Report by Tufts CTSI on

Common Metrics Implementation Program and Evaluation Study

Executive Summary

Tufts | CTSI Tufts Clinical and Translational Science Institute
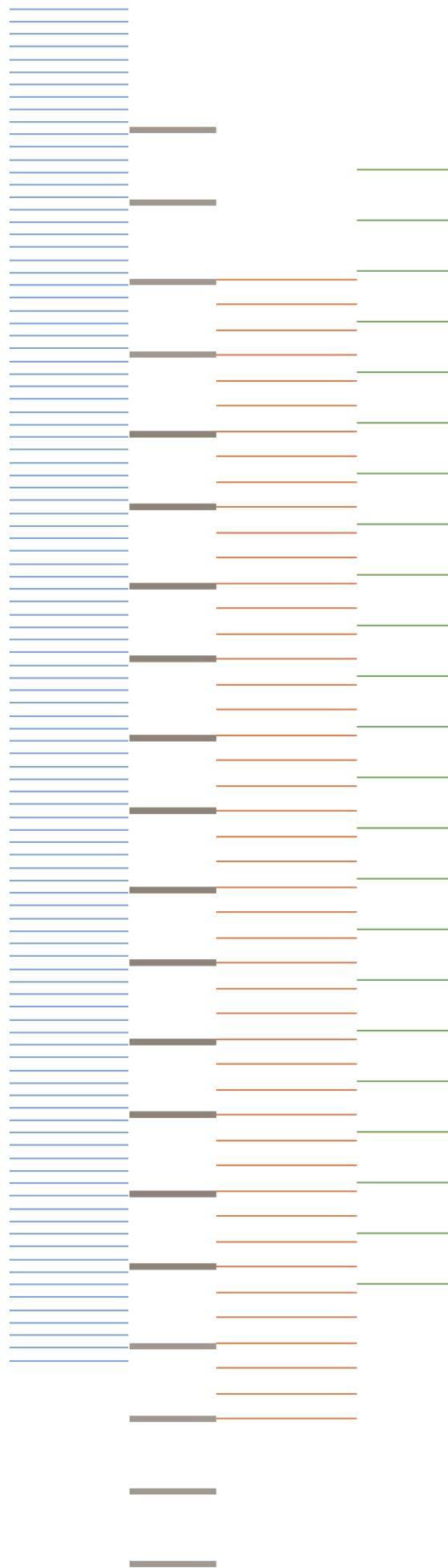
INTRODUCTION

The Clinical and Translational Science Award (CTSA) Program Consortium has been charged with accelerating and improving clinical and translational research on a national scale (https://ncats.nih.gov/ctsa/about/hubs). In the Congressionally mandated 2013 Report on the CTSA Program, the Institute of Medicine (IOM) was very supportive of the program, but recommended that CTSAs become a "*more tightly integrated network that works collectively.*"[1] This included a call for standardized evaluation processes based on measurable strategic goals and uniform, actionable "common metrics" to enhance transparency and accountability in decision-making. This was seen as key for powering needed evolution of CTSAs and their research communities. In response, the National Institutes of Health (NIH) National Center for Advancing Translational Science (NCATS), the home of the CTSA Program, and the CTSA institutional hubs nationally, implemented the Common Metrics Initiative. Using standardized metrics and the principles of the Results Based Accountability (RBA) performance management framework,[2] this initiative aims to improve the strategic management of individual CTSA hubs and of the national CTSA Consortium. This initiative also was seen as an opportunity to develop, demonstrate, and disseminate methods of improving "the science of doing science."

Having a focus of research process improvement, Tufts Clinical and Translational Science Institute (CTSI) was asked by NCATS to implement the Common Metrics and the RBA framework across the CTSA Consortium, and to run an evaluation study of this implementation, for three years, starting September 2015. This report synthesizes results and recommendations both from the Tufts Common Metrics Implementation Program and from the Tufts Common Metrics Evaluation Study. Summary recommendations are included below in the text; more specific recommendations related to each are at the end of the Executive Summary.

The primary goal of this report is to inform decision-making on future directions of the CTSA Program Common Metrics Initiative. However, it also may be of interest to other research groups or networks implementing standardized metrics and performance improvement processes, potentially including other NIH institutes and centers. Additionally, as a perturbation of the complex operations of the CTSA Consortium and its hubs, the responses to, and successes with, the Common Metrics Initiative may provide insights into the characteristics and operation of such clinical and translational science enterprises, and how they may be enhanced.

SELECTING AND DEVELOPING METRICS

The initial development of the metrics was outside the scope of the Tufts projects. However, the Tufts Implementation Program and Evaluation Study highlighted the importance of pilot testing and continuously reviewing metrics, considering local usefulness, and ensuring the validity of aggregating results across hubs.

Metric Development and Ongoing Review

In collaboration with NCATS and CTSA hubs, the Tufts Implementation Team conducted two types of pilot testing. The first three common metrics (Careers in Clinical and Translational Research, Pilot Funding Publications, and Institutional Review Board [IRB] Review Duration) were pilot-tested for data collection feasibility. A fourth metric (Clinical Trial Accrual) was tested for data collection feasibility and for usefulness for strategic management. Although both types of pilot testing identified gaps in metric definitions, the more robust approach also uncovered confusion about metric calculations and challenges to the usefulness of the metrics for strategic management that were important to address before widespread implementation.

Summary Recommendation 1:

Develop metrics using robust pilot testing, and engage stakeholders in ongoing review.

Metric Usefulness

Usefulness to Local CTSA Hubs

Hubs experienced value in implementing Common Metrics, but continued to have concerns about whether the metrics provided enough benefit to justify their effort. Hubs founds value in three ways: 1) implementing a formal structured process if one did not exist previously, 2) using the performance improvement process to enable strategic conversations, and 3) making improvements in processes and immediate outcomes. They also provided external requirements that helped hubs justify recommendations for institutional changes and/or targeted funding.

However, participants in qualitative interviews at more than two-thirds of CTSA hubs expressed concerns about the usefulness of the first three metrics and their overall value relative to the effort expended. Many hubs found it difficult to use the metrics for local improvement, particularly when the metrics did not align with local institutional priorities or addressed topics on which the hub was already performing well.

Usefulness to the National CTSA Consortium

Hubs reported continuing concerns about between-hub variation in how metric data would be collected and computed. To be used for benchmarking, comparison, or aggregate reporting, metric data must be comparable. However, during the period in which Tufts led the implementation process, two indicators raised questions about comparability across hubs. First, for calendar year 2015 metric results, 80% of hubs revised data for one or more metrics after originally entered. This suggested a changing understanding or approach to calculating metric results. Second, there was limited ability to assess data quality or the extent that hubs correctly followed Operational Guidelines because the submission of raw data or contextual information was not required. In addition,

meaningful interpretation of the range of metric values across hubs would require understanding the institutional, hub, and program-level characteristics that shape metric values.

Summary Recommendations 2 and 3:

Maximize usefulness to hubs by selecting metrics that align with local needs.

Maximize usefulness to the National CTSA Consortium by ensuring validity of aggregation and comparison reporting.

IMPROVING PERFORMANCE

Training, Coaching and Technical Assistance

Members of each hub received interactive webinar-based training and participated in small group coaching sessions during implementation of the initial Common Metric of their choice. Additional support included ad hoc individualized coaching, a Help Desk, worksheets to assist in calculating metric results, and exemplar strategic management plans ("Turn the Curve plans"). Small group coaching sessions and regular coach assessments of Turn the Curve plans were effective in gauging team progress, tailoring coaching sessions, and sharing experiences across hubs. Participants appreciated the opportunity coaching sessions provided for peer-to-peer learning and some valued a structured approach to meeting project milestones.

Eighty percent of hubs reported satisfaction with the training, coaching, and technical assistance received, and hubs that used these offerings indicated they facilitated their work. The vast majority of hubs reported that they gained the knowledge and proficiency they needed, or more, to carry out the work of the Common Metrics. However, almost one-third of hubs reported gaining more knowledge and proficiency than was needed, suggesting that hubs perceived different levels of need.

Summary Recommendation 4:

Equip hubs to fully implement each metric and performance management by providing peer-to-peer learning and training, coaching, and assistance for varying levels of experience.

Support for Implementation

Effective Teams

Each CTSA hub formed a core Common Metrics team to oversee data collection and the use of the metric for strategic management. Although team composition was often in line with the Implementation Team's guidance on team membership (e.g., included leadership, project management, and data system and subject matter experts), some hubs appeared to delegate responsibility for Common Metrics implementation to a small number of individuals and/or did not include subject matter experts with specialized knowledge of, or influence over, metric topic(s).

Collecting Data and Developing Strategic Management Plans

Implementing Common Metrics entailed collecting metric data and developing metric-specific strategic management plans. By the end of the evaluation period, the vast majority of hubs self-reported that they had computed metric results according to the Operational Guidelines and had completed activities to understand "the story behind" current performance. However, hubs experienced challenges in completing each element of the RBA performance improvement framework, some specific to the particular element and others more general across the improvement process. On average, hubs developed performance improvement plans for 77% of the metrics, with fewer plans for the IRB Review Duration metric than for the Careers or Pilot Funding Publications metrics.

Identifying Targets for Performance

A number of hubs were interested in identifying benchmarks to help understand their performance and progress, and to prioritize areas for improvement. However, some participants were concerned about lack of comparability of metric results across hubs that would undermine their use for comparing. Some were also concerned about results being used by NCATS to judge hubs' performance.

Disseminating Performance Drivers and Strategies for Improvement

Driver Diagrams, Change Packages, Collaborative Learning Sessions, and publication of hub *Success Stories* were used to disseminate best and promising practices for metric strategic management. Hub teams expressed considerable interest in continuing to share challenges and learning from each other about proven and potentially successful strategies to improve performance for Common Metric topics.

Software Support

Hubs used Clear Impact's proprietary Scorecard software to report their metric values and to document, manage, and communicate strategic management plans. Hubs recognized the value of a common software platform but reported technical limitations that hindered its benefit. The Scorecard software platform was not configured to support data collection, data management, or quality checks for metric results. Also, reporting and analytic features were limited, and hubs indicated the desire for enhanced user experience and visualization capabilities.

ADDRESSING BARRIERS AND SUSTAINING ENGAGEMENT

Not surprisingly, those hubs reporting active engagement in implementing Common Metrics completed more performance improvement activities than those reporting only a compliance-based approach. Many factors could affect a hub's level of engagement.

Resources and Personnel

Availability of resources was the most common reason hubs cited for not completing Common Metrics activities. Of note, the size of a hub's funding award did not fully account for this. Challenges related to investment from home institutions, interrupted funding, alignment of existing systems with the needs of the Common Metrics, and availability of needed personnel and expertise all affected whether hubs could devote sufficient time and resources to fully implement Common Metrics and performance improvement activities.

Local CTSA Program Hub Context

CTSA organizations are heterogeneous in their structures, organizational processes, and experience with metric-based performance improvement. Alignment with technical needs of the Common Metrics Implementation, especially compatibility with local structures, processes, metrics, and experience, facilitated completion of the work. When there was lack of alignment in these areas, more resources were required to conduct the work of the Common Metrics, and this hampered hubs' abilities to adapt and engage in that work. A second type of alignment, compatibility of Common Metrics with existing institutional priorities, also shaped hubs' progress on the work of the Common Metrics.

Local Authority

A CTSA hub leader's position in their home institutional authority structure was important for accessing needed data, affecting improvements, and facilitating stakeholder engagement. Hubs whose leaders did not have line authority over data or processes related to Common Metrics experienced challenges in implementing performance improvement. Drawing on or creating personal relationships to build communication about the topics of the Common Metrics was helpful for gaining buy-in by stakeholders. However, this did not fully compensate for lack of direct authority.

Effective Communication

Communication strategies included website postings, e-newsletter updates, presentations, conference panels, and roundtable discussions, and these appeared to provide an appropriate level of communication for initial implementation efforts. Communication challenges included ensuring that newly participating and existing staff were aware of how to access project resources and received initiative communications. Some hubs also indicated a need for more information about future plans for the Common Metrics Initiative, particularly when weighing the overall value of the initiative.

Summary Recommendation 7:

Sustain engagement by facilitating solutions to barriers due to resources and authority, accounting for hub heterogeneity, and ensuring effective communication.

EXPANDING DATA-DRIVEN DECISION MAKING

The Tufts Implementation and Evaluation efforts revealed many structural and cultural aspects of CTSA organizations that affected hubs' abilities to engage with and complete Common Metrics and performance improvement activities. These findings suggest two opportunities to expand data-driven decision making in the CTSA Program. First, several hubs indicated a desire to learn from other available data and research results in order to inform and enhance the Common Metrics Initiative. These hubs spoke of using clinical and operational data beyond that needed to calculate the metric result to inform the choice of Common Metrics and elucidate drivers of performance.

Second and more broadly, insights about structural and cultural factors affecting implementation of Common Metrics likely apply to other CTSA-wide initiatives as well. Expanding "the science of doing science" approach to governing the CTSA Program would address additional questions more deeply. For example, are there "best practices" for CTSA organizational structures and ways of linking to home institutions that expedite performance improvement or other types of initiatives? How did implementation of Common Metrics affect CTSAs and their relationships with home institutions? Are there metrics or processes that are commonly used across hubs that could inform selection of future Common Metrics or other priority initiatives?

Summary Recommendation 8:

Expand use of data to inform future directions of the Common Metrics Initiative and the CTSA Program.

LEARNINGS ABOUT THE FIRST THREE COMMON METRICS

Implementing the first three metrics and input from hubs during the evaluation study revealed questions about the usefulness of these metrics. As described above, there is an important opportunity to review the metrics with hubs to ensure they are useful relative to required effort and to address any needed modifications to the Operational Guidelines.

Metric: IRB Review Duration

Hubs continued to have questions about, and differences in, how they applied the Operational Guideline definitions. They also were challenged in developing strategic management plans in conjunction with the IRB, which often required crossing organizational boundaries. Usefulness of the IRB metric at the local level varied depending on the number of IRBs, the types of protocols reviewed, the extent of reliance on central IRB models, and the work process of reviews for ethics, feasibility, and budgets/contracts.

Metric: Pilot Funding Publications

Hubs identified two challenges with interpreting and using this metric. First, the metric is cumulative and some hubs reported that this made interpreting metric results at the hub level difficult and not useful for measuring improvement, particularly for hubs with large numbers of pilot awards. Regarding local usefulness, some hubs considered the Pilot metric to be too narrow in scope to capture the goals of their local programs. Specifically, there was disagreement as to whether publication was a primary desired outcome of pilot awards.

Metric: Careers in Clinical and Translational Science

The Operational Guideline provided examples of what it means to be "engaged in research" rather than a required definition, and there was disagreement with, or confusion about, several of the metric exclusion criteria. As a result, hubs used a range of definitions in their data collection, which raised questions about comparability of metric results across hubs. Additionally, a number of hubs reported that the cumulative nature of these metrics made interpretation of metric results difficult at the hub level, and some hubs considered the Careers metric to be too narrow in scope to capture the goals of their local programs.

Summary Recommendation 9:

Make improvements to the first three Common Metrics by clarifying Operational Guidelines and assessing usefulness with hubs.

CLINICAL TRIAL ACCRUAL METRIC PILOT TEST

Both hubs with and without clinical trial management systems (CTMSs) faced numerous challenges collecting the metric data and developing strategic management plans. Only one of eight pilot hubs for this metric was able to assess the accrual ratio for all its eligible trials, and all hubs had difficulty creating a central list of trials at their institution to use for a sampling frame. At many hubs, existing data sources did not align with the inclusion/exclusion criteria and variable definitions in the Operational Guideline, or the hubs did not have the needed variables. The exclusion of trials with fewer than 10 targeted participants removed many otherwise potentially-eligible clinical trials from the sampling frame. Concerns about data quality limited the usefulness of the metric for strategic management.

Summary Recommendation 10:

Use the results and recommendations detailed in the full Accrual Metric Pilot report to determine the future direction of metric implementation.

CONCLUSION

The Tufts Common Metrics Implementation and Evaluation Study generated insights and evidence to assess and reflect on the Common Metrics Initiative. Findings and conclusions speak most directly to future directions of the Common Metrics Initiative, but they can also inform other CTSA Program initiatives and similar networks that plan to embark on implementing shared metrics and performance improvement frameworks.

Summary and detailed recommendations

1	**Develop metrics using robust pilot testing, and engage stakeholders in ongoing review.**

1a For each new Common Metric, conduct a robust pilot test that equally emphasizes feasibility of data collection and usefulness of the metric for local and Consortium-wide strategic management.

1b Consider phased pilot testing (i.e., test data collection first, then strategic management) for metrics for which data quality or feasibility issues are likely.

1c Include requirements for collection and reporting of additional data points in Operational Guidelines to confirm that comparisons across hubs are valid. Ideally, reporting of all underlying data would allow for data audits. Short of that, reporting all data elements used to calculate metric values would support oversight of data quality.

1d Periodically engage hubs in a review of each metric for completeness, clarity, usefulness, and required effort.

2	**Maximize usefulness to hubs by selecting metrics that align with local needs.**

2a Select metrics that better align with local CTSA and home institution needs and priorities. For example:

    i. Consider clustering similar CTSAs to address selected metric topics rather than creating Consortium-wide requirements.

    ii. Acknowledge and communicate to hubs that local priorities can influence performance targets.

3	**Maximize usefulness to the National CTSA Consortium by ensuring validity of aggregation and comparison reporting.**

3a Regularly review metric results for those missing, clearly incorrect or inconsistent with Operational Guidelines and follow-up with hubs.

3b If aggregation or comparisons of hubs' metric results are pursued, ensure results are comparable across hubs. Consider implementing a data coordinating center function with formal data cleaning or auditing processes.

3c Until a full and thorough review of metric results can be performed, acknowledge inconsistencies in data collection when interpreting aggregated reports.

3d To allow meaningful interpretation of metric results and comparison across diverse hubs, collect and report factors relevant to performance, including hub and institutional characteristics.

Summary and detailed recommendations, continued

4	**Equip hubs to fully implement each metric and performance management by providing peer-to-peer learning and training, coaching, and assistance for varying levels of experience.**

4a Provide training and coaching that meets the needs of adult learners with different learning styles and various levels of prior experience in performance management.

4b Add training and coaching on more advanced strategic management concepts and relevant examples as the Common Metrics Initiative matures and participants become more proficient in implementing metrics and creating performance management plans.

4c Provide small group coaching when implementing each new Common Metric, and provide mechanisms to promote peer-to-peer learning and accountability for meeting implementation milestones. Consider extending coaching beyond initial metric implementation for hubs wanting or needing additional support.

4d Provide concrete examples of how to calculate each metric (e.g., metric calculation worksheets) and exemplar strategic management plans to assist hubs to conduct and document their planning.

5	**Support implementation by promoting metric-specific teams, allowing for capacity-building periods, providing accurate benchmarks, and updating performance drivers and best practices.**

5a Encourage the use of metric-specific teams with active subject matter experts who are able to address data issues and strategic management specific to the metric topic.

5b Promote hub-identified facilitators for building effective teams, including identifying one team member who takes ownership of the project and a local champion on the team, and attending to team climate and interactions.

5c Encourage involvement of the CTSA Principal Investigator to provide strategic guidance and oversight, to champion the project, and to facilitate stakeholder engagement.

5d Allow for a capacity-building period prior to mandating collection and reporting of metric data to support revising existing data sources, developing data sources and systems, and training personnel.

5e Encourage hubs to engage partners and subject matter experts outside of the core team to gain a deeper understanding of underlying causes of existing hub performance, and to assist in selecting and implementing improvement strategies.

5f Provide useful, accurate benchmarking data to help hubs better target areas for improvement.

5g If using aggregated hubs' metric results to identify performance benchmarks, acknowledge and describe local reasons for variation in metric results.

5h Provide hubs a repository of best and promising practices, including newly developed and updated Driver Diagrams and Change Packages, to speed and focus development of strategic management plans.

5i Promote peer-to-peer learning and disseminate best and promising practices.

  i. Consider continuing and establishing additional mechanisms for shared learning (e.g., Collaborative Learning sessions) and disseminating best and promising strategies (e.g., publishing hub Success Stories).

  ii. Highlight successful adoption and application of RBA and CTSA Consortium achievements.

Summary and detailed recommendations, continued

6	**Maximize usefulness of the reporting platform by enhancing functionality, visualization options, and user experience.**

6a Expand software features and functionality to support data collection, storage, and quality checks.

6b Enhance user experience (e.g., speed, intuitiveness, number of clicks required to navigate), and improve visualization capability (e.g., create more display options, display multiple metrics simultaneously).

7	**Sustain engagement by facilitating solutions to barriers due to resources and authority, accounting for hub heterogeneity, and ensuring effective communication.**

7a Facilitate solutions to limited resources and personnel and use multiple strategies to account for heterogeneity across hubs. For example:

    i. Consider aligning Common Metrics reporting with other required reporting (e.g., annual reporting).

    ii. Consider an explicit process to weigh the value of a metric with the effort to obtain data.

    iii. Consider a designated budget allocation to support Common Metrics work.

    iv. Use a software platform that does not limit the number of users due to fees.

7b Account for heterogeneity of hub data, processes, and local priorities. For example:

    i. Consider clustering similar CTSAs to address selected metric topics rather than creating Consortium-wide requirements.

    ii. Offer expanded flexibility in choice of performance improvement framework.

7c Maintain realistic expectations about the amount of improvement that can be achieved and the pace of change, particularly when the CTSA leader does not have line authority over the target processes.

7d Promote peer-to-peer learning about successful strategies for affecting change in the home institution.

7e Develop and maintain effective ongoing communication strategies for hub leadership and staff, and particularly new staff.

7f Inform hubs of future directions for the Common Metrics Initiative.

8	**Expand use of data to inform future directions of the Common Metrics Initiative and the CTSA Program.**

8a Use hub data beyond what is needed to implement the Common Metrics (e.g., other clinical and operational data) to inform the selection of metrics and to identify potential drivers of outcomes.

8b Use discussion of the Common Metrics Evaluation results to catalyze a broader conversation about other high impact research projects to drive data-driven decisions related to the structure of CTSAs and the CTSA Program.

Summary and detailed recommendations, continued

9	Make improvements to the first three Common Metrics by clarifying Operational Guidelines and assessing usefulness with hubs.

Metric: IRB Review Duration

9a   Consider clarifying Operational Guideline definitions about inclusion of multiple institutional IRBs.

9b   Collect additional data about the number of IRBs included in the metric calculation and the hub's ability to exclude pre-review activities to inform appropriate metric comparisons across the CTSA Consortium.

9c   Assess with hubs the usefulness of this metric to hubs and the CTSA Consortium given disparate local IRB processes, variation in types of clinical protocols, and concerns about comparability of metric values across hubs.

Metric: Pilot Funding Publications

9d   Consider modifying the metric only to include pilots that have had sufficient time to publish (e.g., one year after pilot conclusion).

9e   Assess with hubs the usefulness of this metric for local improvement, particularly the extent to which the metric reflects local priorities.

Metric: Careers in Clinical and Translational Science

9f   Modify the Operational Guideline to further define and clarify exclusion criteria:

    i.   clarify whether hubs may use additional definitions of "engaged in research,"

    ii.   add exclusion of solely institutionally-funded scholars,

    iii.   clarify definition of "still in training,"

    iv.   add criteria for "lost to follow-up."

9g   Consider modifying the metric to be annual (e.g., percent of 2015 graduates who are in CTR; percent of 2016 graduates who are in CTR, etc.).

9h   Assess with hubs the usefulness of this metric for local improvement, particularly the extent to which the metric definitions reflect local priorities.

Summary and detailed recommendations, continued

10	**Use the results and recommendations detailed in the full Accrual Metric Pilot report to determine the future direction of metric implementation.**

10a Consider providing an infrastructure-building period prior to mandated collection of metric data to allow hubs time to devise and/or revise data sources and systems and data collection and data quality procedures, and train personnel.

10b Modify the metric to be collected prospectively rather than retrospectively to increase its potential usefulness for strategic management, including the ability to identify and intervene in individual trials as needed.

10c Revise the Operational Guideline to address certain multi-site clinical trials (e.g., those of competitive enrollment design) in which key accrual Metric variables are not known.

10d Re-evaluate the exclusion criterion for trials with fewer than 10 targeted participants; consider lowering the cut-off (e.g., to trials with less than five targeted participants).

10e Do not exclude clinical trials of dose-to-toxicity design.

10f Collect and report additional information, including information about the mix of clinical trials at the primary institution or included in the Median Accrual Ratio, to understand how representative the median is of the intended sample.

10g Provide a template of tested survey qeustions and survey considerations.

10h Provide hubs with best or promising practices and strategies for implementing a CTMS to produce metrics.

REFERENCES

1   IOM (Institute of Medicine, now National Academy of Medicine). The CTSA Program at NIH: Opportunities for Advancing Clinical and Translational Research. Washington, DC: The National Academies Press. June, 2013. Accessed from: http://www.nationalacademies.org/hmd/Reports/2013/The-CTSA-Program-at-NIH-Opportunities-for-Advancing-Clinical-and-Translational-Research.aspx

2   Friedman M. Trying Hard Is Not Good Enough: How to Produce Measurable Improvements for Customers and Communities. FPSI Publishing, 2005. ISBN: 1-4392-3786-7

ACKNOWLEDGMENTS