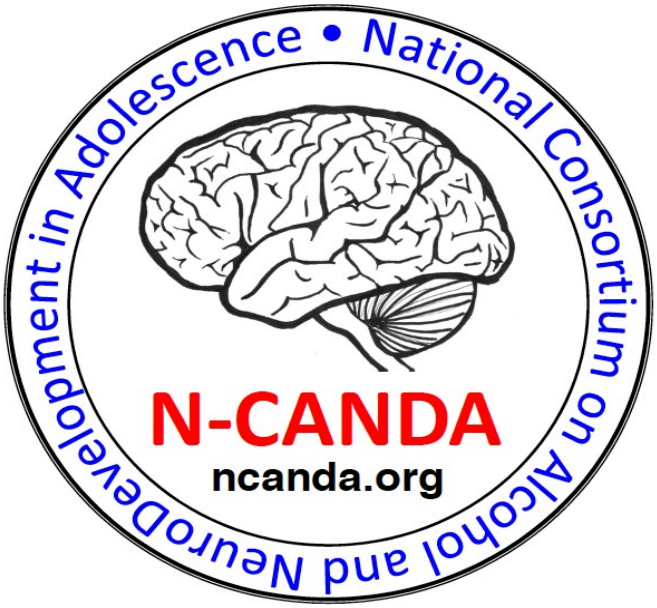
National Consortium on Alcohol and NeuroDevelopment in Adolescence



Data Management Manual

February 27, 2014

The National Consortium on Alcohol and NeuroDevelopment in Adolescence

Funded by the National Institute on Alcohol Abuse and Alcoholism (NIAAA)

Administrative Component

University of California, San Diego (1 U01 AA021695)
Sandra A. Brown, PhD, and Susan F. Tapert, PhD

Data Integration and Informatics Component

SRI International (1 U01 AA021697)
Adolf Pfefferbaum, MD, and Torsten Rohlfing, PhD

Data Collection Sites

Duke University (1 U01 AA021681)
Michael D. DeBellis, MD

SRI International (1 U01 AA021696)
Ian M. Colrain, PhD, and Fiona Baker, PhD

Oregon Health Sciences University (1 U01 AA021691)
Bonnie J. Nagel, PhD

University of California, San Diego (1 U01 AA021692)
Susan F. Tapert, PhD

University of Pittsburgh (1 U01 AA021690)
Duncan Clark, MD, PhD

The latest version of this document is available from
http://www.nitrc.org/docman/view.php/672/1216/NCANDA_Data_Manual.pdf

For inquiries regarding this manual, contact Torsten Rohlfing, torsten@synapse.sri.com

Table of Contents

1 Subject Numbering Scheme

1.1 Consortium ID

Each subject is assigned a "Consortium ID" according to the following scheme:

S-NNNNN-G-C

where:

S ____one-character site ID, assigned by the Administrative Component. Site Ids have been communicated to the Site PIs. **The assignment of IDs to sites is confidential to avoid geographic localization of subjects.**

N_____five-digit subject "sequence" number. Returning subjects are not assigned a new number but keep their previously assigned number.

G_____subject gender code, "M" for male and "F" for female, "P" for phantom (MR imaging only), "T" for "test subjects" (these will be ignored when importing submitted data into the consortium database)..

C ____one-digit check sum computed from the subject's birth date as follows:

Write numerical subject birth date, using 4-digit year, e.g., for 24th October 1998, write:

10/24/1998

Add all digits of the birth date and keep last digit of the result as check sum, C. In the example above,

1+0+2+4+1+9+9+8 = 34

C = 4

Note: Not every software package used by NCANDA to capture raw data allows for this format. See below for how to adapt the primary format to each program.

Note also: Once assigned, subject IDs should never changed. In particular, they should be entered for follow-up visits exactly the same as they were entered for baseline data collection.

1.2 Subject ID Translation for Capture Instruments with Constraints

N-CANDA ID:	A-12345-M-0	B-99999-F-5
<u>Constrained Instruments:</u>		
Screener:	A-12345	B-99999
WebCNP:	A12345M0	B99999F5
Stroop (ePrime):	1123-4510	2999-9925
FrACT:		???
<u>Unconstrained Instruments:</u>		
Delay Discounting:	A-12345-M-0	B-99999-F-5
PASAT:	A-12345-M-0	B-99999-F-5
Interview/SR:	A-12345-M-0	B-99999-F-5
SSAGA (Blaise):	A-12345-M-0	B-99999-F-5

Rules for Adapting the Primary ID

Drop hyphens if the character count constrained
Pad with leading zeroes for fixed-length
If only numbers are allowed:

- Drop hyphens
- Encode site ID as 1 (A) through 5 (E)
- Encode Gender as M=1, F=2, Test=9

2 N-CANDA REDCap Server

The N-CANDA REDCap (Research Electronic Data Capture) server can be accessed by pointing a web browser to the following address:

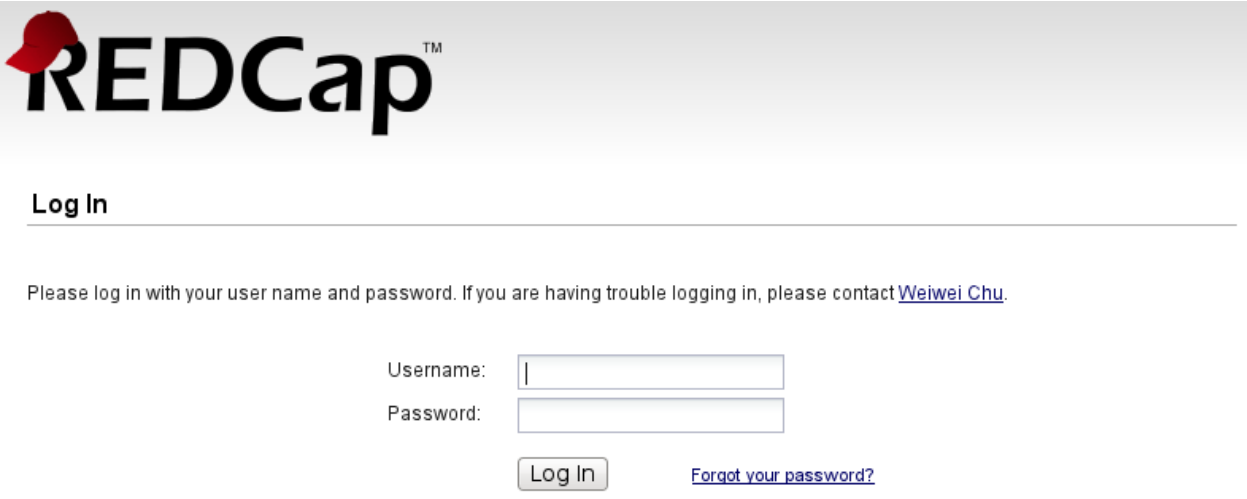
https://ncanda.sri.com/redcap/

In order to access the server, a user account must be created by the Data Analysis Component. To request accounts, the Site PI should send the names and email addresses of the requested users to the Data Analysis Component PI, Torsten Rohlfing, at torsten@synapse.sri.com.

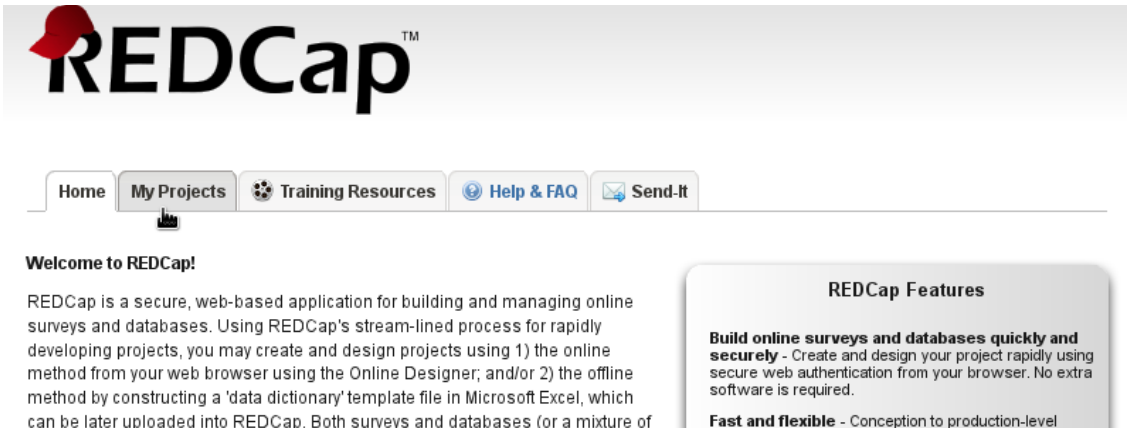
User accounts on the N-CANDA REDCap server should be used only for the following tasks:

1. Entry of subject visits (see next section)
2. Manual correction and editing of data captured by the site laptops (see section below)
3. Query of the compiled dataset imported from the data collection laptops

To connect to the N-CANDA REDCap database, point your web browser to the server address provided above, then enter your user name and password into the login screen:



You will then see the main REDCap screen, where you should click on the “My Projects” tab:



2.1 Summary of Projects

2.1.1 N-CANDA Data Entry

The “N-CANDA Data Entry” project in REDCap is the primary project for the N-CANDA data. This is where subject and visit information are to be entered, and it is also where data imported from other sources (e.g., WebCNP, MRI, data capture laptops) are compiled.

2.1.1.1 Direct Data Entry

The following forms are intended for direct data entry. Any necessary changes should be made directly in these forms (see Section 3 for details):

- Basic Demographics
- Visit Date and Notes
- NP: Ishihara
- NP: Edinburgh Handedness Inventory
- NP: WRAT4 Word Reading and Math Computation
- NP: Grooved Pegboard
- NP: Rey-Osterrieth Complex Figure PDF Uploads
- NP: Rey-Osterrieth Complex Figure Scores
- NP: Modified Gregly-Graybiel Test of Ataxia
- NP: WAIS-IV Coding
- Biological Data: NP Day
- Biological Data: MR Day
- Saliva Samples

2.1.1.2 Data Imported from Site Laptops

The following projects gather data originating from the sites' data capture laptops. Any necessary changes should be made in the corresponding forms in the "N-CANDA Imported from Laptops" project (see Section 5 for details).

- Delayed Discounting, $1000
- Delayed Discounting, $100
- Paced Auditory Serial Addition Test (PASAT)

New data will be imported from the site laptops every night. Any changes made in the "Imported from Laptops" project will also be merged into these forms nightly.

2.1.1.3 Data Imported from Penn WebCNP

The following project gathers data from the WebCNP database at U Penn. Any necessary corrections should be made in the "N-CANDA Imported from PennCNP" project instead (see Section 6 for details).

- Penn WebCNP Summary

New data are downloaded from Penn and merged nightly. Any corrections made in the "Imported from PennCNP" project will also be merged into these forms nightly.

2.1.1.4 Data Imported from MRI Database

The following project gathers information on MRI sessions stored in the N-CANDA XNAT image database system (see Section 7 for details).

- MRI Session Report

The data in this form is updated nightly.

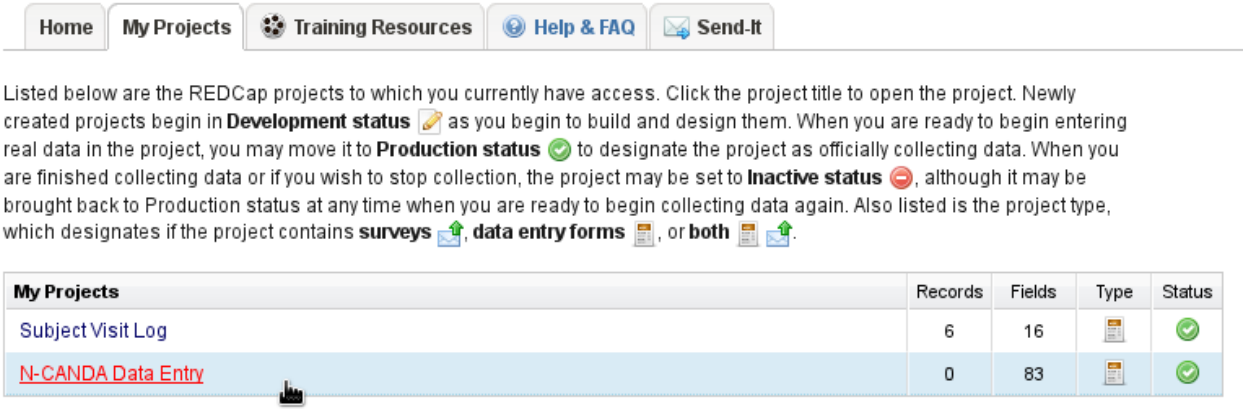
Any necessary changes and corrections should be made directly in the XNAT database, which is reached via https://ncanda.sri.com/xnat. See also the **N-CANDA MRI and Image Management Manual** for XNAT-specific instructions.

3 Data Entry Instructions: Subject Visit Dates and NP Instruments

Currently, the following instruments are to be entered into REDCap, as opposed to being entered into the data collection laptops using LimeSurvey:

- Ishihara
- Edinburgh Handedness Inventory
- WRAT4 Word Reading / Math Computation
- Grooved Pegboard
- Rey-O
- Fregly-Braybiel Ataxia
- WAIS-IV Coding

To enter data for a new subject visit or scores for an existing subjects, select the "N-CANDA Data Entry" project on the REDCap front page:

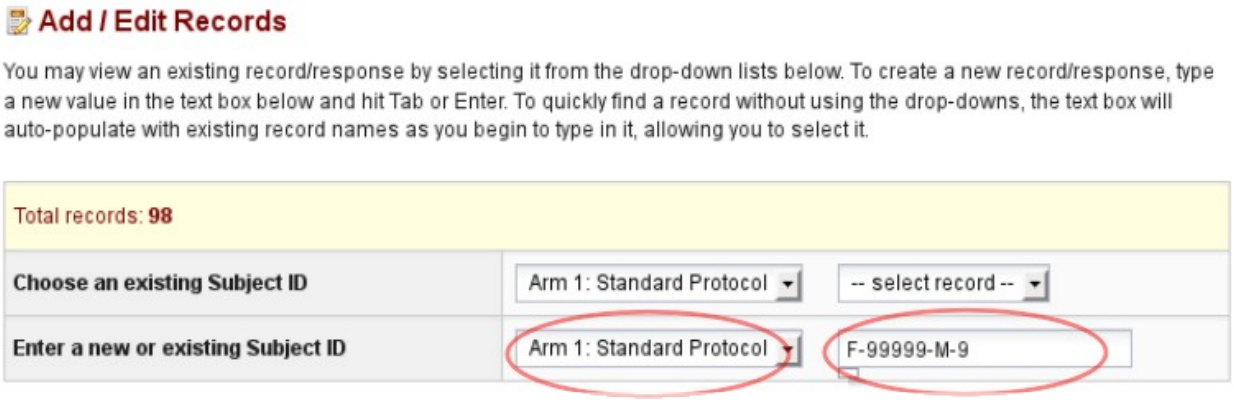


Next, the "N-CANDA Data Entry" project page will appear. To enter new data or edit existing data, find and click on the "Add / Edit Records" link under "Data Collection" in the menu on the left side of the screen.

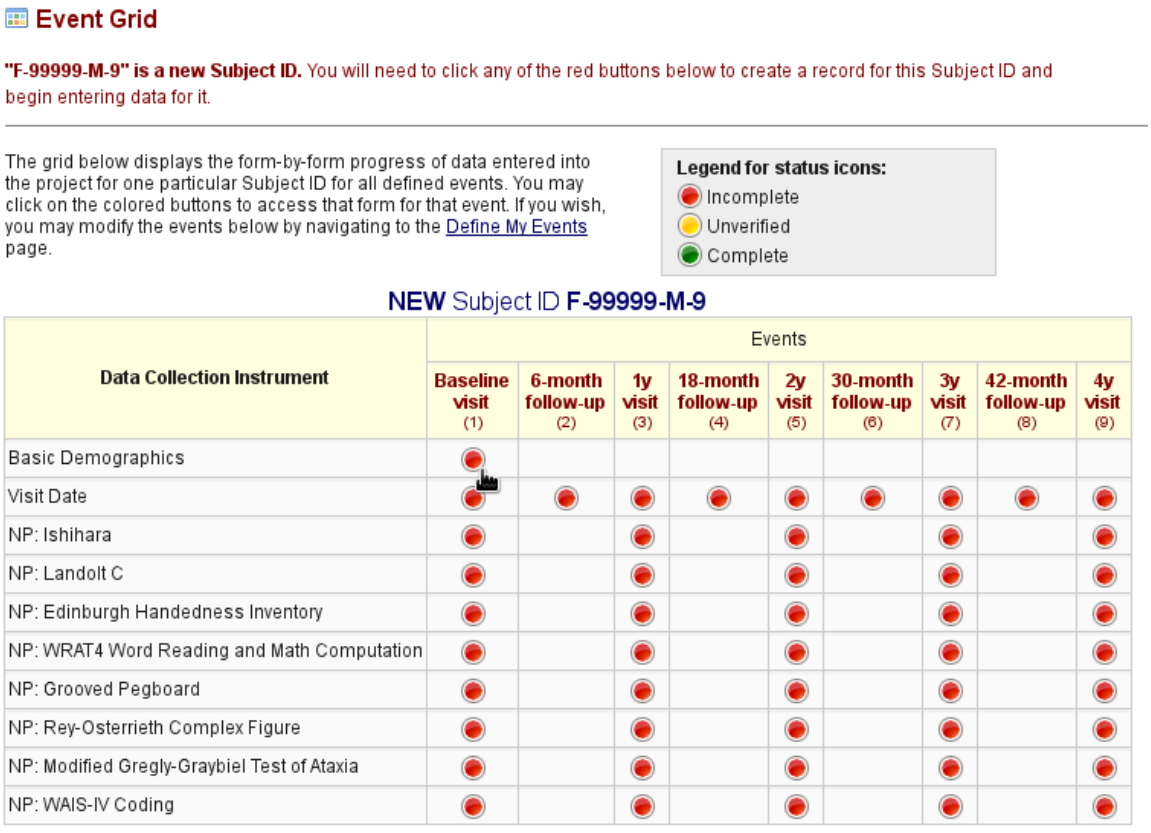
3.1 New Subject Entry

To enter a new subject, type the subject ID into the entry field to "Enter a new or existing Subject ID" in the form shown below. Only well-formed N-CANDA Consortium subject IDs will be accepted.

Make sure that the selection box before the ID entry field remains set to "Arm 1: Standard Protocol."

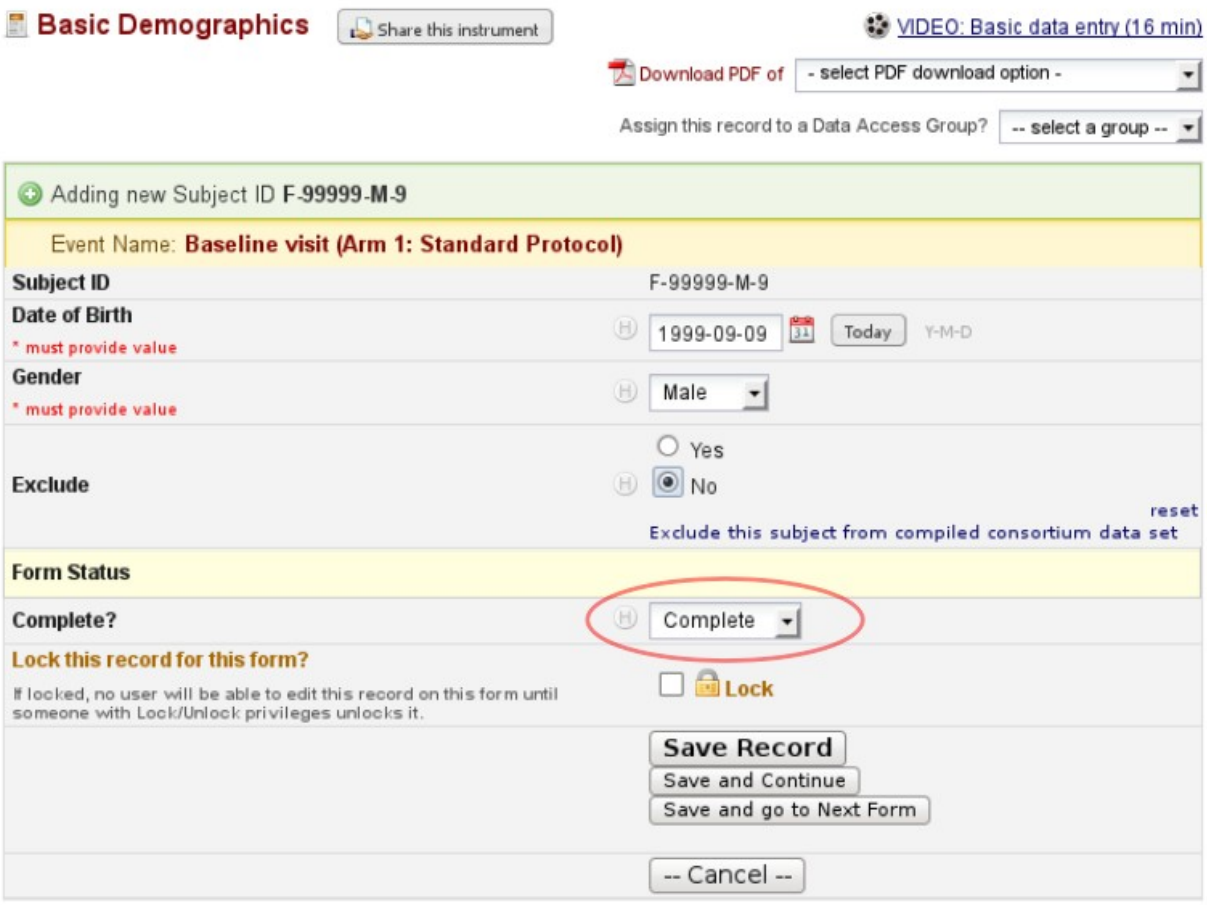


An "Event Grid" (shown below) will now appear, which will show the entry status of all data entry forms, many of which are repeated for some or all of the visits throughout the study.



To start new subject entry, click on the "Basic Demographics" field.

The “Basic Demographics” are subject Date of Birth and Gender. Also select whether this subject should be excluded from the compiled consortium dataset (in case of exclusion, a “Reason for Exclusion” entry field will also appear.)

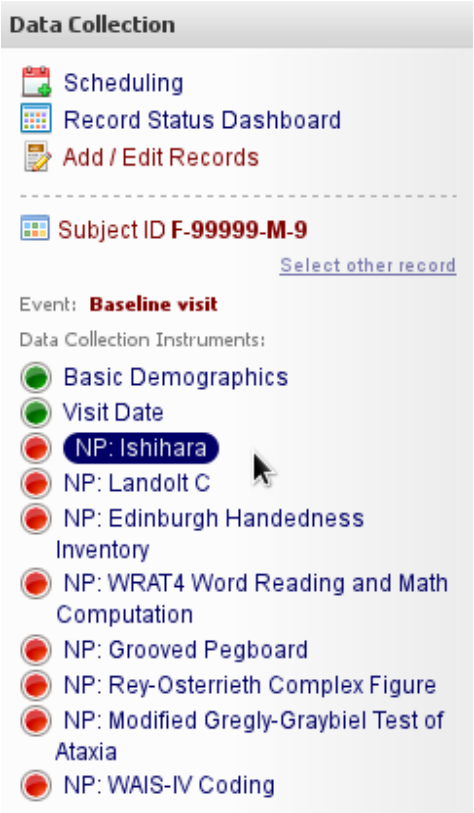


After entering **all information**, select “Complete” in the selection box at the bottom of the form. Then, leave the form by pressing either the “Save Record” button (will take you back to the “Event Grid” or the “Save and go to Next Form” button (this will take you to the entry of the baseline visit date).

In the “Visit Date” entry, **enter the first day on which a subject was seen for a given visit** (here: “Baseline”), **regardless of what instruments (including MRI) were administered that day**.

Note that the subject age is automatically computed.

After entering the visit date, set completion status to “Complete” and leave the entry mask via either the “Save Record” or “Save and go to Next Form” button at the bottom.

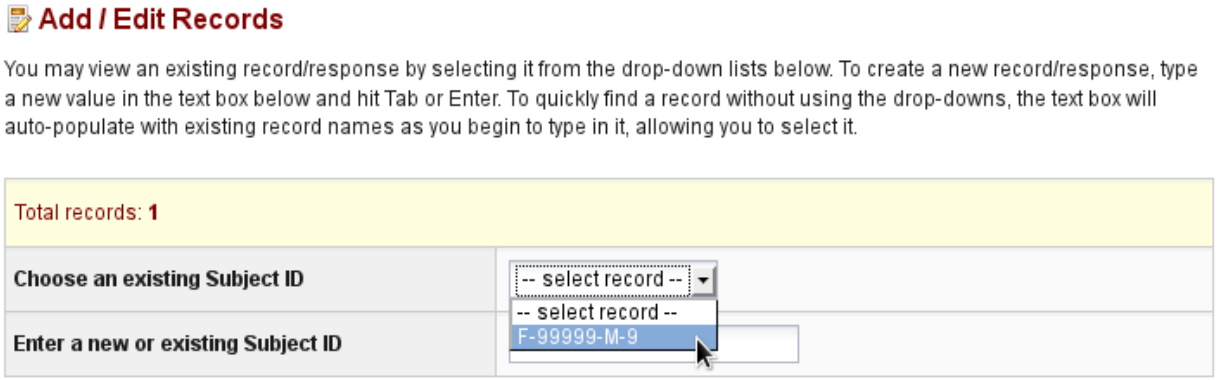


During date entry, the menu on the left side of the screen will show all “Data Collection Instruments” in the “Data Collection” section. Here, completed instruments will be marked in green, missing ones in red, and incomplete ones in yellow.

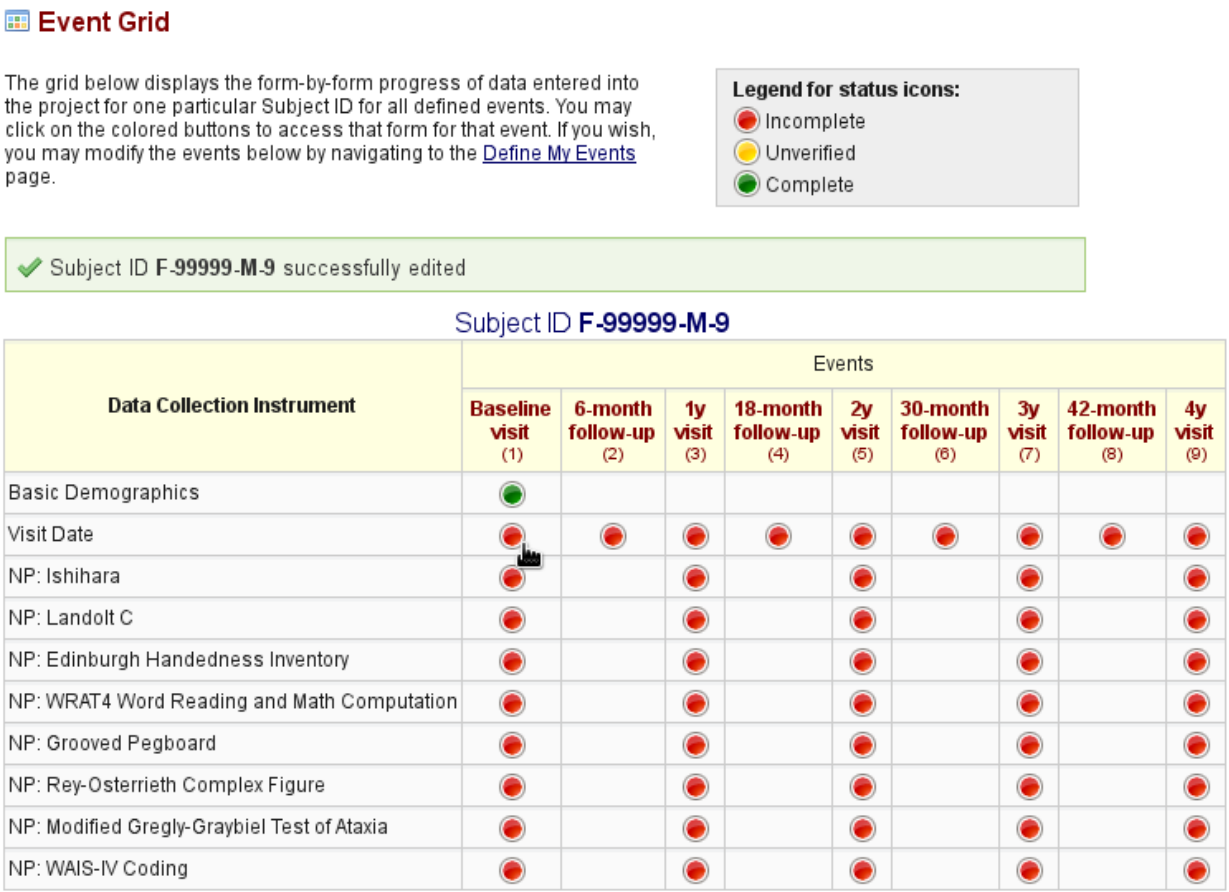
By clicking on one instrument, one can easily switch to the corresponding entry form.

3.2 Edit or Add to Existing Subject

To edit data entered previously or to add new data to an already existing subject, use the "Choose an existing Subject ID" selection box on the "Add / Edit Record" form to select the correct subject:



This will take you to the "Event Grid" for this subject, in which the current entry status for all instruments and entry forms is shown in color. Simply click on the circle representing the correct instrument for the correct visit, which will take you to the corresponding data entry form.



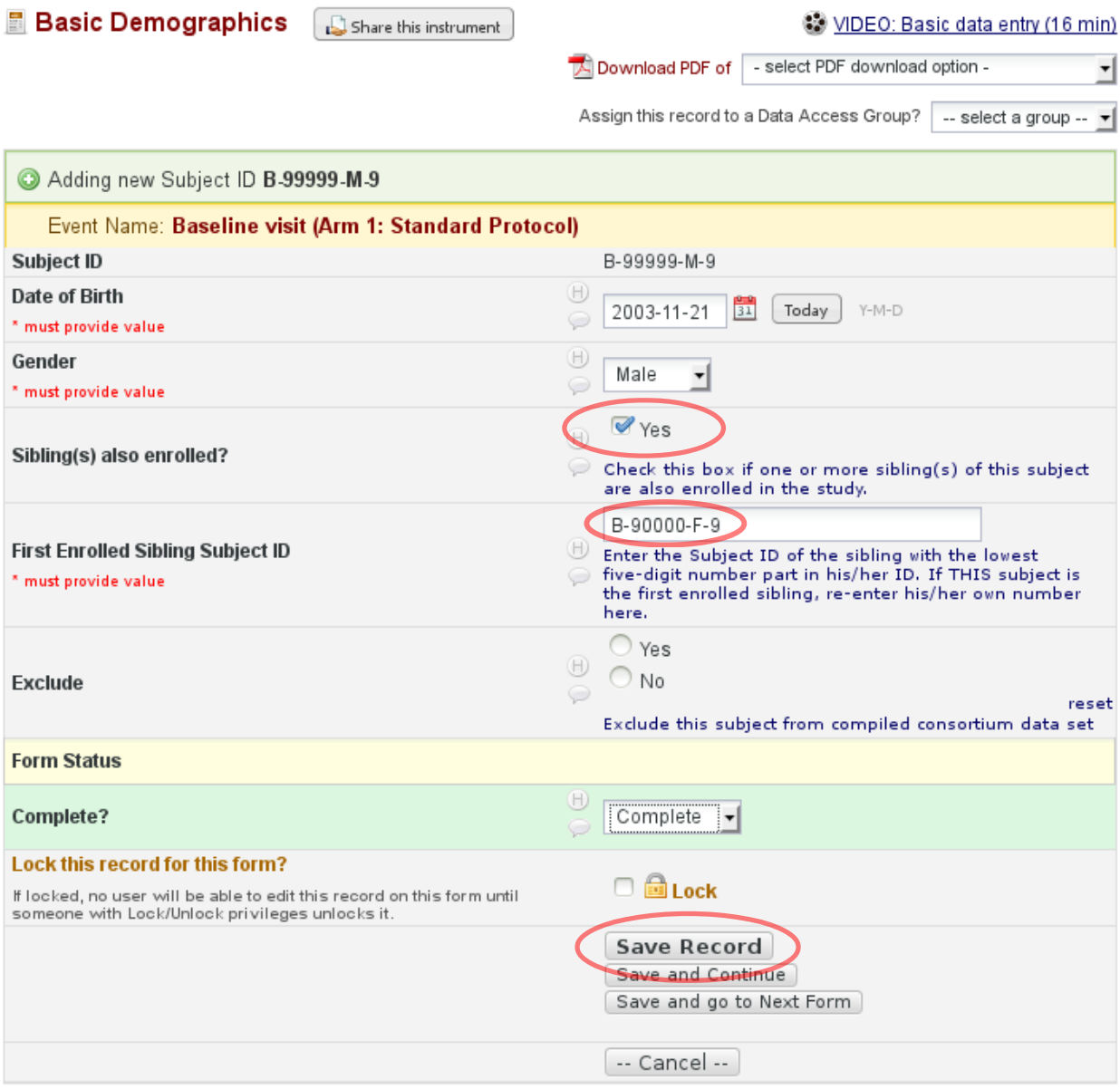
Data Collection Instrument	Events								
	Baseline visit (1)	6-month follow-up (2)	1y visit (3)	18-month follow-up (4)	2y visit (5)	30-month follow-up (6)	3y visit (7)	42-month follow-up (8)	4y visit (9)
Basic Demographics	●								
Visit Date	●	●	●	●	●	●	●	●	●
NP: Ishihara	●		●		●		●		●
NP: Landolt C	●		●		●		●		●
NP: Edinburgh Handedness Inventory	●		●		●		●		●
NP: WRAT4 Word Reading and Math Computation	●		●		●		●		●
NP: Grooved Pegboard	●		●		●		●		●
NP: Rey-Osterrieth Complex Figure	●		●		●		●		●
NP: Modified Gregly-Graybiel Test of Ataxia	●		●		●		●		●
NP: WAIS-IV Coding	●		●		●		●		●

3.3 Adding Siblings of Existing Subjects

When a subject has one or more siblings (brothers or sisters) also enrolled in the N-CANDA study, then this relationship must be marked in REDCap.

To mark a subject (for example "B-99999-M-9"; see screenshot below), as the sibling of another subject, open his or her "Basic Demographics" form (or stay on this form, when you are currently entering a new subject into the database).

Check the **"Sibling(s) also enrolled?"** check box – a new entry field **"First Enrolled Sibling Subject ID"** will then appear. Here, enter the full Subject ID of the first sibling from this family who was enrolled in the study (see below for details on how to determine that ID). Then set the Form Status to "Complete" (or "Unverified," if appropriate) and click the "Save Record" button.



3.3.1 First Enrolled Sibling Subject ID

In general, the Subject ID of the "First Enrolled Sibling" in each family should be easy to determine – out of all siblings, select the ID with the lowest five-digit number part.

For example: A family has the following siblings enrolled in the study: B-00101-M-5, B-00102-M-1, and B-00251-F-6. For all these subjects, the "First Enrolled Sibling" should be entered as "B-00101-M-1", because 00101 is lower than both 00102 and 00251.

Note that for the first enrolled sibling, his or her own number should be entered in the "First Enrolled Sibling Subject ID" entry field.

3.3.2 Using the Fifth Digit to Encode Study Sub-groups

A minor complication arises when sites use the leftmost digit to encode study sub-groups, e.g., a site might decide to enroll some subjects but not others in an overnight sleep protocol. That site might assign subject IDs with first digit "8" to mark the subjects not enrolled in the sleep study.

In this case, we strongly recommend to use a **continuous numbering scheme for the remaining 4 digits** and determine the "First Enrolled Sibling" based on these. As an example, the subject ID assigned following "B-00010-M-5" should have the number part "00011" or "80011," depending on what sub-group the subject is in.

3.3.3 Updating a Subject When Siblings Are Later Enrolled

If siblings of an existing subject are enrolled later in the study, remember to update the first enrolled subject also. That is, when the first sibling of an existing subject is enrolled, re-open that existing subject's "Basic Demographics" form and check the "Siblings also enrolled?" box. After that, enter the existing subject's own ID as the First Enrolled Sibling" ID, because the existing subject was the first sibling enrolled from his or her family.

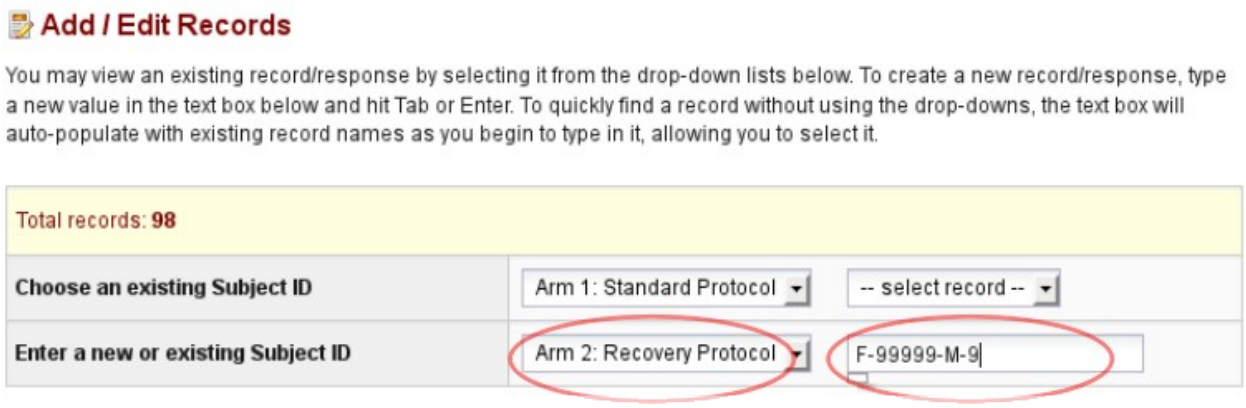
Example: Subject "D-00101-F-2" has been enrolled in the study for a while. Her brother is now enrolled as "D-00351-M-6." The following must be done (not necessarily in this order):

1. Open "Basic Demographics" for D-00101-F-2
    a. Check "Siblings also enrolled?" box
    b. Enter "D-00101-F-2" in the 'First Enrolled Sibling Subject ID" field
    c. Save the updated form
2. Open/create "Basic Demographics" for D-00351-M-6
    a. Check "Siblings also enrolled?" box
    b. Enter "D-00101-F-2" in the 'First Enrolled Sibling Subject ID" field
    c. Save the updated form

3.4 Enrolling a Subject in the Recovery Protocol

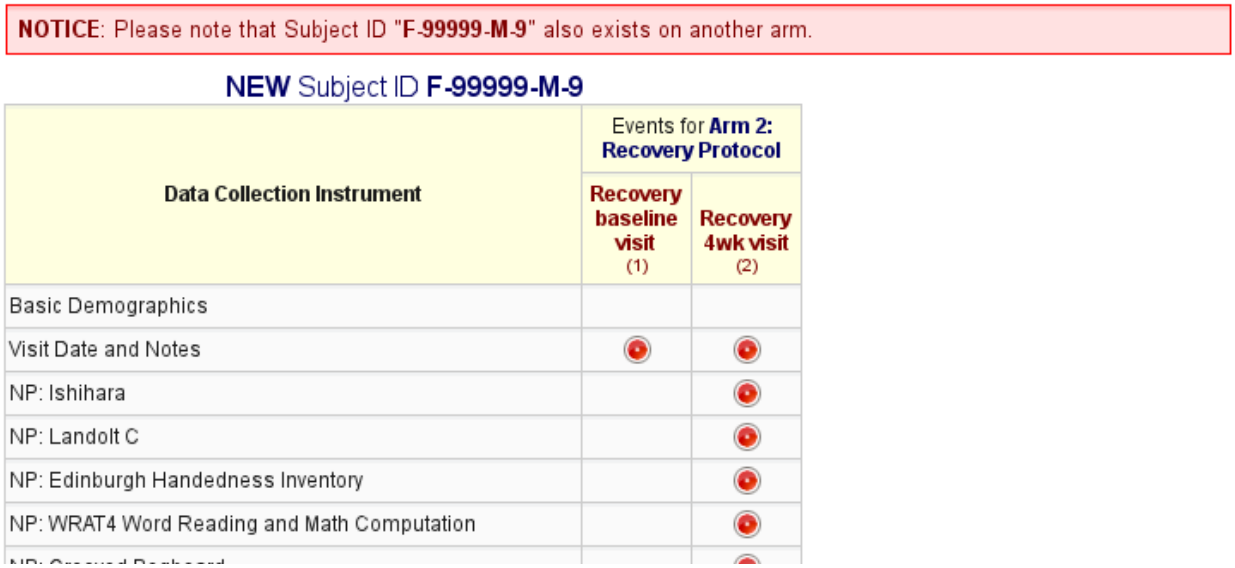
To enroll a subject (new or existing) in the Recovery Protocol, navigate to the "Add / Edit Records" form as described above.

Prior to entering the Subject ID, make sure to select the “Arm 2: Recovery Protocol” option in the selection box to the left of the ID entry field (see screenshot below).



After confirming the Subject ID, a grid with available instruments for the Recovery Protocol for this subject will be shown. Note that these instruments are for visits that are **in addition to the visits in the “Standard Protocol”** for the same subject.

Also note that if the subject has already been entered into the Standard Protocol (as should usually be the case), then there will be a warning shown above the Data Collection Instrument table, informing you that the subject already exists on another “arm” of the study (the Standard Protocol arm). This is expected and does not indicate a problem.



3.5 Enrolling a Subject in the Sleep Protocol

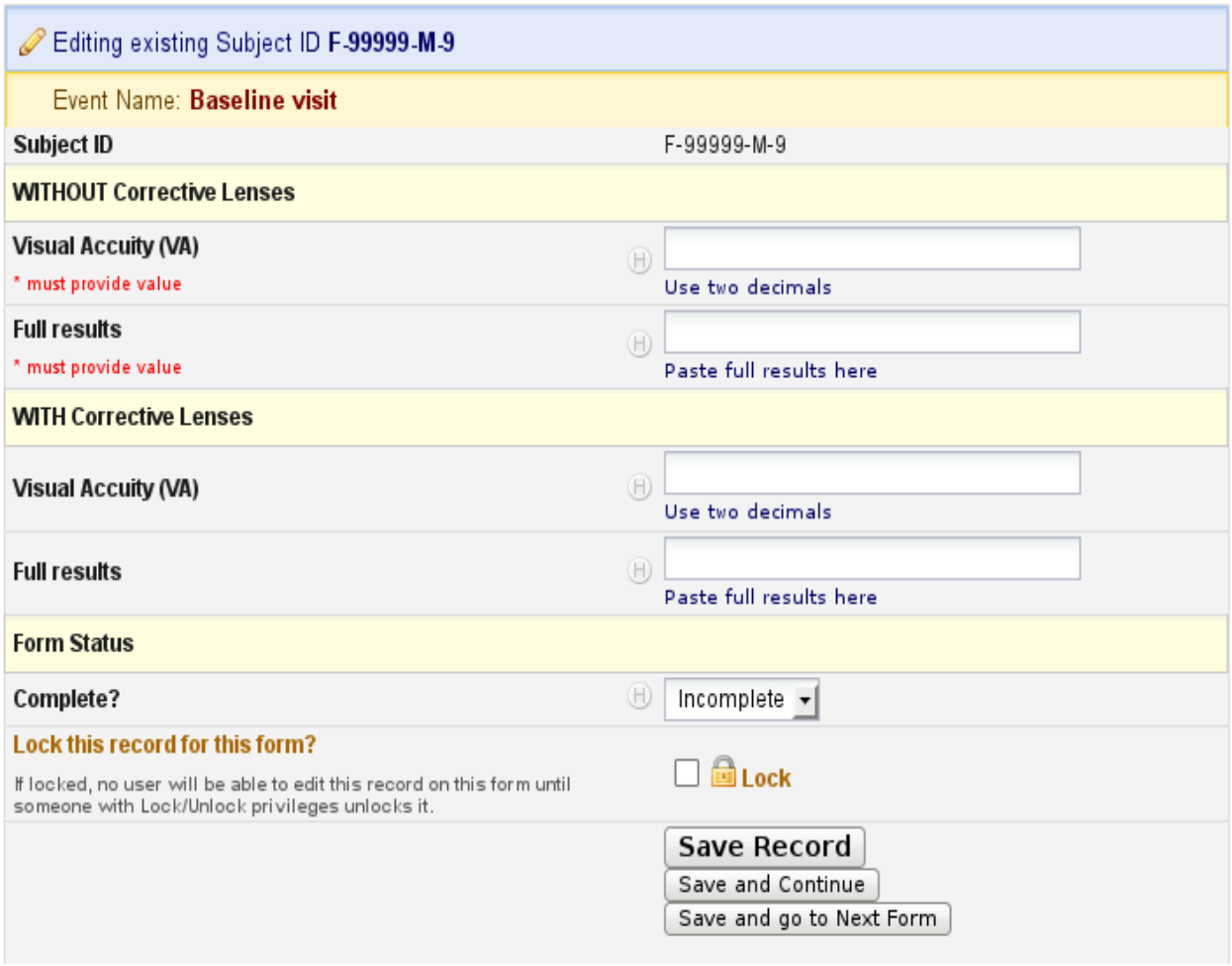
To enroll a subject in the overnight sleep study protocol, proceed exactly as shown above for the recovery protocol, but select the “Arm 3: Overnight Sleep Study” arm of the study.

3.6 Instrument-Specific Instructions

3.6.1 Landolt-C Visual Acuity Test

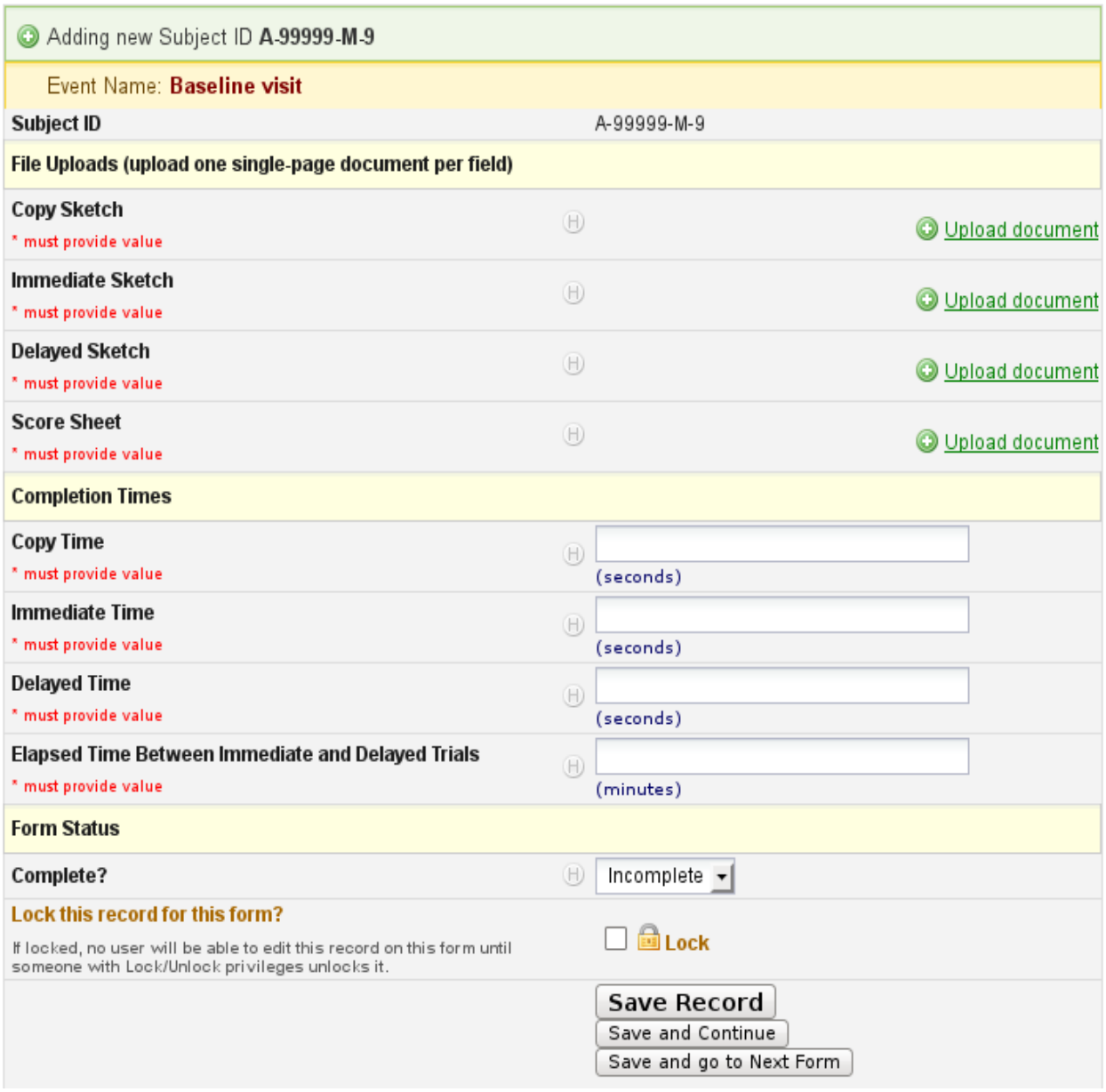
In the Landolt-C entry form, enter results **without corrective lenses for all subjects**, regardless of whether the subject has glasses or contact lenses or not.

For subjects that do use corrective lenses, enter the respective results in the second set of entry fields.



3.6.2 Rey-Osterrieth Complex Figure Times and File Uploads

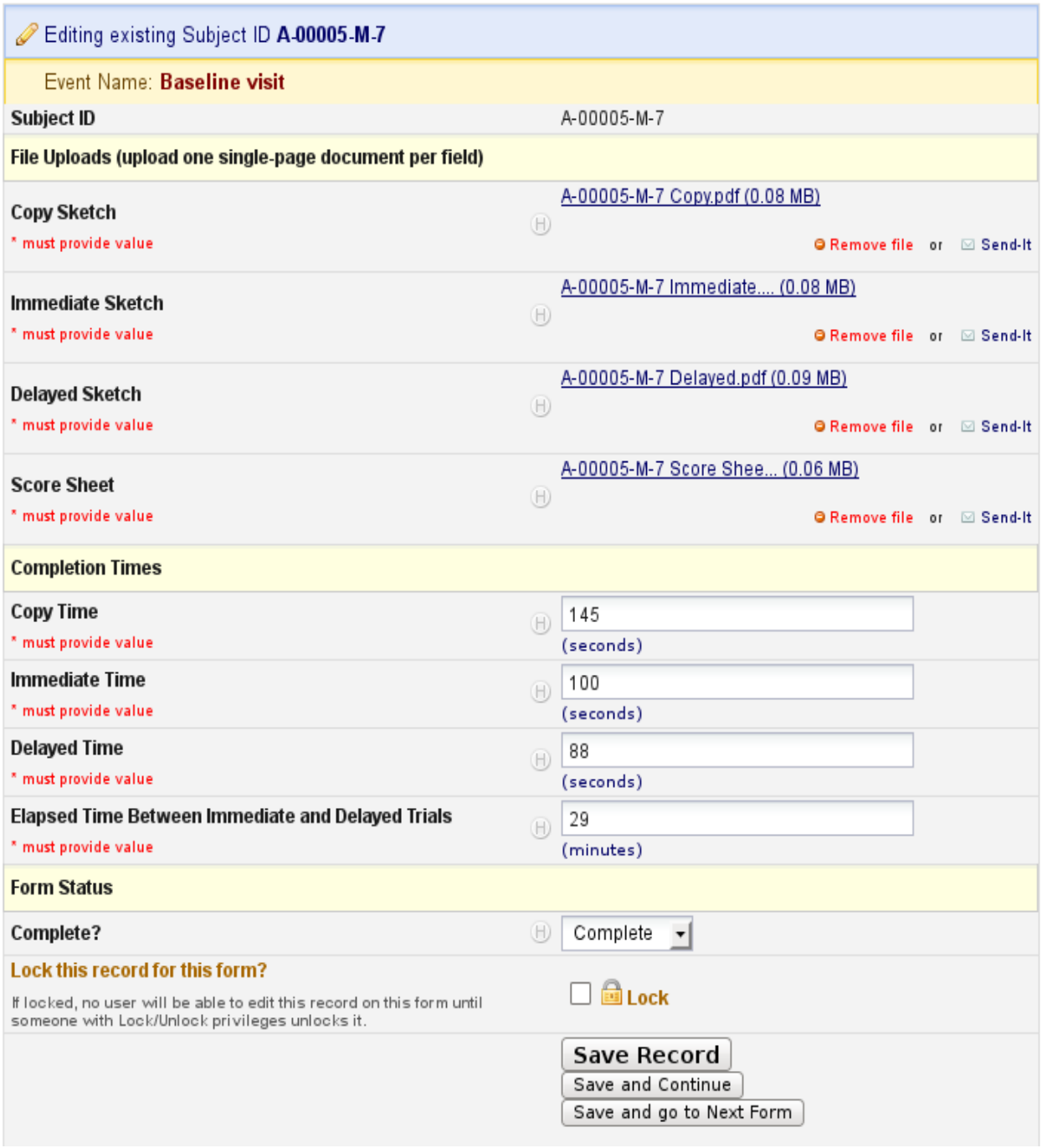
The "Rey-O Times and File Uploads" form (shown below) accepts the data submitted by the N-CANDA sites for the Rey-O instrument. In the top half, the form allows the upload of scanned or photographed pictures of the subject's sketches as well as the original score sheet(s). In the bottom half, the times taken by the subject for each of the sketches should also be entered. The uploaded files and times are then available for second-scoring at SRI.



Each sketch and the score sheet should be uploaded separately in single-page files. All upload and entry fields are mandatory.

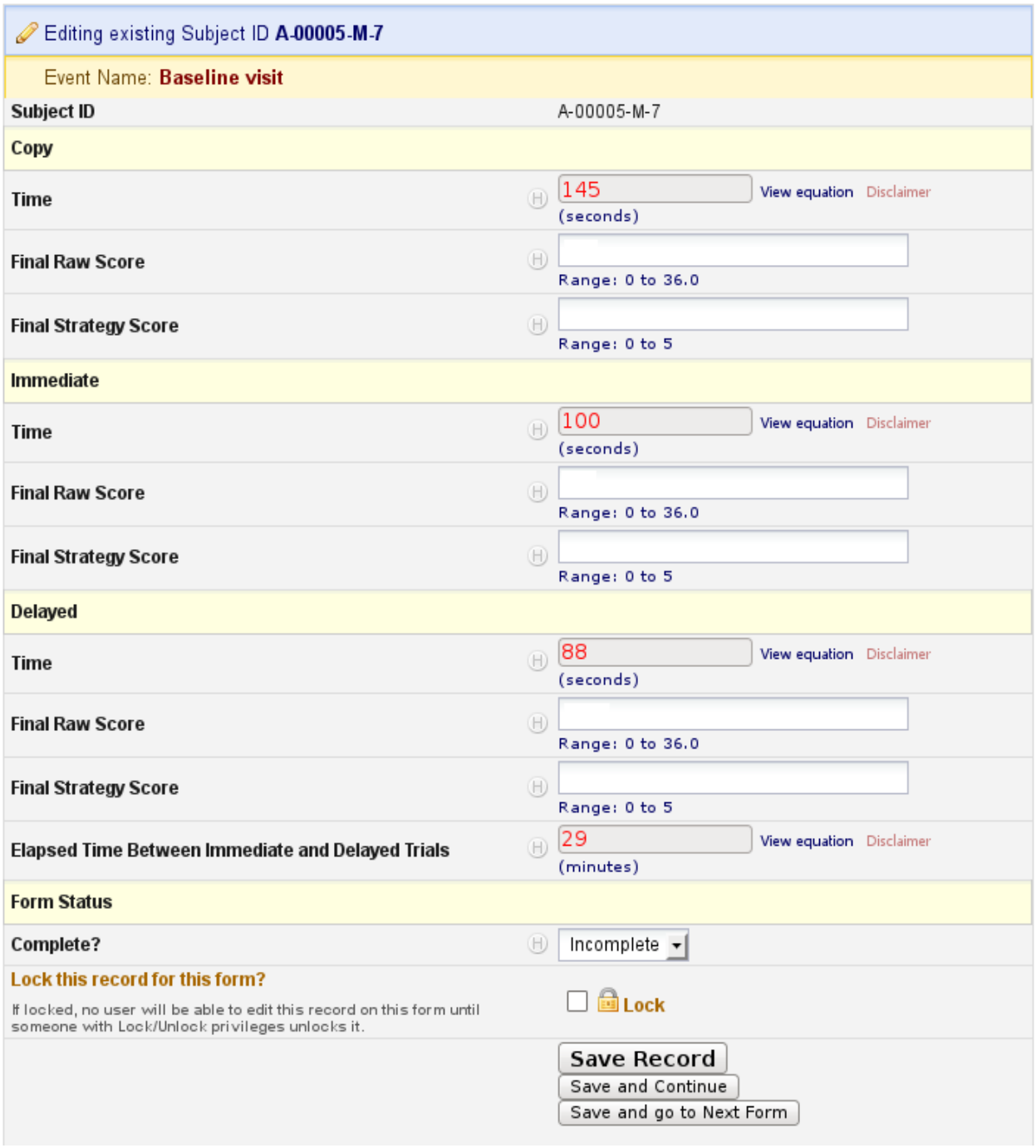
Preferred upload file format is PDF, but other common picture file types, such as JPEG, are also acceptable if creation of PDF files is not feasible. Please negotiate file formats with the person coordinating the second scoring at SRI. (Please do not upload uncompressed files, e.g., TIFF.)

.A properly filled-out form is shown below. Note that the **"Form Status" has been set to "Complete"** by the person entering and uploading the data to indicate that these data are ready for second-scoring.



3.6.3 Rey-Osterrieth Complex Figure Scores

The “Rey-O Scores” data entry form can be edited only by tyhe second-scorer(s) – all other users have only read access. This helps prevent accidental labeling of this form as “Complete” before second scoring has been done. Note that the “Time” fields are simply copies of the respective files in the “Rey-O Times and File Uploads” form (see previous section). Any changes to these variables need to be made in that form.

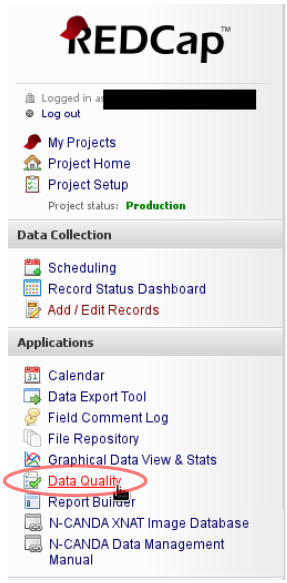


4 Data Quality

4.1 Executing Data Quality Rules

To help collection sites ensure the quality and completeness of their collected data, we have set up "Data Quality" rules in REDCap, which help spot and address problems with the data in the database.

To access the "Data Quality" functions, enter the "N-CANDA Data Entry" project in REDCap. Then, find and click on the "Data Quality" link in the "Applications" menu on the left-hand side of the screen as shown in the screenshot on the right.



You will then be taken to a screen showing a number of "Data Quality Rules" similar to what is shown in the screenshot below:

Data Quality Rules — Execute rules: All | All except A&B | All custom | Clear

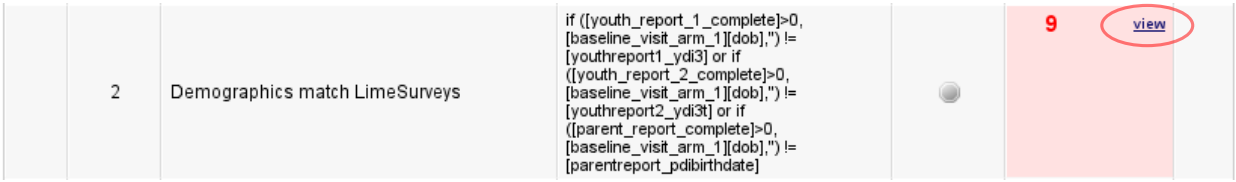
	Rule #	Rule Name	Rule Logic (Show discrepancy only if...)	Real-time execution	Total Discrepancies
	A	Missing values*	-		Execute
	B	Missing values* (required fields only)	-		Execute
	C	Field validation errors (incorrect data type)	-		Execute
	D	Field validation errors (out of range)	-		Execute
	E	Outliers for numerical fields (numbers, integers, sliders, calc fields)	-		Execute
	F	Hidden fields that contain values**	-		Execute
	G	Multiple choice fields with invalid values	-		Execute
	1	Demographics match CNP	if([cnp_instruments(test_sessions)]=1, [baseline_visit_arm_1][dob],'') != [cnp_test_sessions_dob]		Execute
	2	Demographics match LimeSurveys	if ([youth_report_1_complete]>0, [baseline_visit_arm_1][dob],'') != [youthreport1_ydi3] or if ([youth_report_2_complete]>0, [baseline_visit_arm_1][dob],'') != [youthreport2_ydi3t] or if ([parent_report_complete]>0, [baseline_visit_arm_1][dob],'') != [parentreport_pdibirthdate]		Execute
	3	Missing Forms, Delayed Discounting	[visit_date] != "" and ((not [delayed_discounting_1000_complete]>0 and not [dd1000_missing]=1) or (not [delayed_discounting_100_complete]>0 and not [dd100_missing]=1))		Execute

Each rule tests for a specific type of problem. You can either run a specific rule, by clicking the "Execute" button for that rule, or run "All custom" rules by clicking the button labeled accordingly on top of the table.

(Rules A-G, which are pre-defined by REDCap, are not currently relevant for use by the data collection sites; thus, running "All" or "All except A&B" rules does not add any benefit over running only the "Custom" rules.)

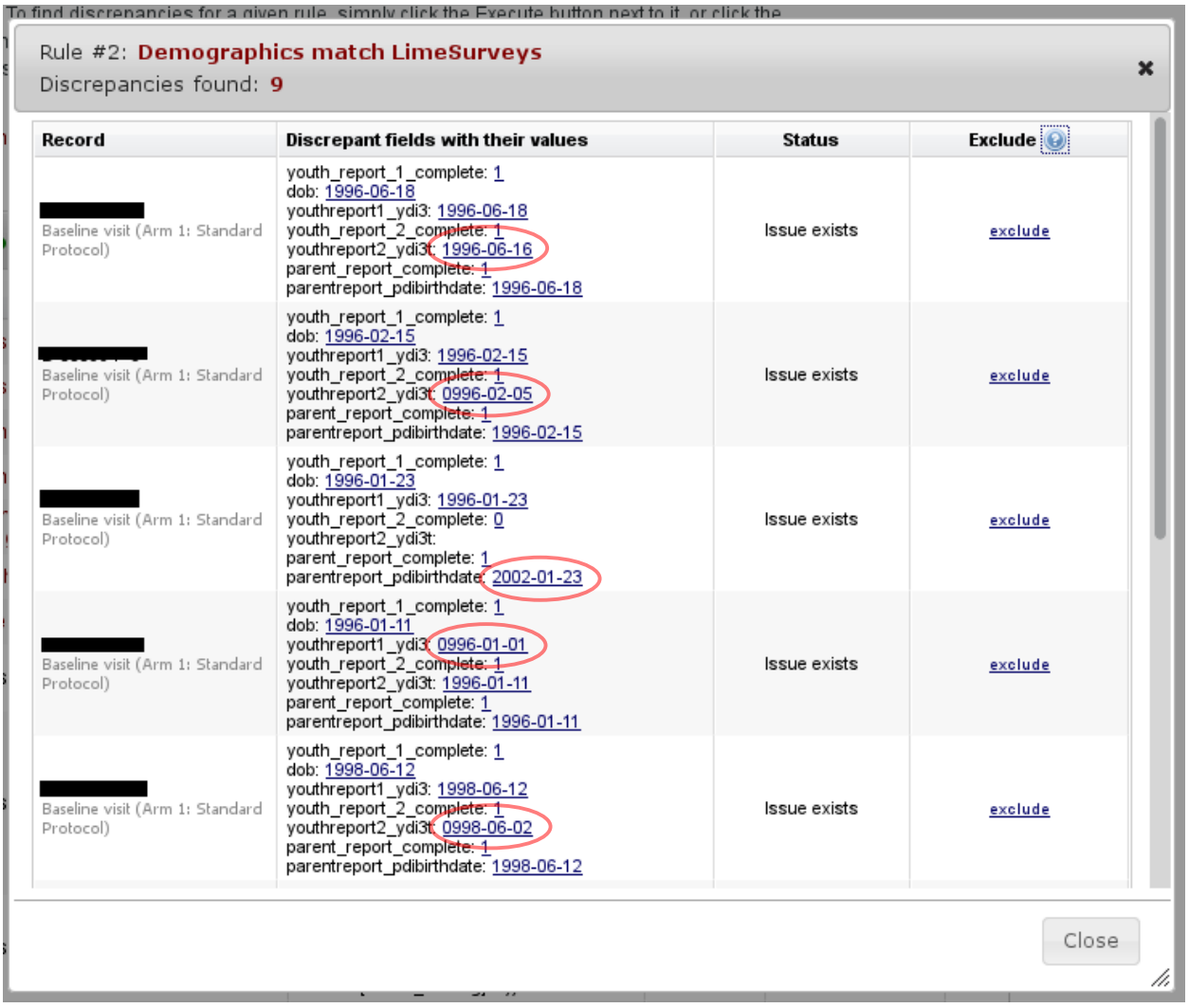
Depending on number of executed rules, their complexity, and the number of records in the database, rule execution may take from a few seconds to several minutes. For each completed rule, the table will then show the number of records violating that rule (or "0" if there aren't any).

As shown below, when rule violations ("Discrepancies") are found, there will also be a "view" link, which should be clicked to inspect the violating records in detail:



	2	Demographics match LimeSurveys	if ([youth_report_1_complete]>0, [baseline_visit_arm_1][dob],'') != [youthreport1_ydi3] or if ([youth_report_2_complete]>0, [baseline_visit_arm_1][dob],'') != [youthreport2_ydi3t] or if ([parent_report_complete]>0, [baseline_visit_arm_1][dob],'') != [parentreport_pdibirthdate]		9 view

After clicking "view" for a rule with discrepancies, a new window will open. Here, for each violating record the "discrepant fields with their values" are shown, and each value represents a link to the entry field for that variable in the violating record. Below is an example showing all records for which at least one of the multiple entries for "subject birth date" in the LimeSurvey data is different from the birth date entered in the "Basic Demographics" form in REDCap.



To find discrepancies for a given rule, simply click the Execute button next to it, or click the

Rule #2: **Demographics match LimeSurveys**
Discrepancies found: **9**

Record	Discrepant fields with their values	Status	Exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	youth_report_1_complete: 1 dob: 1996-06-18 youthreport1_ydi3: 1996-06-18 youth_report_2_complete: 1 youthreport2_ydi3t: 1996-06-16 parent_report_complete: 1 parentreport_pdibirthdate: 1996-06-18	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	youth_report_1_complete: 1 dob: 1996-02-15 youthreport1_ydi3: 1996-02-15 youth_report_2_complete: 1 youthreport2_ydi3t: 0996-02-05 parent_report_complete: 1 parentreport_pdibirthdate: 1996-02-15	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	youth_report_1_complete: 1 dob: 1996-01-23 youthreport1_ydi3: 1996-01-23 youth_report_2_complete: 0 youthreport2_ydi3t: parent_report_complete: 1 parentreport_pdibirthdate: 2002-01-23	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	youth_report_1_complete: 1 dob: 1996-01-11 youthreport1_ydi3: 0996-01-01 youth_report_2_complete: 1 youthreport2_ydi3t: 1996-01-11 parent_report_complete: 1 parentreport_pdibirthdate: 1996-01-11	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	youth_report_1_complete: 1 dob: 1998-06-12 youthreport1_ydi3: 1998-06-12 youth_report_2_complete: 1 youthreport2_ydi3t: 0998-06-02 parent_report_complete: 1 parentreport_pdibirthdate: 1998-06-12	Issue exists	exclude

Close

Note that, just as in this example, not every value shown has to be incorrect. It is your

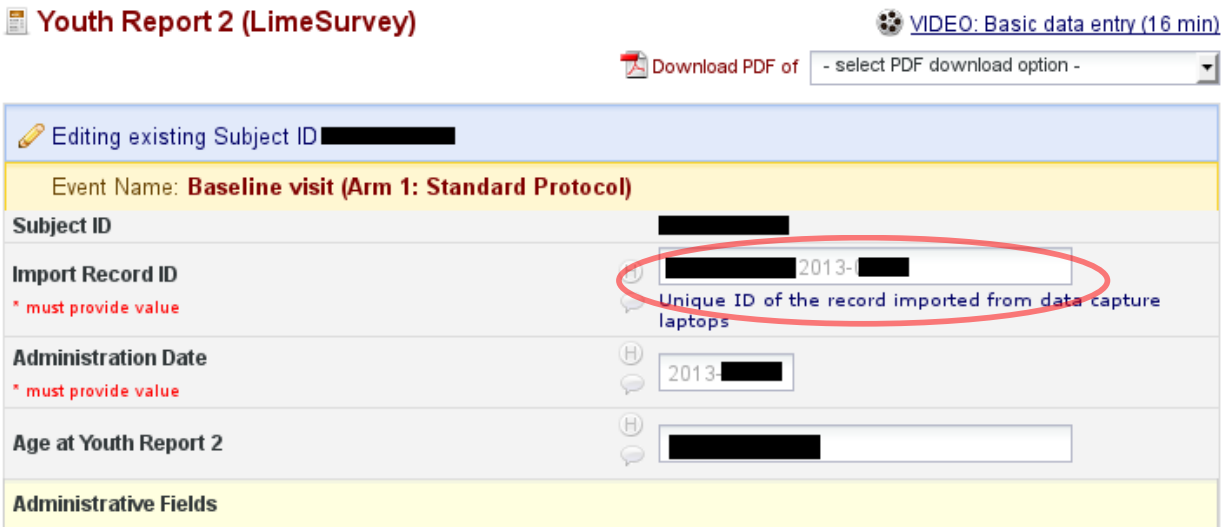
responsibility to find the value that is in violation of the data quality rule and correct it.

Important: Do not click the "exclude" link to exclude a violating record from the list of results! If the detected discrepancy is a false positive, contact Torsten Rohlfing (torsten@synapse.sri.com) to work on a more specific rule.

4.2 Locating and Correcting Data Imported from Laptops

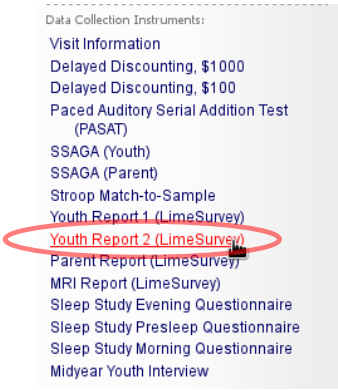
When incorrect data originate from the collection laptops, they cannot be corrected in the "N-CANDA Data Entry" REDCap project. Specifically, after following a link from the results of a failed "Data Quality Rule" execution, you will be taken to the offending form in the "Data Entry" project, but you first need to locate the original imported record and correct the offending data there.

To this end, find the "Import Record ID" field at or near the top of the form in "Data Entry." This Record ID is the identifier that is needed to locate the original record:

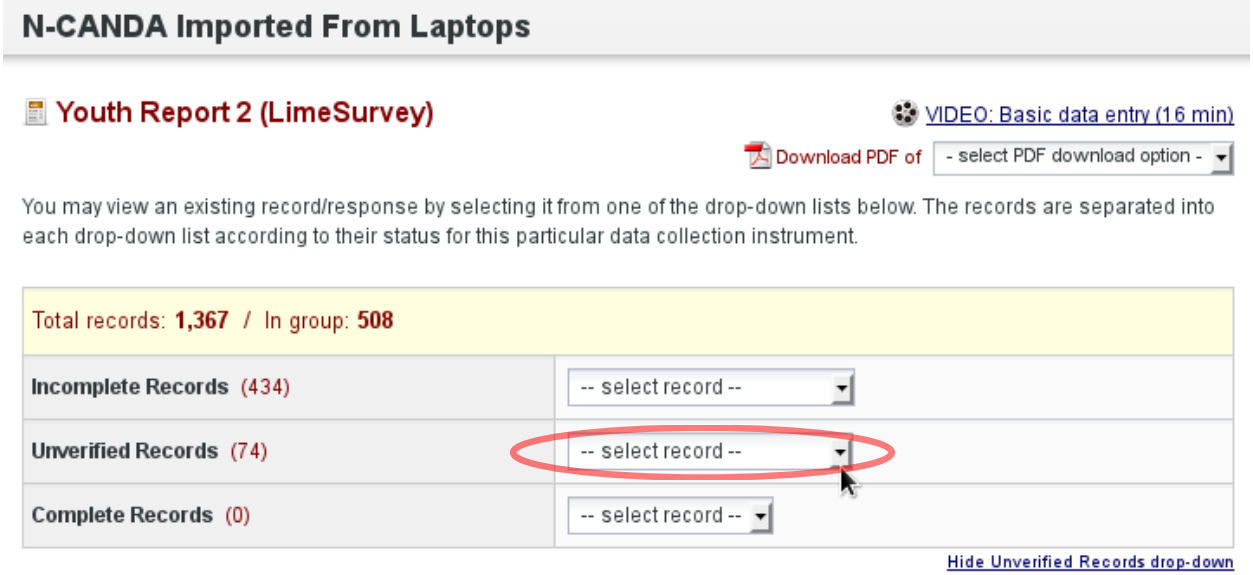


For data imported from the N-CANDA data collection laptops, the original record can be found in the "N-CANDA Imported from Laptops" project. To locate this record, first navigate to the "Imported from Laptops" project via the "My Projects" tab on the N-CANDA REDCap front page.

Next, select the form that you are trying to correct, e.g., "Youth Report 2" in the example above, from the "Data Collection Instruments" list on the left-hand side of the screen as illustrated in the screenshot on the right.



Then, select the desired record from either the "Unverified Records" or "Complete Records" drop-down list:



You should then be taken to the form containing the data imported from the collection laptops. Make any necessary changes there, change the "Form Status" to "Complete" at the bottom and "Save" the form.

Your corrections should appear in the "N-CANDA Data Entry" project within 24 hours. If this is not the case, correct Torsten Rohlfing (torsten@synapse.sri.com) with the Form Name and Record ID.

4.3 Locating and Correcting Data Imported from Penn WebCNP

Corrections to data imported from the U Penn WebCNP system can only be made to correct Subject ID and Collection Date (i.e., they can be used to assign collected data to the correct subject and visit).

Corrections to any other fields, e.g., scores, cannot be made in REDCap – you must contact the WebCNP maintainers directly.

4.4 Identifying and Labeling Uncollected or Lost Data

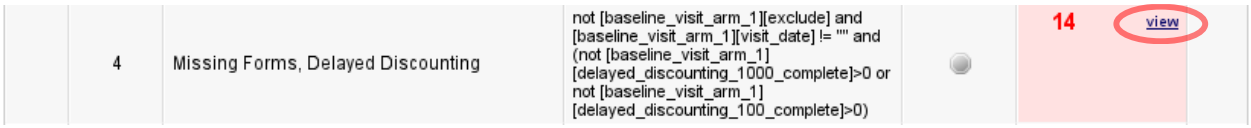
The procedure described in this section requires special access privileges on the N-CANDA REDCap server and cannot be executed by every user.

Many of the Data Quality describes above serve to detect "missing data" in the sense that they create lists of records for which a given form does not contain data. In some cases, this is due to these data genuinely not existing. For example, for subjects over 18 years of age, the "Parent Report" and "Parent SSAGA" may not have been collected. Other times, the data exist but have not been assigned to the correct subject and record (e.g., due to mis-entered subject ID), or they

may not have been imported into REDCap at all, due to problems with the laptop data transmission.

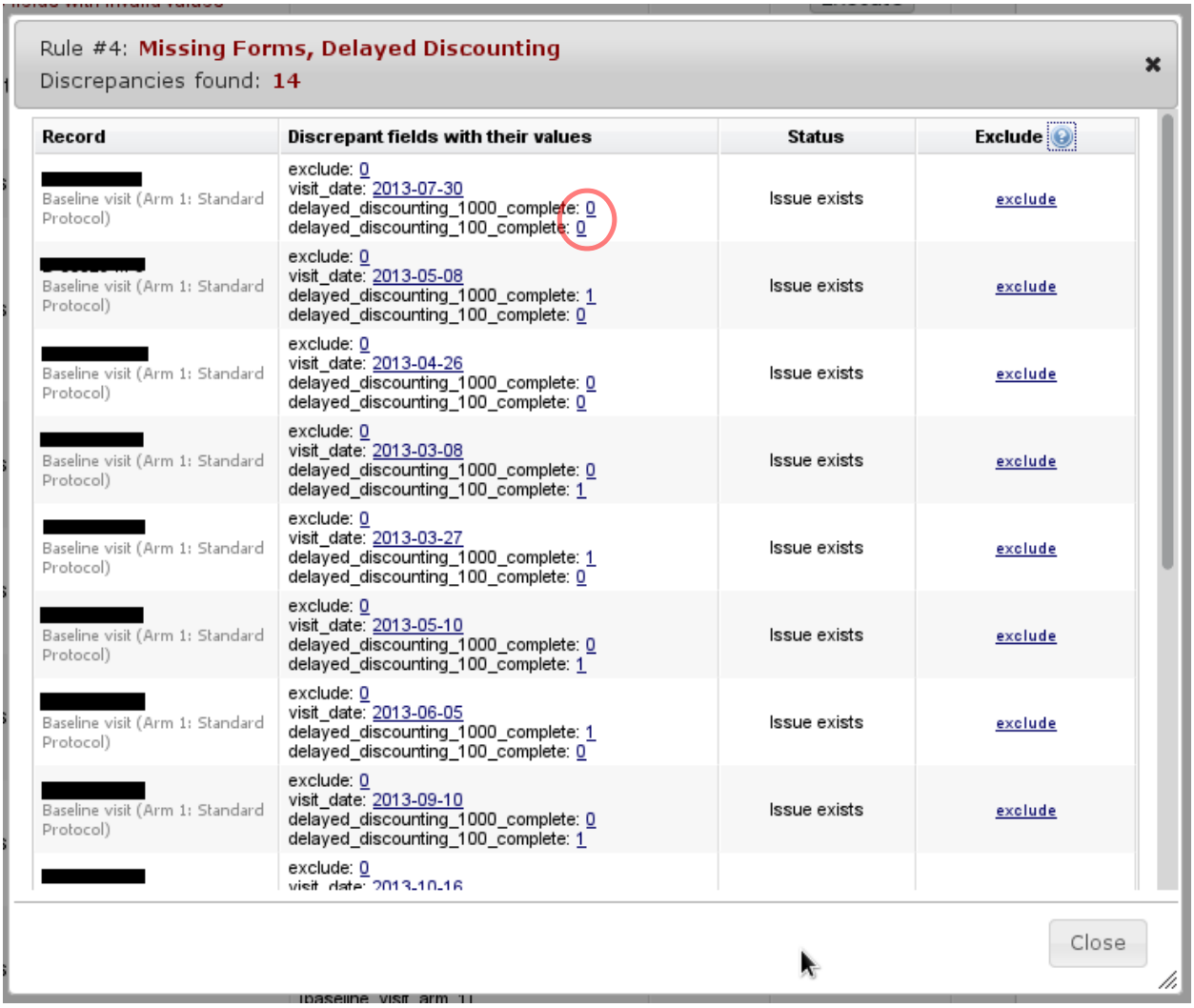
To determine which data truly do not exist, and to make sure all existing data are properly imported and assigned, run one of the "Missing Forms" rules on the "Data Quality" screen as described above. You may also want to run "All custom rules", so that tests are run for missing records in all forms.

After executing a data quality rule for a given form, say as an example for "Delayed Discounting," you will see a display such as the following when unexplained missing records have been found:



	4	Missing Forms, Delayed Discounting	not [baseline_visit_arm_1][exclude] and [baseline_visit_arm_1][visit_date] != "" and (not [baseline_visit_arm_1] [delayed_discounting_1000_complete]>0 or not [baseline_visit_arm_1] [delayed_discounting_100_complete]>0)		14	view

As before, click on the "view" link to open a window showing the list of affected records, where the missing data are those where a "complete" variable is "0" (zero) or does not exist at all:



Rule #4: **Missing Forms, Delayed Discounting**
Discrepancies found: **14**

Record	Discrepant fields with their values	Status	Exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-07-30 delayed_discounting_1000_complete: 0 delayed_discounting_100_complete: 0	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-05-08 delayed_discounting_1000_complete: 1 delayed_discounting_100_complete: 0	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-04-26 delayed_discounting_1000_complete: 0 delayed_discounting_100_complete: 0	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-03-08 delayed_discounting_1000_complete: 0 delayed_discounting_100_complete: 1	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-03-27 delayed_discounting_1000_complete: 1 delayed_discounting_100_complete: 0	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-05-10 delayed_discounting_1000_complete: 0 delayed_discounting_100_complete: 1	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-06-05 delayed_discounting_1000_complete: 1 delayed_discounting_100_complete: 0	Issue exists	exclude
[redacted] Baseline visit (Arm 1: Standard Protocol)	exclude: 0 visit_date: 2013-09-10 delayed_discounting_1000_complete: 0 delayed_discounting_100_complete: 1	Issue exists	exclude
[redacted]	exclude: 0 visit_date: 2013-10-16		

Close

Note that for convenience, several data quality rules simultaneously check for multiple missing forms. A discrepancy will be shown when any one form is missing unexplained.

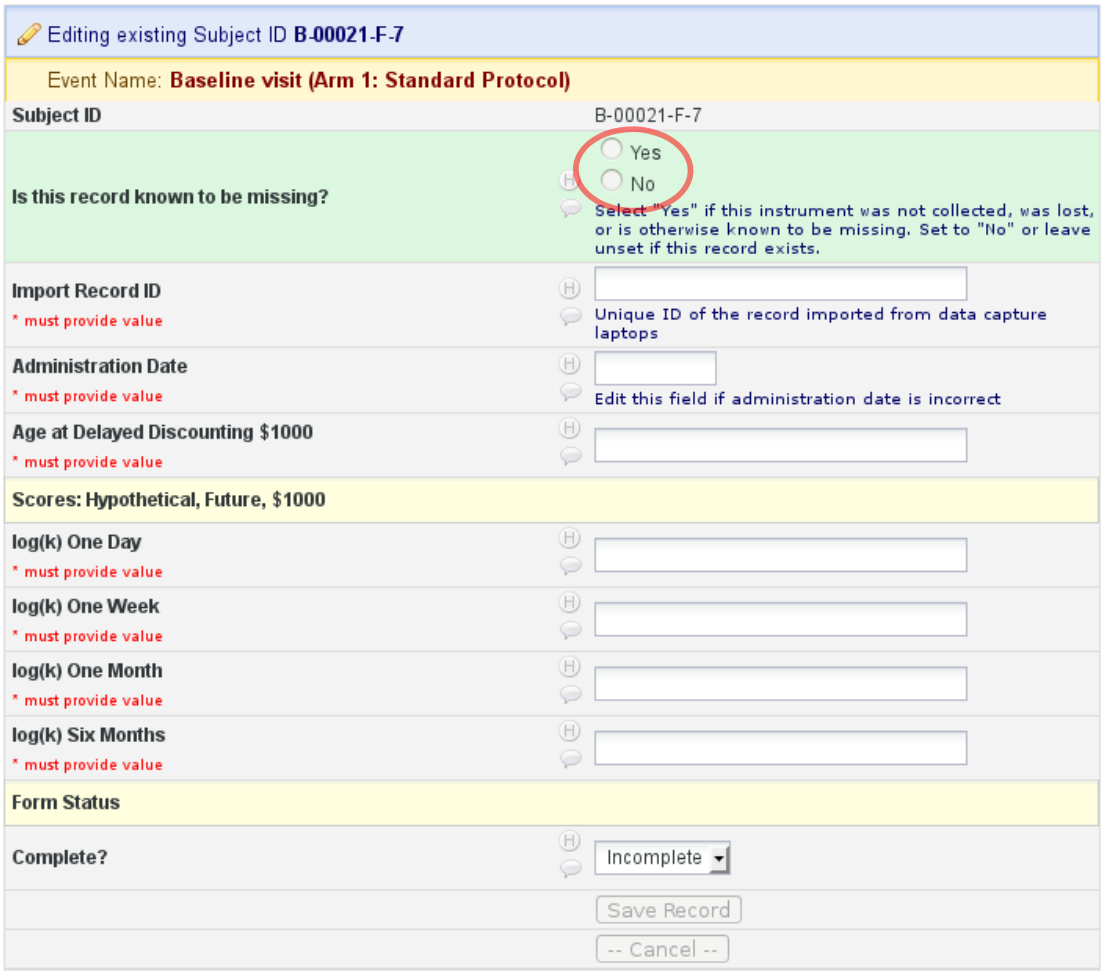
As described above, clicking on a link for the zero "complete" variable will take you to the missing record in the "Data Entry" project.

Here, you have two options, depending on whether the missing record truly does not exist, or has simply not been properly imported and/or assigned.

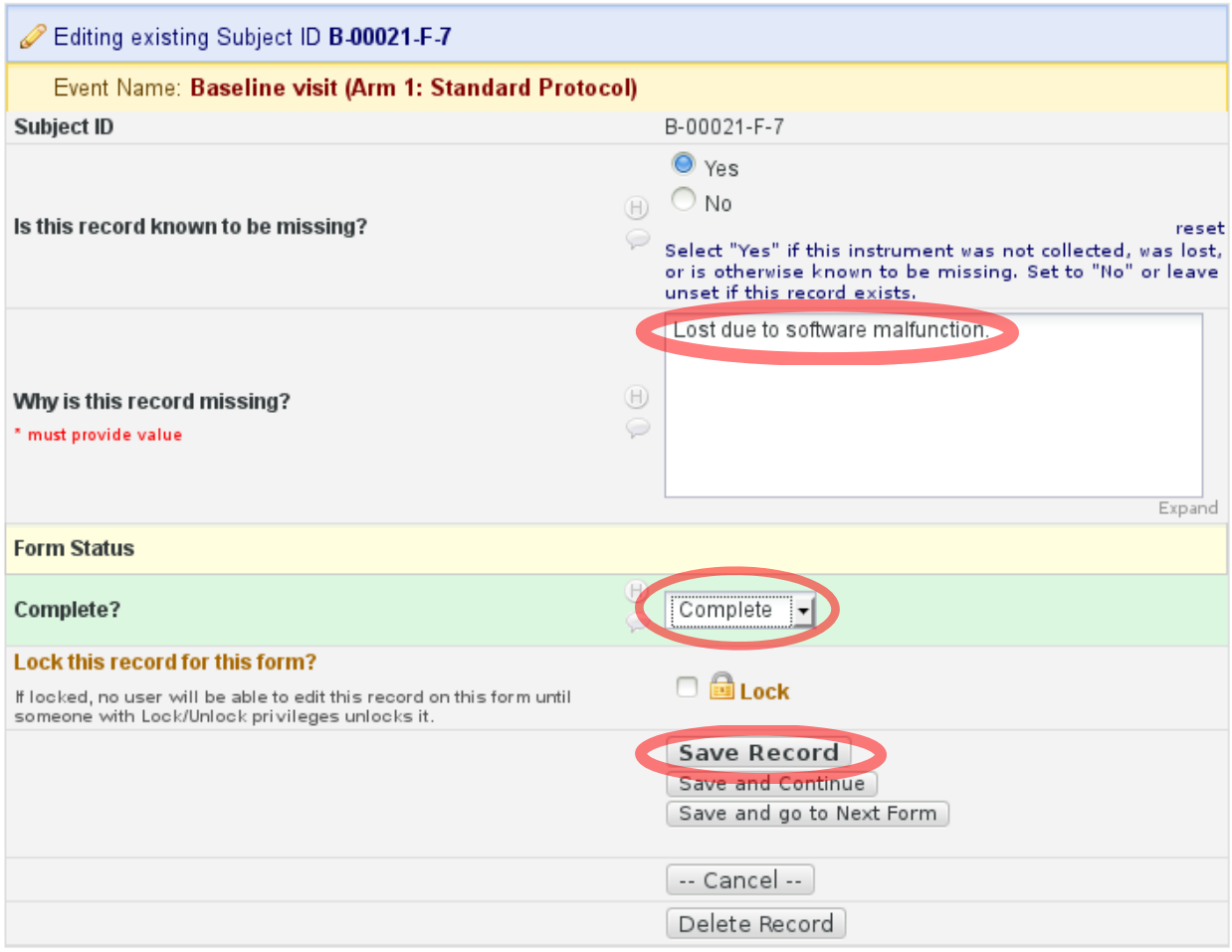
In the latter case, to locate a missing record that is believed to exist, to ensure it is properly assigned to the correct subject and visit, follow the instructions in Section 5.2 below.

Before proceeding with the marking of a non-existent record, make absolutely sure it truly does not exist. Ideally, you should be able to explain the absence of the record (e.g., "not collected due to subject exclusion", or "lost due to software/hardware malfunction"). Once marked as "missing permanently," records will no longer appear in data quality results, and they will no longer be assigned even if they later appear in the "Imported" project database.

Once it has been confirmed that a missing record is permanently missing, navigate to the missing form in "N-CANDA Data Entry" via the data quality results as described above. At the top of the form, you will find a Yes/No toggle labeled "Is this record known to be missing?" as shown below:



Setting this toggle to “Yes” will make the form fields disappear and instead reveal a new entry field, “Why is this record missing?” You are expected and strongly encouraged to provide an explanation here, as shown in the example:



After entering an explanation for the missing record, make sure to se the “Complete?” form status to “Complete.” This will both hide this record from future Data Quality runs and also mark it as “Complete” in the “Data Quality Dashboard.”

To perform the above procedure for marking permanently missing data, you have been assigned edit permissions to forms in the “Data Entry” project which should otherwise not be edited. This is because their contents are assigned from records in the “Imported” projects.

Any changes made to the fields of such forms in “Data Entry” (with the exception of the “Missing?” fields) may be overwritten with data from the “Imported” projects at any time.

5 Data Imported from Laptops

Before we discuss, in detail, the process for uploading, locating, and potentially correcting data captured by the N-CANDA laptops, it is important to remember one principle:

Once data have been submitted from the laptops and uploaded to the server at the N-CANDA Data Integration Component at SRI, they should **never, under any circumstances, be modified on the laptops and re-submitted.**

Any necessary corrections must be made in the "N-CANDA Imported from Laptops" project in REDCap. This applies to all instruments and all types of corrections. **No exceptions!**

The entire data submission and import procedure has been designed based on the assumption that every record submitted from the laptops will be submitted exactly once. Sending the same record repeatedly, with corrections, will in the best-case scenarios simply be ignored. More likely, repeated submission will lead to multiple, conflicting copies of the same record in the REDCap database, which must be avoided.

5.1 Outline

Data captured using the N-CANDA laptops (e.g., LimeSurvey data entry, Delayed Discounting, PASAT, Stroop) are imported and organized into the REDCap database following these steps:

1. The laptops "Upload" all captured data by committing them to the N-CANDA Subversion repository
2. The N-CANDA server periodically (e.g., nightly) retrieves all new and updated files and converts them to CSV format.
3. All generated CSV files for new records (uniquely identified by instrument, subject ID, and administration date) are imported into the "N-CANDA Imported From Laptops" project in REDCap.
4. The subjects and visits entered into the "N-CANDA Data Entry" REDCap project are used to select the correct records from the import.
5. Selected records are copied into the respective forms in the "N-CANDA Data Entry" project.

Note that each record imported from the laptops is identified with a unique identifier, which is a combination of the Subject ID entered and the date of administration. **This identifier is permanent and cannot be edited.** Both Subject ID and administration date are, however, also stored separately for each imported instrument, and these separate values can be edited to correct errors.

5.2 Locating Missing Records

See the example on the right, showing the Record Status (in the "Data Entry" project) for a subject that should have data available for Delayed Discounting with both $1000 and $100 dollar rewards at the Baseline visit. The red status indicator shows that the data for the $100 experiment have not been successfully assigned.

Subject ID **B-00023-M-6**

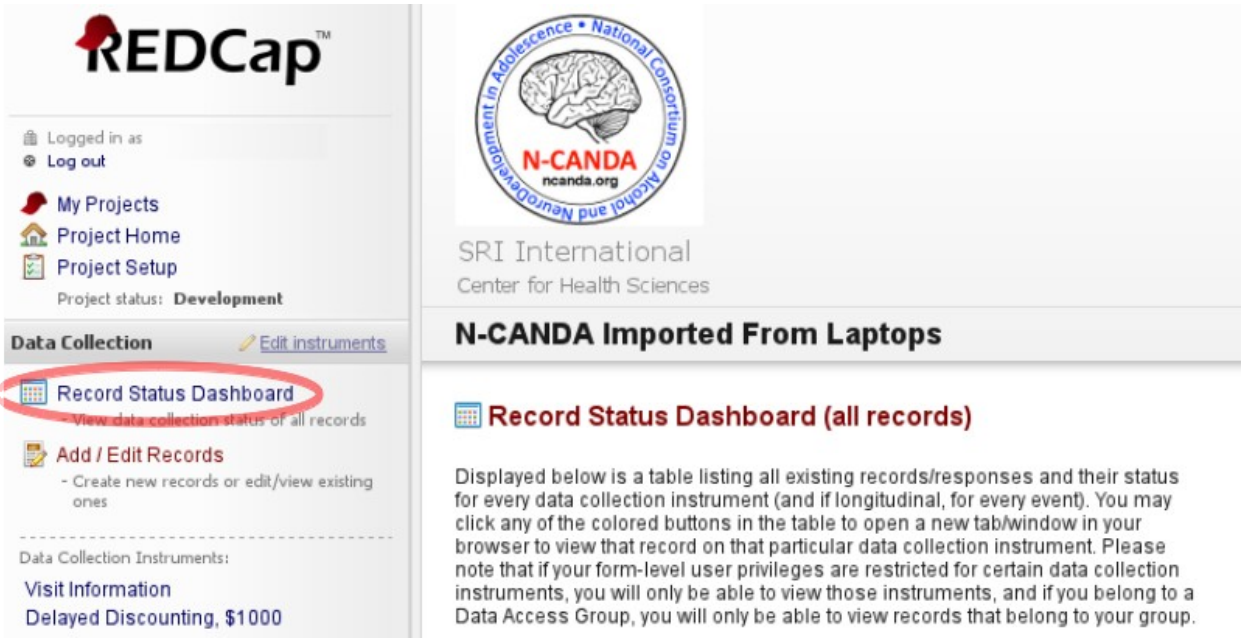
Data Collection Instrument	Baseline visit (1)	6-month follow-up (2)	1y visit (3)	18-month follow-up (4)
Basic Demographics				
Visit Date and Notes				
NP: Ishihara				
NP: Landolt C				
NP: Edinburgh Handedness Inventory				
NP: WRAT4 Word Reading and Math Computation				
NP: Grooved Pegboard				
NP: Rey-Osterrieth Complex Figure PDF Uploads				
NP: Rey-Osterrieth Complex Figure Scores				
NP: Modified Gregly-Graybiel Test of Ataxia				
NP: WAIS-IV Coding				
Biological Data: NP Day				
Biological Data: MR Day				
Saliva Samples				
Delayed Discounting, $1000				
Delayed Discounting, $100				
Paced Auditory Serial Addition Test (PASAT)				
Penn WebCNP Summary				
MRI Session Report				

5.2.1 Locating and Assigning Imported Data

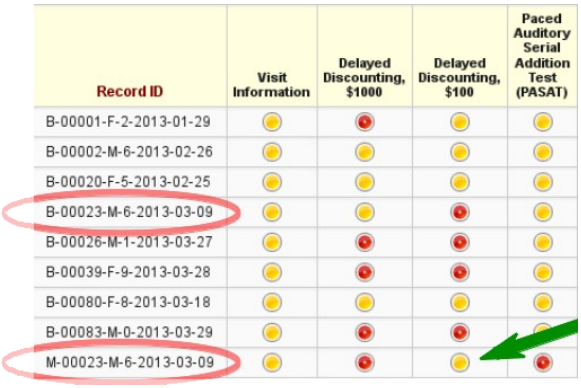
The first step to reconcile imported data with the final, compiled data set is to locate the missing data in the imported records. To this end, from the list of "My Projects", first select the "N-CANDA Imported from Laptops" project:

My Projects	Records	Fields	Type	Status
N-CANDA Imported From Laptops	74	48		
N-CANDA Data Entry	65	305		
N-CANDA Imported from PennCNP	65	923		

Next, select the "Record Status Dashboard" from the menu on the left-hand side of the screen:

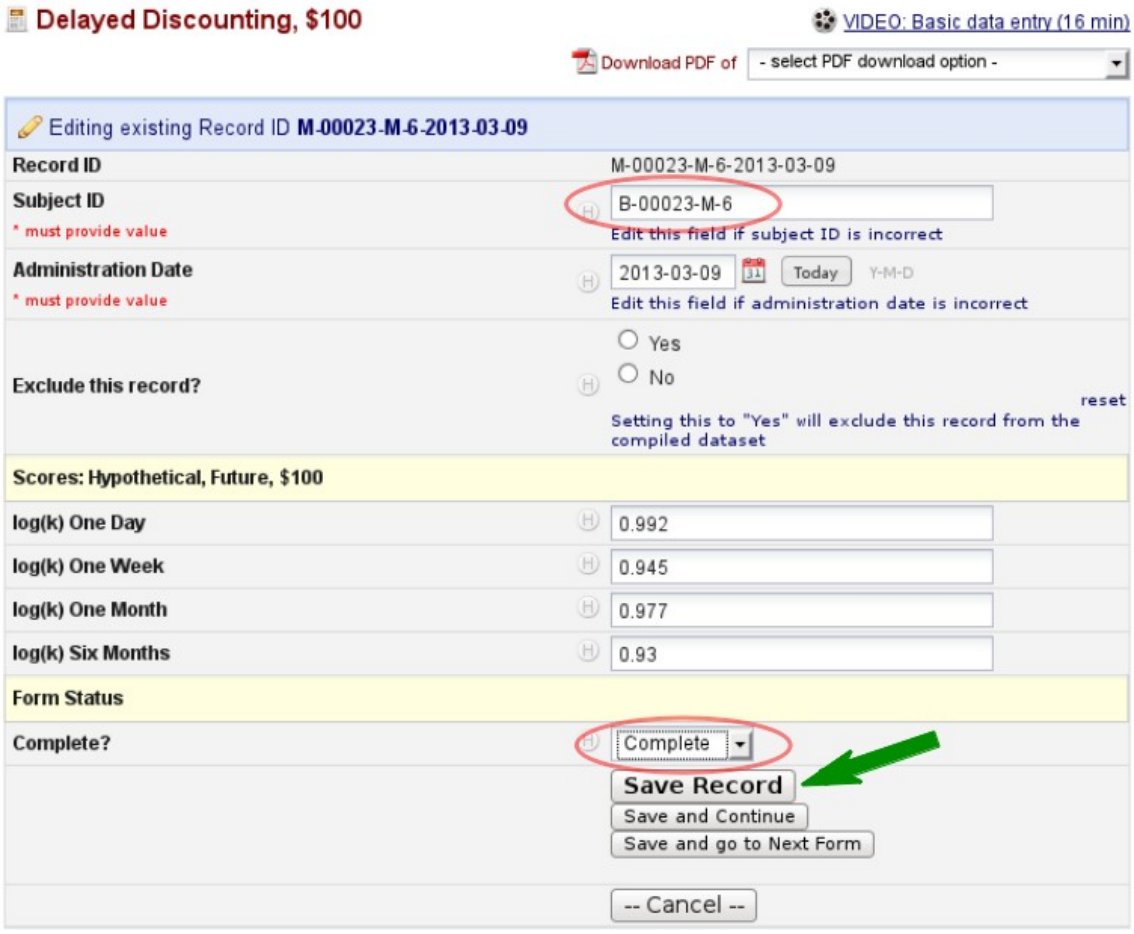


Now locate the missing data in the dashboard – in the example below, the Subject ID has been mistyped for the "Delayed Discounting $100" task. To allow assignment of these scores to the correct subject and visit, the mis-entered Subject ID must be corrected. To this end, click on the Status Indicator of the mislabeled instrument (green arrow).



Record ID	Visit Information	Delayed Discounting, $1000	Delayed Discounting, $100	Paced Auditory Serial Addition Test (PASAT)
B-00001-F-2-2013-01-29				
B-00002-M-6-2013-02-26				
B-00020-F-5-2013-02-25				
B-00023-M-6-2013-03-09				
B-00026-M-1-2013-03-27				
B-00039-F-9-2013-03-28				
B-00080-F-8-2013-03-18				
B-00083-M-0-2013-03-29				
M-00023-M-6-2013-03-09				

Next, enter the correct Subject ID in the instrument data entry form. Afterwards, set the "Complete?" status field to "Complete", which will protect this record from being overwritten by future data imports. Then, click on the "Save Record" button.



Delayed Discounting, $100

VIDEO: Basic data entry (16 min)

Download PDF of - select PDF download option -

Editing existing Record ID **M-00023-M-6-2013-03-09**

Record ID	M-00023-M-6-2013-03-09
Subject ID * must provide value	B-00023-M-6 Edit this field if subject ID is incorrect
Administration Date * must provide value	2013-03-09 Today Y-M-D Edit this field if administration date is incorrect
Exclude this record?	Yes No reset Setting this to "Yes" will exclude this record from the compiled dataset
Scores: Hypothetical, Future, $100	
log(k) One Day	0.992
log(k) One Week	0.945
log(k) One Month	0.977
log(k) Six Months	0.93
Form Status	
Complete?	Complete

Save Record

Save and Continue

Save and go to Next Form

-- Cancel --

The correction procedure is entirely analogous for all other instruments imported from the laptop computers.

Note that if the same incorrect Subject ID has been used for multiple instruments, the **correction must be done for each instrument**. This is necessary because, in theory, every instrument can be administered with a different incorrect Subject ID, so we cannot assume that two instruments administered to the same Subject ID do, in fact, belong to the same subject.

5.2.2 Suggestions for Locating Misnamed Imported Records

The most common reason for missing imported records are mis-entered Subject IDs. To locate the missing records in the “N-CANDA Imported from Laptops” project, you will then have to use the data collection date.

A convenient way for doing just this is as follows:

1. In the “Imported from Laptops" project, click on the instrument in the “Data Collection Instruments” menu on the left.
2. Under "Data Search" select "record_id" as the "Field to search"
3. In "Search query" box, start typing the date. Remember that all dates are in the format YYYY-MM-DD, e.g., July 15, 2013 would be typed as "2013-07-15".
4. You should see a list of matching records from the given date. You may first want to check for any that have an almost-correct Subject ID (but with a wrong digit, check sum, or sex code).

Note that for some instruments (most notoriously ePrime-based Stroop), the Subject ID may look entirely invalid – see Section 5.2.4.2 below for more details. For Stroop in particular, you may use this to quickly narrow the pool of missing records by specifically checking the “Imported” project for Stroop records with seemingly non-sensical Subject IDs but meaningful collection dates. These records are likely the result of mis-entered subject and session numbers.

5.2.3 Record Missing in the “Imported” Project

If a record matching the desired subject and visit date cannot be found, you will need to check whether that record is present on the data capture laptops in the first place. If in doubt, repeat the “upload” procedure for the respective laptops and watch carefully for any errors. See Section 9 below for suggestions how to fix common transmission problems.

If data are being sent by the laptop to the server in general, but a file for the missing record is not included, then you will need to find the correct file on the laptop in its original place (i.e., where the data collection software first created it). Often, mis-entry of the Subject ID will lead to file names that are missed by the data transmission script (e.g., by introducing spaces in the file name). See the following section for hints as to where to locate these files for different instruments.

5.2.4 Locating Data Files on the Collection Laptops

5.2.4.1 Locating LimeSurvey Files

The LimeSurvey files are stored in CSV format (file suffix “.csv”) in the following folder on each Mac laptop:

`/Library/WebServer/Documents/limesurvey/upload/csvdumps`

Each survey type is identified by a numerical code, which is part of the file name. The codes are as follows (note that some surveys have more than one code; the respective files can use either one, depending on what laptop they were collected on):

Baseline Visit	
11584	Youth Report 1 (Part A and B)
12471	Youth Report 2
31627	Parent Report
32869	MRI Report
1y Follow-up Visit	
13947	Youth Report 1 Part A
72223	Youth Report 1 Part B
92874	Youth Report 2
21598	Parent Report
14134	SSAGA, Part 1 (Youth or Parent)
81475	SSAGA, Part 2 (Youth or Parent)
91768	SSAGA, Part 3 (Youth or Parent)
12396	SSAGA, Part 4 (Youth or Parent)
6-monthly Phone Interview	
54587	Midyear Youth Interview
Special Projects: Sleep	
29361 or 82312	Sleep: Evening Questionnaire
96417	Sleep: Morning Questionnaire
34495 or 88371	Sleep: Pre-sleep Questionnaire

5.2.4.2 Locating Stroop Files

Files for the "on-land" (non-MRI) Stroop Match-to-Sample test are created in the following directory on the Windows laptops:

`c:\Users\kcummins\Desktop\NCANDA_StroopMtS_110112`

Most importantly, recall that due to limitations of the ePrime software, Subject IDs are translated to two 4-digit numbers (Subject and Session Number) before being entered into ePrime. These numbers are then back-translated into N-CANDA Subject IDs when they are imported.

Mis-entered subject and session numbers will lead to invalid Subject IDs. It is therefore possible that existing Stroop files are actually imported into REDCap, but you did not find them because they are listed under a seemingly non-sensical ID.

If you know what the mis-entered subject and session numbers were, you may use the example below to determine the Subject ID under which the record was imported into REDCap.

Example:

Subject ID B-00017-M-2 translates to ePrime Subject Number “2000” and Session Number “1712”.

Let's say these were actually entered (mistyped and switched) as ePrime Subject=17, ePrime Session=2.

When determining the N-CANDA Subject ID, the import system proceeds as follows:

1. Pad ePrime subject and session numbers to four digits by adding zeroes on the LEFT, resulting in our example in ePrime Subject=0017, ePrime Session=0002
2. The first digit of the ePrime Subject number codes site as follows: 1=A, 2=B, 3=C, 4=D, 5=E, 6=F, 7=G, 8=H, 9=I, and 0=J. In our example, this results in Site Code "J".
3. The second through fourth digit of ePrime Subject Number and first and second digit of ePrime Session Number code the numerical part of the subject ID, in our example: “01700”
4. The third digit of the ePrime Session Number codes Sex: 1=M, 2=F, 9=T, all other digits "X". In our example, the Sex code is therefore “X”.
5. The last digit of the ePrime Session Number is the N-CANDA Subject ID check sum digit, in our example “2”.

In summary, the mis-entered ID information above leads to the N-CANDA Subject ID

J-01700-X-2.

This is the ID under which the record should be found in REDCap.

Note that the import system determines the ePrime Subject and Session Numbers from the **contents** of the ePrime output file, **not** from its file name. Therefore, fixing the file name (which should not be done in the first place, because it creates duplicate records) will not fix the imported subject ID.

5.2.4.3 Delayed Discounting

Result files for Delayed Discounting should be found in the following directory on the Windows laptops:

```
c:\temp\local
```

5.2.4.4 Blaise SSAGA (Youth and Parent)

The Blaise SSAGA surveys are transmitted in a single database containing all records collected on the given laptop. Therefore, it is impossible for a single record to be missing if other records from the same laptop have been successfully transferred.

If specific records are missing, they are either misnamed (wrong Subject ID), or the respective laptop may be having general data upload problems (see Section 9 for instructions how to fix some of these).

5.2.4.5 PASAT

The PASAT results for all subjects collected on a laptop are transmitted in a single database. Therefore, it is impossible for a single record to be missing if other records from the same laptop have been successfully transferred.

If specific records are missing, they are either misnamed (wrong Subject ID), or the respective laptop may be having general data upload problems (see Section 9 for instructions how to fix some of these).

5.2.5 Imported Record is Correct but not Assigned

Finally, there may be cases where a missing record has the correct Subject ID and is, in fact, present under the correct ID and date in the "Imported from Laptops" project, yet it still does not appear in the "N-CANDA Data Entry" project for the correct subject and visit.

In these cases:

1. Make sure any corrections to the imported record have been done at least 24h ago, imported records are propagated into the "Data Entry" project only once per day (over night). Any corrections become effective only during that operation.
2. Confirm that the date on which the respective instrument was administered is within 30 days of the "Visit Date" entered for that subject and visit in the "Data Entry Project." Only records within 30 days of that date will be accepted as belonging to the same visit. (By extension, also confirm that there is a "Visit Date" entered for the subject and visit in the "Data Entry" project in the first place.)

5.3 Correcting Data Imported From Laptops

Occasionally, it may be necessary to correct the actual data imported from thedata collection laptops, beyond correcting subject IDs or adjusting collection dates for the purpose of visit assignment discussed above.

All changes to imported data must be made in the "Imported from Laptops" project. While all REDCap users have the necessary permissions to change data in that project, it is important that **only designated "Super RAs" or "Site Data Managers" perform these corrections**, because only they have the necessary permissions to protect corrected records from future overwriting.

Note that this correction procedure only applies to data imported directly from the collection laptops, such as: LimeSurvey, Delay Discounting, and PASAT. Data entered directly into REDCap (e.g., vision, screening, and tox screen forms) should be corrected directly in their

respective entry forms.

Note also that the **Blaise SSAGA and Stroop data can not be corrected**. These instruments undergo a scoring procedure of the collected raw data, which can not be repeated after data have been imported into REDCap.

When applicable given these limitations, correction of imported data involves the following steps, which are also detailed below and illustrated using screen shots:

a. Locate the record that needs to be corrected in the "Data Entry" project and determine the record ID of its source in the "Imported From Laptops" project.

b. Set the "Complete" status of the form in "Data Entry" to "Incomplete" or empty. This is important because otherwise the data in the form will not be updated after the correction.

c. Locate the imported record in "Imported From Laptops."

d. Make the necessary corrections in "Imported From Laptops."

e. "Lock" the corrected record in "Imported From Laptops." This is important because otherwise the corrected record may be overwritten in the future with the original, incorrect record as it is stored on the laptop.

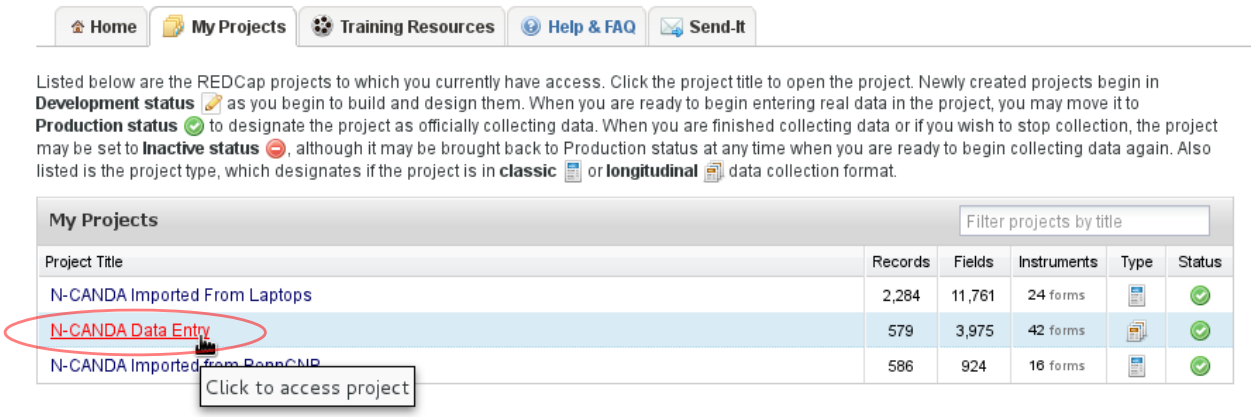
f. "Save" the corrected record in "Imported From Laptops."

After this, the corrected record should appear in the "Data Entry" project within a few hours, or at most one day.

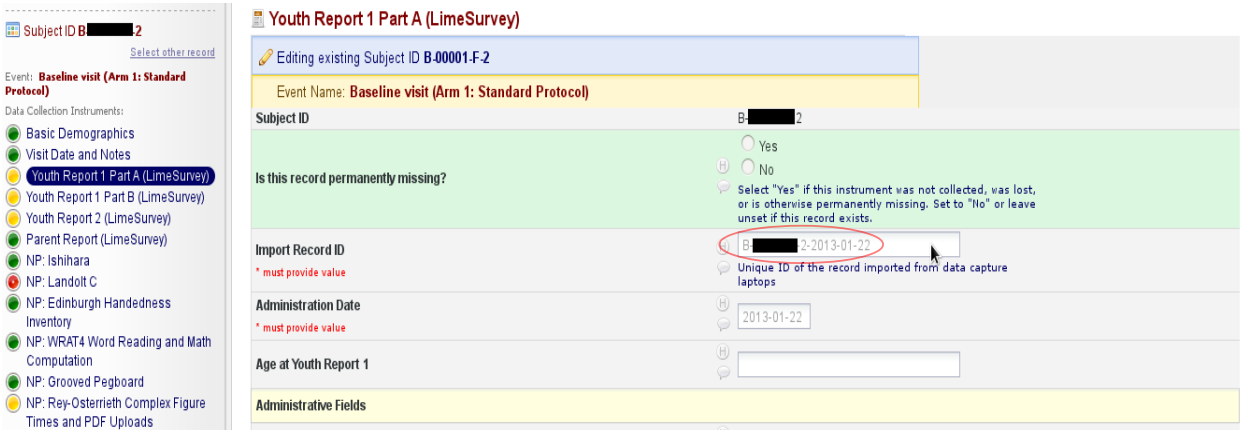
5.3.1 Step-by-Step Illustrated Instructions

Note that you will not be able to perform all necessary steps of this procedure unless you are a designated "Site Data Manager." All other REDCap users will be able to correct the imported data, but **will not be able to protect the corrected data from overwriting** in the future.

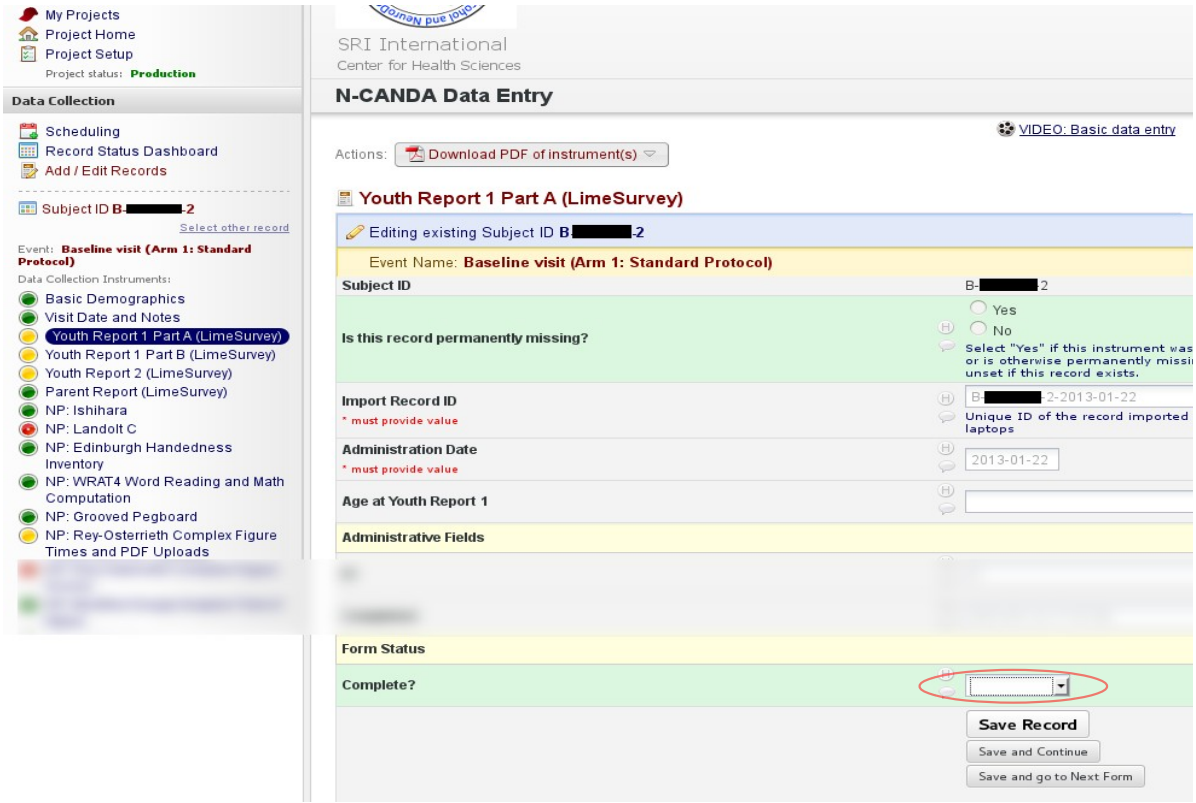
Step 1 – Navigate to the "Data Entry" project



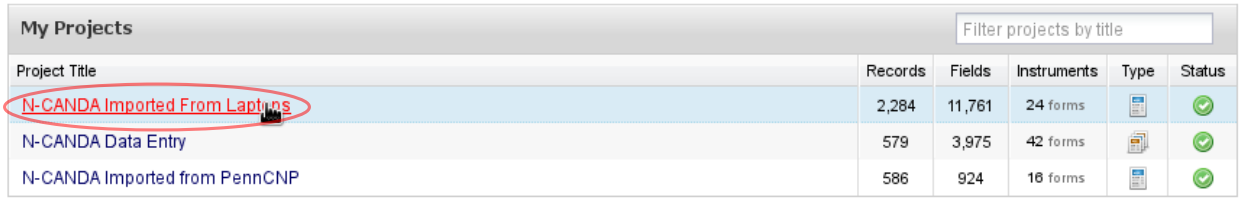
Step 2 – Navigate to the incorrect record in "Data Entry" and determine the "Import Record ID."



Step 3 – Mark "Complete" status of form (this is all the way at the bottom) in the "Data Entry" project as "Incomplete" or empty, then "Save Record"

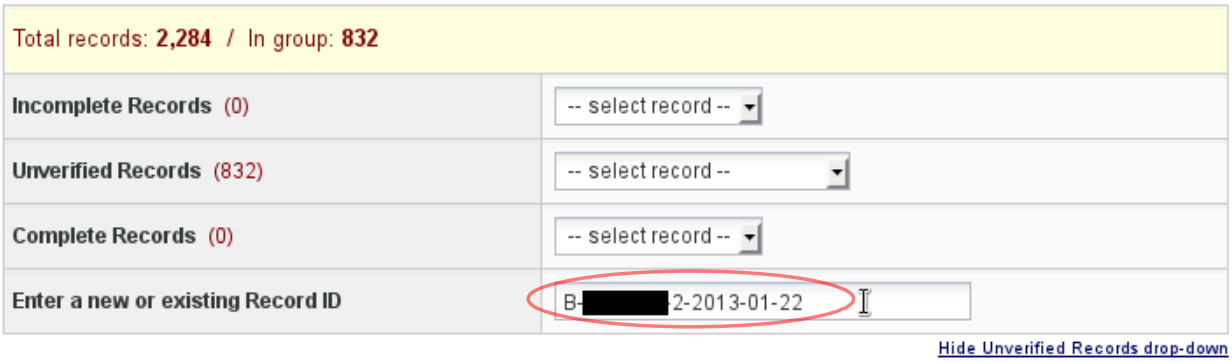


Step 4 – Navigate to the "Imported from Laptops" project

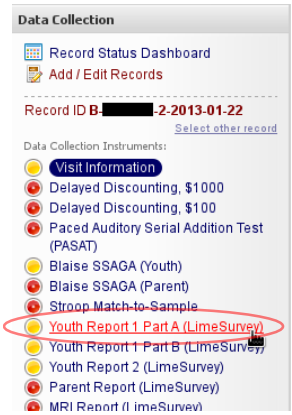


Step 5 – Click on "Add / Edit Records"

Step 6 – Enter previously noted "Import Record ID"

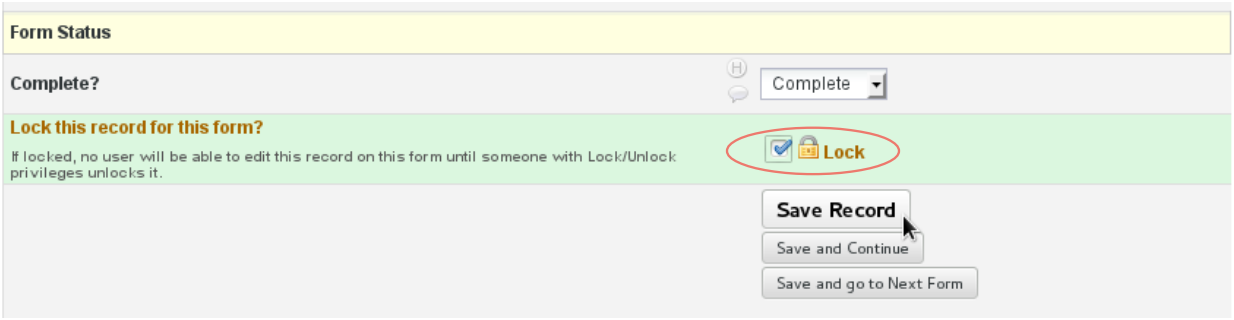


Step 7 – Select form to correct



Step 8 – Make the necessary corrections to the form data

Step 9 – "Lock" the corrected form and "Save Record" (these two are at the very bottom of the form).



The last step is particularly important, because failure to "Lock" the corrected record will leave it vulnerable to overwriting, should it ever be re-imported from the collection laptops.

6 Data Imported from U-Penn WebCNP System

6.1 Outline

Import of scores from the WebCNP system at U Penn involves the following stages:

1. A CSV file with all scores is download from the U Penn server.
2. The downloaded scores are imported into the "N-CANDA Imported From PennCNP" project in REDCap.
3. The subjects and visits entered into the "N-CANDA Data Entry" REDCap project are used to select the correct records from the imported data.
4. Selected imported records are copied into the "Penn WebCNP Summary" form in the "N-CANDA-Data Entry" project.

Note that each record imported from WebCNP is identified with a unique identifier, which is a combination of the Subject ID entered, date of administration, and a unique ID number assigned by the WebCNP system. **This identifier is permanent and cannot be edited.** Both Subject ID and administration date are, however, also stored separately for each imported record, and these separate values can be edited to correct errors.

6.2 Locating Missing Records

See the example below of the Record Status (in the "Data Entry" project) for a subject that should have CNP data available for the Baseline visit. The red status indicator shows that these data have not been successfully assigned.

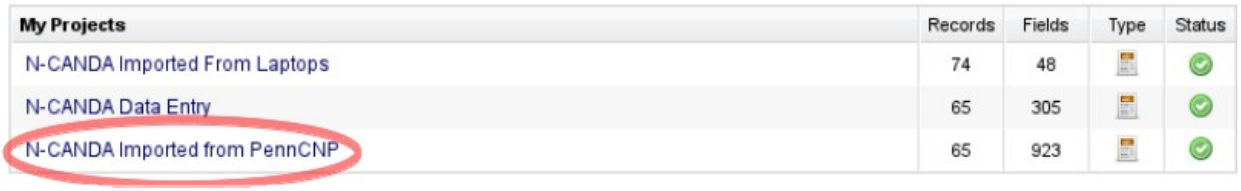
Subject ID **B-00086-M-2**

Data Collection Instrument	Events								
	Baseline visit (1)	6-month follow-up (2)	1y visit (3)	18-month follow-up (4)	2y visit (5)	30-month follow-up (6)	3y visit (7)	42-month follow-up (8)	4y visit (9)
Basic Demographics									
Visit Date and Notes									
NP: Ishihara									
NP: Landolt C									
NP: Edinburgh Handedness Inventory									
NP: WRAT4 Word Reading and Math Computation									
NP: Grooved Pegboard									
NP: Rey-Osterrieth Complex Figure PDF Uploads									
NP: Rey-Osterrieth Complex Figure Scores									
NP: Modified Gregly-Graybiel Test of Ataxia									
NP: WAIS-IV Coding									
Biological Data: NP Day									
Biological Data: MR Day									
Saliva Samples									
Delayed Discounting, $1000									
Delayed Discounting, $100									
Paced Auditory Serial Addition Test (PASAT)									
Penn WebCNP Summary									
MRI Session Report									

To locate the missing data, follow the steps outlined below.

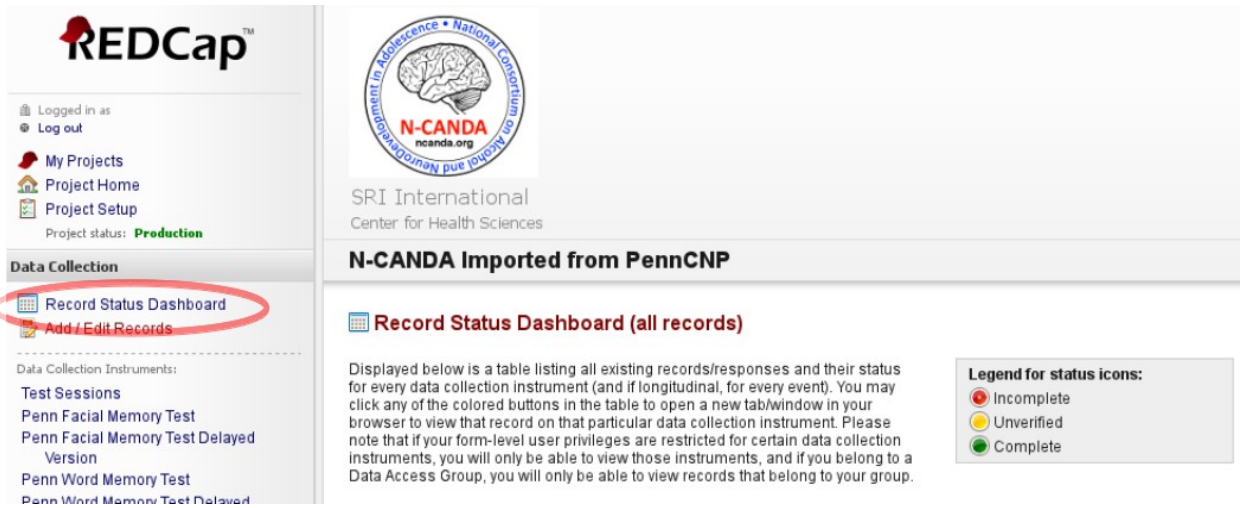
6.2.1 Locating and Correcting Imported Data

The first step to reconcile imported data with the final, compiled data set is to locate the missing data in the imported records. To this end, from the list of "My Projects", select the "N-CANDA Imported from PennCNP" project:



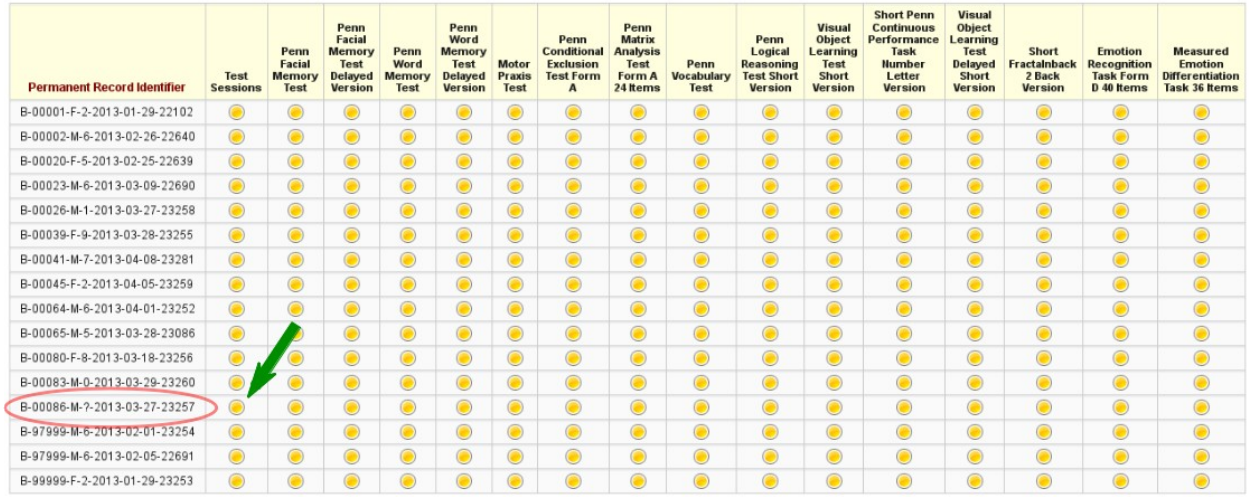
My Projects	Records	Fields	Type	Status
N-CANDA Imported From Laptops	74	48		
N-CANDA Data Entry	65	305		
N-CANDA Imported from PennCNP	65	923		

Next, select the "Record Status Dashboard" from the menu on the left-hand side of the screen:



In the Record Status Dashboard, identify the record containing the missing data. Each record is identified by a) the subject ID as entered into WebCNP, b) the test administration date, and c) the unique record number assigned by the WebCNP system.

If a records is present but not successfully assigned to the correct subject and visit, most likely, the Subject ID was mis-entered (as in the example below).

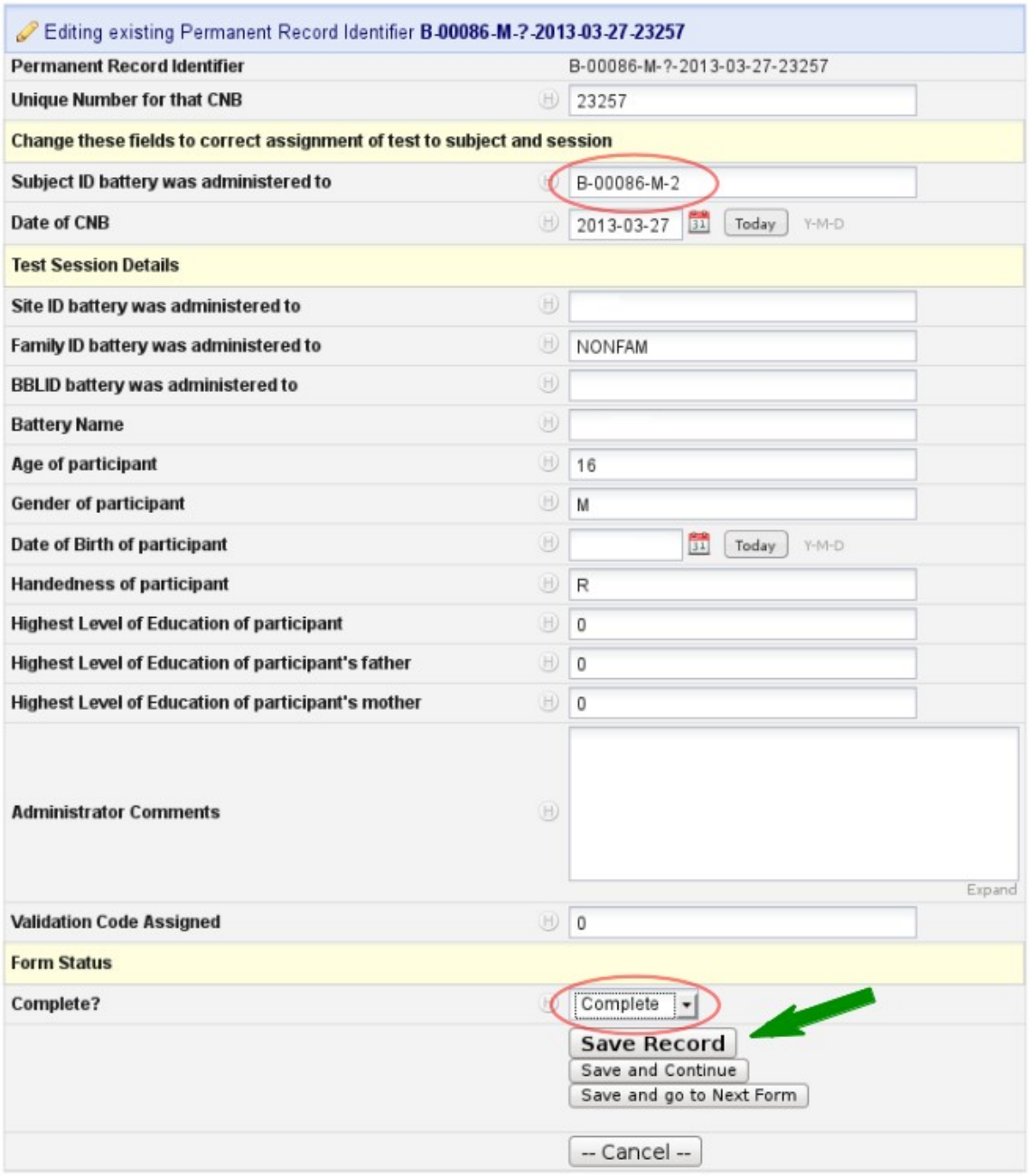


Permanent Record Identifier	Test Sessions	Penn Facial Memory Test	Penn Facial Memory Test Delayed Version	Penn Word Memory Test	Penn Word Memory Test Delayed Version	Motor Praxis Test	Penn Conditional Exclusion Test Form A	Penn Matrix Analysis Test Form A 24 Items	Penn Vocabulary Test	Penn Logical Reasoning Test Short Version	Visual Object Learning Test Short Version	Short Penn Continuous Performance Task Number Letter Version	Visual Object Learning Test Delayed Short Version	Short Fractalnback 2 Back Version	Emotion Recognition Task Form D 40 Items	Measured Emotion Differentiation Task 36 Items
B-00001-F-2-2013-01-29-22102																
B-00002-M-6-2013-02-26-22640																
B-00020-F-5-2013-02-25-22639																
B-00023-M-6-2013-03-09-22690																
B-00026-M-1-2013-03-27-23258																
B-00039-F-9-2013-03-28-23255																
B-00041-M-7-2013-04-08-23281																
B-00045-F-2-2013-04-05-23259																
B-00064-M-6-2013-04-01-23252																
B-00065-M-5-2013-03-28-23086																
B-00080-F-8-2013-03-18-23256																
B-00083-M-0-2013-03-29-23260																
B-00086-M-?-2013-03-27-23257																
B-97999-M-6-2013-02-01-23254																
B-97999-M-6-2013-02-05-22691																
B-99999-F-2-2013-01-29-23253																

Once the missing record has been located, click on its “Test Sessions” status indicator (shown in the previous figure) to correct it.

In the Entry Form that follows, proceed to

1. change the Subject ID to the correct one,
2. change the “Complete?” status to “Complete” (this will protect the corrected record from overwriting), and
3. press “Save Record”



Note that the correction has to be performed only once, in the "Test Sessions" form, and will be applied to all instruments acquired through the WebCNP system for this subject and visit.

6.2.2 Record Missing in the Imported Dataset

If a record matching the desired subject and visit date cannot be found, you will need to check whether that record is present in the WebCNP database at U Penn in the first place. See separate instructions on use of the WebCNP system for more information on how to locate records in their system and how to reconcile data capture laptops if necessary to upload missing data.

6.2.3 Imported Record is Correct but not Assigned

There may be cases where a CNP record has the correct Subject ID etc. But still does not appear in the "N-CANDA Data Entry" project for the correct subject and visit. In these cases:

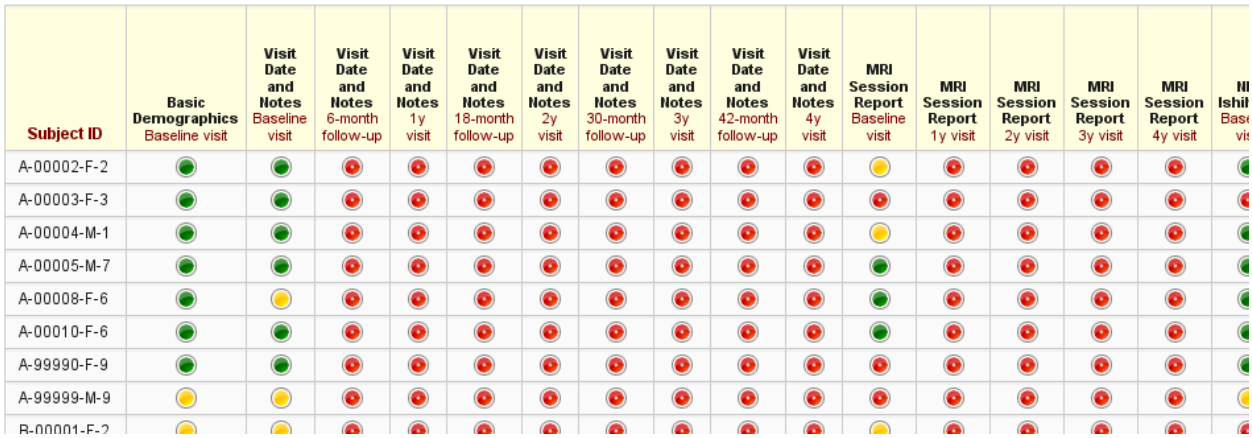
1. Make sure any corrections to the imported record have been done at least 24h ago, imported records are propagated into the "Data Entry" project only once per day (over night). Any corrections become effective only during that operation.
2. Confirm that the date on which the WebCNP tests were administered is within 30 days of the "Visit Date" entered for that subject and visit in the "Data Entry Project." Only records within 30 days of that date will be accepted as belonging to the same visit. (By extension, also confirm that there is a Visit Date entered in the first place.)

7 Cross-Referencing with Imaging Database

Every night, the records in the REDCap database are cross-referenced with those in the N-CANDA XNAT Image Database, which stores the data for all MR imaging series.

For each subject and visit, there is a "MRI Session Report" form in the "N-CANDA Data Entry" project. These forms are read-only and not intended for manual data entry or correction, as they are automatically populated.

In the following figure, note the three different colors in the MRI Session Report (baseline visit), which represent different levels of Quality Assurance:



Subject ID	Basic Demographics Baseline visit	Visit Date and Notes Baseline visit	Visit Date and Notes 6-month follow-up	Visit Date and Notes 1y visit	Visit Date and Notes 18-month follow-up	Visit Date and Notes 2y visit	Visit Date and Notes 30-month follow-up	Visit Date and Notes 3y visit	Visit Date and Notes 42-month follow-up	Visit Date and Notes 4y visit	MRI Session Report Baseline visit	MRI Session Report 1y visit	MRI Session Report 2y visit	MRI Session Report 3y visit	MRI Session Report 4y visit	NI Ishil Base vis
A-00002-F-2																
A-00003-F-3																
A-00004-M-1																
A-00005-M-7																
A-00008-F-6																
A-00010-F-6																
A-99990-F-9																
A-99999-M-9																
B-00001-F-2																

Red – "Incomplete"

> There are no MRI data for this subject and visit in the XNAT imaging database. This could be the result of a mis-entered subject ID in either XNAT or REDCap. It could also be the result of incorrect or missing subject visit date in REDCap (e.g., the visit date entered may be after the MR imaging date, in which case an existing MRI session would not be assigned to the correct visit).

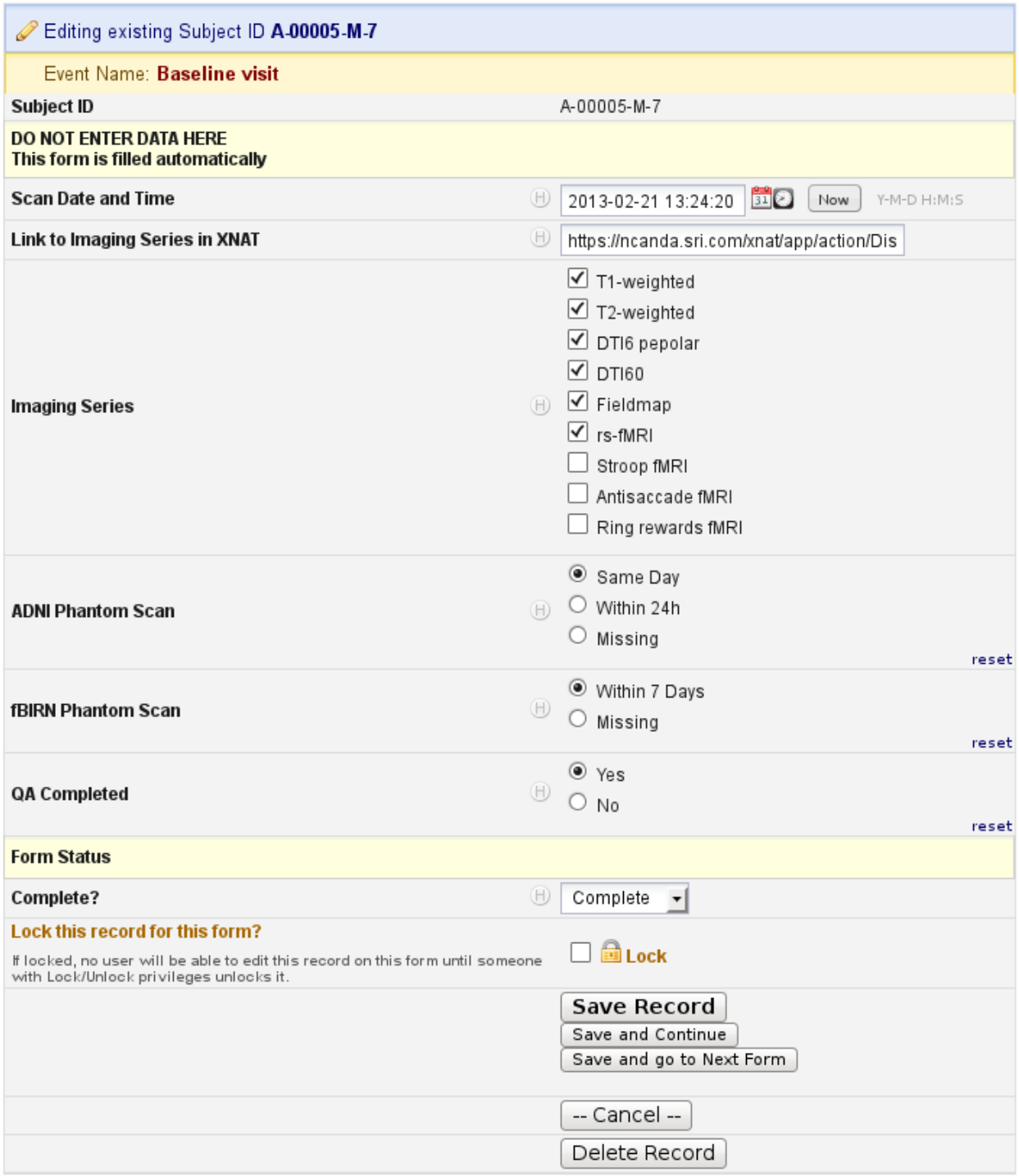
Yellow – "Unverified"

> An MRI session was found, but is still undergoing quality assurance.

Green – "Complete"

> The MRI session has been fully inspected and released for analysis.

The report form for each subject and visit contains the following data:



The “Link to Imaging Series in XNAT” can be copied and pasted into a web browser to access directly (after login) this imaging series.

“QA Completed” denotes whether all scans in the session have been inspected and labeled as “usable” or “unusable.”

8 Study Arms

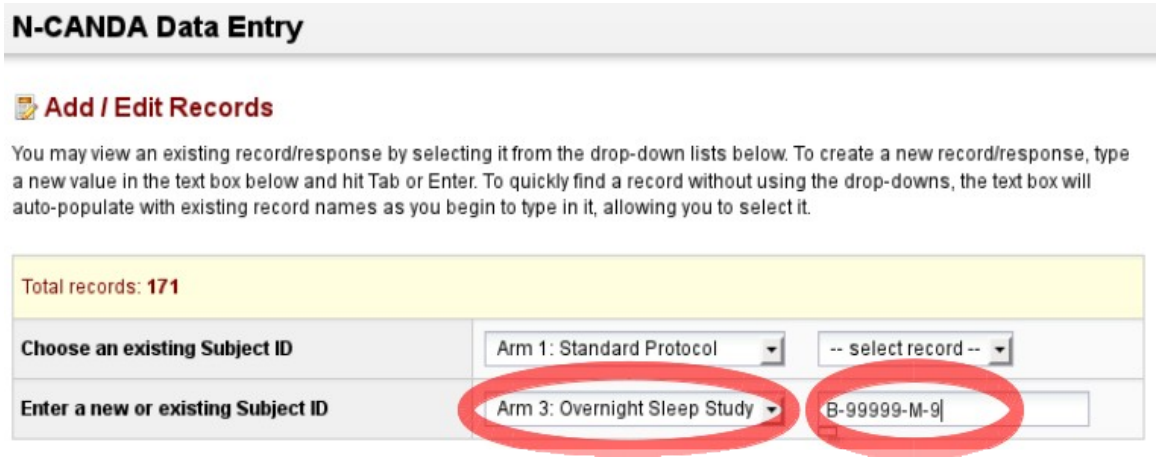
In addition to the Standard Protocol of instruments administered to every subject in the N-CANDA study, there are several optional "arms" that only some subjects will be participating in. Examples of these are the Recovery Protocol ("Arm 2" in REDCap) and the Overnight Sleep Studies ("Arm 3" in REDCap). Below, we describe how to "enroll" a subject into these arms and what data need to be entered for each arm.

8.1 Arm 3: Overnight Sleep Studies

8.1.1 New Subject Entrollment

To enter a subject into the overnight sleep study protocol, choose the "Add / Edit Records" link from the "Data Collection" menu of the "N-CANDA Data Entry" REDCap project.

In the entry screen that follows, first set the study arm for "Enter a new or existing Subject ID" to "**Arm 3: Overnight Sleep Study**." Then, enter the subject ID in the entry field on the right.



NOTE: We strongly suggest to verify before entering a subject into the Sleep Study Arm that this subject has already been entered into the Standard Protocol (Arm 1).

To confirm that the newly enrolled subject is already in the Standard Protocol, verify that REDCap shows the following notice after entering the Subject ID:

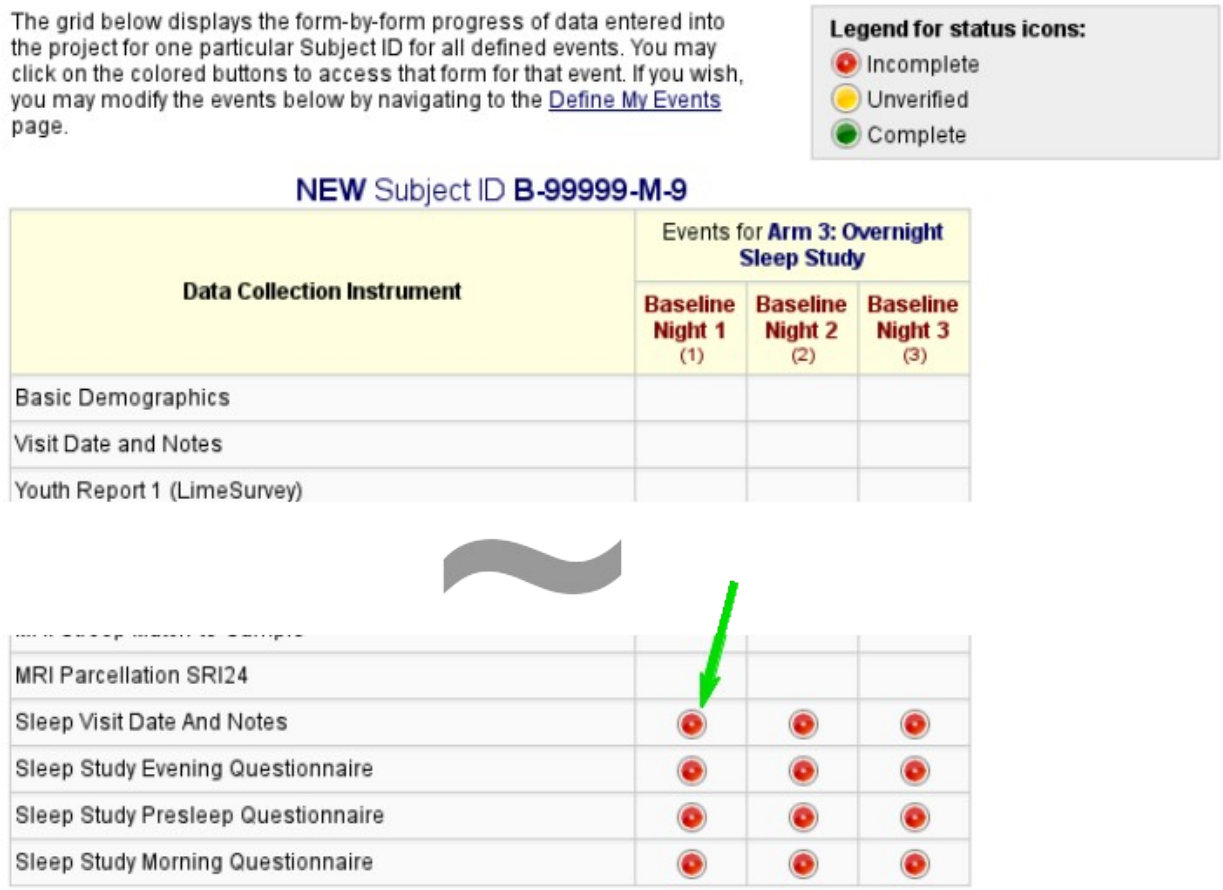
If this notice appears (above the "Event Grid" showing the instruments for each visit), then the subject is already in the Standard Protocol and it is safe to proceed.

If this notice does **not** appear, then the subject should be entered into the Standard Protocol first. It is also possible that the Subject ID for enrollment into the Sleep Study was mistyped, which should be verified.

8.1.2 Sleep Visit Data Entry

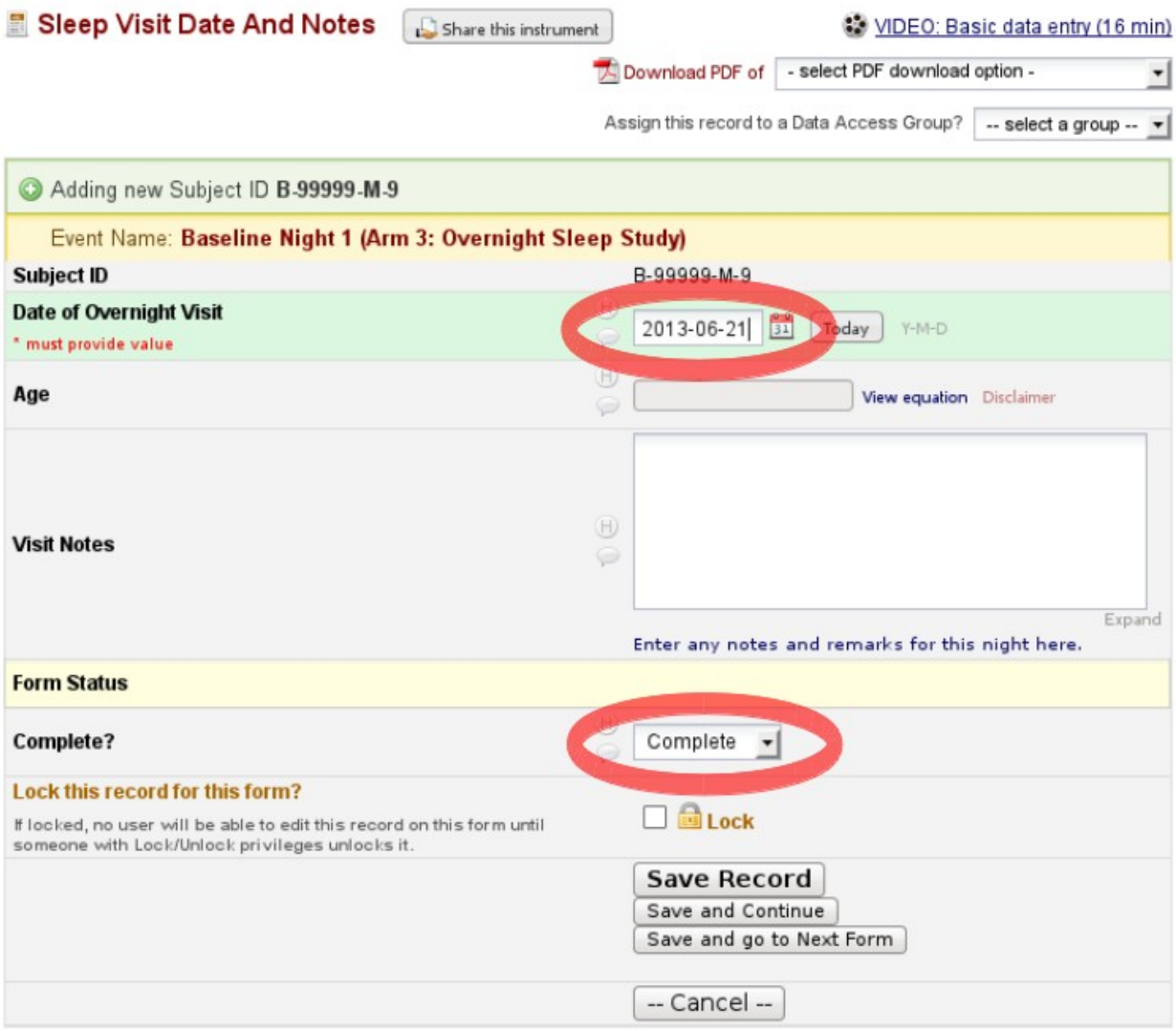
Each "Visit" of the sleep study comprises three separate sub-events, one for each overnight stay. For each night, the date must be entered manually. This allows the system to assign data imported from the data collection laptops (e.g., surveys) to the correct subject and visit. Requiring the manual entry of the date for every one of the three nights of a visit accommodates cases in which the nights may not be on directly successive dates.

To enter data for a night of one visit, find the sleep-related instruments at the bottom of the "Event Grid" displayed after entering the Subject ID:



Date entry for one visit is begun by clicking on the status icon for the "Sleep Visit Data and Notes" instrument (e.g., see green arrow in picture above for Night 1 of the Baseline visit).

The form for entering visit date (and notes, if there are any) is quite simple (see below). The only required field is the visit date itself:



Subject age at that date is computed automatically. After entering the visit date, set the Form Status to "Complete" and press the "Save Record" button.

8.1.3 Confirming Data Imports

Sleep-related data imported from the data capture laptops (e.g., LimeSurvey instruments) should appear about 24h after submission from the laptops. Arrival of these data is indicated by the Form Status display in the Event Grid of existing subjects:

If data from the laptops do not appear within a few days, the standard instructions for locating missing data apply (see Section 5.2 above).

Note that unlike the instruments in the Standard Protocol (Arm 1) of the project, the instruments in the Overnight Sleep Study arm require exact date matches for assignment to the correct visit and night.

Specifically, the Evening Questionnaire and Presleep Questionnaire must be **dated exactly as the date entered** for that visit and night. The Morning Questionnaire must be **dated exactly one day later**.

9 Common Problems and Solutions

9.1 Laptop not sending data to Subversion (SVN) repository

When the data upload is interrupted (e.g., by closing the upload window before the message appears that it is okay to do so), often the local copy of the data to be transferred remains in a "locked" state. This prevents all future uploads from the same laptop.

To fix this problem, follow these steps:

1. Log into the affected laptop
2. Open a terminal window ("Cygwin Terminal" on the Dell machines; "Terminal" on the Macs)
3. **On a Mac**, run the following command (preferably copy & paste into terminal window):

```
cd /Users/ncanda/Documents/ncanda_data

svn cleanup

svn update
```

4. **On a Dell**, run the following command:

```
cd Documents/ncanda_data

svn cleanup

svn update
```

If any of the above commands returns an error, please send an exact copy of that error to torsten@synapse.sri.com.

Note that it **is safe to perform the above procedure** even if the actual problem is different from the ones that it is designed to fix.