changes in SQ readings for a specific project, records for that project in the kXR_BU_SQ00 datasets are used for baseline visit values, and then longitudinal changes from those values are found by using values from the same variable in the follow-up kXR_BU_SQxx datasets which have the same values for ID, SIDE and READPRJ.
    - Merging together only records for Project 15 provides longitudinal data for baseline thru 48-month visits.
    - Merging together only records for Project 37 (or 42), typically provides longitudinal data only for baseline, 48-month, 72-month and 96-month visits (although if the participant had no 48-month radiograph, the closest previous visit with one was used, usually the 36-month visit).
- Overlap between reading projects in this dataset:
    - Unlike the last version of this dataset, many knees have two records in this dataset.
    - Knees with no 72-month and no 96-month visit knee radiographs only have one record (in Project 15).
    - Knees which were scored KLG=2 at baseline only have one record (in Project 15) even if they had 72-month or 96-month visit X-rays, <u>except</u> for a random selection of knees which were KLG=2 at baseline in which 72-month and 96-month visit x-rays were read in Project 37 (or 42).
    - Knees which were scored KLG=3 at baseline in Project 15 only have one record (in Project 15) even if they had 72-month or 96-month visit X-rays, <u>except</u> for a random selection of knees which were KLG=3 at baseline in which 72-month and 96-month visit x-rays were read in Project 37 (or 42).

- Knees which were scored KLG 0 or 1 at baseline and then either remained KLG 0 or 1 at all visits up to 48 months or who developed KLG=2, but were still JSN=0 at 48 months (in Project 15) will have one record in Project 15 and will have a second record for Project 37 (or 42) if they had 72-month and/or 96-month knee X-rays.
- Separate datasets (kXR_SQ_Rel_BUxx) of reliability data, where knees read for Project 15 were reread, are available to assess reliability of KLG, JSN and other IRF scores.

- Definitions of **incident radiographic OA**:
  - In both Project 15 and Project 37 (or 42), we have collected data which can be used to create two different definitions of incident radiographic OA. See Appendix A for further information about how these are defined and used in the OAI Outcomes dataset (OUTCOMES99).
  - Since data for incidence of radiographic knee OA are now in the OUTCOMES99 dataset, variables VxxXRNWKL2 and VxxXRNW2N which were previously in the kXR_BU_SQxx datasets have been unreleased since they are redundant and difficult to use.

1.5 Variables and reading methods
See the dataset documentation file *kXR_SQ_BUxx_Contents.pdf* in the compressed documentation file for a complete list of all the variables in the dataset, their SAS variable names, descriptive variable labels and attributes.

Variables assessed for all projects by this vendor include:

- Kellgren and Lawrence (K&L) grades(1);
- individual radiographic features (IRFs) such as osteophytes and joint space narrowing in specific anatomic locations, based on published atlases(2).

Any additional project-specific variables are described below in the section related to that project (Section 2 for Project 15, Section 3 for Project 37 (or 42))

The reading methods and variables are similar in the different projects, and are outlined below. Any project-specific methods are described in the section related to that project, including:

- radiographic features of OA scored at each time point;
- adjudication procedures;
- criteria for selecting the knees to analyze.

Readers. In general, two expert readers independently assessed each X-ray, blinded to each other's reading and to a subject's clinical data. Baseline and follow-up X-rays were scored while being viewed simultaneously and with the readers partially blinded to chronological order of the images. In Project 15, the baseline visit X-ray was known, but follow-up images were randomly ordered. In Project 37 (or 42), the baseline visit, and the last previously read visit (usually 48-month visit) were known, but the 72-month and 96-month visit images were randomly ordered. In Project 15, all features were scored de-novo, but in Project 37 (or 42), only the 72-month and 96-month visits were scored de-novo, and scores for the baseline and 48-month visit (or closest previously scored visit) were provided to the readers. The readers were allowed to change KLG, JSN and other IRFs from the two previously scored visits, if they felt that this was necessary based on new information provided by the extended duration of follow-up now being viewed.

Adjudication. Pre-specified discrepancies between readers for selected variables were adjudicated in a consensus session with blinded films viewed simultaneously and a third reader participating. Details of the various discrepancies requiring adjudication are defined for each project later in this document, but typically include disagreements about presence/absence of radiographic OA or worsening of joint space narrowing (JSN) grade. The values in this dataset are the final, adjudicated readings. When a variable was not adjudicated for discrepancies, the value in the dataset is from the senior reader, an experienced musculoskeletal radiologist.

Non-integer grades for joint space narrowing (JSN). Non-integer JSN variables at a follow-up visit represent the same OARSI grade as a prior visit, but differ by a fraction of an integer to denote definite progression within an OARSI grade. This means that non-integer grades of x.2, x.4, x.6 and x.8 are used, and indicate whether JSN has narrowed, but by less than a whole OARSI grade. An example of this would be a knees scored JSN=2 at baseline, JSN=2.2 at 12-month follow-up, JSN=2.4 at 24-month follow-up, JSN=2.4 at 36-month follow-up, and JSN=3 at 48-month follow-up. This would indicate that the knee has an OARSI JSN grade=2 at all visits, except the 48-month visit, but that the readers had agreed that further narrowing had occurred between baseline and 12-months, and between 12-months and 24-months, and that there was no narrowing between 24-month and 36-months. There was then also narrowing between 36-months and 48-months, and that the 48-month visit OARSI grade was 3. For more detail about the method see Felson *et al* (4).

IMPORTANT NOTE: These non-integer grades SHOULD NOT be used as equivalent to specific fractional amount of joint space narrowing (i.e. 2.2 to 2.4 does not represent two tenths of a grade change) but rather indicate only a perception of a qualitative increase in JSN within the same OARSI grade compared to another time point.

Other individual radiographic features. Apart from Kellgren and Lawrence grades and JSN grades, which are read in all knees at all time points, various other IRFs are read longitudinally only when the participant has KLG≥2 on at least one time point in either knee. When these IRFs were not read the missing data is indicated by a special missing value as indicated below.

Missing data. Missing data can occur for a variety of reasons. In the individual kXR_SQ_BUxx datasets, SAS special missing values are used to indicate the following:

.P if data is missing due to a prosthesis/knee replacement.

.T if data is missing due to technical reasons (e.g. poor image quality).

.A if the data is not expected (e.g.: some V00XR…may be missing if the participant has KLG<2 in both knees at all time points).

Missing data can also occur after merging data from different visits and reading projects. For example, a participant may have V00XR…, V03XR…, V05XR…, and V06XR… scores but if that person did not attend their 12-month visit, they would have no data in the kXR_SQ_BU01 dataset, but in a merged dataset containing longitudinal data from baseline thru 48-month visits, that participant would have missing values for their V01XR… variables.

2. Methods specific to Project 15

2.1 Image type
Fixed-flexion bilateral knee radiograph. See the "*Radiographic (X-ray) Manua*l" for the acquisition protocol.

2.2 Time points
Baseline, 12-month, 24-month, 36-month and 48-month visits.

2.3 Measurement methods
For all readings performed for Project 15, the readers know the true baseline visit X-ray, but were blinded to the chronological order of follow-up visit X-rays, with grouped images from all available time points viewed simultaneously. The readers were also blinded to the true OAI ID of the participant and hence to any existing clinical or radiological data about the participant.

Reading Procedures for Kellgren and Lawrence and Joint Space Narrowing Grades
Each participant's X-rays were initially scored by the primary reader for Kellgren and Lawrence Grades (KLG)[1] and OARSI Joint Space Narrowing (JSN) grades[2] at each time point. Non-integer grades (X.2, X.4, X.6 and X.8) for joint space narrowing at a visit represent the same OARSI grade compared to another time point with grade X, but denoting a definite difference in JSN within a grade between time points. If some of the participant's X-rays had previously been read for Project 06, the reader was presented with the results from that previous reading, but was allowed to modify the scores.

If the primary reader scored both knees as KLG 0 or 1 at all visits, these images were then read by a second reader to confirm the absence of radiographic OA (all KLG=0 or 1) at all visits. If this was confirmed, then the primary readers KLG and JSN scores are the only readings available for the participant for Project 15. If the second reader did not confirm the absence of radiographic OA for a participant, then the two readers performed a consensus reading of KLG and JSN grades for the participants X-rays. If no OA at all visits was confirmed by consensus, the consensus results became the final reading data for the participant for Project 15.

If the primary reader scored either knee as KLG≥2 at any visit, or if the above consensus reading resulted in a KLG≥2 at any visit, then the images were read by a second reader who independently scored KLG and JSN grades, again using non-integer grades for JSN to represent a definite change from baseline. For participants previously read for Project 06, the second reader was also presented with the result of those previous readings, but was allowed to modify the scores.

Following this secondary reading of KLG and JSN grades, specific discrepancies between the two readers required further adjudication and the adjudication procedures are described in Appendix C.

Non-integer JSN grades representing "within-grade" change
For this project, readers know the true baseline visit X-ray, and when they consider that a definite change has occurred at a blinded follow-up visit, but that the change is not large enough for the follow-up visit to be scored using a different OARSI JSN grade, then non-integer grades (X.2, X.4, X.6, X.8) have been used to identify such narrowing. The actual magnitude of that non-integer value does not indicate the amount of narrowing, and only indicates that changes have occurred between visits. So that the following longitudinal series of scores (for example), represent equivalent time course of JSN changes:

V00=1	V01=1.2	V03=1.4	V05=2	V06=2
V00=1	V01=1.4	V03=1.8	V05=2	V06=2

where the score is grade 1 at baseline (V00), 12-month (V01) and 24-month (V03) follow-up visits and grade 2 at 36-month (V05) and 48-month (V06) follow-up visits, but the readers have recorded that the joint space becomes more narrow between V00 and V01 and also between V01 and V03, and then again between V03 and V05, and that by V05 the narrowing was enough to be scored grade=2. No further narrowing had occurred at V06.

Readings of additional IRFs in participants with definite radiographic OA

For any participants for whom the primary and secondary reader agreed on presence of radiographic OA (or for whom the adjudication process confirmed presence of radiographic OA) in either knee at any time point, all the X-rays for that participant were then read by the senior reader for additional IRFs (e.g.: osteophytes, sclerosis, chondrocalcinosis).

A flowchart illustrating the reading protocol in more detail is provided in Appendix C.

2.4 Variables

Variables are prefixed VxxXR, and KLG and OARSI JSN grades are available for all participants in this project at all available time points. Scores for other IRFs (osteophytes, subchondral sclerosis, cysts and attrition) are available only in participants with definite radiographic OA at least one knee at one (or more) of the time points.

2.5 Sample

Project 15 now contains data for all OAI participants for whom knee X-rays were acquired at two or more visits. The following table gives some demographic information about the participants with data currently released for Project 15:

Project 15 sample: Distribution of Race by Sex

	White or Caucasian	Non-White	Total
Male	1,597	282	1,879
Female	2,004	605	2,609
Total	3,601	887	4,488*

*Race data missing for 4 participants

Project 15 sample: Distribution of Age by Sex

	Age (years)				
	45 to 49	50 to 59	60 to 69	70 to 79	Total
Male	240	705	476	461	1,882
Female	270	861	893	586	2,610
Total	510	1,566	1,369	1,047	4,492

3. Methods specific to Project 37 (or 42)

3.1 Image type
Fixed-flexion bilateral knee radiograph. See the "*Radiographic (X-ray) Manua*l" for the acquisition protocol.

3.2 Time points
Baseline, 48-month, 72-month and 96-month visits, although if a knee did not have a 48-month visit the closest previous visit (usually the 36-month visit) was used.

3.3 Measurement methods
Almost all knees in this project had previous scores available from baseline thru 48-month visits from Project 15, except for a few participants who did not have knee X-rays acquired between baseline and the 72-month and 96-month visits. For those people only Project 37 (or 42) readings exist.

For Project 37 (or 42), images and scores from baseline and the latest previously scored visit (usually the 48-month follow-up visit) were provided to the readers, in known chronological order. If no 48-month visit knee radiograph was available, the closest previous visit (usually the 36-month visit) was used. The readers were then presented with the 72-month and 96-month visit knee radiographs blinded to chronological order. Images from all four time points were viewed simultaneously, along with the scores from the 2 earlier visits from Project 15. The readers were also blinded to the true OAI ID of the participant and hence to any existing clinical or radiological data about the participant.

Reading Procedures for Kellgren and Lawrence and Joint Space Narrowing Grades
Each participant's X-rays were independently scored by the primary and secondary reader for Kellgren and Lawrence Grades (KLG) and OARSI Joint Space Narrowing (JSN) grades at the 72-month and/or 96-month time points. Non-integer grades (X.2, X.4, x.6 and X.8) for joint space narrowing at a visit represent the same OARSI grade compared to another time point with grade X, but denoting a definite difference in JSN within a grade between time points. For the participant's X-rays that had previously been read for Project 15, the reader was presented with the results from that previous reading, but was allowed to modify the scores.

Specific discrepancies between the two readers required further adjudication and the adjudication procedures are described in Appendix E.

Non-integer JSN grades representing "within-grade" change
For this project, readers know the true baseline visit X-ray and the latest previously scored visit, and when they consider that a definite change has occurred at a blinded follow-up visit, but that the change is not large enough for the follow-up visit to be scored using a different OARSI JSN grade, then non-integer grades (X.2, X.4, X.6, X.8) have been used to identify such narrowing. The actual magnitude of that non-integer value does not indicate the amount of narrowing, and only indicates that changes have occurred between visits. So that the following longitudinal series of scores (for example), represent equivalent time course of JSN changes:

V00=1	V06=1.2	V08=1.4	V10=2
V00=1	V06=1.4	V08=1.8	V10=2

where the score is grade 1 at baseline (V00), 48-month (V06) and 72-month (V08) follow-up visits and grade 2 at the 96-month (V10) follow-up visit, but the readers have recorded that joint space becomes more narrow between V00 and V06 and also between V06 and V08, and that by V10 the narrowing was enough to be scored grade=2.

Readings of additional IRFs in participants with definite radiographic OA
For any participants for whom the primary and secondary reader agreed on presence of radiographic OA (or for whom the adjudication process confirmed presence of radiographic OA) in either knee at any time point, all the X-rays for that participant were then read by the senior reader for additional IRFs (specifically: osteophytes and chondrocalcinosis).

3.4 Variables
Variables are prefixed VxxXR, and KLG and OARSI JSN grades are available for all participants in this project at all available time points. Scores for other IRFs (osteophytes, subchondral sclerosis, cysts and attrition) are available only in participants with definite radiographic OA at least one knee at one (or more) of the time points.

Please note that although some participants are coded READPRJ=42, they are in fact participants in Project 37. Users should recode these participants from READPRJ=42 to READPRJ=37.

3.5 Sample
Project 37 (or 42) contains data for all knees with X-rays acquired at the 72-month and/or 96-month visit and that were KLG 0, 1, or KLG 2 with JSN=0 at 48 months, and therefore eligible to develop incident radiographic OA with JSN (KLG≥2 with JSN≥1) after the 48-month visit. These knees were read to provide a complete dataset of incident radiographic OA over the entire 96-month follow-up of OAI. Project 37 (or 42) also includes a small number of knees for participants who only have baseline along with 72-month and/or 96-month visit knee X-rays. These participant's X-rays had never been previously read but are now included in these datasets.

Project 37 (or 42) also contains readings for all knees with 72-month and/or 96-month which were KLG 0 or 1 at baseline and had already developed KLG≥2 with JSN≥1 by 48 months to determine rates of progression over the 48-month to 96-month follow-up period.

A random selection of knees which were KLG 2 or 3 at baseline also have 72-month/96-month visit readings available in Project 37 (or 42).

The following table gives some demographic information about the participants with data currently released for Project 37 (or 42):

Project 37 sample: Distribution of Race by Sex

	White or Caucasian	Non-White	Total
Male	904	122	1,026
Female	1055	231	1,286
Total	1,959	353	2,312*

*Race data missing for 2 participants

Project 37 sample: Distribution of Age by Sex

	Age (years)				
	45 to 49	50 to 59	60 to 69	70 to 79	Total
Male	151	421	260	195	1,027
Female	154	473	424	236	1,287
Total	305	894	684	431	2,314

4. References

1. Kellgren JH, Lawrence JS. Atlas of Standard Radiographs. Oxford; 1963.
2. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15 Suppl A:A1-56 PMID: 17320422 http://dx.doi.org/10.1016/j.joca.2006.11.009
3. Felson DT, Niu J, Guermazi A. Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. Ann Rheum Dis 2011; 70: 1884-1886. PMID: 21908453 PMCID: PMC3653624 http://www.dx.doi.org/10.1136/ard.2011.155119
4. Felson DT, Nevitt MC, Yang M, Niu J, Torner JC, Lewis CE, et al. A New Approach Yields High Rates of X-Ray Progression in Knee Osteoarthritis (OA).J Rheumatol 2008. PMID: 18793000 PMCID: PMC2758234

Appendix A. Assessing incidence and progression of radiographic OA.

A.1 Variables and Definitions of Incident Radiographic OA

Although users can use the data from the kXR_SQ_BUxx datasets for individual time points to examine incidence and progression of radiographic knee OA themselves, we suggest that they use the variables in OUTCOMES99 dataset. In that dataset, we provide information on the incidence of radiographic OA (ROA) in terms of two definitions:

1. Development of KLG≥2 from a knee that was KLG 0 or 1 at baseline, regardless of the presence/absence of joint space narrowing (OARI JSN can be 0 at the visit of incidence).
2. Development of KLG≥2 from a knee that was KLG 0 or 1 at baseline in which OARSI JSN score≥1 in either compartment at the visit of incidence.

Definition #1 can be considered the traditional definition of incident OA (requiring only development of a definite osteophyte), and definition #2 can be considered a modified definition which requires both presence of a definite osteophyte and definite JSN at the visit of incidence. Further information and examples of these definitions of incident radiographic OA are described by Felson *et al*[1].

Notice that all knees that meet definition #2 at a specific visit, also meet definition #1 at that visit, although they may have met definition #1 at an earlier visit if the presence of a definite osteophyte preceded development of definite JSN.

The variables for the traditional (KLG≥2) definition of radiographic OA (definition #1) are:

Indicator for visit of incidence (KLG ≥2 with or without JSN):

V99ERXIOA (right knee) and **V99ELXIOA** (left knee): indicate that the knee remained KLG 0 or 1 throughout the follow-up period, or that the knee was already KLG≥2 at baseline, or the first visit at which KLG≥2 (regardless of JSN score).

Indicator for last visit without ROA:

V99ERXNOA (right knee) and **V99ELXNOA** (left knee): indicates the last visit at which the knee was KLG 0 or 1, which is usually the visit before which incidence occurred.

The variables for the modified (KLG≥2, JSN≥1) definition of radiographic OA (definition #2) are:

Indicator for visit of incidence (KLG≥2 with JSN):

V99ERXIOAN (right knee) and **V99ELXIOAN** (left knee) indicate the first visit at which KLG≥2 **and** JSN ≥ 1, or whether the knee was KLG ≥2 and JSN ≥ 1 at baseline, or whether the knee remained KLG 0, 1 or KLG 2 with JSN=0 throughout the follow-up period.

Indicator for last visit without ROA+JSN:

V99ERXNOAN (right knee) and **V99ELXNOAN** (left knee) indicate the last visit at which the knee was KLG 0 or 1 or for which KLG=2 and JSN=0, which is usually the visit before which incidence occurred.

Please note that since readings were performed blinded to chronological order (see section A.3 for details), there are some knees that develop incident radiographic OA at early follow-up visits, but which do not meet the matching definition for incident radiographic OA at later time points. For such situations the indicator for the last visit without radiographic OA will be later than the first visit at which radiographic OA occurred.

Table A.1 shows the frequencies with which incident OA is found in the current readings, showing that definition #1 has higher rates of incidence than definition #2 (as expected). Also notice that there are 4056/4796 OAI participants with data on incident ROA in the OUTCOMES99 dataset. The remaining 290 participants had no follow-up visit knee X-rays to read.

Table A1. Frequency Tables of the variables for incident ROA from OUTCOMES99 by the 2 different definitions

(1) Traditional Definition : KLG ≥2 with or without JSN:

Variable Label: Summary incident TF ROA KL >= 2 (calc)	Variables:	**RIGHT KNEE** V99ERXIAO	**LEFT KNEE** V99ELXIAO	**TOTAL** (knees)
1: KL≥2 at BL		2032	1934	3966
2: no incident KL >= 2		2103	2200	4303
3: incident KL>=2 at 12 mo		97	99	196
4: incident KL>=2 at 24 mo		45	49	94
5: incident KL>=2 at 36 mo		57	54	111
6: incident KL>=2 at 48 mo		40	34	74
7: incident KL>=2 at 72 mo		84	71	155
8: incident KL>=2 at 96 mo		48	66	114
		4506	4506	9013

(2) Modified Definition: KLG ≥2, JSN ≥1:

Variable Label: Summary incident TF ROA KL >= 2 with JSN (calc)	Variables:	**RIGHT KNEE** V99ERXIAON	**LEFT KNEE** V99ELXIAON	**TOTAL** (knees)
1: KL>=2 with JSN at BL		1620	1506	3126
2: no incid KL>=2 with JSN		2567	2684	5251
3: incident KL>=2 w/JSN at 12m		71	85	156
4: incident KL>=2 w/JSN at 24m		51	45	96
5: incident KL>=2 w/JSN at 36m		47	44	91
6: incident KL>=2 w/JSN at 48m		43	34	77
7: incident KL>=2 w/JSN at 72m		59	54	113
8: incident KL>=2 w/JSN at 96m		48	55	103
		4506	4506	9013

Defining incidence for knees with data from both Project 15 and 37 (or 42): As described previously, some knees had their baseline through 48-month visit knee radiographs read for Project 15 and then had their later visits (usually paired readings of baseline, 48-month, 72-month and 96-month visits) read for Project 37 (or 42). Kellgren and Lawrence Grade and JSN Grades from both projects were examined and the earliest visit at which incident ROA was found from either project was used.

To illustrate this method, consider the RIGHT KNEE of participant with ID 9061666 and the Kellgren and Lawrence Grades from Projects 15 and 37, and in this knee, JSN=1 at all visits in both projects, except medial compartment JSN=2 at 96 months in Project 37.

READPRJ	**ID**	**SIDE**	**V00XRKL (BL)**	**V01XRKL (12m)**	**V03XRKL (24m)**	**V05XRKL (36m)**	**V06XRKL (48m)**	**V08XRKL (72m)**	**V10XRKL (96m)**
15	9061666	**1: Right**	1	1	1	1	1	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook
37	9061666	**1: Right**	1	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	2	2	3

In Project 15 (READPRJ=15), the knee is scored KLG=1 at all visits. But in Project 37, the readers altered the 48-month visit score to KLG=2 and by 96-month visit KLG=3.

So although the knee was scored no incident ROA (i.e.: KLG=1) at all visits in Project 15, the knee was scored KLG=2 at 48-month visit in Project 37 so the earliest visit of incidence is the 48-month visit for this participant's RIGHT knee:

9061666, Right knee: V99ERXIOA="6: incident KL>=2 at 48 mo"

Also for this knee, since medial compartment is JSN=1 at 48-months:

9061666, Right knee: V99ERXIOAN="6: incident KL>=2 w/JSN at 48 mo"

A.2 Variables and Definitions for Progression of OA

As can be seen from the example of the right knee of participant 9061666, between 72-month and 96-month visits, the KLG score changed from KLG=2 to KLG=3 and the medial JSN score changed from JSN=1 to JSN=2. This "whole-grade" change in OARSI JSN score in a knee with radiographic OA can be considered JSN progression. As pointed out in section 1.3 of the kXR_SQ_BUxx dataset documentation, OAI also scores "within-grade" changes in OARSI JSN Scores which users may want to consider as progression(2). We have therefore defined a new set of outcomes variables that differentiate whether a knee has narrowed based on two definitions, one based on whether the OARSI JSN score had worsened by a whole grade or more since baseline and the 2nd definition based on whether the OARSI JSN score had worsened by a partial grade or more since baseline

These four variables are related to whether the OARSI JSN score worsened by a whole grade or more:

Indicator for the first visit at which JSN score was a whole grade or more larger than at baseline

V99ERJSFW (right knee) and **V99ELJSFW** (left knee): whether a knee remained at the same OARSI JSN score at all visits, or (when it has narrowed by a whole grade or more) the visit by which that narrowing first occurred.

Indicator for last visit at which OARSI JSN score was the same as at baseline

V99ERNJSLW (right knee) and **V99ELNJSLW** (left knee): indicate the latest visit at which the OARSI JSN score was the same as at baseline, ignoring partial grade changes.

Indicator for the last visit at which JSN score was a whole grade or more larger than at baseline

V99ERJSLW (right knee) and **V99ELJSLW** (left knee): indicate the latest visit at which the OARSI JSN score was a whole grade or more larger than at baseline.

Indicator for which tibio-femoral compartment narrowed by a whole grade of more

V99ERJSTFW (right knee) and **V99ELJSTFW** (left knee): indicate the tibio-femoral compartment in which the worsening of JSN score by a whole grade or more first occurred.

The variables V99ERJSFW for the right V99ELJSFW for the left can be used to easily determine whether a knee joint space has narrowed compared to baseline and by which visit. In general, it is expected that the last visit at which the OARSI JSN score was the same as the baseline score (V99ERNJSLW, V99ELNJSLW) would be earlier than the first visit at which the JSN score has worsened (V99ERJSFW, V99ELJSFW).

Knees where V99ERNJSLW is later than V99ERJSFW are those right knees which JSN initially narrows but later widens; the variable V99ERJSLW allows users to determine whether a knee narrows, then widens and then narrows again. The same can be done for the left knees using V99ELNJSLW, V99ELJSFW and V99ELJSLW. For knees which narrow and then widen again, perhaps only those that remain narrowed by their latest visit may be considered as progressing. Users should develop their own methods for determining how to handle knees which narrow from baseline to a follow-up visit, but then widen again at a later time point.

These four variables are related to whether the OARSI JSN score worsened by a partial grade or more:

Indicator for the first visit at which JSN score was a partial grade or more larger than at baseline
V99ERJSFP (right knee) and **V99ELJSFP** (left knee): whether a knee remained at the same OARSI JSN score at all visits, or (when it has worsened by a partial grade or more) the visit by which that narrowing first occurred.

Indicator for last visit at which OARSI JSN score was the same as at baseline
V99ERNJSLP (right knee) and **V99ELNJSLP** (left knee): indicate the latest visit at which the OARSI JSN score was the same as at baseline, including any partial grade worsening as a change.

Indicator for the last visit at which JSN score was a partial grade or more larger than at baseline
V99ERJSLP (right knee) and **V99ELJSLP** (left knee): indicate the latest visit at which the OARSI JSN score was a partial grade or more larger than at baseline.

Indicator for which tibio-femoral compartment narrowed by a partial grade of more
V99ERJSTFP (right knee) and **V99ELJSTFP** (left knee): indicate the tibio-femoral compartment in which the worsening of JSN score by a partial grade or more first occurred.

Similar to the whole JSN change variables, there are variables that allow users to determine whether a knee remained narrowed throughout the remaining follow-up visits, or both narrowed and widened at follow-up visits.

Table A.2 shows the numbers of knees showing progression from baseline using both definitions of progression.

Table A2. Frequency Tables of the variables for JSN Progression from OUTCOMES99 by the 2 different definitions

(1) JSN worsens by a whole grade or more from baseline:

Variable Label:		**RIGHT KNEE**	**LEFT KNEE**	**TOTAL**
First visit with JSN Progression (whole grade or more)	Variables:	V99ERJSFW	V99ELJSFW	(knees)
0: Baseline [these knees do not progress]		3623	3709	7332
1: 12-month		203	166	369
3: 24-month		105	115	220
5: 36-month		108	96	204
6: 48-month		111	113	224
8: 72-month		84	70	154
10: 96-month		79	64	143
		4313	4333	8646

(2) JSN worsens by a partial grade or more from baseline:

Variable Label:		**RIGHT KNEE**	**LEFT KNEE**	**TOTAL**
First visit with JSN Progression (partial grade or more)	Variables:	V99ERJSFP	V99ELJSFP	(knees)
0: Baseline [these knees do not progress]		3390	3447	6837
1: 12-month		351	334	685
3: 24-month		163	170	333
5: 36-month		120	116	236
6: 48-month		117	115	232
8: 72-month		91	81	172
10: 96-month		81	70	151
		4313	4333	8646

IMPORTANT NOTE: Knees which are JSN=3 in either medial and/or lateral tibio-femoral compartment at baseline are considered end stage; they are not eligible for progression and have missing values for the JSN progression outcome variables. When a knee progressed in both medial and lateral compartments, the compartment with narrowing (compared to baseline) at the earliest follow-up visit was indicated as the involved compartment. If both compartments narrowed at the same time, the compartment which was most narrowed at baseline was indicated. For the few knees in which both compartments narrowed at the same time and had the same JSN score at baseline, both compartments are indicated as the involved compartments.

Table A.3 shows the frequencies of which compartment(s) narrowed from baseline using (a) the whole grade or more definition of progression and (b) the partial grade or more definition of progression.

Table A3. Frequency Tables of the variables from OUTCOMES99 for which compartment showed JSN Progression

(1) JSN worsens by a whole grade or more from baseline:

Variable Label:		**RIGHT KNEE**	**LEFT KNEE**	**TOTAL**
TF Compartment with initial JSN Progression (whole grade)	Vars:	V99ERJSTFW	V99ELJSTFW	(knees)
0: None [these knees do not progress]		3623	3709	7332
1: Lateral		217	185	402
2: Medial		467	434	901
3: Medial and Lateral		6	5	11
		4313	4333	8646

(2) JSN worsens by a partial grade or more from baseline:

Variable Label:		**RIGHT KNEE**	**LEFT KNEE**	**TOTAL**
TF Compartment with initial JSN Progression (partial grade)	Vars:	V99ERJSTFP	V99ELJSTFP	(knees)
0: None [these knees do not progress]		3390	3447	6837
1: Lateral		280	237	517
2: Medial		632	640	1272
3: Medial and Lateral		11	9	20
		4313	4333	8646

A.3 Effects of reading blinded to chronological order on apparent regression of KLG and JSN score
When the baseline through 48-month knee radiographs were read for Project 15, the baseline X-ray was identified but the other time points were presented in random order to the readers. As a result, knees were sometimes given a score for KLG or IRFs that were lower than the scores assigned for earlier time points, an occurrence that is much less likely to happen when X-rays are presented to readers in known chronological order. When this does occur, it frequently results from substantial differences in radiographic position between time points.

As an example, knees that were scored as incident radiographic OA using the traditional definition (incident definite osteophytes KLG≥2 with or without JSN) at 24 months may have been scored KLG 0 or 1 at 36 and/or 48 months. The summary variables provided for incident OA in the OUTCOMES99 dataset (V99ERXIOA and V99ELXIOA, right and left knees respectively) in this example would have the knee as incident radiographic OA at 24 months and do not take into account the KLG 0 or 1 scores at subsequent time points.

As of version 6 the OUTCOMES99 dataset includes new variables (V99ERXNOA and V99ELXNOA for right and left knee respectively), that indicate the last visit at which the knee DID NOT meet the definition for incident ROA. These can be used to determine whether a knee with incident radiographic OA “wobbled”

between “incident” and “not incident” status subsequent to the first time point it was scored as incident for the traditional definition (incident KLG≥2 with or without JSN).

Usually, the last visit at which the knee did not meet the incident ROA criteria would be a visit prior to the visit of incidence. “Wobblers” may be identified as knees where the last visit of non-incidence occurs AFTER the visit of OA incidence.

There are analogous variables for the modified definition of incident ROA that also requires presence of JSN (incident KLG≥2 with JSN≥1). For this modified definition, V99ERXIOAN indicates the visit of incidence in right knees and V99ERXNOAN indicates the last visit of non-incidence. For left knees, there are two equivalent matching variables V99ELXIOAN and V99ELXNOAN.

“Wobblers” are usually knees scored as incident osteophytes KLG≥2 without JSN (which were previously designated as incident KLG 2N), but “wobbling” also sometimes occurred in knees that had been scored as incident osteophytes KLG≥2 with JSN≥1 at one time point but were scored as KLG 0, KLG 1 or incident osteophytes without JSN (KLG=1, JSN=0) at one or more later time points.

IMPORTANT NOTE: When analyzing incident OA, users may decide not to count these “wobbling” knees as incidence endpoints, either in their main analyses or in sensitivity analyses. These incidence "wobblers" also vary in the number of discrepant time points following the first incident time point, and users may wish to base decisions about individual knees based on the number of such discrepant time points or the last available KLG score.

When the 72-month and 96-month knee radiographs were read for incident OA in Project 37 (or 42), readers were presented with the baseline and latest previous visit (usually 48-month) X-rays in known order, along with their existing KLG and JSN scores, followed by the unscored 72- and 96-month X-rays in random order. The readers had the option of changing their previous reading of the earlier time points to be consistent with how they now intended to score the later unread time points. This sometimes resulted in knees that were previously scored as incident OA by 48 months being rescored as not incident by 48 months.

Appendix B contains information and examples illustrating how the OA status of a knee at any particular visit, or over the entire 96 months of follow-up in OAI, can differ between Projects 15 and 37 (or 42) due to differences in the reading methods and how this can affect outcomes when data from the two projects are merged.

A.4 References

1. Felson DT, Niu J, Guermazi A. Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. Ann Rheum Dis 2011; 70: 1884-1886. PMID 21908453 PMCID PMC3653624 http://www.dx.doi.org/10.1136/ard.2011.155119.
2. Felson DT, Nevitt MC, Yang M, Niu J, Torner JC, Lewis CE, et al. A New Approach Yields High Rates of X-Ray Progression in Knee Osteoarthritis (OA).J Rheumatol 2008. PMID: 18793000 PMCID: PMC2758234

Appendix B. Special Consideration for Use of Radiographic OA Incidence Data from Projects 37 (or 42) & 15

B.1 Knees with different radiographic knee OA Status in Project 37 (or 42) and Project 15
As has been pointed out elsewhere in this document, all knees in OAI with follow-up X-rays had their baseline through 48-month visit X-rays read for Kellgren and Lawrence Grade (KLG) and OARSI Joint Space Narrowing (JSN) grades in Project 15. In contrast, Project 37 (or 42) involved reading 72-month and 96-month knee X-rays only for knees that were KLG 0 or 1 or KLG 2 with JSN=0 at 48 months (or the last time point with SQ reading data) from Project 15. In Project 37 (or 42) knees were read paired with baseline and (usually) 48-month visit X-rays (with their existing scores provided to readers). Readers for Project 37 (or 42) were allowed to alter scores for previously read visits if they considered that necessary based on the new information from the additional time points.

This means that in the 3,834 knees read for Project 15 that also had readings for Project 37 (or 42), the radiographic OA status at specific visits may differ between projects. Apart from minor disagreements such as KLG=0 in one project vs KLG=1 in the other project, only 67 knees were affected by an altered radiographic knee OA status. These broke down into 4 different categories:

(1) 40 knees with incident KLG≥2 (JSN=0) by 48-months in Project 15 which are scored KLG 0 or 1 at both baseline and 48-months in Project 37 (or 42); these subdivide into:
    a. 29 remain KLG 0 or 1 at 72 months and 96 months in Project 37 (or 42);
    b. 11 incident KLG≥2 (JSN=0) or KLG≥2 (JSN≥1) at 72 months or 96 months in Project 37 (or 42).
(2) 17 knees with no ROA (i.e. KLG<2) at baseline through 48-months in Project 15 but KLG≥2 (JSN=0) or KLG≥2 (JSN≥1) at 48 months, 72 months and 96 months in Project 37 (or 42).
(3) 5 knees with incident ROA using the “traditional” definition (KLG≥2, JSN=0) by 48 months in Project 15, but in Project 37 (or 42) were scored incident ROA using the “modified” definition (KLG≥2, JSN≥1) at 48 months.
(4) 1 knee scored incident KLG≥2 (JSN≥1) in Project 15, but KLG=2 (JSN=1) at baseline in Project 37 (or 42).
(5) 4 knees scored KLG=2 at baseline in Project 15, but scored KLG=1 at baseline in Project 37 (or 42).

The ID and side of the knees in each category are listed in Table B.2

For these knees, when they have incident OA in either project, it is the earliest visit from either project that is recorded in the OUTCOMES99 summary variables.

For all right knees in the incident ROA categories above, V99ERXIOA records the earliest visit of incidence of KLG≥2 and if the knee had visits with KLG≥2 and JSN≥1, the earliest visit at which that occurred is recorded in V99ERXIOAN. Similarly, for left knees, V99ELXIOA records the earliest visit of incidence of KLG≥2 and if the knee had visits with KLG≥2 (JSN≥1) the earliest visit at which that occurred is recorded in V99ELXIOAN.

B.1.1 Knees scored incident ROA in Project 15 but no ROA scored in Project 37
The 19 knees in category #1(a) above are knees where the OUTCOMES99 dataset records the earliest visit of incidence using the Project 15 readings, but in which later visits at 72 month and 96 month have scores for no ROA from Project 37. For these 19 knees scored KLG 0 or 1 in Project 37, the variables V99ERXNOA (if right knee) or V99ELXNOA (if left knee) take the values “9: 72-month” or “10: 96-month” to indicate that the last visit with no ROA is later than the visit at which they were scored incident.

These 19 knees are considered “wobblers” for incident ROA since they become incident at an intermediate time point and then were scored as no ROA at later visits. See Appendix A.3 for further discussion of “wobblers.”

The data below shows the KLG and JSN scores for participant 9850845, right knee, which belongs to category 1(a). JSN=0 at all visits for this knee. The KLG values “wobbled” in and out of incidence in Project 15, and remained KLG<2 even at 96 months in Project 37:

READPRJ	ID	SIDE	V00XRKL (BL)	V01XRKL (12m)	V03XRKL (24m)	V05XRKL (36m)	V06XRKL (48m)	V08XRKL (72m)	V10XRKL (96m)
15	9850845	1: Right	0	2	0	2	2	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook
37	9850845	1: Right	0	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	0	0	0

Knees such as this, where the last visit that does not have ROA is later than the first visit at which ROA was scored, are often ones which need careful consideration regarding their ROA status in any analyses.

B.1.2 Knees scored incident ROA in Project 15 but with incidence after 48 months in Project 37
For the 7 knees in category #1(b), the “traditional” KLG≥2 definition is met at the 48-month visit in Project 15, but for these knee Project 37 data has KLG<2 at 48 month with incidence at 72 month or 96 months. For these knees, the first visit of incidence recorded in the OUTCOMES99 dataset is the earliest (48 month) visit based on Project 15 readings.

In the example below the KLG scores for one of these knees (9560051, left knee) are shown. This knee was scored JSN=0 at all visits prior to the 72-month visit, in both projects, and then JSN=2 at 72 months and 96 months.

READPRJ	ID	SIDE	V00XRKL (BL)	V01XRKL (12m)	V03XRKL (24m)	V05XRKL (36m)	V06XRKL (48m)	V08XRKL (72m)	V10XRKL (96m)
15	9560051	2: Left	1	1	1	1	2	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook
37	9560051	2: Left	1	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	.: Missing Form/Incomplete Workbook	1	3	3

For this knee, the data in OUTCOMES99 about incident radiographic OA using the traditional (KLG≥2) definition is:

9560051, Left knee: V99ELXIOA=”6: incident KL>=2 at 48 mo”

Using the modified (KLG≥2, JSN≥1) definition of incident radiographic OA, which requires JSN≥1 at the visit of incidence, this knee has:

9560051, Left knee: V99ELXIOAN=”7: incident KL>=2 w/JSN at 72 mo”

B.1.3 Knees with no ROA scored by 48 months in Project 15, but incident ROA at 48 months in Project 37
The 15 knees in category #2 above were scored KLG 0 or 1 at all time points up to 48 months in Project 15. However, when early visit X-rays were paired with 72-month and 96-month visit X-rays, the readers in Project 37 scored the knees incident ROA at 48 months.

These include knees which were previously categorized as no ROA in the first 48 months of the study, which now are categorized in the OUTCOMES99 summary variable as incident ROA (by either the traditional or modified definition) at the 48-month visit based on the Project 37 reading.

B.1.4 Knees which did not meet the "modified" definition of incident ROA in Project 15, but do in Project 37
The 4 knees in category #3, above, were scored incident ROA in Project 15 with JSN scores of 0 at all visits; but when images were read along with 72-month and 96-month visit radiographs for Project 37, both of the readers agreed that the knees had definite JSN at 48 months and therefore in Project 37, they now met the "modified" definition of incident ROA (with KLG≥2 and JSN≥1) at 48 months.

B.2 Comparing the visit of incident ROA with the last visit without ROA
As has been previously pointed out, the visit at which incident ROA was first scored is recorded in the OUTCOMES99 dataset and is almost always later than the last visit at which the knee did not have ROA.

The table (B.1) below, shows the relationship between first visit of incidence and the last visit with no ROA for the traditional (KLG≥2) definition of ROA (KLG≥2 with or without JSN) for right knees as a cross-tabulation of variables V99ERXIOA and V99ERXNOA.

Equivalent tables can be created for LEFT knees and also for the modified (KLG≥2, JSN≥1) definition of ROA which required JSN≥1 at the visit of incidence.

Table B.1 Showing the relationship for RIGHT knees, between the visit of first incident KLG≥2 (with or without JSN) (V99ERXIOA) against last visit KLG<2 (V99ERXNOA). Cells with a **black background** are knees considered "wobblers", since they are recorded KLG<2 after their first visit KLG≥2.

V99ERXIAO Summary incident TF ROA KL >= 2 (right knee)	**V99ERXNOA** Last visit KL < 2 (right knee)							
	0: Baseline	1: 12-month	3: 24-month	5: 36-month	6: 48-month	8: 72-month	10: 96-month	Total
2: KL0-1 at BL, follow-up X-rays but no incident KL >= 2	0	104	97	113	353	46	1390	2103
3: KL0-1 at BL, incident KL2 at 12 mo	**76**	0	**5**	**1**	**5**	**0**	**10**	97
4: KL0-1 at BL, incident KL2 at 24 mo	5	**36**	0	**2**	**0**	**0**	**2**	45
5: KL0-1 at BL, incident KL2 at 36 mo	2	5	**45**	0	**1**	**3**	**1**	57
6: KL0-1 at BL, incident KL2 at 48 mo	1	1	1	**27**	9*	**1**	**0**	40
7: KL0-1 at BL, incident KL2 at 72 mo	0	0	1	5	**73**	0	**5**	84
8: KL0-1 at BL, incident KL2 at 96 mo	1	2	0	3	7	**35**	0	48
Total	85	148	149	151	448	85	1408	2474

* For these 9 knees the visit of incidence differs between projects 15 and 37,resulting in the anomalous situation of the 48 month visit being both the time point of incidence and the last time point with KL <2

There are some important points to note about Table B.1:

(1) The top row shows that there are 2,103 right knees that remain KLG 0-1 at all time points, but some of them only have short length of follow-up (e.g.: 104 have 12 months of follow up, 97 only have 24 months of follow up). There is however a large number of knees (1,390) that remain KLG 0 or 1 throughout the entire 96 months of follow up of OAI.

(2) The knees in cells with a **black background** are "wobblers" which are incident ROA meeting the traditional definition (KLG≥2 with or without JSN), but are then scored KLG 0 or 1 at a later time point (a total of 35 right knees). Although some of these are due to inconsistencies between readings from Project 15 and Project 37, many of them have data from Project 15 alone. Even though these knees are KLG 0 or 1 at their last visit, a visit of incidence is also recorded in the OUTCOMES99 dataset.

(3) Knees on the diagonal with a grey background in the cell are ones where are no missed visits between KLG<2 and KLG≥2.

(4) Knees to the lower left of the table are ones where there are skipped visits with no knee X-ray to score.

For an example of skipped visit, the single right knee in the lower left cell of the table represents participant 9112400 (right knee) and this person only had knee radiographs acquired at baseline and 96-month follow-up visits. Knees in this category in the table are ones where there is a lot of uncertainty as to the actual time at which ROA developed.

Table B.2 Listing knees in which radiographic knee OA status defined from Project 37 KLG/JSN grades differs from the status defined from Project 15 KLG/JSN grades (see Page 20 for further details about the 5 categories described).

(1a) incident KLG=2(JSN=0) by 48 months in Project 15 which are scored KLG 0 or 1 at all visits to 96 months in Project 37

ReleaseID	SIDE
9009927	2: Left
9049447	2: Left
9102297	1: Right
9112976	1: Right
9187595	2: Left
9191533	1: Right
9200057	1: Right
9224983	2: Left
9250756	2: Left
9291210	1: Right
9301641	2: Left
9309716	2: Left
9329345	2: Left
9376430	2: Left
9393127	2: Left
9409501	2: Left
9426421	2: Left
9462266	2: Left
9522615	1: Right
9566429	1: Right
9585175	2: Left
9616919	2: Left
9619405	1: Right
9689083	1: Right
9692923	2: Left
9719999	2: Left
9850845	1: Right
9901658	2: Left
9918304	2: Left

(1b) incident KLG=2(JSN=0) by 48 months in Project 15 which are scored KLG 0 or 1 at all visits to 48 months in Project 37, then KLG≥2 at 72m/96m in Project 37

ReleaseID	SIDE
9034812	1: Right
9218438	1: Right
9254853	2: Left
9369982	1: Right
9412037	2: Left
9560051	2: Left
9674454	1: Right
9717977	1: Right
9717977	2: Left
9723032	1: Right
9854620	2: Left
9034812	1: Right

(2) incident KLG=2 (JSN>=1) by 48 months in Project 15, but which was scored KLG=2 with JSN=1 at baseline in Project 37

ReleaseID	SIDE
9505140	2: Left

(3) prevalent KLG=2 at baseline but scored KLH=1 at baseline in Project 37

ReleaseID	SIDE
9156214	1: Right
9440307	2: Left
9524666	2: Left
9659155	1: Right

(4) no ROA (KLG<2) baseline through 48 months in Project 15 which are scored KLG=2 (JSN=0) or KLG≥2 (JSN≥1) at 48 months, 72 months and 96 months in Project 37

ReleaseID	SIDE
9061666	1: Right
9218916	1: Right
9268052	1: Right
9305065	2: Left
9361281	2: Left
9369286	2: Left
9409198	1: Right
9566495	2: Left
9596248	2: Left
9646127	2: Left
9660697	2: Left
9683704	1: Right
9804177	1: Right
9804177	2: Left
9809967	1: Right
9816138	1: Right

(5) incident KLG=2 (JSN=0) by 48 months in Project 15, but which were scored KLG≥2 (JSN≥1) 48 months in Project 37

ReleaseID	SIDE
9218438	2: Left
9289020	2: Left
9429115	1: Right
9444196	1: Right
9585175	1: Right

B.3 Four different strategies for merging data from Project 15 and Project 37 (or 42)
Rather than using the OUTCOMES99 dataset, there may be situations where users want to merge together data for knees that have readings in both Project 37 (or 42) and Project 15. An example of this may be if a user wants to determine changes in KLG between the 24-month and 72-month visits (two time points that do not usually have reading data from a single project).

The following are 4 possible methods for doing this:

(1) Use Project 15 data and add 72-month and 96-month scores from Project 37 (or 42)
(2) Use Project 37 (or 42) data and insert missing data using Project 15 scores
(3) Use the "worst" (maximum) score from Projects 15 and 37 (or 42)
(4) Use "worst" change score from baseline from Project 15 and Project 37 (or 42)

In general, incident ROA status in OUTCOMES99 agrees with method #3, and it is only for the 45 knees listed in Table B.2 where different merging strategies will affect whether knees meet either the traditional (KLG 2) or modified (KLG≥2, JSN≥1) definitions of incident ROA at specific visits.

JSN progression status in OUTCOMES99 agrees with method #4 since JSN score worsening within a particular project indicates progression; for the outcomes variables, we determine the first visit by which worsening has occurred in this manner.

The only other situation where users need to be careful about how to merge data is when they want to select knees which are radiographically normal (KLG=0) at a visit. Depending on which strategy is used for merging data, knees may switch between KLG=0 and KLG=1 at specific visits. For this situation, it may be best to use option #3 for merging data since it would guarantee: that at any specific visit, a knee with KLG=0 in the merged data will definitely not have KLG=1 at that visit in either of the 2 reading projects.

The same issue will occur when users want to select knees which have JSN=0 or JSN=1 at a particular visit, but this can generally be addressed in the same manner as KLG selection since knees which are KLG=0 at a particular visit would also be JSN=0 at that visit.

This is a list of the 42 knees which were scored KLG=0 at baseline in one reading project and KLG=1 at baseline in the other reading project:

ID	SIDE
9001897	2: Left
9002316	2: Left
9043446	1: Right
9043945	1: Right
9066677	2: Left
9197064	1: Right
9204538	1: Right
9226514	1: Right
9226514	2: Left
9232259	1: Right
9237974	1: Right
9250756	2: Left
9264781	1: Right
9268724	2: Left

ID	SIDE
9327968	1: Right
9346242	1: Right
9362264	2: Left
9362660	2: Left
9373786	1: Right
9416851	2: Left
9493869	2: Left
9544188	2: Left
9621208	1: Right
9626197	1: Right
9631230	1: Right
9634064	1: Right
9635414	1: Right
9656390	2: Left

ID	SIDE
9689083	1: Right
9702463	1: Right
9739273	1: Right
9753580	2: Left
9795669	1: Right
9807724	2: Left
9808818	1: Right
9816138	1: Right
9853562	1: Right
9875794	2: Left
9882107	2: Left
9887094	2: Left
9905276	2: Left
9973569	2: Left

These 23 knees are ones which are KLG=0 at baseline in both projects, but were scored KLG=1 at a follow-up visit in Project 15 and KLG=0 at the same follow-up visit in Project 37 (or vice versa):

ID	SIDE
9066677	1: Right
9068197	2: Left
9100862	1: Right
9125977	1: Right
9127495	2: Left
9131911	1: Right
9287295	2: Left
9300796	2: Left

ID	SIDE
9330186	2: Left
9341449	1: Right
9532533	2: Left
9551556	2: Left
9558611	1: Right
9606490	1: Right
9609732	2: Left
9621886	2: Left

ID	SIDE
9679111	1: Right
9762412	1: Right
9789297	2: Left
9804177	2: Left
9896743	1: Right
9914944	1: Right

Appendix C. Project 15: Reader disagreements and adjudication procedures

Adjudication of discrepancies for a knee between the two readers was done on the key variables listed below (only the variables that the readers did not agree on were adjudicated). When a variable is not adjudicated, the value in the dataset is from the primary (and most senior) reader.

For each discrepancy between the primary and a second reader that was identified, a third expert reader reviewed all the scores from both readers and the grouped radiographs. If the third reader agreed with either of the primary or second reader's scoring, then the agreed upon scoring was determined to be final. If the third expert reader did not agree with either reader, the three readers attended an adjudication session in which a consensus scoring was obtained.

1. Kellgren and Lawrence Grades (KLG) (KLG<2 (No OA) vs. KLG≥2 (Definite OA)) and KLG progression

The readings were adjudicated if the primary and second reader did not agree on:

- OA vs no OA status at any time point
- Change in OA status based on KLG, and direction of change, between any pair of time points. Change in OA status is defined as present when there is a change from KLG 0-1 to KLG 2-4, or vice versa, between time points.
- Change in KLG, and direction of change, between any pair of time points.
    - KLG change was not adjudicated if both readers scored KLG change in the same direction but disagreed on the amount of change (e.g. one reader scored an increase in 1 grade; the other reader scored an increase of 2 grades and there was no change in OA status).

2. Joint Space Narrowing (JSN) change

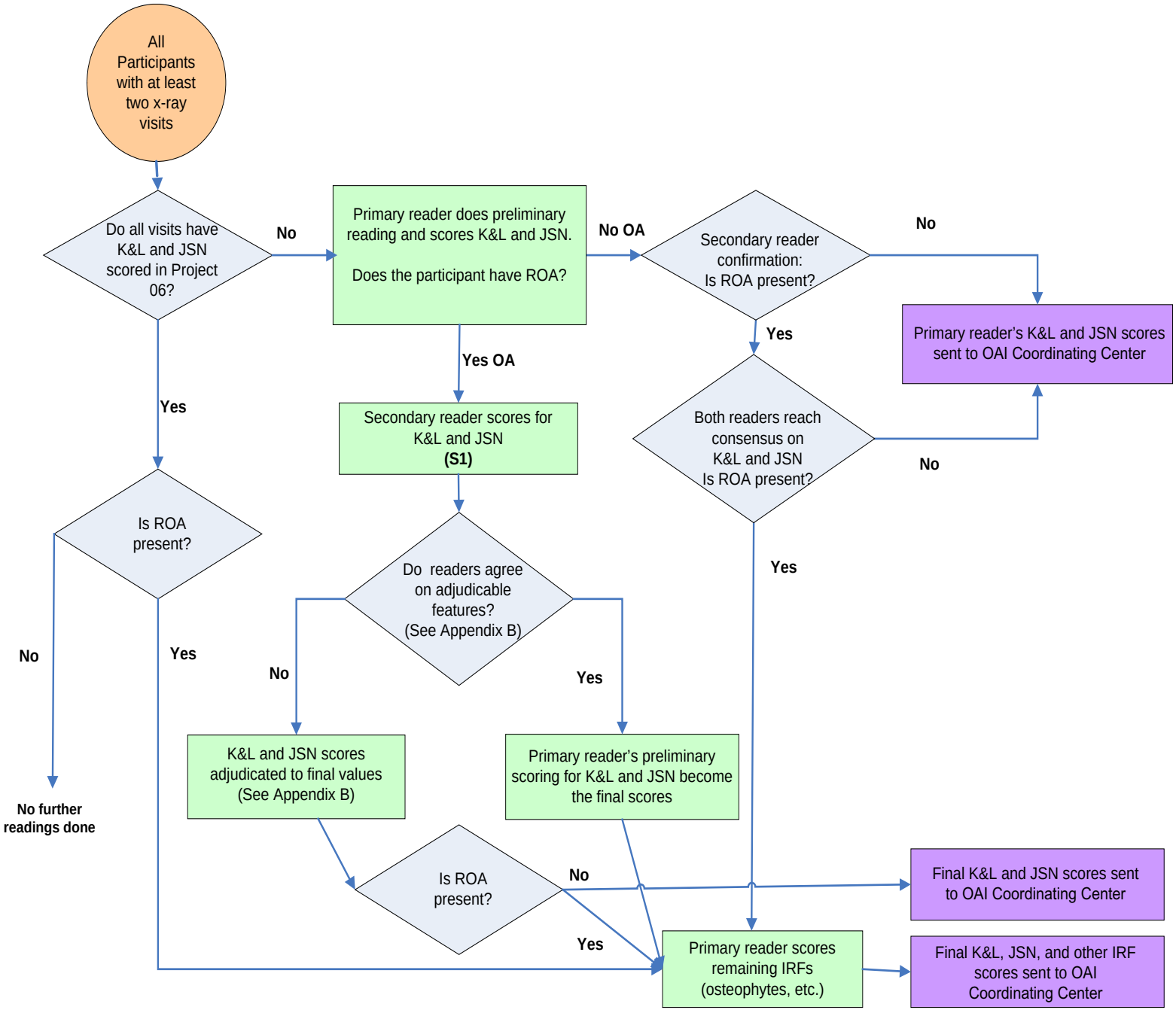
The readings were adjudicated if the two readers did not agree on:

- Change in JSN, and direction of change, evaluated in the medial and the lateral compartments separately, between any pair of time points. Change in JSN grade is defined as present when there is a change >0 in JSN score (in either direction) between two time points. Fractional "within-grade" changes in JSN grade were considered a change.
    - JSN change was not adjudicated if both readers scored JSN change in the same direction but disagreed on the amount of change (e.g. one reader scored a change of 0.2 grades and the other reader scored a change in the same direction but of a greater magnitude).

Appendix D. Project 15: Flowchart of X-ray reading process

The figure on the next page outlines the manner in which participant's knee X-rays were read for Project 15. All participants in the OAI cohort with at least two fixed flexion knee radiographs were examined (top left). Participants with all their X-rays previous read in project 06 already had adjudicated joint space narrowing (JSN) and Kellgren and Lawrence Grades (K&L) grades available, and for those in which radiographic OA was present complete IRFs (osteophytes, etc) were then scored at all time points (represented by the very left side, and bottom row of the figure). All remaining participants then proceeded through the flow outlined in the remainder of the figure, until either only K&L and JSN grades, or a more complete set of radiographic features, were sent to the OAI Coordinating center (represented by the boxes on the right hand side of the figure).

Appendix D. Project 15: Flowchart of X-Ray Reading Process.
-continued-



Updated 3-29-2011

Appendix E. Project 37 (or 42): Reader disagreements and adjudication procedures

Adjudication of discrepancies for a knee between the two readers was done on the key variables listed below (only the variables that the readers did not agree on were adjudicated). When a variable is not adjudicated, the value in the dataset is from the primary (and most senior) reader.

For each discrepancy between the primary and a second reader that was identified, the two readers attended an adjudication session in which a consensus scoring was obtained.

1. Kellgren and Lawrence Grade (KLG <2 (No OA) vs. KLG ≥2 (Definite OA)) and KLG progression

The readings were adjudicated if the primary and second reader did not agree on:

- OA vs no OA status at any time point
- Change in OA status based on KLG, and direction of change, between any pair of time points. Change in OA status is defined as present when there is a change from KLG 0-1 to KLG 2-4, or vice versa, between time points.
- Change in KLG, and direction of change, between any pair of time points.
    - KLG change was not adjudicated if both readers scored KLG change in the same direction but disagreed on the amount of change (e.g. one reader scored an increase in 1 grade; the other reader scored an increase of 2 grades and there was no change in OA status).

2. Joint Space Narrowing (JSN) change

The readings were adjudicated if the two readers did not agree on:

- Change in JSN, and direction of change, evaluated in the medial and the lateral compartments separately, between any pair of time points. Change in JSN grade is defined as present when there is a change >0 in JSN score (in either direction) between two time points. Fractional “within-grade” changes in JSN grade were considered a change.
    - JSN change was not adjudicated if both readers scored JSN change in the same direction but disagreed on the amount of change (e.g. one reader scored a change of 0.2 grades and the other reader scored a change in the same direction but of a greater magnitude).