National Aeronautics and Space Administration

CAN SIMULATION CREDIBILITY BE IMPROVED USING SENSITIVITY ANALYSIS TO UNDERSTAND INPUT DATA EFFECTS ON MODEL OUTCOME?

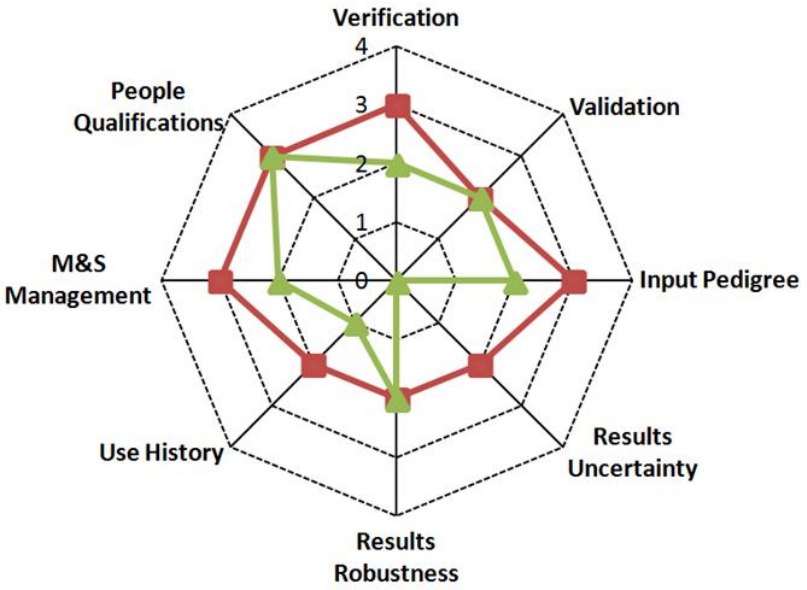
J. Myers[1], M. Young[2], D. Goodenow[1], A. Keenan[2], M. Walton[2], L. Boley[2]

1. NASA Glenn Research Center
2. Wyle Science, Technology & Engineering Group

Its Really About Model *Credibility*!
Achieving a high level of belief or trust in the model

- **NASA-STD-7009**
  - Standard for Models and Simulations (M&S)
- **Eight Factors**
  - Verification
  - Validation
  - ***Input Pedigree***
  - Results Uncertainty
  - Results Robustness
  - Use History
  - M&S Management
  - People Qualifications

Generic STD-7009 Credibility Assessment



Credibility Assessment Factors		Evidence			Technical Review		Factor Score	Weighted Subfactor Score	Overall Score	Sufficiency Threshold
		Score*	Weight+	Threshold*	Score*	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	1	0.15		
7	M&S Management	2	0.05	3	N/A	N/A	2	0.10		
8	People Qualifications	4	0.05	3	N/A	N/A	4	0.20		

* Maximum = 4; where 0=insufficient evidence and 4=highest fidelity/rigor achievable

+ Minimum = 0.05, maximum = 0.25 and sum of all weights must equal 1.0

Problem with "Input Pedigree" Ranking

- **A complex model integrating many factors can have a wide array of input pedigree scores across the range of model parameters**
  - Source, quality of the data, type of data

- **From NASA-STD-7009: Input Pedigree (Ranking) involves the evaluation of *all data that is used as input for the current M&S results*. It includes not only data that is unique to the model, but also data that is produced by other simulations.**

- **Implication: Model and Simulation Input Pedigree Ranking is equivalent to the *lowest input pedigree* of the model parameters (i.e. logical AND).**

Question:

- **In practice, the Input Pedigree ranking of complex multi-parameter models tends to be low, resulting in lower than expected model and simulation credibility.**

- **Question: Is this evaluation of credibility true for all applications and simulations of the model?**

- **Considerations:**
  - Application credibility differs from the model credibility with respect to input pedigree.
  - Application credibility can be more precisely determined using Sensitivity analysis (also known as Results Robustness) to address relative importance of input parameters.

Sensitivity Analysis Methodology

- **Saltelli: "Sensitivity Analysis is the study of how variation in the output of a model can be apportioned, qualitatively or quantitatively, to different sources of variation (input) and how the given model depends upon the information fed into it."**

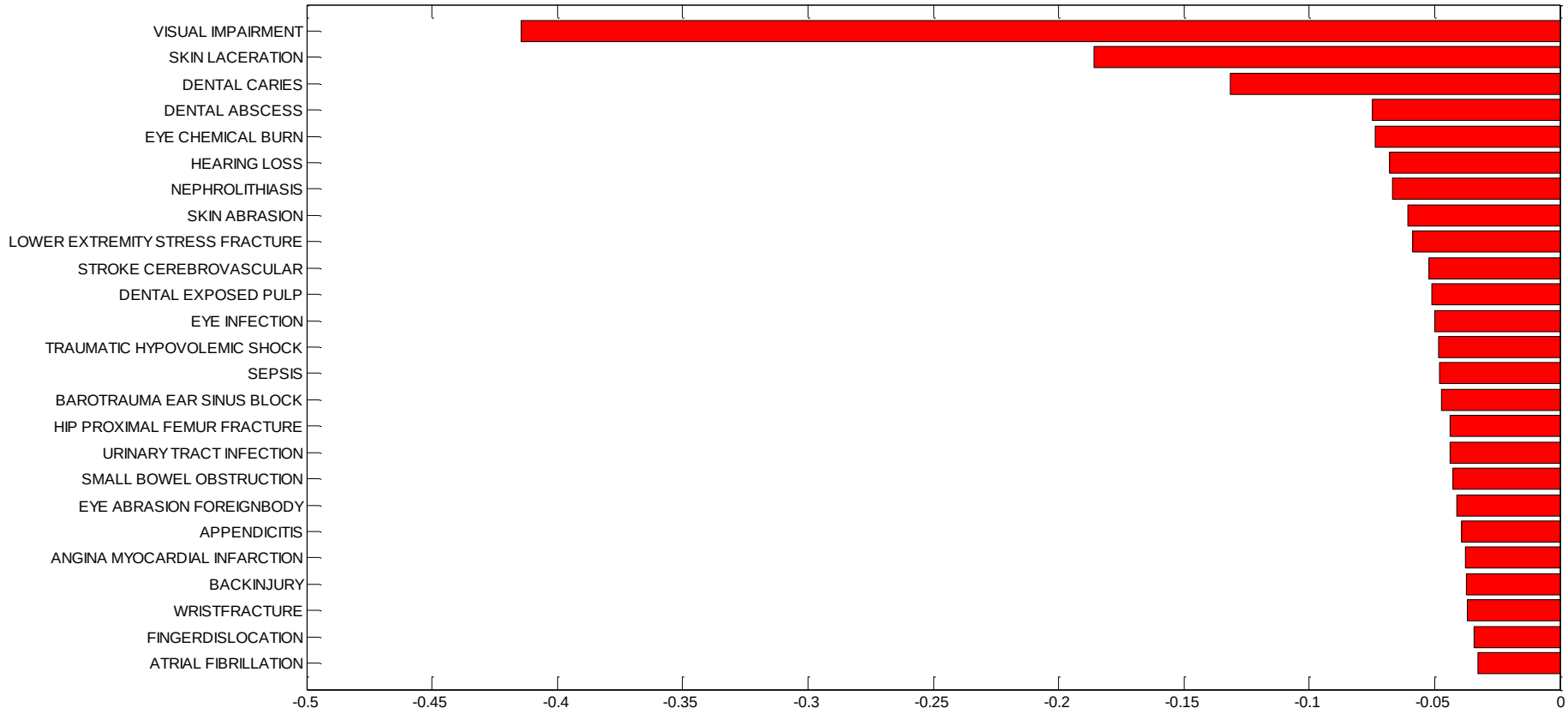
- **Partial Rank Correlation Coefficient (PRCC) Analysis**
  - Provides the linear relationships between two variables (one input parameter and one output parameter) when all linear effects of other variables are removed after rank transformation.
  - Rank Transformation: transforms non-linear monotonic relations to linear.

- ***KEEP IN MIND*** **the difference in a contributing parameter and a sensitive parameter**
  - Many parameters contribute substantially to the mean output of the model
    - Low sensitivity may indicate a "DC-signal effect" over the range of model application and parameter variance
  - Parameter variance affecting model output (magnitude and variance) indicates a sensitive parameter

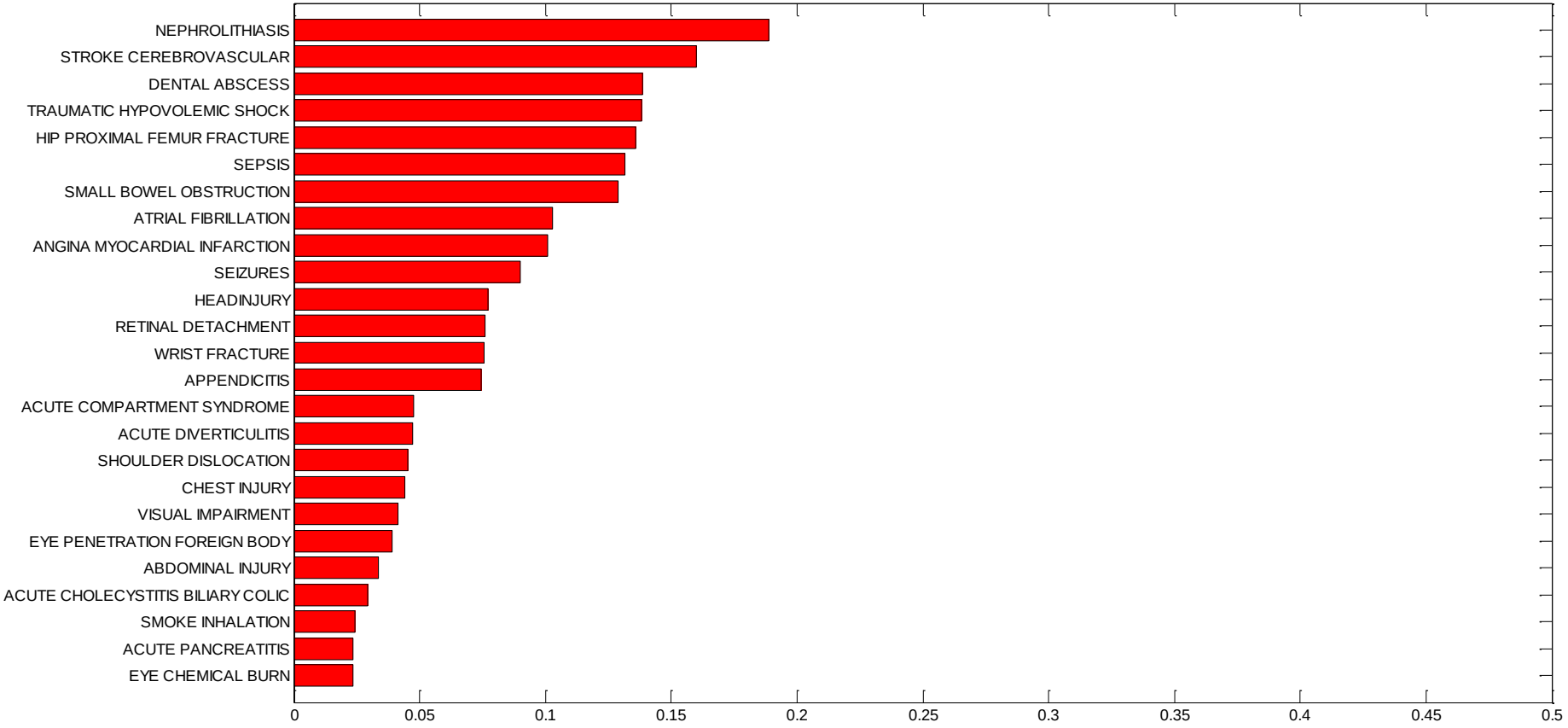
Using IMM for ISS mission 6 crew members

- **IMM Provides probabilistic analysis of 99 medical condition occurrences and impact to mission outcomes**
- **Context : 6-month, 6 crew (M-4, F-2), ISS med kit, and evacuation is possible**
- **Output:**
  - Crew Health Index (crew available time – time lost due to medical events)
  - Probability of consideration of evacuation

Sensitivity Analysis Results: Crew Health Index for 6 Month Mission to the International Space Station



Sensitivity Analysis Results: Consideration of Evacuation for 6 Month Mission to the International Space Station



Interpretation: Weighted Averaging

- Weighted averaging of the sensitivity scores uses the PRCC output (W_{PRCC})to "weigh" parameter pedigree (IP_i)
- For this analysis assume
  - One input parameter ranked at 1, the rest are 3
  - Originally a strict "Std-7009" Input Pedigree Score = 1

$$IP_{total} = \frac{\sum W_{PRCC,i}\, IP_i}{\sum W_{PRCC,i}}$$

Eye Chemical Burn - IP_{total}	100 Conditions	25 Most Influential Conditions
CHI	2.944	2.918
Consideration of Evacuation	2.772	2.935

VIIP - IP_{total}	99 Conditions	25 Most Influential Conditions
CHI	2.963	2.959
Consideration of Evacuation	2.691	2.545

Interpretation: Threshold Specification

- **Set a threshold value below which the contribution of the parameter is considered to be small**
  - Pedigree of data below this level is not considered in the assessment of simulation input pedigree
- **Assume conservative threshold to be a PRCC =1/n = 0.01 and a more lenient threshold to be PRCC = 2/n = 0.02**
  - n = the number of parameters
- **For this analysis assume**
  - One input parameter ranked at 1, the rest are 3
  - Strict “Std-7009” Input Pedigree Score = 1

Threshold Analysis Results

Conservative Threshold +/-0.01

CHI

	W_{PRCC}	IP - Con	IP - LCon
VIIP	-0.41	1	1
Eye Chem Burn	-0.07	1	1
Dental Carries	-0.13	1	1

Less Conservative Threshold +/-0.02

Consideration of Evacuation

	W_{PRCC}	IP - Con	IP - LCon
VIIP	0.04	1	1
Eye Chem Burn	0.02	1	3
Dental Carries	0.002	3	3

Conclusions

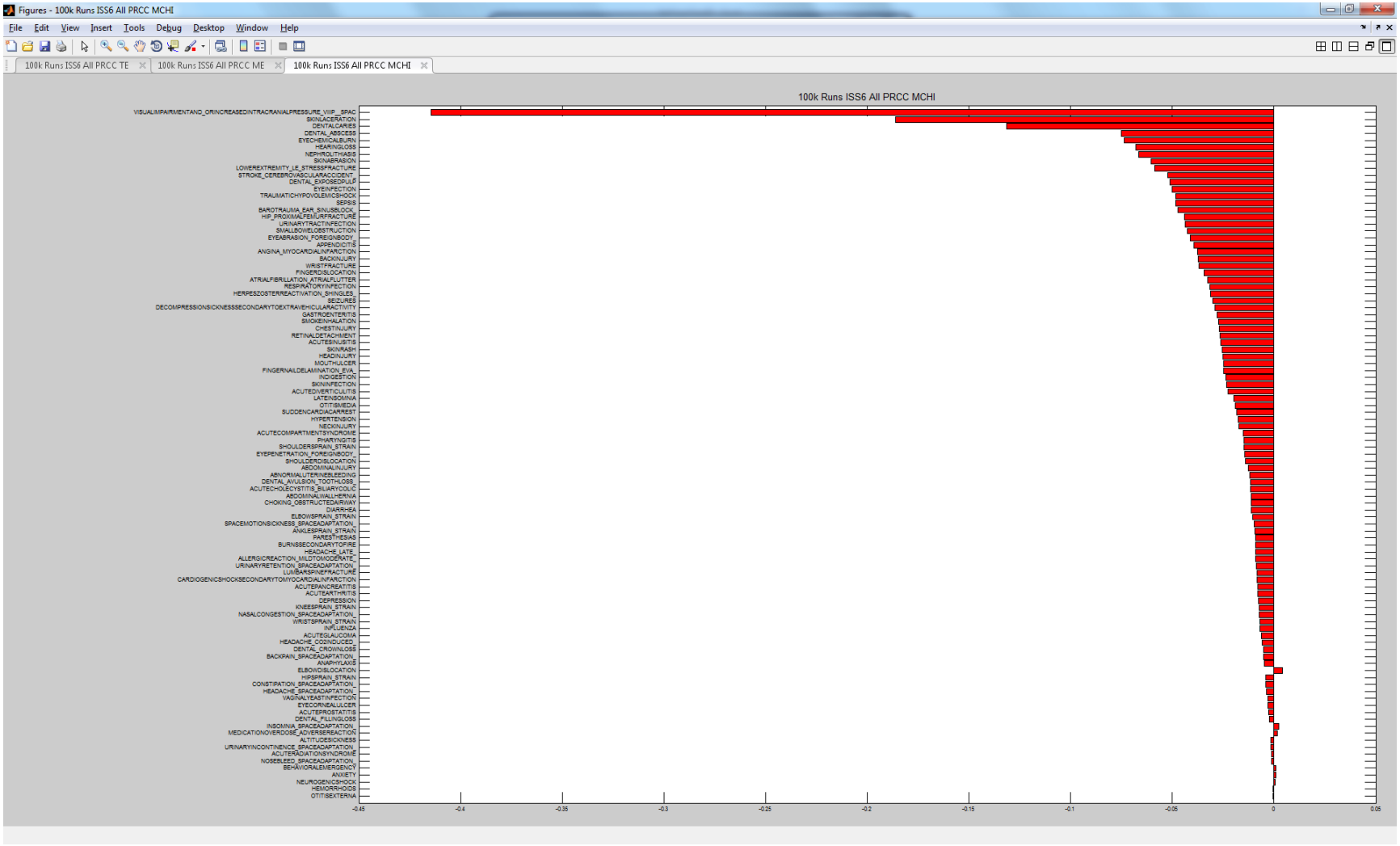
- **Proof of concept analysis indicates**
  - Sensitivity analysis can be used to better characterize the input pedigree ranking of specific simulations
  - One analysis is insufficient to characterize the input pedigree ranking for all possible analyses – Case by case basis
- **Method used to evaluate input pedigree requires further investigation**
  - Current sample approaches are reasonable but not optimized
  - Quantification should be repeatable and consistent with other communications components of 7009 to maintain decision making information quality
- **Future efforts will address**
  - More appropriate definitions of cutoff thresholds, allowing a distinction in numerical noise and statistical significance

INTEGRATED MEDICAL MODEL OVERVIEW

Thank you!

Questions?

CHI – All Conditions



Consideration of Evacuation – All Conditions