2018-2019 Mid-Term Credibility Plan Review

Presented by: Jerry Myers (NASA GRC)

On behalf of the Committee on Credible Practice of Modeling & Simulation in Healthcare

https://simtk.org/home/cpms

cpmsinhealthcare@gmail.com

Mid-term Credibility Plan Review

- Feedback from the 2018 IMAG-MSM meeting identified the need for more consistent communication of model credibility status

- In Response CPMS Initiated “Mid-term Review” of credibility status based on the CPMS Ten Simple Rules

Goals :

- Improve communicating credibility plan progress
- Opportunity to practice communicating credibility
- Relate credibility activities to model contextual use
- Evaluate tools to promote credibility communication

Ten Simple Rules

Ten Simple Rules	
R1 - Define context clearly	***R6 - Document adequately***
R2 - Use appropriate data	***R7 - Disseminate broadly***
R3 - Evaluate within context	***R8 - Get independent reviews***
R4 - List limitations explicitly	***R9 - Test competing implementations***
R5 - Use version control	***R10 -Conform to standards***

https://www.imagwiki.nibib.nih.gov/content/committee-credible-practice-modeling-simulation-healthcare-description

Request to IMAG-MSM U01 Grantees

Utilizing the CPMS Ten Simple Rules for Model and Simulation Credibility

- Details on credibility plan actions
- Description of information gained by each action
- Plans for the next reporting cycle
- Summary activities table classified within the CPMS TSR
- Issues/concerns in achieving the standard of credibility
- Identify other factors that contribute to credibility

Review Scope

The reviewers:

DID NOT

Assess the implemented credible practice or research progress of the M&S projects

DID

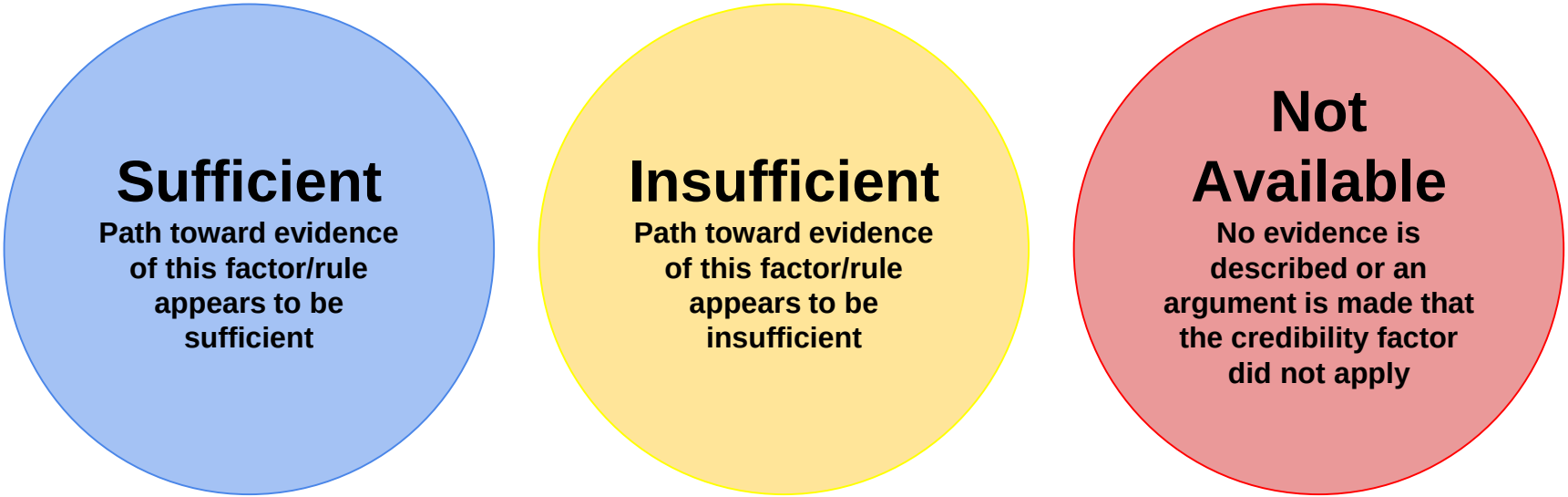
Opinion: communication sufficiency of credibility plans and accomplishments

Identify areas of improvement of the reporting / review process

Recommend credibility topic areas for discussion in IMAG/MSM community

CPMS TSR

Reviewers Scoring Rubric For Each TSR



Ancillary evidence and provided development history could be considered in assessing sufficiency of communicating content in each TSR factor

Review Overview

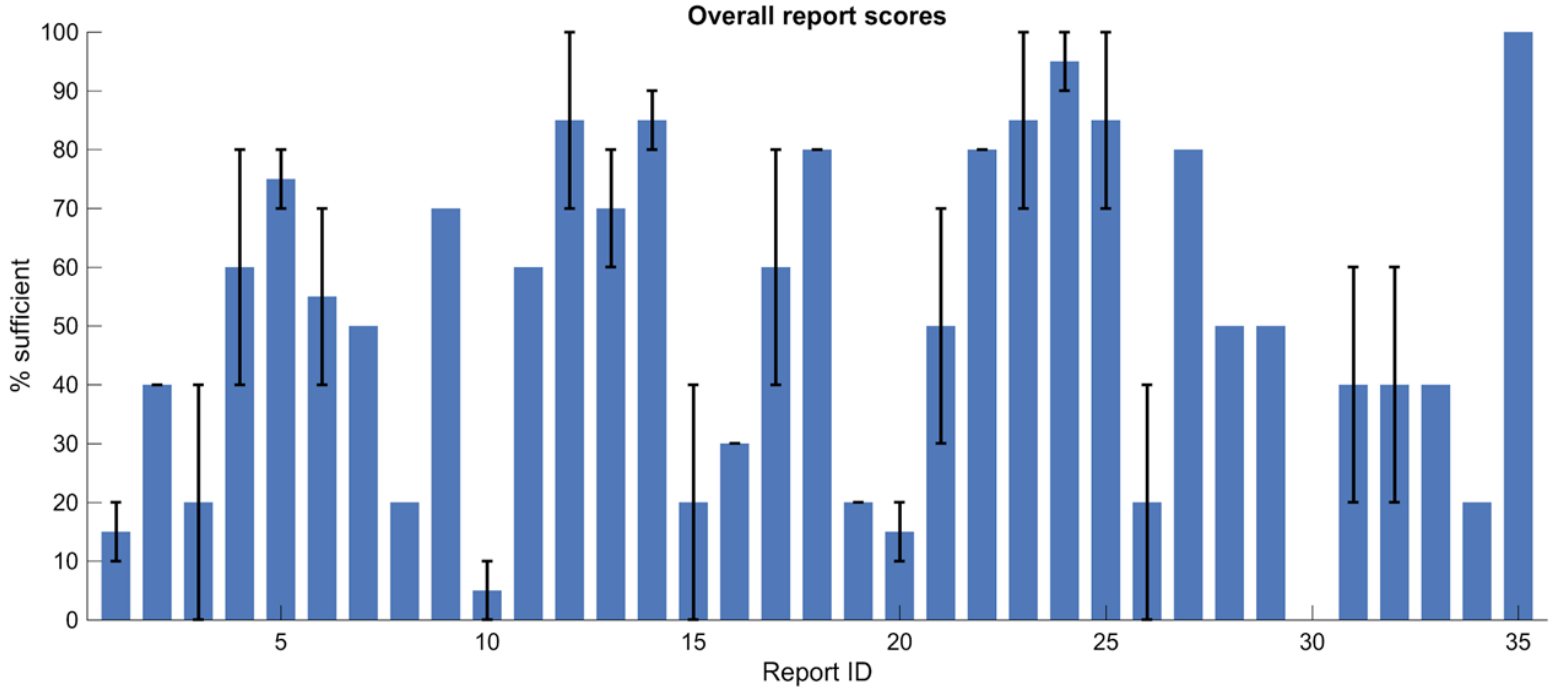
35 Credibility Plan Mid-Term Updates Submitted

100% of submissions with 1 review
85% of the submissions with 2 reviews

11/20/2018: Outbrief IMAG-MSM Steering Committee
12/18/2018: All feedback posted to IMAG Wiki

Scores: % of Rules Marked as “Sufficient”

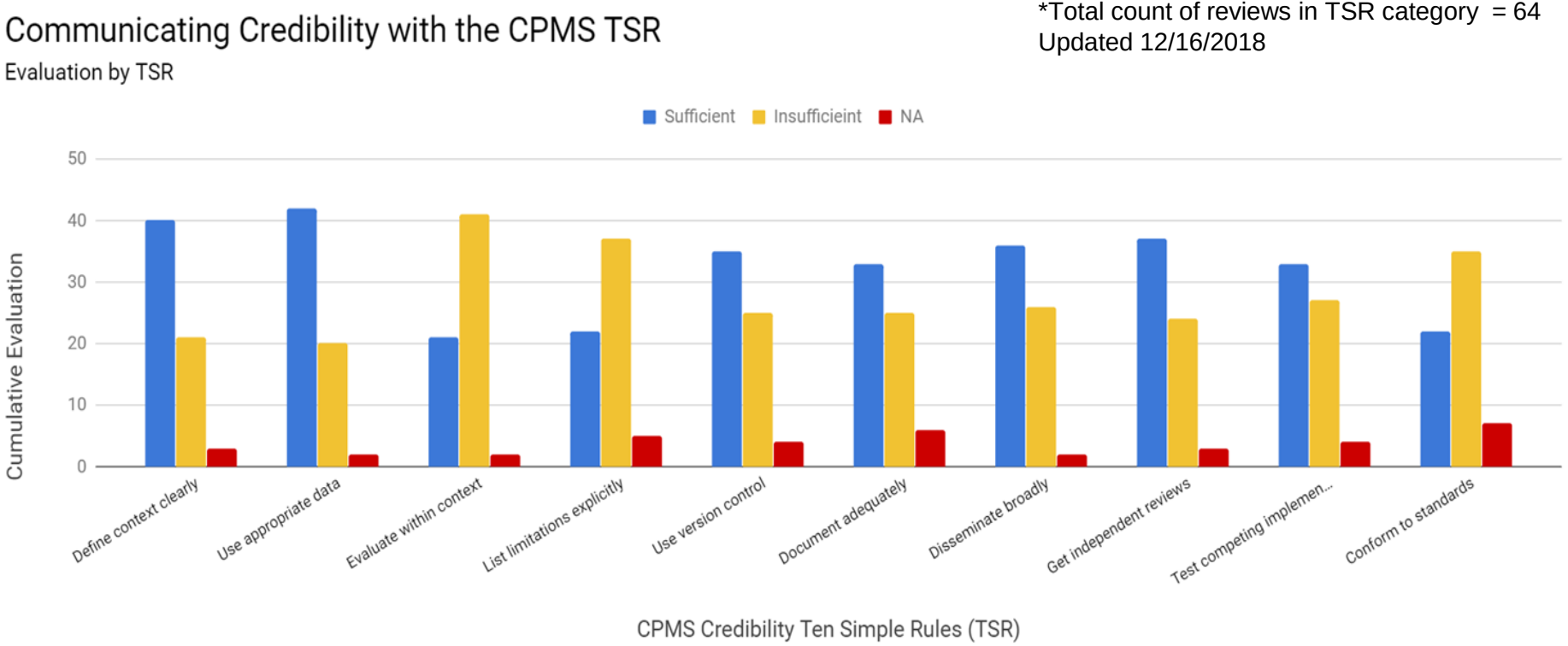
Total count of reviews in TSR category = 60



Average % sufficient scores, error bars denoting the full range

Final Scores by TSR

*Total count of reviews in TSR category = 64
Updated 12/16/2018



Distribution of Scores Based on Reviewer Agreement

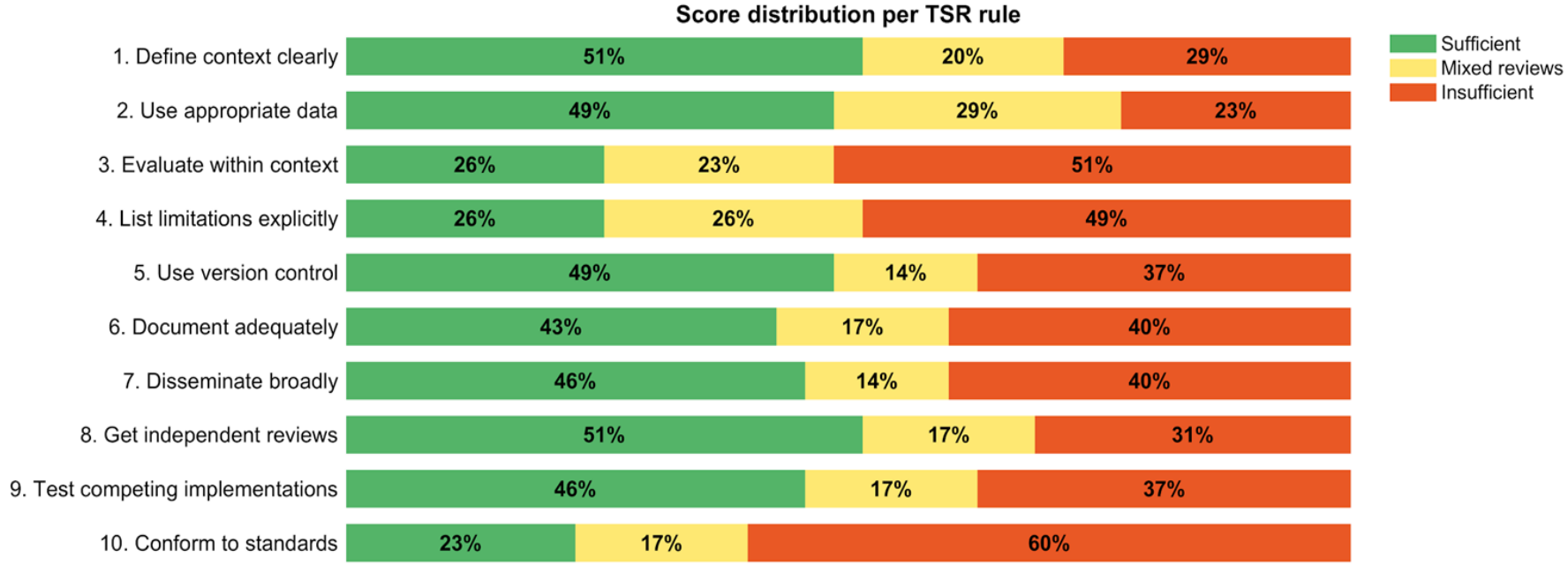


Chart illustrates agreement between reviewers in evaluating the TSR credibility category.
Yellow indicates the 2 reviewers split on evaluation

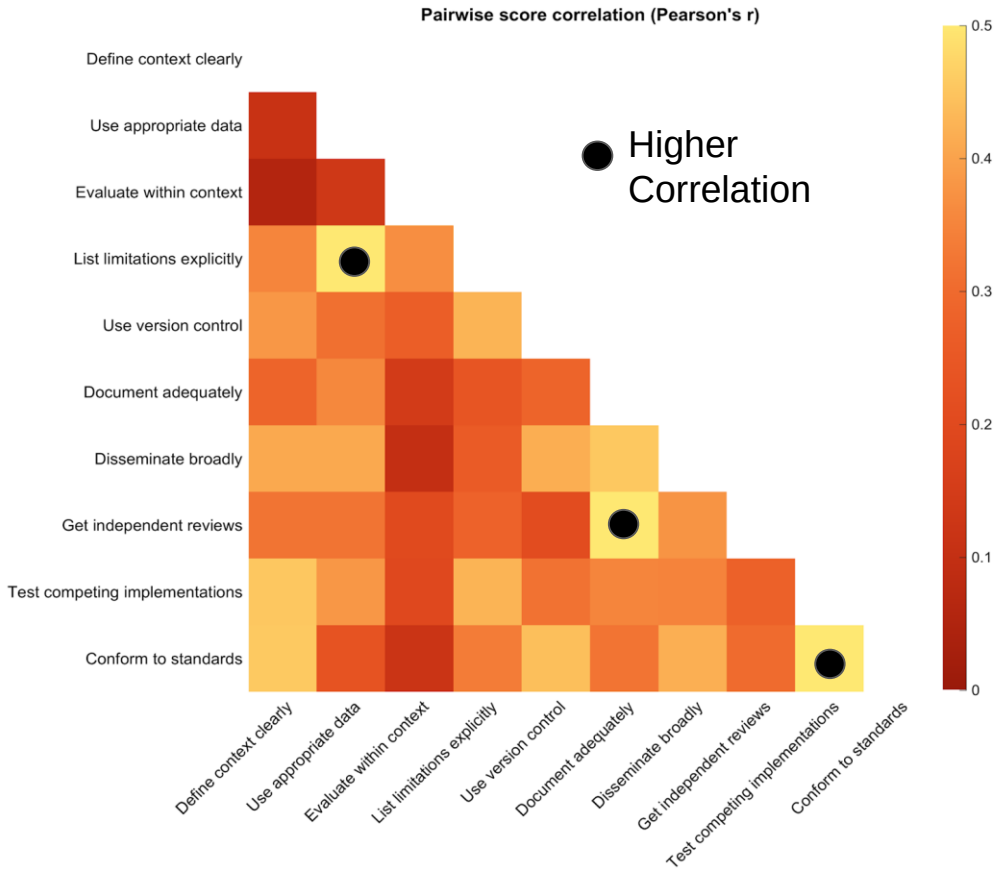
Cross-Correlation between TSR

Correlated TSRs

List Limitations Explicitly /
Use Appropriate Data

Get Independent Reviews /
Document Adequately

Conform to Standards /
Test Competing Implementations



Reviewer Observations

- Appreciative and Impressed with each PI team
- Still areas where improvement can be achieved

- PI experienced challenges in providing detail
  - Project or plan implementation started recently
  - PI had difficulty with template and/or instructions
  - PIs reported on project progress, not model credibility activities
  - PIs had difficulty articulating planned versus completed activities

TSR Satisfactory Description

The more detailed the information provided in each TSR category, the more likely a reviewer would score a provided description as "***sufficient to allow credibility to be assessed***"

...although these were often accompanied by comments and caveats from the reviewers.

Reviewer Observations cont.

- PIs grappled with differences in reporting
  - limitations of the modeling and limitations of data acquisition
  - documentation of assumptions and key decisions, not just code and data
  - conforming to standards vs. internal best practices
  - evaluate within context (V, V, & UQ) and test competing implementations
  - how some rules apply to their projects

Clarification - Its okay to suggest that a rule does not apply or that a particular credibility activity is not being pursued as long as evidence supports the position.

Recommendations to IMAG-MSM Community

Provide PI guidance / opportunities on approaches to “Documentation” and “Dissemination”

- Several PIs illustrated detailed documentation and dissemination approaches
- Many PIs rely on peer publications to meet documentation and dissemination
- Others described a ”post it online” approach, without discussing curation

CPMS concern: Reliance on peer publication or uncurated posting as the primary means of providing evidence in these categories may be unsatisfactory to the user community

Recommendations to IMAG-MSM Community

Define criteria and facilitate opportunity for implementing independent reviewer processes and what exactly is acceptable to sufficiently "include the user community"

PIs described a range of approaches

- Teaming of independent research groups, providing continuous reviews
- Seeking internal reviews within one's organization but not on the project
- Hiring external consultants to provide the review
- Journal publication peer review

CPMS Concern: If not addressed early in grant it may become unachievable at a level commensurate with community's credibility expectations.

Recommendations to CPMS

Implement process improvements of mid-term credibility review exercise

- Standardize input format to improve ability to provide applicable information
- Improve directions on breadth and depth of requested information
- Enhance definition of CPMS-TSR, to improve relatability and remove jargon
- Provide examples communicating model credibility information in each TSR
- Develop and implement TSR rubrics estimating TSR compliance level

Target to have all these by 9/2019, including a TSR Papers - outlining the TSR development and implementation process

IMAG – MSM 2019 One – on – One Session
Individual Consults and Some group led discussions

Acknowledgements: CPMS Reviewers

A. Drach, Ph.D.
Co-Chair
U. of Texas

J. Myers, Ph.D.
Co-Chair
NASA

A. Erdemir, Ph.D.
Past Co-Chair
Cleveland Clinic

L. Mulugeta, M.Sc.
Past Co-Chair
InSilico Labs LLC

W. Lytton, M.D.
Kings County Hosp.
Downstate Med. Cntr

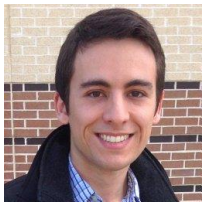
M. Horner, Ph.D.
ANSYS, Inc.

T. Morrison, Ph.D.
FDA

J. Ku, Ph.D.
Stanford U.

R. Vadigepalli, Ph.D.
Thomas Jefferson U.

Invaluable contributions to meta-analysis of all collected data

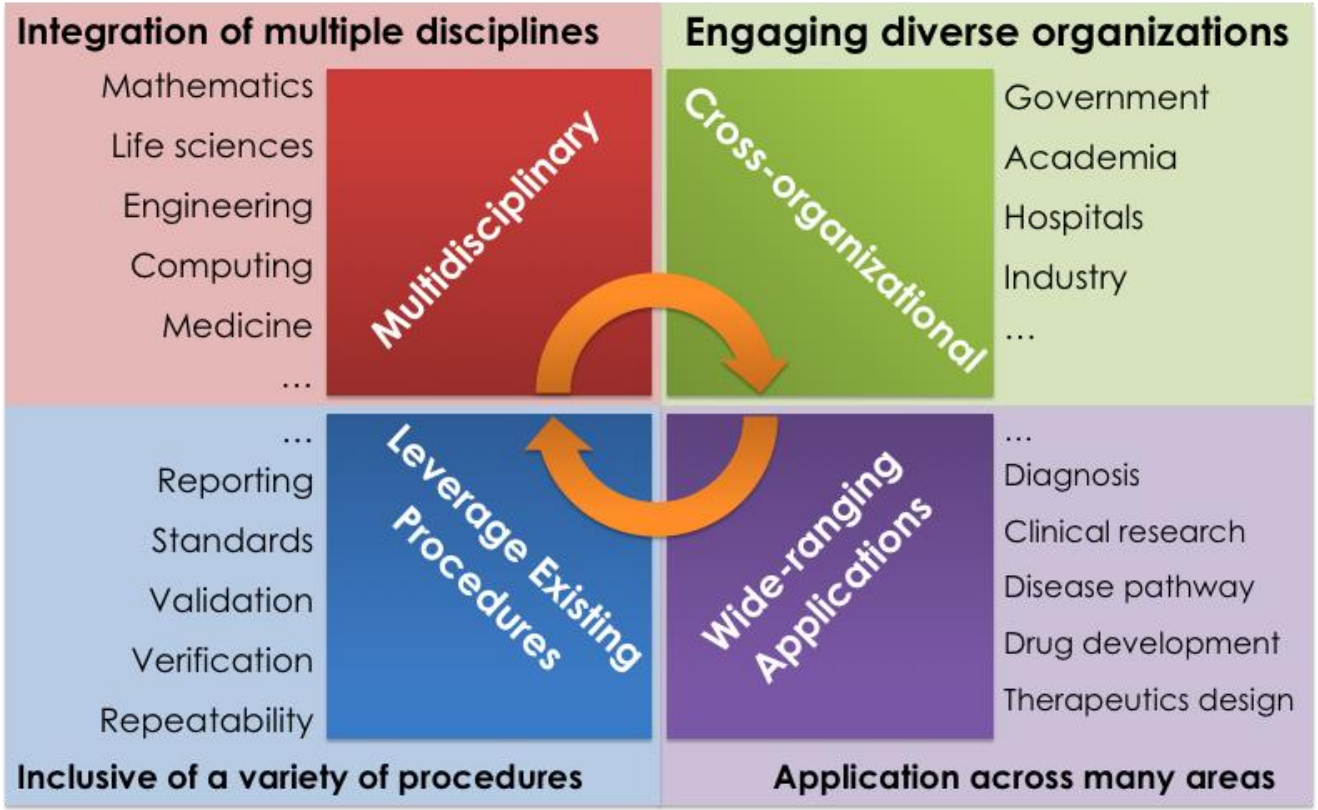


Bruno Rego, PhD candidate
University of Texas at Austin

Amir Khalighi, PhD
University of Texas at Austin

QUESTIONS?

Feedback can be found at https://www.imagwiki.nibib.nih.gov/content/multiscale-modeling-u01-projects



CPMS TSR Rubric - Draft

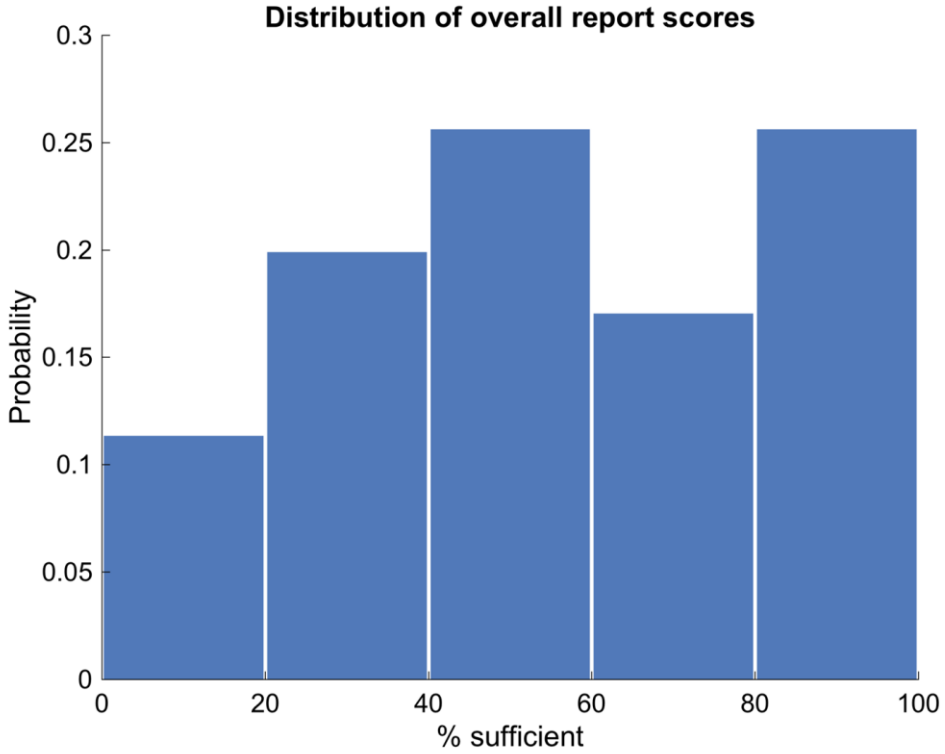
Outreach Capability

Outreach to Application-domain experts that may not be M&S practitioners	Outreach to M&S practitioners that may not be Application-domain experts	Outreach to Application-domain specific M&S practitioners	Outreach to Application-domain specific M&S practitioners	None/Too little
Comprehensive	Extensive	Adequate	Partial	Insufficient

Compliance Level

Note: Specific interpretation being tailored to each TSR rule

Distribution of “Sufficient” Rate Scores



PI Observations

- “***Model credibility is best evaluated by the unbiased user who needs the information coming out of the model the most. If there is a way to identify these people during the model building process and solicit their feedback more regularly, that would be very beneficial to ensuring model credibility.***”
- There seems to be gaps between the standard developed and that of medical doctors for clinical applications.
- ***Involve MDs to bridge the gaps between modelers and MDs***
- Face-to-face visit beyond annual IMAG meeting