Informatics Common Metrics Pilot Final Report

Submitted 3/30/18

Submitted by: CLIC Common Metrics Implementation Team
**Table of Contents**

I. Executive Summary ............................................................................................................... 2  
II. Development of the Metric Prior to Pilot Phase ................................................................. 3  
III. Metric Description ............................................................................................................. 4  
IV. Pilot Implementation Overview & Timeline ...................................................................... 4  
V. Hub Assessment & Pilot Sites Selection ............................................................................ 4  
VI. Pilot Implementation Activities ......................................................................................... 6  
VII. Development of Training Materials, Tools, and Resources ............................................. 7  
VIII. Scorecard Use and Completion ....................................................................................... 9  
IX. Post Pilot Activities ........................................................................................................... 10  
X. Results: Post Pilot Assessment and Interview Calls ............................................................ 11  
XI. Preparation to implement Consortium-wide .................................................................... 14  
XII. Lessons Learned for future pilots .................................................................................... 16  
XIII. Proposed next steps & timeline .................................................................................... 16  
XIV. Appendices ..................................................................................................................... 18
I. Executive Summary

The goal of this Informatics Common Metric is to enhance collaboration and opportunities for discovery through interoperable data. Informatics provides critical tools, methods and resources to accelerate translational research. The forthcoming Informatics Common Metric aims to improve discovery opportunities within and among CTSA Program hubs through metrics that encourage data repository quality and harmonization across hubs.

A team of informatics and evaluation experts developed the metric definition and Operational Guidelines. The Common Metrics Implementation team at the Center for Leading Innovation and Collaboration (CLIC) began the pilot activities in August 2017 through March 2018. Sixteen hubs participated in the pilot, representing diversity such as: hub size, single vs. multi-institution, and data model (i2b2/ACT, i2b2/TriNetX, OMOP, PCORnet).

Based on post-pilot assessment including hub interviews, a number of key findings were identified to inform the recommendations for Consortium-wide implementation. These included: (1) Providing scripts to the hubs was a critical addition to the metric that was viewed positively by the hubs, as it reduced duplicative efforts. (2) The amount of script editing required varied by data model and hub. (3) Informatics Metric was seen as valuable to all although new to some hubs. (4) Not all hubs had data for all required domains, (e.g. observations) although all the domains were seen as valuable. (5) Age of the data was a significant concern to the hubs – the value of including data older than 10 years was questioned. (6) For some hubs influence on improving their data warehouse was limited as not all hubs view themselves as able to influence the management and/or collection of the clinical data. (7) RBA process was viewed as valuable to strategic management although there were challenges with the Clear Impact Scorecard. (8) Hubs indicated it would be helpful to have guidance on metric team formation and engagement.

Prior to Consortium scale-up, CLIC suggests the following recommendations: (1) Modify the Operational Guidelines to refine data collection and clarify scope of data. (2) Expand CLIC technical support to the hubs including facilitating access to script subject matter expert. (3) Develop Implementation Guide specific to the metric. (4) Market the Informatics metric to demonstrate value and future potential. (5) Revise the structure of the Scorecard to reduce redundancies.

Communication of the pilot findings and network scale-up plans will be presented to key Consortium stakeholder groups over the next couple of months. Scale-up implementation to the Consortium will start upon approval from the NCATS leadership.
II. Development of the Metric Prior to Pilot Phase

The metric development team worked with the Informatics Domain Task Force (iDTF) to develop this metric. (See Appendix 1: List of ICM Development Team Members)

This team generated the metric Operational Guidelines (OG). (See Appendix 2: Informatics Common Metric (ICM) OG). A significant amount of pre-work was undertaken by this group to streamline the work required by each participating hub. To that end, in preparation for the pilot, common data models were identified and scripts were developed and pretested for each. These scripts were for pilot hubs to query their enterprise data warehouse as stipulated in the Operational Guidelines. Below is a brief description of the data model selection and script development.

Data Model Selection: The iDTF initially solicited information about which data models were used with high frequency within the Clinical and Translational Science Awards (CTSA) Program. This resulted in the identification of the OMOP, PCORnet and i2b2/ACT data models. In addition, many hubs had developed partnerships with the company TriNetX to access data from their enterprise data warehouse. Through TriNetX, metric data may be provided to the hubs using the i2b2 and OMOP data models.

Script Development Decision Process

- Scripts were designed for use by hubs to query their enterprise data warehouse to generate the metric data.
- Experts in each data model collaborated to develop a GitHub location to develop the script based on the operational guidelines; scripts were pre-tested and modified before approved for use during the pilot
- This was an iterative process that further engaged iDTF members

Development Team agreed that:

- If hubs wanted to utilize another i2b2 data model (not ACT or through TriNetX) they would be responsible for acquiring the data from their enterprise data warehouse:
  - The hubs would need to develop their own script to query their enterprise data warehouse
  - Hubs are encouraged to share the script on the NCATS CTSA Metrics GitHub site (https://github.com/ncats/CTSA-Metrics); however, there is no assurance that other hubs would be able to validate the script due to differences in the hub’s enterprise data warehouse and/or data model

- If hubs wanted to use any other data model:
  - Hubs wishing to use other data models should suggest additions (along with a list of other adopting hubs) to CLIC, who will coordinate with the iDTF for consideration.
After significant iterative discussions involving the iDTF and the Steering Committee, an agreed upon approach to assess interoperability was determined inclusive of 9 domains to include in the initial rollout of this metric. The domains are presented in the OG and in Section VI, but the focus was to assess the use of standard nomenclature (such as LOINC for laboratory results) within the clinical data warehouse.

III. Metric Description

Informatics is the study and practice of creating, storing, finding, manipulating and sharing information (See Appendix 3: Informatics Information Sheet). To accelerate translation, researchers need to be provided with access to a broad range of data (electronic health records, omics, imaging, genetics, behavioral, etc.). These data can come from different sources such as clinical databases, research datasets, sensors, mobile technology, patient generated data, and publicly available data sets. The sharing and pooling of data within and across CTSA Program hubs requires that data be represented in a format that can be queried and adheres to commonly accepted standards. A longer-term goal of the CTSA Program is to harmonize data and data standards so that a query written by any site can be run unaltered against all CTSA Program data repositories. Before we can complete such a data harmonization, we first need an understanding of what types of data are being collected, managed, and stored in each hub’s clinical data repository and how much of this data is in a standard format.

This Common Metric will provide a baseline scan of the level of coverage of the types of data that each hub should have in their clinical data repository. The purpose of this common metric is to identify clinical data gaps and opportunities for improvement. This common metric will improve local as well as network capacity to efficiently use data to conduct research. Improving the CTSA Program’s clinical research data ecosystem can enhance the effectiveness of collaborative initiatives within and outside of the CTSA Program to provide tools to identify patient cohorts (i2b2/ACT, SHRINE, All of Us, PCORI).

IV. Pilot Implementation Overview & Timeline

In August 2017, the CLIC Common Metric Implementation Team assumed responsibility for the metric pilot activities. CLIC and the metric development team worked closely on the planning of the rollout pilot activities, the types of communications to occur to the various Consortium stakeholders, and the timeframes of the pilot activities. See Appendix 4: ICM Pilot Timeline for rollout details and activities.

V. Hub Assessment & Pilot Sites Selection

a. Pre-pilot Implementation Landscape Hub Assessment

Prior to starting the pilot, the Informatics Development Team, in collaboration with the CLIC Common Metrics team, solicited information to assess:
The diversity of the types of data models that hubs use for their enterprise data warehouse,

- Of the data models approved by the Informatics DTF (iDTF) for use to report on the Informatics Metric, which model(s) the hubs are using currently,
- Of the approved data model(s), which model the hub will be using to report on this metric, and
- Anticipated effort to acquire the data for this metric.

Sixty-two responses were submitted.

For complete results, see Appendix 5: Landscape Assessment Results

b. Interviews with Potential Pilot Hubs

The CLIC CM team conducted 21 interviews out of the 29 hubs that indicated an interest in participating in the pilot, to assess readiness for participation in the pilot, and to assist with the final pilot site selection. Calls lasted about 30 minutes per hub, and the following questions were asked:

- When was your CTSA first funded? Where are they in the current cycle? Would you describe your hub as a single or multi-institution?
- Confirm the data model intended to use?
- How long have you been using/working with this model?
- What is your interest/motivation to serve as a pilot?
- Describe how developed your informatics team is & functions? Who will be involved in this pilot? How familiar are you with the Scorecard system and RBA framework?
- How has your CTSA program worked to address the data interoperability and data quality?
- How active has your hub been on the iDTF?

c. Final Selection for Pilot

Sixteen sites were selected to participate in the pilot out of the 21 interviews conducted. The selection criteria were: data warehouse diversity; hub size; whether single or multi-institution; and the experience level of informatics involvement. (See also Appendix 6: ICM Pilot Sites)
VI. Pilot Implementation Activities

Scorecard Build and Data Collection: The ICM Pilot Scorecard was built so that pilot hubs could enter their hub’s data in a standard format. As noted above the list of domains was generated by the Informatics Domain Task Force (iDTF). The actual metric results in eight scores of data domain-based values calculated as a percentage. (Count of unique patients with the standard value [numerator] / count of unique patients in the clinical data repository [denominator]). An additional domain was assessed as a Yes/No (present/absent) rather than as a percentage.

a. Domains (1-8) and other data (9-16) included in each hub’s ICM Pilot Scorecard:

1. Count of unique patients with and age/DOB value
2. Count of unique patients with administrative sex value
3. Count of unique patients with LOINC ID value
4. Count of unique patients with RxNorm ID value
5. Count of unique patients with an ICD 9/10 or SNOMED value
6. Count of unique patients with an ICD 9/10 or CPT Procedure Code
7. Count of unique patients with a free text data
8. Presence of observations or absence of observations (Y=1, N=0)
9. Count of unique patients within the clinical data repository
10. Percentage of patients with administrative sex value
11. Percentage of patients with an age or date of birth value
12. Percentage of patients with free text data
13. Percentage of patients with ICD 9/10 or CPT Procedure value
14. Percentage of patients with ICD 9/10 or SNOMED value
15. Percentage of patients with LOINC ID value
16. Percentage of patients with RxNorm value

b. Turn the Curve Plan (TTC)
Based on findings and consistent with the approach used for the Common Metrics Initiative, Pilot hubs were expected to follow Results Based Accountability approaches for their strategic management to create “Turn the Curve” (TTC) plans. These were to be entered into the Scorecard as well. The TTC plan consists of the following sections:

- The Story Behind the Curve – analysis of the factors that contributed to the data reported (background, limiting factors, facilitators, etc.)
- Partners - identification of existing and potential new partners who have a role in improving the metric
- What Works – what has been done, or could be done to address the contributing factors, and improve the metric
- Strategies – brainstorm of ideas, including potentially outrageous “no limits” ideas in addition to established best practices
- Actions – specific, implementable strategies and actions to move forward, including steps to completion, timeline, and person(s) responsible

Two additional pieces of information were requested

- Operational Guidelines - potential ways to refine the Informatics Metric Operational Guidelines
- Inclusion / Exclusion Decisions – description of both included and excluded data at the hub (i.e. patients; years of data reported; specific partners/collaborators; etc.)

- Hubs were asked to develop one TTC Plan for all eight domains.

c. **Worksheet Upload:** Hubs filled out and upload an excel worksheet with actual numerators and denominators, not just percentages

VII. **Development of Training Materials, Tools, and Resources**

To support the rollout of the ICM pilot, CLIC developed a number of tools and resources, including a cloud-based collaborative drive, webinar slides and recordings, step-by-step instructions for using the Clear Impact Scorecard, and Question and Answer pages to document material covered during the technical assistance calls.

a. **Cloud-based Collaborative Drive:** Access to this is limited to the sixteen hubs participating in the pilot, and the ICM Development Team. It houses materials, tools and resources related to the ongoing Informatics Common Metric (ICM) pilot, including:
- Slide sets
- Recorded webinars
• Q & A documents from technical assistance calls
• Resources specific to the ICM pilot
• ICM data worksheet

b. Webinar slides: CLIC developed webinar slides for five technical assistance webinars, and three post pilot follow up webinars.

• Slide decks are housed on cloud-based collaborative drive, along with recordings of the webinars, and Q&A sheets where appropriate.

c. Resources: To support the ICM pilot sites’ use of the pilot Scorecard, CLIC staff developed a resource with step-by-step instructions for entering both quantitative and qualitative data into the Scorecard.

• Includes screenshots of each step, and instructions
• Available to all pilot hubs on the collaborative drive. (see Appendix 7: Informatics Metric Pilot Scorecard Resource)

d. Technical Assistance (TA) Webinar Calls: CLIC developed and hosted five TA webinar calls to support the rollout of the ICM pilot (For additional detail, see Appendix 8: TA Calls and Support) Representatives from the Development Team were present on these calls.

• Kick-off Call:
  o Co-delivered by Redonna Chandler (representing NCATS)
  o Review of Common Metrics Initiative goals
  o Use of the RBA framework for strategic management and development of the Turn the Curve plans,
  o Overview of informatics and the Informatics Common Metric
  o Operational Guidelines review for the pilot
  o Detailed review of pilot process (data sources, data collection, and frequency of reporting),
  o Review of expectations and timelines for the participating hubs
  o Questions and discussion period

• RBA & Scorecard Refresher Call:
  o Focus on the use of the RBA framework
  o The five steps of a Turn the Curve Process
  o A step-by-step walk through of the new ICM pilot Scorecard.
  o Specific next steps with timelines for running the data model scripts and to prepare to discuss script output process on the next scheduled calls
  o Questions and discussion period

• Operational Guidelines & Script Output Calls (2):
  o In advance of the calls, CLIC sent emails to all hubs asking them run their data model scripts, and to be prepared to report out on the following questions, based on each hub’s script output:
Did you have any issues with using and/or interpreting the Operational Guidelines? Do you have any recommendations for modifications?

How easy/difficult was it to implement the script. What modifications were needed, if any?

Did you experience any challenges to managing or interpreting the script output? What resources were required?

- Hubs shared responses by data model group, so that script-specific or data model-specific issues could be discussed.

- **Scorecard & Turn the Curve Plans Call:**
  - Focus on what hubs needed to meet the deadlines for data entry and development of the Turn the Curve plans.
  - Prior to this call, CLIC polled the hubs to identify priority topics for discussion, then facilitated the call in an open discussion format.

**VIII. Scorecard Use and Completion**

- **Informatics Metric Pilot Scorecard TA:** As discussed in Sections VI and VII (above) CLIC built a new Scorecard specifically for the ICM pilot, and the ICM pilot Scorecard was discussed on all of the Technical Assistance webinar calls. In addition, the CLIC staff also responded to e-mail and phone inquiries from pilot hubs, as the hubs were running their scripts and completing their Scorecard entry.

  - Most ICM pilot hubs were able to enter their information into the Scorecard with few issues.
  - At some hubs, individuals with access to the Scorecard had not participated on the technical assistance calls.
  - Several hubs did not use the “ICM Pilot Scorecard” at all, but rather created new sections in their “Common Metrics” Scorecard. This caused some delay and confusion.
  - CLIC staff initiated/responded to over 65 e-mails, and 5 technical assistance phone calls to facilitate, monitor, track and encourage completion of the ICM pilot Scorecard entry.
  - 3 hubs have incomplete sections of the Turn the Curve (TTC) plans, (two hubs are missing “Actions” and one is missing a response in “Operational Guidelines”) despite multiple reminders, and follow up calls.
  - There is some redundancy in the sections of the TTC plans, as hubs included “Actions” in the “Strategies” section, or the “Story Behind the Curve.”

- **Scorecard Results:**

  - Average percent of Performance Measures reported by data model: Table 1 (below) shows the completeness of the data by performance measure and by data model.
### Table 1: Average Percent of Records Performance Measures Reported by Data Model

<table>
<thead>
<tr>
<th>Informatics Performance Measures</th>
<th>i2b2ACT</th>
<th>i2b2TriNetX</th>
<th>OMOP</th>
<th>PCORnet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of patients with administrative sex value</td>
<td>99.2</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Percentage of patients with an age or date of birth value</td>
<td>99.8</td>
<td>100.0</td>
<td>100.0</td>
<td>99.7</td>
</tr>
<tr>
<td>Percentage of patients with free text data</td>
<td>0.0</td>
<td>34.9</td>
<td>3.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Percentage of patients with ICD 9/10 or CPT Procedure value</td>
<td>52.6</td>
<td>71.5</td>
<td>65.1</td>
<td>72.3</td>
</tr>
<tr>
<td>Percentage of patients with ICD 9/10 or SNOMED value</td>
<td>75.8</td>
<td>81.7</td>
<td>68.1</td>
<td>82.1</td>
</tr>
<tr>
<td>Percentage of patients with LOINC ID value</td>
<td>39.7</td>
<td>69.8</td>
<td>44.6</td>
<td>39.8</td>
</tr>
<tr>
<td>Percentage of patients with RxNorm value</td>
<td>53.4</td>
<td>24.6</td>
<td>41.4</td>
<td>46.8</td>
</tr>
</tbody>
</table>

- **Count (x 100,000) of All Patients in Data Repositories by Data Model:** Figure 1 (below) shows the number of all patients in pilot hubs’ data repositories, by data model. The number of patients in hub data repositories ranges from 589,903 – 5,663,620.

![Figure 1: Count (x 100,000) of All Patients in Data Repositories by Data Model](image)

See also Appendix 9: Number of Unique Patient Records in Repository by Hub

### IX. Post Pilot Activities

a. **Assessment Process:** In order to learn more about the ICM pilot processes, hub specific experiences and to inform recommendations for next steps and Consortium-wide scale-up of the Informatics Common Metric, CLIC solicited feedback from the hubs, by developing and distributing the ICM Post Pilot Assessment, and then scheduling group follow up calls with hubs. (See Appendix 10: ICM Post Pilot Assessment)

- Development & Administration of Assessment:
  - Post Pilot Assessment developed by CLIC, with input from the ICM Development Team
  - Built in RedCap™.
  - Asked hubs to respond to questions and to give feedback in the following eight areas: Data model choice; Script processes;
Operational Guidelines; Hub implementation team; RBA process and training process; Scorecard usability; Usefulness of the Informatics Common Metric at the hub; and Pilot training and support.

- Only one assessment per hub was submitted, although the assessment was designed for multiple responders
- CLIC sent hubs a .pdf of the assessment questions, in order for hubs to determine hub staff member who could best respond to the sections
- All 16 ICM pilot hubs completed the assessment in a timely fashion

- **Interviews: Post Pilot Data Model Group Calls:**
  - Upon receiving the assessment results, CLIC staff scheduled three follow up calls with the hubs, organized by data model used
  - i2b2 group included hubs that used i2b2-ACT as well as hubs that used i2b2 TriNetX
  - Calls were structured around issues identified during the survey and the calls, to garner feedback on possible recommendations for changes to Operational Guidelines and for scale-up procedures

For additional details, see Appendix 11: Summary Post Pilot Prompts & Responses

X. **Results: Post Pilot Assessment and Interview Calls**

a. **Post Pilot Assessment of Pilot Framework – Feasibility, Quality, and Usability:**

- **Feasibility of Scripts:** Scripts were provided to hubs for the OMOP, PCORnet and i2b2/ACT data models, which enabled standardized, automated query against the data repository (or repositories) at each of the pilot hubs. For hubs that choose to use the approved data models that are also utilized by TriNetX, TriNetX provided the metric data to hubs directly. Scripts were developed, tested, and approved collaboratively by the Informatics Common Metric Development Team and the iDTF. CLIC elicited feedback from the ICM pilot hubs in the post pilot assessment, and found:
  - Of the six i2b2-ACT and TriNetX hubs
    - 100% i2b2-ACT hubs needed to do some editing of the scripts
    - 100% i2b2-ACT hubs ran the scripts 2-4 times to get final data
    - TriNetX does not have access to free text clinical notes
      - One TriNetX hub separately queried their data warehouse to measure completeness
    - 100% expressed confidence in the results
  - Of the four OMOP hubs
    - 75% needed to do some editing of the script
    - 75% ran the scripts 2-4 times to get final data
    - 100% of hubs were “very confident” in their results
  - Of the six PCORnet hubs
- 83% needed to do some editing of the script
- 83% ran the scripts 2-4 times to get final data
- 33% do not include notes in their data repository, so no data could be retrieved for that domain
- 100% of hubs were “very confident” in their results
  - Of all 16 pilot hubs:
    - 50% “Not sure” if they will use the same data model they used for the pilot, when this metric is implemented Consortium-wide

- Quality of Data Completeness: Feedback from the ICM pilot hubs in the post pilot assessment, included:
  - Did hubs have complete data for all 8 domains?
    - 38% Yes
    - 38% Somewhat
    - 25% No
  - Hub comments on data completeness:
    - We did not have data to support certain domains such as notes
    - We had results for each domain; but our lab doesn’t use LOINC code for testing results so result was zero
    - Scripts had 'null' for some domains
    - TriNetX does not have access to free text clinical notes
    - No notes are entered in our PCORnet data mart at this time.
    - We don’t have notes data in our data repository
    - In support of CDRN initiatives, our instance only contains subsets of the full data available (for some domains) in the internal Research Data Warehouse.

- Usability of the Scorecard: Feedback from respondent hubs:
  - 56% - Scorecard entry was time-consuming
  - 38% - Scorecard was difficult to navigate
  - 63% - Sections of the Scorecard were redundant
  - 70% - Using the Scorecard added value to the pilot process

- Usability of the RBA process: Among respondent hubs:
  - RBA training was an issue for many
    - 12 hubs indicated ≤50% of ICM team trained in RBA
    - 4 hubs indicated ≥51% of ICM team trained in RBA
  - 88% - RBA process was useful overall
  - 87% - Agreed/Strongly Agreed RBA process helped organize planning
  - 69% - Agreed/Strongly Agreed RBA process helped identify aspects/components of the issue/process not previously considered
  - 94% - Agreed/Strongly Agreed RBA process helped organize reporting

- Usability of the cloud-based collaborative space:
o 12 of the 16 hubs indicated that they had accessed resources from the collaborative drive
o All 12 hubs that accessed the drive indicated that accessing the drive was easy or somewhat easy

- Composition & engagement of team: Hubs were asked to describe the process of building and engaging the hub’s ICM implementation team.
  - Most identified team members based on institutional roles
    - Informatics
    - Evaluation
    - Data model ontology
  - Other specific expertise identified as potentially needed in future
    - Script expertise
    - RBA expertise
    - Scorecard expertise/access

b. **Interview Feedback: Key Points from Data Model Group Calls:** Comments from the hubs are shown in italics. (for a full list of questions asked during interview calls, see Appendix 12: Informatics Metric Post Pilot Group Call Questions)

- *Keep “Notes” domain – likely will be useful in the future*
- Hubs favor running an “all-in” script (entire data repository) followed by a second run covering shorter time period (e.g. last 10 years)
  - To help demonstrate improvements
  - To generate more useable data
  - Useful for benchmarking
  - PCORnet hubs only load past 10-12 years into their CDM
  - Data requests rarely/never go back more than 5 years

- Consortium-wide implementation should have the same restrictions as the pilot (one model/institution) – not combining data across different models/CDW or institutions as is allowed in the Operational Guidance
  - A change post pilot could decrease reproducibility and increase work
  - Hubs would like to be able to do this (run more than one model), but not as a requirement to report
  - Perhaps ask hubs to limit to databases/warehouses used to support research
  - Hubs with multiple health care systems will find it valuable to combine models in future
  - Could be optional

- Do not add a “stretch metric” (such as genomic data) at this time (adding a stretch metric was a suggestion from the post pilot assessment)
  - Want the first run of the ICM to be successful across majority of hubs
  - For hubs with advanced informatics capacity they could add/test a stretch metric
  - Could be an IDTF discussion (data quality; data expansion)

- Hub suggestions for Consortium-wide implementation improvement
An ICM Implementation Guide would be very helpful

- Include information on balancing informatics, administration and evaluation
- Guidance - Metric Team Development (Informatics/IT/RBA/Scorecard access)
- Possibly apply a maturity model for hubs with advanced informatics capacity
- Guidance - What is Informatics?
- Guidance - What is the TTC plan development process?

- Continue to provide scripts to hubs - saves hubs a lot of work
- Give us feedback on how we did at developing TTC plans

XI. Preparation to implement Consortium-wide

a. Key Findings:
   - The scripts are generally effective tools
   - Age of data is a significant concern for hubs
     - Limits ability to show improvement over time if denominator is entire data repository
     - Relevance – data requests rarely go back > 5 years
   - Aggregate reporting of data is of interest to a majority of hubs
     - Visualizations of hub data in relation to other de-identified hubs
   - The RBA process was generally seen as useful to the hubs
     - Increase access to training/resources
   - Not all hubs have influence on the Clinical Data Warehouse data (elements included, quality control) so their ability to make improvements in the metric may be limited as it may require a longer term approach to garner institutional buy-in.
   - Several hubs had issues with Clear Impact Scorecard
     - Redundancy – several places to enter same information
     - Time consuming
     - Difficult to navigate
   - Informatics Metric - new to some hubs, but seen as valuable by all
   - Not all hubs had data for all domains
     - For some domains the value was zero
   - Great variability of experience in running scripts – some hubs needed to engage subject matter expert with script expertise

b. Recommendations:
   - Modifications to Operational Guidelines:
     - Clarify the definition of “observation”
     - Describe how to interpret the script output
       - Usefulness/meaningfulness of the output (%)s put in appropriate institutional context
     - Add to inclusion/exclusion criteria
- Hubs to run script on full data repository
- Hubs to run script again on a more recent time frame (i.e. past 5-10 years)
  - Clarify expectations for achieving data completeness where script limitations exist
    - One TriNetX hub separately queried their data warehouse to measure completeness because TriNetX did not have access to free text clinical notes
- Changes to processes: What to do the same/differently from the pilot
  - Clarify who from the hubs should be included in the training and TA calls
  - Keep “Notes” field even though not all hubs have data for it
    - Will likely be useful in the future
  - Be specific about what years are included in the data
  - Develop feedback/visualization process for how to feed data back to the hubs (that aligns with CMI visualization approach being deployed for other metrics)
  - Foster collaboration between hubs
    - Continue to support the cloud-based collaborative space for hubs
    - Design to facilitate hub interaction
    - Use to support collaborative discussions on TTC plans
    - Develop a wiki/forum
- Script Support:
  - Make script developers available for TA/consultation
  - Keep GitHub to support script discussions and share versions
- RBA Training:
  - Clarify expectations for team members’ training/familiarity with RBA
  - Extent of RBA training needed may depend on role at the hub
- Scorecard Improvement and Support:
  - Clarify how to enter data if the response is null (missing vs. 0)
  - Navigation - clarify where to put what
  - Eliminate redundant sections/fields
- Develop ICM Implementation Guide, including:
  - Definitions of key terms (i.e. observations, etc.)
  - Overview of Informatics
  - Overview of RBA Process
  - Guidance on TTC plan development
    - Examples/templates
  - Guidelines for formation and engagement of Metric Team
    - Informatics expertise
    - RBA/Scorecard expertise
- Script expertise
  
- Marketing the Informatics Metric moving forward
  - Demonstrate/reiterate value of Informatics metric and future potential
  - Build buy-in for the Informatics metric especially in hubs where control rests outside of hub’s direct influence
  - Foster collaboration between and among the hubs

  c. Other considerations for further discussion:
     
     - Implications of changing data models from one reporting period to the next on how data are interpreted over time
       - Several hubs indicated they would likely use a different data model in the future; some may shift to different data models over time;
     
     - Variability in hubs’ Informatics capacity
       - Determine options for adding stretch metrics for hubs with greater Informatics capacity (i.e. genomics data, others) or engaging them to develop the next domains to test

XII. Lessons Learned for future pilots

a. Pre-test Scripts or tools:
   
   - Pretest scripts or tools during metric development
   - Utilize internal expertise within Metric Development Team to write scripts or develop tools

b. Engage Subject Matter Experts:
   
   - Maintain engagement with Development Team throughout the pilot
     - Participation in TA calls and webinars
   - Engage partners with specialized expertise – will vary by metric

c. Provide Samples and Templates:
   
   - All in one place
   - Example TTC plan(s)
   - Data worksheets
   - Step-by-step Scorecard resources
   - Job aids (Informatics; RBA; etc.)

XIII. Proposed next steps & timeline

a. Stakeholder Communication: CLIC will engage in communications with all stakeholders to describe Pilot results, lessons learned, and recommendations for Consortium-wide implementation. CLIC will present pilot findings and recommendations to the following stakeholder groups, and will elicit feedback from each:
   
   - Metric Development Team
   - Common Metrics Executive Committee
   - NCATS Project Officers
• Common Metrics Steering Committee
• CTSA Evaluators Group

Upon receipt of final approval from NCATS DCI Leadership, CLIC will move forward to Consortium-wide implementation, by providing metric-related training and technical assistance to hubs as described below. (See also Appendix 13: New Metric Implementation Timeline)

b. **Training & Technical Assistance:** Once the Metric’s Pilot Report is complete, and any changes to the Operational Guidelines have been agreed to by the Development Team, CLIC will prepare training and technical assistance materials related to the ICM Operational Guidelines and the ICM Scorecard.

- **Scheduling Training Events:** To protect time going forward, to maximize efficiencies, and to streamline scheduling communications with hubs, CLIC will establish regular call dates and times for:
  - Training webinars/webinettes
  - Office Hours (via Zoom)

- **Webinars/Webinettes:** Each hub will be invited to attend a series of trainings, delivered via live webinars, shorter-format webinettes, and conference calls, some with pre-event assignments before sessions.
  - Webinars/webinettes will be:
    - Recorded
    - Made available on cloud-based collaborative space
  - Topics will include, but are not limited to:
    - RBA/TTC
    - Engaging stakeholders
    - Scorecard
    - Metric calculations, etc.

- **Office Hours:** CLIC staff will host regular and periodic Office Hours.
  - Monthly Office Hours topics: Data entry; Evaluation;
  - Periodic Office Hours topics: RBA; Scorecard; Metric calculations; others as needed/requested

- **Technical Assistance Consultations:** Technical Assistance (TA) consultations will be scheduled at the request of individual hubs (or for all as needed) and will include topics such as:
  - Interpreting and applying the Operational Guidelines
  - Identifying and overcoming barriers to data collection
  - Using of the RBA framework
  - Using the Scorecard software
  - Progress, challenges, and barriers in applying RBA
- Collecting Common Metric data according to the Operational Guidelines
- Developing and completing Turn the Curve Plans

- **Q & A and FAQ:** As questions are received by CLIC via e-mail or phone calls, they will be answered by the appropriate CLIC team member (based on topic), with consultation from NCATS and Development Team as needed. Questions and answers will be tracked using the Questions Inventory. If the questions and answers are broadly applicable and relevant to all hubs they will be added to the FAQ section of the CLIC website. If additional training or TA needs are identified through this process, then training/TA materials may be developed as well.

- **Peer-to-Peer Technical Consultation:** CLIC will work with NCATS and CTSA program hubs to identify appropriate candidates for provision and receipt of peer-to-peer TA related to new Metric implementation. CLIC will coordinate and facilitate peer-to-peer TA calls between hubs as necessary.

  - **Trial - Hub Self-Assessment:** CLIC proposes a trial of a Hub Self-Assessment process where hubs would use an Assessment Criteria rubric to perform a self-assessment of their progress toward applying the RBA Framework, and could then discuss their self-assessment scores with CLIC on individual calls as needed. These calls could also provide CLIC with potential topics for training webinars or webinettees, as needs and gaps may be identified.

**XIV. Appendices:**

- Appendix 1: List of ICM Development Team Members
- Appendix 2: Informatics Common Metric (ICM) OG
- Appendix 3: Informatics Information Sheet
- Appendix 4: ICM Pilot Timeline
- Appendix 5: Landscape Assessment Results
- Appendix 6: ICM Pilot Sites
- Appendix 7: Informatics Metric Pilot Scorecard Resource
- Appendix 8: TA Calls and Support
- Appendix 9: Number of Unique Patient Records in Repository by Hub
- Appendix 10: ICM Post Pilot Assessment
- Appendix 11: Summary Post Pilot Call Prompts & Responses
- Appendix 12: ICM Post Pilot Group Call Questions
- Appendix 13: New Metric Implementation Timeline
Appendix 1: List of ICM Development Team Members

<table>
<thead>
<tr>
<th>Member</th>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administrator</td>
<td>Thomas Fogg</td>
<td>Rochester</td>
</tr>
<tr>
<td>Administrator / Associate Director of Bioinformatics</td>
<td>Elizabeth Wood</td>
<td>Weill Cornell</td>
</tr>
<tr>
<td>Evaluator</td>
<td>Kristi Holmes</td>
<td>Northwestern</td>
</tr>
<tr>
<td>Evaluator</td>
<td>Patrick Barlow</td>
<td>University of Iowa</td>
</tr>
<tr>
<td>PI</td>
<td>Bob Clark</td>
<td>UTHSC at San Antonio</td>
</tr>
<tr>
<td>PI</td>
<td>Jiajie Zhang</td>
<td>UTHSC at Houston</td>
</tr>
<tr>
<td>Subject Matter Expert</td>
<td>Justin Starren</td>
<td>Northwestern</td>
</tr>
</tbody>
</table>
**DRAFT CTSA Common Metric (CM) Operational Guideline: Interoperable Clinical Data Availability and Completeness**

<table>
<thead>
<tr>
<th>Template Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Operationalized Metric Title</td>
<td>Interoperable clinical data availability and completeness</td>
</tr>
<tr>
<td>2. Rationale</td>
<td>To accelerate translation, researchers need to be provided with access to a broad range of data (electronic health records, omics, imaging, genetics, behavioral, etc.). These data can come from different sources such as clinical databases, research datasets, sensors, mobile technology, patient generated data, and publicly available data sets. The sharing and pooling of data within and across CTSA Program hubs requires that data be represented in a format that can be queried and adheres to commonly accepted standards. A longer-term goal of the CTSA Program is to harmonize data and data standards so that a query written by any site can be run unaltered against all CTSA Program data repositories. Before we can complete such a data harmonization, we first need an understanding of what types of data are being collected, managed, and stored in each hub’s clinical data repository and how much of this data is in a standard format. This Common Metric will provide a baseline scan of the level of coverage of the types of data that each hub should have in their clinical data repository. The purpose of this common metric is to identify clinical data gaps and opportunities for improvement. This common metric will improve local as well as network capacity to efficiently use data to conduct research. Improving the CTSA Program’s clinical research data ecosystem can enhance the effectiveness of collaborative initiatives within and outside of the CTSA Program to provide tools to identify patient cohorts (i2b2/ACT, SHRINE, All of Us, PCORI).</td>
</tr>
<tr>
<td>3. Operational Specification</td>
<td>Level of availability and completeness of the baseline types of data in a standard (CTSA-interoperable) format within a clinical data repository at the CTSA Program primary institution (hub).</td>
</tr>
</tbody>
</table>
## Appendix 2: Informatics Common Metric (ICM) OG

<table>
<thead>
<tr>
<th>4. Technical Description</th>
<th>The list of data types has been generated by the Informatics Domain Task Force (iDTF); Additional types of data will be considered in the future.</th>
</tr>
</thead>
</table>

### Data Scope: Data from centralized primary resources available to researchers at the primary hub institution. If patients can be unambiguously counted across multiple data repositories, the aggregate counts may be reported for the hub.

### Key Definitions:

- **Clinical Data Repository**: For the purposes of the CTSA Program and this metric, a clinical data repository is a standards-based repository of clinical and/or research data (at the individual subject, client, or patient level) available to investigators for research at a CTSA Program hub. This generic term includes clinical data repositories and enterprise data warehouses.

- **Standard Value**: The rules for values of data elements in a database; will ensure consistency.

- **Data Model**: A data model is a descriptive design of data and its relationships. The data model organizes data elements and standardizes how they relate to one another and to properties of the real world entities. Data models approved for use by the informatics Domain Task Force (iDTF) for this metric include OMOP, PCORnet, i2b2/ACT, or other i2b2 data model. Hubs wishing to use other data models should suggest additions (along with a list of other adopting hubs) to the coordinating center, who will coordinate with the iDTF for consideration.

- **OMOP Data Model**:
  - Documentation for the OMOP CDM is found on Github ([link](link)) and the OMOP CDM Wiki ([link](link)). The OMOP Common Data Model is an open-source, community standard for observational healthcare data that is managed by Observational Health Data Sciences and Informatics (OHDSI). OHDSI has been established as a multi-stakeholder, interdisciplinary collaborative to create open-source solutions that bring out the value of observational health data through large-scale analytics. The OHDSI collaborative includes all of the original OMOP research investigators, and will develop its tools using the OMOP Common Data Model. Learn more at
Appendix 2: Informatics Common Metric (ICM) OG

https://www.ohdsi.org/. OMOP CDM contains person, condition, drug, procedure and visit information, and provider and cost information. This will support health economics use cases and medical treatment outcome studies, including medical device safety, comparative effectiveness and healthcare quality. For more information see https://github.com/OHDSI/CommonDataModel/wiki/Background

- PCORnet Common Data Model:
  - The PCORnet CDM (see http://www.pcornet.org/pcornet-common-data-model/) is based on the FDA Sentinel Initiative Common Data Model and has been informed by other distributed initiatives such as the Health Care Systems Research Network, the Vaccine Safety Datalink, various AHRQ Distributed Research Network projects, and the ONC Standards & Interoperability Framework Query Health Initiative. The PCORnet CDM leverages standard terminologies and coding systems for healthcare (including ICD, SNOMED, CPT, HCPSC, and LOINC) to enable interoperability with and responsiveness to evolving data standards. See: https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model/sentinel-common-data-model

- i2b2/ACT Data Model:
  - Informatics for Integrating Biology and the Bedside (i2b2) was an NIH-funded National Center for Biomedical Computing (NCBC) based at Partners HealthCare System. The i2b2 NCBC developed a scalable informatics framework that is designed to bridge clinical research data and the vast data banks arising from basic science research in order to better understand the genetic bases of complex diseases. i2b2 was designed for cohort identification. The i2b2 framework employs a simple, yet powerful data model. It consists of facts and dimensions. A fact is the piece of information being queried, and the dimensions are groups of hierarchies and descriptors that describe the facts. The i2b2 database utilizes a star schema that consists of one fact table surrounded by numerous dimension tables. Facts in i2b2 are observations about a patient, including things like diagnoses, demographics, laboratory results, etc. See: https://www.i2b2.org/.
Appendix 2: Informatics Common Metric (ICM) OG

- Accrual to Clinical Trials (ACT) is a network of 21 CTSA Program hubs that use a common informatics platform for cohort discovery. See: https://www.act-network.org/node/29
- Use of Data Models with TriNetX:
  - TriNetX is a global health research network connecting healthcare organizations (including 25 CTSA Program hubs), biopharma and contract research organizations. The TriNetX platform enables cohort identification and hypothesis generation based on clinical data that can currently be sourced from a common data model (i2b2, OMOP, NAACCR, etc.), flat files, or via NLP of narrative documents. See: https://www.trinetx.com/

| Metric Description: Data to be collected |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Data Domain     | Standard Value  | Numerator       | Denominator     | Metric           |
| Patient         | N/A             | count of unique patients with an age/DOB value | Count of all patients in the Data Repository | % patients with age/DOB value |
| Patient         | Administrative Sex | count of unique patients with administrative sex value | Count of all patients in the Data Repository | % patients with administrative sex value |
| Labs            | LOINC ID        | count of unique patients with a LOINC ID value | Count of all patients in the Data Repository | % of patients with LOINC ID value |
| Medications / Drugs | RxNorm ID       | count of unique patients with a RxNorm ID value | Count of all patients in the Data Repository | % of patients with RxNorm ID value |
| Conditions / Diagnosis | ICD 9/10 or SNOMED | count of unique patients with an ICD 9/10 or SNOMED value | Count of all patients in the Data Repository | % of patients with ICD 9/10 or SNOMED value |
## Appendix 2: Informatics Common Metric (ICM) OG

<table>
<thead>
<tr>
<th>Procedures</th>
<th>ICD 9/10 CPT</th>
<th>Count of unique patients with an ICD 9/10 or CPT Procedure Code</th>
<th>Count of all patients in the Data Repository</th>
<th>% of patients with ICD 9/10 or CPT Procedure value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes / Narrative</td>
<td>N/A</td>
<td>count of unique patients with free text data</td>
<td>Count of all patients in the Data Repository</td>
<td>% of patients with free text data</td>
</tr>
<tr>
<td>Observations</td>
<td>N/A</td>
<td>Presence of Observations or Absence of Observations</td>
<td>N/A</td>
<td>Presence of Observations or Absence of Observations</td>
</tr>
</tbody>
</table>

### 5. Metric Type, Score(s), Numerator and Denominator Statements or Continuous Variable Statement, Inclusion/Exclusion Criteria

**In addition to the metrics, hubs will be asked to report:**
- Data model they used to report this metric: [OMOP, PCORnet, i2b2/ACT, or other i2b2 data model]
- How many unique patients are represented in the data model.

**This metric results in eight scores of data domain-based values (see Table above):**

**Metric Type: Percent**

**Metric:** Percent of unique patients with the standard value: [%] (see Table above):

To calculate this metric for each data type domain use the following:
- Count of unique patients with the standard value (numerator)
- Count of unique patients within the clinical data repository (denominator)
- % of unique patients with the standard value (% = [numerator]/[denominator])

**Inclusion/Exclusion Criteria:**
- To be explored further during the pilot phase
6. Data Sources & Methods of Data Collection

<table>
<thead>
<tr>
<th>Data Sources: CTSA Program hub clinical data repository. Aggregate counts may be reported for the hub and clinical affiliates, if individuals in the repository can be unambiguously identified and counted across individual data repositories.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method of Data Collection:</strong></td>
</tr>
<tr>
<td>- Queries/scripts will be provided to hubs for the OMOP, PCORnet and i2b2/ACT data models. These queries/scripts will enable standardized automated query against the data repository (or repositories) at each of the CTSA hubs. The scripts will be developed, tested, and approved collaboratively by the Informatics Common Metric Development Team with the iDTF. Data model scripts: <a href="https://github.com/ncats/CTSA-Metrics">https://github.com/ncats/CTSA-Metrics</a></td>
</tr>
<tr>
<td>- For hubs that choose to use the approved data models that are also utilized by TriNetX, TriNetX will provide the metric data to your hub directly for incorporation into your Scorecard. [Note: TriNetX supports the CTSA Program Common Metric for Informatics Solutions. As end users do not have direct access to the in-memory databases on their local or hosted appliances, TriNetX has created a report that can be requested by the individual CTSA Program hubs. A designated user from the site should send an email to <a href="mailto:CTSA@trinetx.com">CTSA@trinetx.com</a> requesting the CTSA Program Common Metric report. TriNetX will then create the report and send it to the requesting site admin.]</td>
</tr>
<tr>
<td>- For hubs that choose to use i2b2 with their own hub-specific data model, that hub will be responsible for generating the script for their own purposes of reporting the required metric data. These hubs are encouraged to share the script on the GitHub site, however, there is no assurance that other hubs would be able to validate the script.</td>
</tr>
<tr>
<td>- Data models approved for use by the informatics Domain Task Force (iDTF) for this metric include OMOP, PCORnet, i2b2/ACT, or other i2b2 data model. Hubs wishing to use other data models should suggest additions (along with a list of other adopting hubs) to the coordinating center, who will coordinate with the iDTF for consideration.</td>
</tr>
</tbody>
</table>
### Appendix 2: Informatics Common Metric (ICM) OG

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7. Frequency of Data Collection and Reporting</strong></td>
<td>Data collected 1x a calendar year. Reported 1x a year.</td>
</tr>
<tr>
<td><strong>8. Unit of Analysis</strong></td>
<td>Data will be collected within each hub from the proposed data type (domain) level and reported at the hub level.</td>
</tr>
<tr>
<td><strong>9. Notes/Comments</strong></td>
<td><strong>Note:</strong> Although 8 metrics will be reported, hubs should develop only 1 turn-the-curve plan. <strong>How can we use this to strategically manage the hub?</strong> <strong>Strategic Management:</strong> This metric demonstrates the extent to which data are complete and use standard approaches. The metric will allow CTSA Program hubs to take the first step necessary to enable a hub to view the breadth and depth of data and through local strategic management activities and turn the curve exercises, hubs will be able to identify factors behind clinical data gaps (e.g., technological, financial, unavailable?) and opportunities for strategic investment of time, funds, and/or focus. <strong>Aspirational Considerations:</strong> This metric will evolve over time to enhance the completeness of the data repositories across the CTSA Program consortium and incorporate additional types of data within repositories as the CTSA Program, in collaboration with the iDTF, finds useful and appropriate. <strong>Glossary:</strong> <a href="https://goo.gl/fJBTPt">https://goo.gl/fJBTPt</a></td>
</tr>
</tbody>
</table>
CTSA Program Informatics Common Metric: Enhancing collaboration and opportunities for discovery through interoperable data

How does Informatics help us realize the vision of the CTSA Program?
The CTSA Program is growing to become a collaborative and interoperable national research network that can leverage resources across multiple information systems and apply unique expertise within our institutions to connect research to health care, resulting in better health through research.

Why an Informatics Common Metric?
Informatics provides critical tools, methods and resources to accelerate translational research. The forthcoming Informatics Common Metric aims to improve discovery opportunities within and among CTSA Program hubs through metrics that encourage data repository quality and harmonization across hubs.

What data will we be required to report?
Hubs will be asked to provide data about their local repository such as the total number of unique patients and prevalence of standardized domain-specific data to describe the quantity and comparability of data in the local repository.

Specific Goals of the Informatics Common Metric
- Facilitate the interoperability of research data models through standards-based data repositories.
- Facilitate sharing of the repository’s assets for discovery.
- Encourage use of standards-based data models, not bound to a specific technology, but harmonized with other organizations, agencies, and initiatives (i.e., PCORI, ONC, FDA, and the Trial Innovation Network).
- The long-term goal is machine-readable, interoperable data that adhere to the FAIR data principles.

How does Informatics support data interoperability in the CTSA Program?

Data Standards: compatible research systems and use of standard terminologies to enable data harmonization
Data Integration: integrating different types of data from different sources for discovery and improved health
Data Access & Data Sharing: ability to query across sources and organizations and respond to diverse queries; enable data access, integration, and processing
Data Quality: ensure data are fit for purpose, provide benchmarking for new tools and algorithms
Data Security: user friendly infrastructure to assist investigators in ensuring the security of their data

New opportunities for strategic management for your hub and for the CTSA Program Consortium
A consensus baseline value will reflect a minimal set of clinical research data for each CTSA Program hub, enabling comparisons and identifications of synergies and gaps across the CTSA Program. This metric will provide continuous improvement for the CTSA Program and hubs by:

1. Enhancing interoperability by broadening the range of standardized data types in a hub’s data repository
   - Data types include demographics, diagnoses, labs, medications, procedures, etc.
2. Offering each hub opportunities for strategic management of their data repository
   - Enabling comparisons with other institutions thereby facilitating prioritization of repository expansion

Tentative Timeline
- Summer 2017: Development and testing of the metric in collaboration with the Informatics Domain Task Force
- Late Summer 2017: Pilot the metric
- Late Fall 2017: Metric finalized and introduced to CTSA Program hubs
- Ongoing: communication and support for hub Common Metric Implementation Teams

What is ‘FAIR’ data?
- Findable: data are assigned a globally unique and eternally persistent identifier.
- Accessible: data are retrievable by their identifier using a standardized communications protocol.
- Interoperable: data use vocabularies that follow FAIR principles.
- Re-usable: data have a plurality of accurate and relevant attributes.
Frequently Asked Questions:

What is Informatics?
The study and practice of creating, storing, finding, manipulating and sharing information. The American Medical Informatics Association (AMIA) indicates that “Informatics is the intersection between the work of stakeholders across the health and healthcare delivery system who seek to improve outcomes, lower costs, increase safety and promote the use of high-quality services.”

What is a Data Repository?
A data repository (for the purposes of the CTSA Program) is a structured collection of clinical and/or research data available to investigators for research at a CTSA Program’s primary institution.

What is the rationale for the informatics Common Metric to focus on data repositories?
In order to accelerate translation, researchers need access to a broad range of previously collected data (electronic health records, omics, imaging, genetics, behavioral, etc.). These data can come from different sources such as clinical databases, research datasets, sensors, mobile technology, social media, patient generated data and publicly available data sets. The sharing and pooling of data across CTSA Program hubs requires that data be represented in a format that may be queried and adheres to commonly accepted standards.

What is the current state and existing challenges with data repositories in the CTSA Program hubs?
- Lack of harmonization with costly support of multiple research models (PCORI, Sentinel, OMOP)
- Lack of semantic and syntactical consistency across the CTSA Program Consortium
- Deep phenotyping and other aspirational research activities are limited by data models, limited data elements, and a lack of quality control
- Lack of ability to engage and support parallel initiatives such as Accrual to Clinical Trials (ACT) Network, that are focused on cohort discovery

What is FAIR data?
FAIR is a set of guiding principles to make data Findable, Accessible, Interoperable, and Re-usable to optimize data for use by humans and machines. Learn more at the FORCE11 FAIR Data Principles website and the FAIR Principles Working Detailed Document.
- Findable: data are assigned a globally unique and eternally persistent identifier.
- Accessible: data are retrievable by their identifier using a standardized communications protocol.
- Interoperable: data use vocabularies that follow FAIR principles.
- Re-usable: data have a plurality of accurate and relevant attributes.

Where can I learn more about the Common Metrics and how data collection, reporting, and strategic management have been accomplished from previous metrics?
Please visit the Common Metrics Initiative website:
- Common Metrics FAQs
- Implementation and Training Materials
- Clear Impact Scorecard Information
- Common Metrics Collaborative Learning Sessions

I have questions and suggestions – who can I contact?
Questions and suggestions can be emailed to: Erica Rosemond, PhD at Erica.Rosemond@nih.gov.
**Appendix 4: Informatics Common Metric Pilot Timeline**

<table>
<thead>
<tr>
<th>Informatics Metric Pilot Testing</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pilot Test Preparation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finalize Informatics Metric Operational Guideline for pilot; post draft OG on Tufts website</td>
<td>Aug</td>
<td>Sept</td>
</tr>
<tr>
<td>Issue Email and Survey to PIs from CLIC; provide info via email that survey and call for pilots has been released (to IDTF, administrators and the evaluators listservs and through the newsletter)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survey due date (2 weeks)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conduct interviews with interested sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Select pilot sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set up Scorecard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set up triage process for TA/support questions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present to CM Executive Team: Dev Slides</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create training materials for OG, TTC, RBA, and Scorecard</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Conduct Pilot Test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kick-off Webinar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call 1: RBA/Scorecard/TTC Training Review Webinar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call 2: OG &amp; Script Output (2 calls)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call 3: RBA &amp; TTC Plans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data collection and entry, TTC plan development (hubs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monitor data entry into Scorecard</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical assistance and support, 1-1 consult (CLIC, Dev team)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Track challenges, solutions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assess feasibility, quality and usability of metric and OG for data collection and strategic management</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Post-Pilot Activities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-pilot Assessment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key Informant Interviews per data model group</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reporting / Communications</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progress Reports: Development team &amp;/or Calls w NCATS CM Lead</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM EC Meeting - CLIC update</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>CTSA Program Newsletter - Update on ICM Pilot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>informatics Full DTF Meeting (pilot updates)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTSA Program Evaluators Call</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Summary Report to Metric Dev Team</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Informatics Common Metric Pre-Pilot Landscape Assessment

University of Rochester Center for Leading Innovation and Collaboration (CLIC)

Funded by the National Center for Advancing Translational Sciences at the National Institutes of Health, Grant U24TR002260

Released 1.31.18
Outline

• The Informatics Common Metric (ICM)
• Assessment
Informatics Common Metric:
Clinical Data Repository Characterization

Reporting on 8 common data domains within a clinical data repository

Each common data domain will include the following*:

• Numerator: Count of unique patients with the standard value
• Denominator: Count of unique patients within the clinical data repository

Metric = % of unique patients with the standard value

*except for the observations data domain where the metric is presence or absence of an observation

Operational Guidelines: (please feel free to provide comments on this page)
https://docs.google.com/document/d/1X3HDV_t_NX1BPX1mhZSKBP4aoUAlspyeZHDpwS4T2s/edit?usp=sharing
## Informatics Common Metric: Clinical Data Repository Characterization

<table>
<thead>
<tr>
<th>Data Domain</th>
<th>Standard Value</th>
<th>Numerator</th>
<th>Denominator</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>N/A</td>
<td>count of unique patients with a age/DOB value</td>
<td></td>
<td>% patients with age/DOB value</td>
</tr>
<tr>
<td>Patient</td>
<td>Administrative Sex</td>
<td>count of unique patients with administrative sex value</td>
<td></td>
<td>% of total with sex value</td>
</tr>
<tr>
<td>Labs</td>
<td>LOINC ID</td>
<td>count of unique patients with a LOINC ID value</td>
<td>Count of all patients in the Data Repository</td>
<td>% of patients with LOINC ID value</td>
</tr>
<tr>
<td>Medications / Drugs</td>
<td>RxNorm ID</td>
<td>count of unique patients with a RxNorm ID value</td>
<td></td>
<td>% of patients with RxNorm ID value</td>
</tr>
<tr>
<td>Conditions / Diagnosis</td>
<td>ICD 9/10 or SNOMED</td>
<td>count of unique patients with an ICD 9/10 value</td>
<td></td>
<td>% of patients with ICD 9/10 value</td>
</tr>
<tr>
<td>Procedures</td>
<td>ICD 9/10 or CPT</td>
<td>count of unique patients with an ICD 9/10 or CPT procedure value</td>
<td></td>
<td>% of patients with ICD 9/10 or CPT procedure value</td>
</tr>
<tr>
<td>Notes / Narrative</td>
<td>N/A</td>
<td>count of unique patients with free text data</td>
<td></td>
<td>% of patients with free text data</td>
</tr>
<tr>
<td>Observations</td>
<td>N/A</td>
<td>Presence of Observations or Absence of Observations</td>
<td>N/A</td>
<td>Presence of Observations or Absence of Observations</td>
</tr>
</tbody>
</table>
Informatics Common Metric: Data Models

• The iDTF initially solicited information about which data models were used with high frequency within the CTSA Program

• This resulted in the identification of the OMOP, PCORnet and i2b2/ACT data models

• In addition, many hubs had developed partnerships with the company TriNetX to access data from their enterprise data warehouse
  • Through TriNetX, metric data may be provided to the hubs using the i2b2 and OMOP data models
Informatics Common Metric: Process of Acquiring Data for the Metric

• Scripts are used to query a hub’s enterprise data warehouse to acquire the metric data

• If hubs want to utilize another i2b2 data model (not ACT or through TriNetX) they will be responsible for acquiring the data from their enterprise data warehouse:
  • The hubs would need to develop their own script to query their enterprise data warehouse
  • Hubs are encouraged to share the script on the NCATS CTSA Metrics GitHub site (https://github.com/ncats/CTSA-Metrics); however, there is no assurance that other hubs would be able to validate the script due to differences in the hub’s enterprise data warehouse and/or data model

• If hubs wanted to use any other data model:
  • Hubs wishing to use other data models should suggest additions (along with a list of other adopting hubs) to CLIC, who will coordinate with the iDTF for consideration
Assessment

• Prior to starting the pilot, the Informatics Development Team, in collaboration with the CLIC Common Metrics team, solicited information to assess:
  • The diversity of the types of data models that a hub primarily uses for their enterprise data warehouse,
  • Of the data models approved by the Informatics DTF (iDTF) for use to report on the Informatics Metric, which model(s) the hubs are using currently,
  • Of the approved data model(s), which model the hub will be using to report on this metric, and
  • The expectation of effort to acquire the data for this metric.

• 62 responses were provided
Assessment: What is the **Primary** Enterprise Data Warehouse Model that hubs are using?

- Approximately half of the hubs are using the approved data models for their **Primary** Enterprise Data Warehouse.
- Of those hubs using other data models, there are various other models in use, some developed in-house.
Assessment: What are the other **Primary Enterprise Data Warehouse Models**?

<table>
<thead>
<tr>
<th>DataDirect</th>
<th>[Hub] Data Warehouse for research and Oracle Health Foundation for Health System</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM Health Care data model</td>
<td>Health Catalyst</td>
</tr>
<tr>
<td>Locally developed dimensional model</td>
<td>Home developed</td>
</tr>
<tr>
<td>Custom currently. Moving to OMOP.</td>
<td>Health Catalyst</td>
</tr>
<tr>
<td>In house developed model</td>
<td>Epic Caboodle, Constellation, Cosmos</td>
</tr>
<tr>
<td>Atomic clinical data warehouse using a relational model</td>
<td>We don’t have a warehouse at the moment</td>
</tr>
<tr>
<td>We have an i2b2 instance that is greatly-expanded beyond what's needed for ACT.</td>
<td>i2b2/other</td>
</tr>
<tr>
<td>Production Zone data model which is a superset / source of the other implemented models and supports [hub] use cases.</td>
<td>Home grown CRDW</td>
</tr>
<tr>
<td>EDW is based on Epic but contains many other data sources so more of a virtual model.</td>
<td>We are currently considering i2b2/ACT</td>
</tr>
<tr>
<td>We have raw data from source EHR and research systems (i.e., Epic, Allscripts, Genomic Information System, CTMS, REDCap) that we subsequently transform into OMOP, PCORnet, i2b2, and TriNetX formats.</td>
<td>We custom conform to both PCORnet &amp; i2b2/TriNetX</td>
</tr>
<tr>
<td>Local i2b2</td>
<td>Home-grown system based on ideas from Carecast and Centricity</td>
</tr>
<tr>
<td>Relational (Netezza) star-schema, subject domain roll-ups. OMOP, CDM, ACT, TriNetX derived from the EDW.</td>
<td>For [Hub] Health, it is data model developed by warehouse team led by Chief Data Officer. For PCORnet CDRN it is PCORnet (14M patients). For ACT, it is i2b2/ACT. We have multiple repositories for multiple purposes.</td>
</tr>
<tr>
<td>Oracle Health Data Warehouse Foundation, OMOP</td>
<td>Custom Integrated data repository with data marts for PCORnet, i2b2/ACT and i2b2/TriNetX.</td>
</tr>
<tr>
<td>We have a custom data model that serves as our primary clinical research dw</td>
<td>ORACLE, PCORNet has the most data, I2b2 is run off Oracle and could be used</td>
</tr>
</tbody>
</table>

*[Hub] Research Data Warehouse model*

*Internal data warehouse*
Assessment: What **approved** data models does your hub use?
Assessment: What approved data models does your hub anticipate using for reporting the ICM?

Note: Those hubs that select to use another i2b2 data model will have to develop their own script to query their clinical data warehouse to report on this metric.
Assessment: Which approved data model does your hub anticipate using for reporting the ICM? (by hub size)
Assessment: What is the anticipated level of effort to run the script and acquire the ICM data?
Assessment: What is the anticipated level of effort to run the script and acquire the ICM data? (by hub size)
Questions?

CLIC Website: Common Metrics Page: https://clic-ctsa.org/common_metrics

Common Metrics Help Desk: https://clic-ctsa.org/common_metrics/help-desk
# Appendix 6: ICM Pilot Sites

<table>
<thead>
<tr>
<th>Data Model</th>
<th>Hub</th>
<th>Hub Size</th>
<th>Single / Multi-Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>I2b2/ACT</td>
<td>Indiana CTSI</td>
<td>Med</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>University of Florida</td>
<td>Small</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>University of Pittsburg</td>
<td>Large</td>
<td>Single</td>
</tr>
<tr>
<td></td>
<td>University of Rochester</td>
<td>Small</td>
<td>Single</td>
</tr>
<tr>
<td>I2b2/TriNetX</td>
<td>Icahn School of Medicine at Mt. Sinai</td>
<td>Med</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>University of Mass Center for CTS</td>
<td>Small</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>Weill Cornell</td>
<td>Med</td>
<td>Multi</td>
</tr>
<tr>
<td>OMOP</td>
<td>Columbia University</td>
<td>Large</td>
<td>Single</td>
</tr>
<tr>
<td></td>
<td>Einstein-Montefiore</td>
<td>Small</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>UC Irvine</td>
<td>Small</td>
<td>Single</td>
</tr>
<tr>
<td>PCORnet</td>
<td>Medical College of Wisconsin</td>
<td>Small</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>Ohio State CCTS</td>
<td>Small</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>UCLA</td>
<td>Large</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>University of Chicago</td>
<td>Med</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>University of Kansas Medical Center</td>
<td>Small</td>
<td>Multi</td>
</tr>
<tr>
<td></td>
<td>Washington University Institute of CTS</td>
<td>Large</td>
<td>Multi</td>
</tr>
</tbody>
</table>
The Center for Leading Innovation and Collaboration (CLIC)

ICM Scorecard Data Entry Instructions
Home Screen – There are three ways to access the Informatics Common Metric Scorecard
From the Home Screen, click “Scorecard Objects,” then click “Scorecards”
A list of all of your hub’s Scorecards will display. Select “Informatics Common Metric Pilot”
Alternatively, use the Search bar, and begin typing the name of the Scorecard you seek, then select from the best options shown.
The third way, is to use the drop-down menu from the black bar at the top of the page.
The Scorecard has eight Performance Measures for which your hub will enter data.
To enter data, click on the “PM” icon and select “Add Data Values” from the drop-down menu.
Click in the “Choose a measure to enter data values for...”
From the drop-down menu, select the numerator or denominator.

Once the denominator (Count of all patients in the data repository) is entered for the first Performance Measure, it will populate into all others automatically.
Selecting the numerator measure brings up the data entry options for that measure. You will see your hub’s data model listed as well.
Click on the “Past” green plus sign to open the data entry windows.
Enter your data in the “Actual Value” box for 2017. Leave the “Target Value” blank – then click “Save”
You will be returned to the data entry window and see a warning message in **red**, that not all of the values have been reported for 2017.
Click again in the “Choose a measure to enter data values for” and the measure not yet recorded will be highlighted.
As before, click on the “Past” green plus sign to open the data entry window.
Enter the value for the measure in the “Actual Value” box for 2017. Leave the “Target Value” blank; then click “Save”
Scorecard will calculate the percentage, then click “Save”
Your Scorecard will look like this with the data completed.
Click on the gray and white plus sign to the left of a performance measure to reveal the point on the graph for the data entered.
After the data is entered for all Measures, click on the text of the last Performance Measure to access the Turn the Curve Plan section.
In the lower right, you will see the “File Attachments” box where you upload your completed worksheet.
Click “Upload File” to upload the worksheet
Enter the information for the Turn the Curve Plan, and click “Save”
Contact Us if You Have Questions!

- help@clic-ctsa.org
- 585.276.4817
- clic-ctsa.org/contact
### Appendix 8: Technical Assistance Calls and Support

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Attendance</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/17/17</td>
<td>Kick-off Call</td>
<td>16/16 Hubs</td>
<td>Several questions were asked about the selection and use of one data model for the pilot, as some hubs are using more than one data model across multiple organizations. CLIC staff and the ICM Development Team clarified that hubs were being asked to select one data model to use for the pilot, but stressed that hubs could expand and collect data for additional models after the pilot.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CLIC Development Team</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>45 participants total</td>
<td></td>
</tr>
<tr>
<td>11/29/17</td>
<td>RBA &amp; Scorecard Refresher</td>
<td>15/16 Hubs</td>
<td>Questions generated were fairly basic, related to confirming deadlines and expectations, and reiterating that the ICM pilot Scorecard was separate from the Common Metrics scorecard being used for the first three Common Metrics.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CLIC Development Team</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>35 participants total</td>
<td></td>
</tr>
<tr>
<td>12/11/17</td>
<td>Operational Guidelines and Script Output</td>
<td>16/16 Hubs</td>
<td>The OMOP group indicated that their script ran well, although 2 hubs indicated that having a script for Oracle, and not just SQL would be helpful.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CLIC Development Team</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>30 participants total</td>
<td></td>
</tr>
<tr>
<td>12/12/17</td>
<td>14/16 Hubs</td>
<td>14/16 Hubs</td>
<td>The PCORnet group indicated that their script ran well, and needed few modifications. The i2b2 ACT hubs had more varied responses related to running their script. One hub reported no difficulty, while 3 others had problems (internal permissions issues and a table name that didn’t exist) one reported they were not sure that the Oracle script run ever finished, and the third reported that their administrator had to debug the script line by line.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CLIC Development Team</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>26 participants total</td>
<td></td>
</tr>
</tbody>
</table>
The i2b2 TriNetX hubs asked TriNetX to run their scripts, and each reported that the run went quickly with no issues. There was much discussion about the “observations” performance measure, and it was determined that the measure would be updated to clearly ask for and give a yes or no response as to whether or not there were observations present, not how many, or what was said.

Questions generated were primarily related to deadlines and timelines. CLIC reiterated that there is a separate Scorecard for the ICM pilot, and hubs would not be using the “Common Metrics” Scorecard.

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Attendance</th>
<th>Notes</th>
</tr>
</thead>
</table>
| 1/9/18| Scorecard and Turn the Curve Plans | 13/16 Hubs CLIC Development Team 33 participants total | Identified priority topics and topics sent by hubs included:  
  • Developing the “What Works,” and “Strategies” sections of the Turn the Curve plan  
  • Planning for integrating the ICM at the hub  
  • Interpreting the script output.  
  • “How will Informatics Common Metrics be harmonized with CTSA ACT metrics? Should CM pilot sites be involved in the harmonization determination?”  
  • “Are there other data points that might help in interpretation of data across all CTSAs (e.g., age of data warehouse)?”  
  • “Should the Guidelines specify a time period (e.g., last or first month of the calendar year) in which the metrics should be run by all sites? (would help standardize data collection).” |
## Appendix 9: Number of Unique Patient Records in Repository by Hub

<table>
<thead>
<tr>
<th>Row Labels</th>
<th>i2b2ACT</th>
<th>i2b2TriNetX</th>
<th>OMOP</th>
<th>PCORnet</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUHS</td>
<td>-</td>
<td>-</td>
<td>5,663,620</td>
<td>-</td>
<td>5,663,620</td>
</tr>
<tr>
<td>Einst</td>
<td>-</td>
<td>-</td>
<td>1,634,971</td>
<td>-</td>
<td>1,634,971</td>
</tr>
<tr>
<td>Indi</td>
<td>5,561,532</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5,561,532</td>
</tr>
<tr>
<td>ISMMS</td>
<td>-</td>
<td>-</td>
<td>2,775,523</td>
<td>-</td>
<td>2,775,523</td>
</tr>
<tr>
<td>MCOFWGH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1,097,615</td>
<td>1,097,615</td>
</tr>
<tr>
<td>OSU</td>
<td>-</td>
<td>-</td>
<td>951,403</td>
<td>951,403</td>
<td></td>
</tr>
<tr>
<td>Pitt</td>
<td>3,336,710</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3,336,710</td>
</tr>
<tr>
<td>UCDOFM</td>
<td>-</td>
<td>-</td>
<td>589,903</td>
<td>589,903</td>
<td></td>
</tr>
<tr>
<td>UCI</td>
<td>-</td>
<td>-</td>
<td>597,192</td>
<td>-</td>
<td>597,192</td>
</tr>
<tr>
<td>UCLA</td>
<td>-</td>
<td>-</td>
<td>1,954,724</td>
<td>1,954,724</td>
<td></td>
</tr>
<tr>
<td>UFI</td>
<td>1,608,741</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1,608,741</td>
</tr>
<tr>
<td>UK</td>
<td>-</td>
<td>-</td>
<td>2,298,089</td>
<td>-</td>
<td>2,298,089</td>
</tr>
<tr>
<td>UMassW</td>
<td>-</td>
<td>2,374,744</td>
<td>-</td>
<td>-</td>
<td>2,374,744</td>
</tr>
<tr>
<td>UR</td>
<td>1,172,544</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1,172,544</td>
</tr>
<tr>
<td>WMCU-CTSA</td>
<td>-</td>
<td>2,087,101</td>
<td>-</td>
<td>-</td>
<td>2,087,101</td>
</tr>
<tr>
<td>WU</td>
<td>-</td>
<td>-</td>
<td>1,664,677</td>
<td>1,664,677</td>
<td></td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td>11,679,527</td>
<td>4,461,845</td>
<td>10,671,306</td>
<td>8,556,411</td>
<td>35,369,089</td>
</tr>
</tbody>
</table>
CLIC - CTSA Informatics Metric Post-Pilot Assessment

The CLIC Common Metrics Team and the Informatics Common Metric Development Team thank you for your participation in the Informatics Common Metric Pilot.

We would like your input about the metric and the processes associated with this pilot to inform the next steps to implement this metric across the CTSA Program.

Within the e-mail you received with the link to this assessment, you also received a pdf file of the assessment questions, to guide your decisions about who at your hub can best answer the assessment sections.

Please share your comments on the following areas:

Section A: Data model choice
Section B: Script process
Section C: Operational Guidelines
Section D: Hub implementation Team
Section E: RBA process and training process
Section F: Scorecard usability
Section G: Usefulness of this Informatics common metric at the hub
Section H: Pilot training and support

We would like to receive ONE response per hub.

Once you have started the assessment, if others at your hub are also going to complete sections, be sure to click "Save and return later," and a return code will pop up on the screen. Copy the return code, and share the code and the assessment link with anyone else at your hub who will be completing sections. When you, or others go back to the assessment once it has been started, click the blue link "Returning?" at the top right of the page to continue to input your responses.

Please submit ONE response to this assessment by COB Wednesday February 28, 2018. This assessment is expected to take your team 30 - 45 minutes to complete. The assessment will inform discussions of the groups that are working with the same data model. These discussions will take place in the month of March.

Thank you in advance for your feedback. If you have questions, please contact Bobbi McCaffery at: roberta_mccaffery@urmc.rochester.edu.

Thank you!

Your Institution

☐ Albert Einstein College of Medicine - Montefiore Health
☐ Columbia University
☐ Icahn School of Medicine at Mount Sinai
☐ Indiana University - Purdue University at Indianapolis
☐ Medical College of Wisconsin
☐ Ohio State University
☐ University of California at Irvine
☐ University of California at Los Angeles
☐ University of Chicago
☐ University of Florida
☐ University of Kansas Medical Center
☐ University of Massachusetts
☐ University of Pittsburgh
☐ University of Rochester
☐ Washington University
☐ Weill Medical College of Cornell University
Section A: Please provide information about the choice of data model for this metric pilot:

This part of the assessment was completed by: Name / Title

A1. When your hub initially volunteered to be an Informatics Metric Pilot site, which data model did your hub propose to use for the pilot?

- OMOP
- PCORnet
- i2b2 ACT
- i2b2 TriNetX
- Other i2b2 data model

A2. Did you use this data model for the pilot?

- Yes
- No

A2a. If No, Why?

A2b. Which data model did you use for the pilot?

- OMOP
- PCORnet
- i2b2 ACT
- Worked with TriNetX to acquire the metric data

A2c. Why?

A3. When this metric is implemented across the CTSA Program, will you use the same data model that you used for the pilot?

- Yes
- No
- Not sure

A3a. If no or not sure, why?
A3b. If no, what model would you plan to use?

- OMOP
- PCORnet
- i2b2 ACT
- We will work with TriNetX to acquire the data for the metric
- We plan to use another i2b2 data model and will develop our own script to acquire the data for this metric
- Other

A4. Other comments related to your hub's choice of data model for this metric pilot:

---

Section B: Please provide information on the process of using the data model script and the data that was acquired:

This part of the assessment was completed by: Name / Title

B1. Was it easy to access the script from the NCATS CTSA-Metrics GitHub site?

- Yes
- No
- Not Applicable - we requested TriNetX to provide the metric data to our Hub

B1a. If No, why?

B2. Did your team have any other challenges working with the NCATS CTSA-Metrics GitHub site?

- Yes
- No

B2a. If Yes, would a short tutorial on GitHub be useful?

- Yes
- No
- Maybe

B2b. If Yes or Maybe, does your team have any suggestions as to how to make this easier?

B3. Did you post questions and/or information on the GitHub site?

- Yes, posted questions
- Yes, posted information
- Yes, posted both questions and information
- No, did not post to GitHub
B3a. If you posted questions and/or information, how easy was it to post information on the GitHub site?

Not at all easy
Somewhat easy
Easy
Very Easy

B3b. Was the question resolved within 1-2 business days?

- Yes
- No

B3c. If No, was there a reason for the length of time it took to have the question resolved?

- Yes
- No
- Don't know

B3d. If Yes, what was the reason?

B4. Considering the first time the script was downloaded from GitHub and run, how much editing was required in order to run the script?

- 1 person / day
- 2-5 person / days
- 1-2 person / weeks
- 1 person / month
- None
- Not applicable, TriNetX provided the metric data to our hub

B5. If any edits had to be provided to the script, who did you work with to edit the script?

- Someone from our Metric Team
- Someone outside our Metric Team

B5a. If you worked with someone from outside your Metric Team, who did your team work with to edit the script?

B5b. After the edits were incorporated into the script, how much effort did it take to run the revised script and receive the metric data results?

- 1 person / day
- 2-5 person / days
- 1-2 person / weeks
- 1 person / month

B6. Please provide the amount of time that it took to run the final script and acquire the metric data:

B7. What is the title/role/position of the person who ran the script?
B7a. Was this person a member of your Metric Team?

○ Yes
○ No

B8. How many times did you have to run the script to get the final set of data that you entered into the Clear Impact Scorecard?

○ 1
○ 2 - 4
○ 5 - 9
○ 10+

B9. Did you have complete data for all eight domains?

○ Yes
○ No
○ Somewhat

B9a. Please explain:

B10. How confident is your hub that the output accurately reflects the status of the model you tested?

○ Not at all confident
○ Moderately confident
○ Somewhat confident
○ Very confident

B10a. Did you get the number you expected?

○ Yes
○ No

B10b. If No, please explain:

B10c. Does your hub trust the data generated?

○ Yes
○ No

B10d. If No, please explain:

B11. Were the final results consistent with other Informatics work/results done by your Hub?

○ Yes
○ No
○ Somewhat
○ No other results to compare to

B11a. Please explain:
B12. Does your team have suggestions for a more effective way of exchanging information about scripts across the consortium?

☐ Yes
☐ No

B12a. If Yes, please describe:

B13. Other comments related to the process of using the data model script and the data that was acquired:

---

**Section C: Please provide information related to the Operational Guidelines**

This part of the assessment was completed by Name / Title:

C1. Were the Operational Guidelines clear?

☐ Yes
☐ No

C1a. If No, what suggestions do you have for improving the Operational Guidelines?

C2. Is there anything your hub would like to see added to the Operational Guidelines?

☐ Yes
☐ No

C2a. If Yes, Please elaborate:

C3. The inclusion/exclusion criteria were easy to define at our hub.

Not at all easy  ☐  Somewhat easy  ☐  Easy  ☐  Very Easy  ☐
C4. Given that the Operational Guidelines provide for the opportunity to aggregate counts across the Hub and clinical affiliates (only if individuals in the repository can be unambiguously identified and counted across individual data repositories), would your Hub plan to report aggregate counts at the time of the Consortium-wide implementation?

- Yes
- No
- Maybe

C5. Other comments related to the Operational Guidelines:

---

Section D: Please provide information on the process of building your hub's implementation team

This part of the assessment was completed by Name / Title:

D1. Please describe your Informatics Metric Pilot team in terms of titles, roles, expertise, etc.:

D2. Were there other individuals or specific areas of expertise for which your hub wanted to engage individuals as team members?

- Yes
- No

D2a. If Yes, please explain:

D3. How were your hub's team members identified?

D4. What strategies did your hub use to engage team members?

D5. Other comments related to the process of building your hub's implementation team:
Section E: Please provide information on the RBA process and training

This part of the assessment was completed by Name / Title:

E1. What proportion of team members completed (or previously completed) Results Based Accountability (RBA) training?

- [ ] 0 - 25%
- [ ] 26% - 50%
- [ ] 51% - 75%
- [ ] 76% - 100%

E2. Was your team clear on the purpose(s) of the different components of the RBA process?

- [ ] Yes
- [ ] Somewhat
- [ ] No
- [ ] Unsure

E3. How easy was it to interpret the numerical data generated by the script?

- [ ] Not at all easy
- [ ] Somewhat easy
- [ ] Easy
- [ ] Very easy

E4. How easy was the process to develop the Story Behind the Curve?

- [ ]

E5. How easy was the process to develop the list of Partners?

- [ ]

E6. How easy was the process to develop the Strategies?

- [ ]

E7. How easy was the process to develop the What Works section?

- [ ]

E8. The RBA process helped organize our planning

- [ ] Strongly Disagree
- [ ] Disagree
- [ ] Agree
- [ ] Strongly Agree

E9. The RBA process helped identify aspects or components of the issue/process we had not considered

- [ ]

E10. The RBA process helped organize our reporting

- [ ]

E11. Overall, the RBA process was a useful approach to strategic management

- [ ] Yes
- [ ] Somewhat
- [ ] No
- [ ] Unsure

E12. Other comments related to the RBA process and training:
**Section F: Please provide information on Scorecard usability**

This part of the assessment was completed by Name / Title:

<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1. The Clear Impact Scorecard was time consuming</td>
<td>×</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>F2. The Clear Impact Scorecard added little value to the process</td>
<td>×</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>F3. The Clear Impact Scorecard was difficult to navigate</td>
<td>×</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>F4. The Clear Impact Scorecard was redundant in some parts</td>
<td>×</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

F5. Other comments related to Scorecard usability:

---

**Section G: Please provide information related to the usefulness of this Informatics Common Metric at the hub**

This part of the assessment was completed by Name / Title:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Somewhat</th>
<th>No</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1. Overall, this Metric provided a useful way to inform quality improvement initiatives related to Informatics</td>
<td>✗</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2. This Metric fits well with other Informatics related initiatives at your hub.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G3. This Metric fits well with other CTSA evaluation activities at your hub.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

G4. Other comments related to the usefulness of this metric at your hub:
Section H: Please provide information on the Pilot training and support

This part of the assessment was completed by Name / Title:

H1. One goal of the Common Metrics Initiative is to increase engagement between and across the hubs. What suggestions does your hub have to increase engagement about Informatics?

H2. Were there any challenges your hub experienced, about which you have not yet had a chance to comment?
   ○ Yes
   ○ No

H2a. If Yes, please explain:

H3. Were the Informatics Common Metric Pilot webinars helpful?
   ○ Yes
   ○ No

H4. Would your hub have preferred using video conferencing for the Informatics Metric Pilot Webinars?
   ○ Yes
   ○ No
   ○ No preference
   ○ We do not have video conferencing equipment

H5. Would your hub have preferred to have set “Office Hours,” during which the CLIC and Metric Development Team staff were available for assistance?
   ○ Yes
   ○ No

H6. Did anyone from your team go to the collaborative space to access webinar recordings or other resources?
   ○ Yes, accessed webinar recordings
   ○ Yes, accessed resources
   ○ Yes, accessed both webinar recordings and resources
   ○ No, did not use the collaborative space

H6a. If Yes, how easy was it to access the collaborative space?

   Not at all Easy
   Somewhat Easy
   Easy
   Very Easy

H7. Other feedback related to Pilot training and support:

H8. Overall comments and/or feedback on the Informatics Common Metric Pilot:
<table>
<thead>
<tr>
<th>Prompt</th>
<th>OMOP Group</th>
<th>i2b2 Group</th>
<th>PCORnet Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Should we eliminate the “Notes” domain?</td>
<td>Majority response: keep the “Notes” domain as it will be used in the future. • OMOP tool for this is not fully developed, but when available would be useful</td>
<td>Majority response: eliminate the “Notes” domain. • Not planning to add notes • Hubs did not make notes available to TriNetX, so was not part of TriNetX script • Not rushing to populate this</td>
<td>Majority response: notes are not part of the CDM for PCORnet, but it may be added in the future • Question as to whether PCORnet hubs will have it in their data mart?</td>
</tr>
<tr>
<td>What if hubs were asked to run an “all-in” script (entire data repository) followed by a shorter time period (last 10 years) with annual updates?</td>
<td>Majority response: hubs in favor of this change. • To help demonstrate improvements • To generate more useable data • Useful for benchmarking</td>
<td>Majority response: hubs in favor of this change. • What showed is not really indicative of where we are • It would be good to have the last 5-6 years • Only if it’s not too much work</td>
<td>Majority response: hubs only load the past 10-12 years into their CDM • CLIC will need to ask PCORnet hubs about this prior to roll out</td>
</tr>
<tr>
<td>Many sites want to combine data across different models/CDW or institutions - should the initial roll-out have same restrictions as the pilot, or differ?</td>
<td>Majority response: keep it as it was in the pilot. • Could decrease reproducibility and increase work • Hubs want to be able to do this, but not as a requirement</td>
<td>Only one i2b2 hub response: • Perhaps ask hubs to limit to databases/warehouses used to support research • Combining the data could mean a great deal of additional work</td>
<td>Majority response: for first year, keep the same as the pilot. • Hubs with multiple health care systems - will be valuable to combine models in future • Does it have to be 1 or the other? Could it be optional?</td>
</tr>
<tr>
<td>Should we add a “stretch metric” such as genomic data?</td>
<td>Majority response: do not add for all. • Possibly add for hubs that are further along with informatics • We want the first run of the ICM to be successful across the majority of hubs</td>
<td>Majority response: do not add for all. • Start with what was tested in the pilot and consider adding a stretch metric later on</td>
<td>Majority response: group is PCORnet CDM, and is at their mercy about what can be loaded • Some hubs may be changing over to OMOP, and would like to do a stretch metric</td>
</tr>
<tr>
<td>Prompt</td>
<td>OMOP Group</td>
<td>i2b2 Group</td>
<td>PCORnet Group</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>• Perhaps this could be an IDTF discussion? (data quality; data expansion)</td>
<td>• This would have to be beneficial for the hubs &amp; how CLIC will report to the hubs</td>
<td>• Is there a maturity model that could be applied – for hubs that are further along with informatics</td>
<td>• Hubs loved having scripts written for them, would suggest expanding this model to TINs/RICs if possible – saves hubs a lot of work</td>
</tr>
<tr>
<td>• Include more information on balancing informatics, administration and evaluation at the hubs to respond to the metric</td>
<td>• This metric is much more technical than the others, and very dependent upon informatics/IT folks at the hubs – don’t always speak the same language</td>
<td>• Guidance for persons doing the TTC plan on what informatics is</td>
<td>• It would be helpful to have guidance on who on the Metric Team should have Scorecard access</td>
</tr>
<tr>
<td>• Metric Team Development may need some guidance here</td>
<td>• Metric Team Development may need some guidance here</td>
<td>• Guidance for informatics folks on what the TTC plan is</td>
<td>• An ICM Implementation Guide would be very helpful</td>
</tr>
<tr>
<td>Anything we should do differently during the roll out?</td>
<td></td>
<td>• Guidance on Metric Team Development would be helpful</td>
<td>• Hubs want feedback on how they did developing their TTC plan</td>
</tr>
<tr>
<td>Retesting of the Data Model Scripts: Did the script work?</td>
<td>The OMOP script worked, and looked the same</td>
<td>Most i2b2-ACT hubs ran the script</td>
<td>Majority response was that the script ran “perfectly”</td>
</tr>
<tr>
<td></td>
<td>• Some hubs had line-specific feedback &amp; were asked to post on GitHub site</td>
<td>• Some hubs had line-specific feedback &amp; were asked to post on GitHub site</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• One hub’s i2b2 admin is on leave – will run the script ASAP</td>
<td>• One hub’s i2b2 admin is on leave – will run the script ASAP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Pitt script is good – looks better than the original</td>
<td>• Pitt script is good – looks better than the original</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 12: Informatics Metric Post Pilot Group Call Questions

Based on hub feedback – Operational Guidelines issues:

- Should we eliminate the “Notes” domain?
- What if hubs were asked to run an “all-in” script (entire data repository) followed by a shorter time period (last 10 years) followed by annual updates?
- Many sites want to combine data across different models/CDW or institutions - should the initial roll-out have the same restrictions as the pilot, or differ?
- Should we add a “stretch metric” such as genomic data?

Based on hub feedback – other issues:

- Are there special considerations based on hub size?
- What “pitfalls” should we watch out for in roll-out?
- What should we do the same as in the pilot?
- What should we NOT do the same as in the pilot?

Other recommendations discussed:

- Moving from pilot to implementation
  - Develop a Metric Implementation Guide, including:
    - Operational Guidelines
    - Instructions on developing the Turn the Curve plan
    - Definitions, examples and templates
  - RBA Training for entire Metric Team
- Re-testing of the scripts
  - Did it work?
  - If not – please comment on the GitHub
Appendix 13: New Metric Implementation Timeline

New Metric Implementation Timeline

Month 1
- Schedule regular dates/times for training webinars & Office Hours
- Prepare materials for OG/Scorecard release webinars

Month 2
- Deliver OG/Scorecard release webinars
- Receive questions & concerns from hubs
- Start coaching & TA sessions
- Begin regular Office Hours

Months 2-3
- Prepare training/TA materials for questions & concerns webinars
- Deliver questions & concerns webinars
- Prepare topic webinnetes (RBA; Metric fractions; etc.)

Months 2-3-4
- Deliver topic webinnetes
- Continue regular Office Hours

Months 4-5-6
- Data checking and TTC Plan review
- Continue regular Office Hours
- Hub feedback to CLIC

Month 6 and Beyond
- Routine maintenance for hubs