Configuring the NextGen Validation Report Set

By Ben Fouts, RCHC Data Analyst

Special Session for NextGen/Relevant Users

Redwood Community Health Coalition, August 12, 2020

Agenda

1. Discussion of consistent data entry definitions
2. Validation report design and the relationship between Transformers and Importers
3. Using Value Sets
4. Problem list validation reports
5. Lab-related validation reports
6. The Cancer Exclusion Validation Report
7. The Incomplete or Non-Standard Demographic Data Report

Instruction Manual

Instructions for Using the Relevant Validation Report Set

(NextGen Edition, Version 2)

REDWOOD HEALTH COMMUNITY COALITION

Serving Sonoma, Napa, Marin & Yolo Counties

Author: Ben Fouts, Informatics
Redwood Community Health Coalition
1310 Redwood Way, Petaluma, California 94954
support@rchc.net

To Copy the Reports to Your Instance of Relevant

- See Appendix B of the Instruction Manual
- Before actually running the report in your instance, it is recommended that you first copy the code to DataGrip so you can test it.
- Note that you cannot run the validation reports in the RCHC instance. They exist there only for you to COPY.
- Document any changes you made to the report so they can be recreated if a new version is released

Discussion of Consistent Data Entry Definitions

Entering data in a uniform manner…

Pulling data in a uniform manner

The Benefits of Consistent Data Entry Definitions

- Data of a particular data type should also be pulled in the same way by Relevant.
- The data types are, for example, diagnosis, labs, images, medications and vaccines.
- Therefore, Transformers for each data type should be designed in a similar manner and pull and transform the data in a uniform way.
- The identification and transformation of data should be harmonious with standard data-entry procedures and staff training

The Benefits of Consistent Data Entry Definitions

- Standardizing the way that data is being pulled helps with the accuracy of the data that is used in all sections of Relevant, like for QMs, reports, Care Gaps, etc.
- Data from Relevant should be a reflection of what can be seen in NextGen. In other words, the data in Relevant makes sense and is believed to be “true”

The Benefits of Consistent Data Entry Definitions

- Taking a consistent approach also helps with staff training.
- Ideally, there should be a standard set of instructions or procedures that all staff should be trained on and must follow. Transformers are built on this premise.
- The validation reports should follow this standard approach and thus display records where this approach is not used. Once recognized, these records should be corrected so they can be accepted by the Transformers and incorporated into the standard QMs, reports, etc.
- The instructions should be a combination of data entry recommendations from NextGen and also the health-center specific workflows

The Benefits of Consistent Data Entry Definitions

- Keep in-mind that these kinds of procedures/instructions and staff training should come FIRST and then Relevant Transformers and other objects designed based on them.
- In an ideal world, Relevant should not “chase” multiple non-standard ways that staff are entering data
- Staff should be told what to do and then Relevant reports should monitor the quality of data entry and then feedback should be given to staff in a positive but corrective manner

The Benefits of Consistent Data Entry Definitions

- For example, all labs should be ordered in THIS section of NextGen with THESE fields completed in THIS way.
- When the results come back, they should be reviewed in THIS section of NextGen with THESE fields completed in THIS way.
- Therefore, a lab can be defined as being “complete” by looking at the data in defined fields.
- The **lab validation report** shows QM-related lab records that are partially complete, but are missing key information. This report can be used to ensure data quality and also to give feedback to staff responsible for entering the data

The Benefits of Consistent Data Entry Definitions

- A particular data set should be pulled the same way for each reason.
- For example, “patients with diabetes.”
- Where applicable (in combination with other characteristics), the same group of patients with diabetes should be identified for all uses.
- In Relevant, the best approach is through a Transformer (i.e. relevant_diabetics).
- Similar code should be used in the Transformer as used for other similar Transformers that pull patients with chronic diseases.

Inconsistent Transformer Definitions

It was difficult to design some of the validation reports because:

- Often, Transformers that pull a particular data type (like the sub-set of Transformers that pull QM-related labs) do so in different ways. For example, A1c and LDL labs are pulled in different ways (i.e., using different criteria from different tables in Relevant).
- Between health centers, Transformers pull data in different ways so that one approach is not possible. For example, A1c labs are pulled in 3 different ways for the 3 NextGen health centers

Recommendation: Standardize Everything!

- Start with the data entry procedures, instructions and trainings. Are these documented at your health center? Are they updated? Do they make sense?
- Based on the documentation, have the Transformers and Importers been designed to capture data from the proper sources (i.e., the appropriate tables and fields identified in the documentation)?
- Do all Transformers/Importers pull the same data type in the same way? Can they be standardized to match whatever ‘Golden Standard’ for data entry you have documented?
- Look at the Validation Reports. Are they pulling the data in the same way that the Transformers/Importers pull the data?

Following National Standards for Data Definition

- The Validation Reports use national standards for defining data types
- Each of these data types are defined by codes:
  - Diagnoses (ICD)
  - Procedures (CPT)
  - Labs (LOINC)
  - Medications (RxNorm)
  - Vaccines (CVX)

National Standards

- Codes are defined by Value Sets provided by national organizations
- The measure definitions refer to particular Value Sets
- The validation reports use Value Sets
- Value Sets containing all of the codes exist in the Relevant Staging Database on the table cqm_value_set_codes
- There are validation reports that display these codes and the corresponding data from NextGen
- We will discuss these reports in more detail later in this presentation

Report Customization

- One set of Validation Reports exists. They have not been customized for each health center
- Customizing must be done by the health centers
- Consider the ideas for standardization mentioned in the previous slides. After standardization is complete, then review the code in the validation reports and adjust accordingly.

Report Customization

- Health centers have complete control over their validation reports. So, feel free to change anything you like in order to make the reports as useful as possible.
- But that also gives you the right to not align them with standards, or to customize them based on the potential variety of different approaches to data extraction that currently exist in your system
- For example, the lab validation report pulls labs in a standard manner. A health center can certainly re-code the report using seven different ways of pulling seven different labs. However, now might be an opportunity to think about and standardize things!

Working With Validation Report Code

- It is suggested that you copy the report code to DataGrip and look for any references to tables and fields that are not “found.” In most cases, a table or field might have a slightly different name.
- Furthermore, you might need to cast key fields (the ones involved with JOINs) so they connect together properly. It seems like the patient id is stored as text on some tables and as an integer on others.

Working With Validation Report Code

- More complicated configuring issues are discussed in the section “Custom Set-up at Health Center” of each report in the instruction manual
- It is recommended that you document the customization you do to the reports and why you did them. Therefore, if a new version is released, you will be able to re-create the changes.

Relationship Between Transformers, Importers, and Validation Reports

Validation reports are built in the Relevant Staging Database but sometimes need data only available from Importers

Outline of the Relevant Data Pipeline

Relevant Component

Raw Tables From EHR or Other Sources
- For example, enc, patients, doctors, etc. (generally names with all lower-case letters)

Transformers
- Based on the raw tables. Flexible in design. SQL begins with "CREATE TABLE…" with a new table name.
- Normally named with Create or Build (standard ones have "relevant_" in the name). Can add new columns of data (standard ones "relevant_" in the name)

The collection of raw tables and tables created by the Transformers is the **STAGING DATABASE**

Importers
- Based on the tables in the Staging Database. Rigid in design. SQL begins with SELECT statements.
- Provides data for a set of tables standardized for all health centers
- Name is capitalized

Importers populate the tables in the **ANALYTICS DATABASE** (aka, PRODUCTION DATABASE)

Importers provide a standard set of data from all health centers that populate the **RCHC AGGREGATE DATABASE**

Example

vitals

Build relevant_bmi

relevant_bmi

BMI Measurements

bmi_measurements

bmi_measurements

Data Extraction Tools

Reports
- Can be based on the Staging Database or the Analytics Database
- Generally record-based but then summarized with Pivot Tables
- Can be run with parameters (for example, a flexible measurement period)

Standard Measures
- Based on the Analytics Database
- Track measures over time

Data Grip
- Used for design and testing Transformers, Reports and Standard Measures
- Features tables and data from the Staging and Analytics Databases
- Can connect to other data sources (for example, dental databases, lists of Partnership patients, etc.).

Validation Report Design

- The Validation Report set has been designed in the Staging Database because it needs to compare raw data to data transformed by the Transformers.
- The Staging Database contains a different set of tables than the Analytics database.
- In the Relevant data model, data is pulled by the Transformers and then Standardized into particular formats by the Importers.

Validation Report Design

- However, note that there are some Importers that contain code to pull particular data directly, instead of related Transformers
- Operationally, this works for QMs because QMs are standardized to work off of the Analytics Database (i.e., from Importer data). Once the data is in the Analytics Database, it does not matter if it was created originally in the Staging or the Analytics Database.
- However, this poses a problem if you want to use data in a report or Care Gap based out of the Staging Database but the data only exists in the Analytics database.

Validation Report Design

- Therefore, for some validation reports, there is a note to COPY the Importer code into a TEMPORARY TABLE so that it is available for the report.
- Check to see if your health center has the code in an Importer. If so, replace the default code.
- In a few examples I have seen, all the NextGen health centers have the code in an Importer, so maybe this was how Relevant originally designed it.

Validation Report Design

If the code is in the Importer, copy it to the TEMPORARY TABLE (replace the default code in the report).

If the code actually exists in a Transformer, then you can

1. Have the TEMPORARY TABLE just pull records from the Transformer in order to preserve the table name and field names needed in later parts of the code, or
2. Swap out the TEMPORARY TABLE name for the Transformer name in your code and use it directly

Using Value Sets in Relevant

And applying them consistently

Consider Using Value Sets in Transformers

- The advantage to using Value Sets is that the latest set is always available for use by the QMs.
- There is no need to update the codes in the system. However, if you do not use the Value Sets, then you do now know if you are using the most recent and correct codes.
- From one study that RCHC did comparing the 2018 to the 2019 Value Sets, the codes for at least one Value Set changed for seven of the thirteen UDS Quality Measures.

How Codes are Used in NextGen

- You should investigate the codes in your system.
- Which tables are they stored? (You can look at how the Validation Reports are coded to see my best guess.)
- How do the codes get in NextGen? Did they come with the software? Do you have to enter them?
- Is it possible to manually update codes in the system?

Example of the eCW Model

- In eCW, vaccine codes must be manually added. Therefore, the report “RCHC List All Vaccines NG” is used to list all vaccines, along with the CVX codes in the system, and then link those codes to the Value Sets.
- The report output is filtered in two ways:
    1. Vaccines with QM-related CVX codes are filtered and the names verified.
    2. Vaccines without CVX codes are filtered and names are scanned for QM vaccines that should have a CVX code. An appropriate CVX code is then added in the system

Adding Value Sets to Transformers

- Consider adding the Value Sets directly to the Transformers instead of using wild cards for names, or listing the codes manually.
- You should always be using the correct codes for the reporting year. Relevant provides Value Sets with those codes when the new QMs are released.
- If you do not use the Value Set codes, then you need to have some system to manually check at least once per year that each Transformer is picking up the correct items.
- How often are new labs, vaccines, medications, etc. added to the system? How sure are you that the current way your Transformers identify these items ALSO pick up the newly added items?

General List of Value Set Codes

RCHC List of QM Value Set Codes NG

- Simply lists all of the codes used by all of the QM-related Value Sets
- Is not dependent on anything in NextGen. Therefore, it is unknown if any of the codes were actually used

Codes Used in NextGen Within a Measurement Period

- RCHC List of QM Lab Names and Attributes NG
- RCHC List of QM Medications NG
- RCHC List of QM Vaccines NG

- Displays the item name in NextGen corresponding to the Value Set as well as the number of times it was used and the last use date

All Items Used Within a Measurement Period

- RCHC List All Lab Names and Attributes NG
- RCHC List All Medications NG
- RCHC List All Vaccines NG
- RCHC List All Diagnosis Codes NG

- These labs display the Value Set (if applicable), the number of times it was used, and the last use date.
- These reports can be used to identify items that are somehow NOT picked up by the Value Sets but SHOULD be

Problem List Validation Reports

Original Model is From eCW

Defining a Patient With a Chronic Disease

Validation reports:

- RCHC Problem List Validation Report NG
- RCHC Persistent Asthma Diagnosis Validation Report NG

Relies on how diagnosis codes are treated by the Transformers and Importers

Model for Defining a Patient With a Chronic Disease Borrowed From eCW

- The question is, can this model be applied to NextGen?
- With the tables and fields that exist in NextGen, is it logically possible to follow the model?
- Do the procedures, instructions and staff training follow the model?

eCW Chronic Disease Model

- A patient with a chronic disease is identified by a standard diagnosis code on the “Problem List”
- The Problem List is considered the single source of truth
- Codes on the Problem List can be added when there is a clinical diagnosis (i.e., an evidence-based diagnosis) and removed when it is no longer needed (i.e., if the condition has resolved or if there was a mistake)

eCW Chronic Disease Model

- This approach minimizes the number of false-positives and false-negatives in the denominator of a measure

- False-positive: a patient who is considered to have a chronic disease by the report but is NOT considered to have the disease in the clinic, and thus is NOT being treated for that disease according to guidelines
- This patient will count AGAINST the health center in the reported measure

eCW Chronic Disease Model

- False negative: a patient who is considered to NOT have a chronic disease by the report but is considered TO HAVE the disease in the clinic, and thus may be treated for that disease according to guidelines
- This patient will not be included in the denominator and numerator of the measure, and therefore cannot have the opportunity contribute positively to the reported measure

Using the Problem List in eCW

- In eCW, the Problem List is front and center on the main patient screen
- Problems that have chronic diseases automatically trigger certain actions for the care team to perform, like labs, education, medication checks, etc.
- Actions may be prompted by Relevant Care Gaps, alerts in eCW, or discussion in the pre-clinic huddle.
- The point is, everybody relies on the Problem List for “truth” and there are validation reports that are useful in listing patients who are potentially false negatives or false positives.

Investigation in NextGen

- Investigate how chronic disease patients are identified in NextGen
- Are Transformers identifying patients with chronic diseases in the same manner as the clinic teams?
- Which tables in Relevant contain the data for which screens or sections in NextGen

Three Sources of Diagnosis Codes

In the Production Database,

- patient_problems
- encounter_diags
- patient_diagnosis

- The Validation Report “RCHC Diagnosis Location Report NG” displays the number of unduplicated patients that can be identified from all three locations for each diagnosis-related QM Value Set

Transformers

- Many existing diagnosis-based Transformers use encounter diagnosis as a basis for identifying patients with chronic disease
- This will tend to add false-positives to your denominator
- An encounter diagnosis cannot be “removed” if it is wrong or if the disease resolves
- Assumed that all historical encounter diagnoses are not readily and obviously displayed on the patient record. Is this really the best source?

Problem List Validation Reports

- Two reports follow this model
- Problem list table: patient_problems
- Encounter diagnosis tables: encounter_diags and patient_diagnosis

RCHC Problem List Validation Report NG

Diagnosis groups:

- Diabetes
- Hypertension
- IVD
- ASCVD
- Depression/Bipolar

RCHC Persistent Asthma Diagnosis Validation Report NG

- Persistent Asthma, which is the diagnosis that is pulled by the Asthma QMs
- Intermittent Asthma, not pulled by the asthma report
- Some patients have both Persistent and Intermittent Asthma on the Problem List
- Some patients are using a long-term acting asthma medication associated with Persistent Asthma but do not have Persistent Asthma on the Problem List

Diagnosis Definitions

- Chronic Diseases and other conditions used by the Quality Measures (for example, for exclusions) are defined by Value Sets
- There is a list of Value Sets in the Instruction Manual
- You can also see the individual codes that comprise the Value Sets with the report QM Value Set Codes

Lab-Related Validation Reports

Reports that list information about labs and attributes

General Comments

- One lab test can have many results
- The LOINC code is associated with the lab result name, not the lab test name
- lab_results_obr_p.test_desc = lab test name
- lab_results_obx.result_desc = lab result name

Observed Level of Consistency

- As mentioned previously, it seems like labs are identified in different ways between health centers, and also among Transformers within a single health center
- Labs are evaluated as “complete” in different ways between health centers, and also differently among Transformers within a single health center

Lab Completion Validation Report

- RCHC Incomplete Lab Validation Report NG
- Displays labs that may be incomplete and so may not be picked up by the Transformers
- Labs are identified by Value Set (LOINC) codes
- Assumes that a “complete” lab meets particular standardized criteria
- This report may be customized depending on the documented recommended procedure in your health center for the completion of labs.
- Transformers should also be standardized to follow this official and recommended procedure

Standards for a Completed Lab

1. A lab date is present, which is a Collected Date OR a Result Date
2. AND a lab result is present, which is any entry in the observed value field (the exception is for A1c and LDL labs, where the entry into this field must be a number)
3. AND no lab results or lab comments that indicate the lab was not actually performed (e.g., “cancelled” or “not adequate” etc. See code for all options)
4. AND the observed result status is equal to ‘F’
5. AND lab test status equal to “OrderCompleted” or “Final”
6. AND no delete flags on any of the associated lab tables

Cancer Exclusion Validation Report

For use with the three cancer screening QMs

Cancer Exclusions

- The RCHC Data Standards and Integrity Committee agreed on recommendations for the standard manner that patients with cancer exclusions should be documented in the health record.
- See Appendix A of the instructions for these recommendations
- Transformers should be designed to identify patients with cancer exclusions through diagnosis codes on the Problem List or key words in surgical history
- The validation report is looking for patients with partial cancer exclusion descriptions in Surgical History

Report Set-Up

- The code will evaluate the text in Surgical History
- The report will exclude patients if the Transformers/Importers already recognize the patient as having an exclusion
- Normally, it would be expected that the code for the exclusion would be in a Transformer, but sometimes it is in an Importer
- As mentioned previously, output tables from Importers are not available in the Staging Database

Report Set-Up

- Therefore, code from certain Importers must be copied to Temporary Tables in the validation report code.
- The code is broken down into Parts (identified in the headers before each Temporary Table).
- In the default version of the report, Parts 2 and 3 contain copies of Importer code, but check your instance to see if the exclusion code for all three cancer types exists in Importers or Transformers

Incomplete or Non-Standard Demographic Data Report

For UDS and QIP demographics

UDS and QIP Demographic Data

Data Elements:

- Age
- Zip code
- Language
- Race
- Ethnicity
- Gender
- Sexual orientation
- Poverty level
- Insurance

Purposes of the Report

1. Record-level data with suggestions for correcting missing or potentially erroneous fields
2. Patterns in the output can help you to make tweaks to the Transformers to better capture and process the data

Main Idea

- The report displays the raw data compared to the “transformed” data that comes from the Transformer
- Question: is the Transformer working as expected for all records? Use the report to display those records that should be checked.
- The problem might lie with the record (e.g., a mistake that can be fixed) but you might also find some data that shows that Transformer needs to be tweaked.

Customizing

- You may need to tweak the code a little if you find that it is displaying records that, when you check them in NextGen, are actually okay.
- Feel free to make a second version of the report that only shows records with particular characteristics that must be changed. You can edit the report to display the columns you want and the records you want.
- ID numbers are displayed so code developers can investigate unusual data on the Transformer level. These columns do not really need to appear on a version of the report for, say, front desk people.

Understanding the Code

- This report will probably be customized by each health center
- For example, different NextGen health centers have different ways of evaluating gender identity. There is a temporary table (temp_gender) that uses code from a Relevant Transformer to define gender identity.
- There are also Temporary Tables (last_pov_temp) and (relevant_insurance_enrollments_temp) that have code from Importers because the report is built in the production database

Understanding the Code

- The default version of the report looks for 9 kinds of problems. See the Report Description section of the instructions for more detail
- The Temporary Table “action_items_temp” evaluates the data and determines if there is a problem. Records with a perceived problem are displayed by the report along with a message in the column “action_needed_agg” saying what needs to be checked
- If you want to customize what kind of records are displayed by the report (removing classes of records, adding classes of records, tweaking how a ‘problem’ is evaluated, changing the problem description, etc.) this is the place in the code to do it

Questions?