



NASA-STD-7009 Guidance Document for Human Health and Performance Models and Simulations

Marlei Walton¹, Lealem Mulugeta², Emily Nelson³, and Jerry Myers³

¹Wyle Science, Technology, and Engineering Group, Houston, TX, ²Universities Space Research Association, Division of Space Life Sciences, Houston, TX, ³NASA Glenn Research Center, Cleveland, OH

INTRODUCTION

Rigorous verification, validation, and credibility (VV&C) processes are imperative to ensure that models and simulations (M&S) are sufficiently reliable to address issues within their intended scope. The NASA standard for M&S, NASA-STD-7009, was a resultant outcome of the Columbia Accident Investigation Board to ensure M&S are developed, applied, and interpreted appropriately for making decisions that may impact crew or mission safety. The NASA-STD-7009 Guidance Document is being developed to augment the governing standard and handbook to provide information, tools, and techniques applicable to the probabilistic and deterministic biological M&S more prevalent in human health and performance (HHP) and space biomedical research and operations.

PHILOSOPHY

Inherent in this guidance is the understanding that the application of many of these human health and systems M&S is to provide insight and information to areas where such information is lacking, versus for design purposes. The key is that a) all parameters may not be known a priori, and b) the fundamental relationships between and among parameters may not be known. Thus in many cases, the M&S are truly research efforts just to generate one simulation. This lack of specificity in the M&S is not a reason for the developer or customer to reduce the rigor in assessing model credibility. Quite the opposite—the more the models “are plastic”, the more rigor the developer must take and the customer must expect in order to adequately quantify the understanding of model output application. By communicating a complete understanding as possible of the model's effective abstraction of the real world human health system, including level of validation and parameter sensitivities, the model becomes credible to the decision maker and an integral part of their decision making process.

DISTINCTIVE FEATURES

Three areas of the NASA-STD-7009 Guidance Document that we consider unique from the governing standard are:

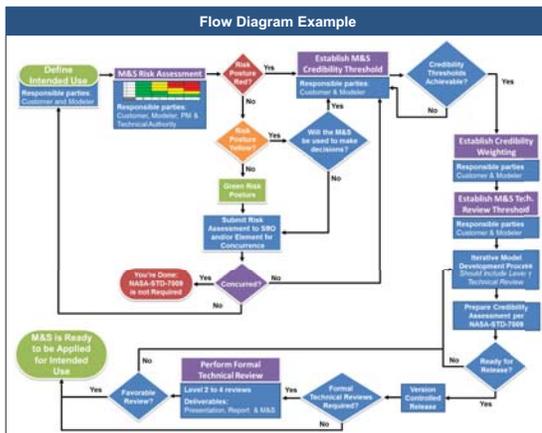
1. Credibility assessment weighting factors- different for probabilistic and deterministic models
2. Criteria for technical review- including not only the details of technical review but also who should be involved at each level
3. Personnel roles and responsibilities

REFERENCES

[1] NASA-STD-7009: Standard for Models and Simulations, 2008. Washington, DC: NASA.

RESULTS

APPLICATION



External Community

Results of the Interagency Modeling and Analysis Group (IMAG) discussion panel and NASA's approach to credibility assessment motivated the establishment of the "Committee for Developing Credible Multiscale Models for Healthcare".

The Academy of Science is currently holding a series of meetings on the V&V and Uncertainty Quantification of complex models and NASA has been asked to contribute to the biomedical modeling portion.



As a direct consequence of a presentation given NIH/IMAG regarding how NASA uses the NASA-STD-7009 to vet biomedical models, the Food and Drug Administration is heavily leveraging 7009 to develop a new standard for "Verification and Validation of Computational Modeling of Medical Devices".

The FDA regularly consults with HRP modeling teams in the development of this new standard and NASA has a presence on the ASME V&V40 Sub-committee that is working with the FDA to develop the standard for "Verification and Validation of Computational Modeling of Medical Devices".



Getting It Right: Better Validation Key to Progress in Biomedical Computing - Bringing models closer to reality

Groundwork laid by the Digital Astronaut Project and Integrated Medical Model was featured in the 2012 fall issue (10/19/12) of the Biomedical Computation Review magazine and lauded as a "Comprehensive Validation" method.

DISTINCTIVE FEATURES

Credibility Assessment Weighting Factors

Credibility Assessment Factors	Evidence Score	Evidence Weight	Evidence Threshold	Tech. Review Score	Tech. Review Weight	Tech. Review Threshold	Factor Score	Overall Score	Sufficiency Threshold
1 Verification	2	0.20	3	2	3	2	0.40	1.75	2.54
2 Validation	2	0.25	2	2	3	2	0.50		
3 Input Pedigree	2	0.10	3	2	3	2	0.20		
4 Results Uncertainty	0	0.10	2	0	3	0	0.00		
5 Results Robustness	2	0.10	2	2	3	2	0.20		
6 Use History	1	0.15	2	N/A	N/A	2	0.15		
7 M&S Management	2	0.05	3	N/A	N/A	2	0.10		
8 People Qual.	4	0.05	3	N/A	N/A	3	0.20		

Overall Credibility

$$= \sum_{i=1}^5 (\text{Evidence Weight} \times \text{CAS_Score}_i + \text{Tech. Review Weight} \times \text{CAS_Score}_i) \times \text{CAS_Weight}_i + \sum_{i=6}^8 \text{CAS_Score}_i \times \text{CAS_Weight}_i$$

To prevent artificial gain to the contribution of overall score, if the defined threshold is surpassed, the combined sufficiency threshold should increase proportionately.

Factor Weight (Proposed)	Deterministic	Probabilistic
1 Verification	0.2	0.075
2 Validation	0.25	0.15
3 Input Pedigree	0.1	0.275
4 Results Uncertainty	0.1	0.2
5 Results Robustness	0.1	0.15
6 Use History	0.15	0.15
7 M&S Management	0.05	0.05
8 People Qualifications	0.05	0.05
TOTAL	1.0	1.0

0.05 < W < 0.25

Criteria for Technical Review

Level 4 – Favorable external peer review accompanied by independent factor evaluation.

Suggested process: Stand-up review with non-advocate committee accompanied with hands-on use and evaluation of the M&S by committee members using their own benchmarks to score the M&S performance within the intended use.

Pass criteria: A favorable review by at least 75% of the committee is required to achieve this level.

Level 3 – Favorable external peer review.

Suggested process: Stand-up review with non-advocate group accompanied with mechanism to gain hands-on insight of inner workings of M&S. May request to review the source code.

Pass criteria: A favorable review by at least 75% of the committee is required to achieve this level.

Level 2 – Favorable formal internal peer review.

Suggested process: Stand-up review internal review team to score model's performance accompanied with hands-on use and evaluation using their own benchmarks to score the M&S performance within the intended use. An independent factor rating is not required. May request to review the source code.

Pass criteria: A favorable review by at least 75% of the committee is required to achieve this level.

Level 1 – Favorable informal internal peer review.

Suggested process: Technical interchange meetings or document reviews at major mile-stones of the M&S phases as defined in the project schedule.

Pass criteria: A favorable review by at least 75% of the committee is required to achieve this level.

Level 0 – Insufficient evidence.

Personnel Roles and Responsibilities

Qualifications of the people involved in the development and implementation of the M&S should be evaluated based on two criteria:

1. What is the primary expertise of the personnel based on their academic training and years of experience in the field?
2. How well do the personnel's academic and experience match with the task which they have been assigned within the M&S activity.

NASA-STD-7009 categories of personnel:

Developers – Establish the fundamental principles and mathematical abstractions of the model. Responsibility is scientific and technical application of various principles to provide a means of creating relevant simulations. Should have a strong background in fundamental and applied mathematics, physics and computational sciences. Responsible for credibility and validation of the model.

Operators – Execute the model to perform a simulation. Generally the least technical but most familiar with using the model.

Analysts – Define the initial conditions and boundaries of a simulation, and review and interpret the results of the simulation. Responsible for the credibility and validation of the simulations (not the model). Tend to be subject matter experts within the specific area which is being simulated.

A team member may hold more than one of these three roles within the M&S process. However, that individual's level of qualification to accomplish that task must be evaluated appropriately.

LESSONS LEARNED

- Establishing M&S credibility starts before model development.
- M&S credibility includes modeling team with end user and/or customer at all stages of development and implementation.
- Lack of specificity in the M&S is not a reason for the developer or customer to reduce the rigor in assessing model credibility.
- Successful M&S will have ongoing credibility assessments throughout life of model; it is a continuous process.