National Aeronautics and Space Administration inistration

Integrated Medical Model Verification, Validation, and Cre ibility

Marlei Walton[1], Eric Kerstman[2], Millennia oy[1], Ronak Shah[2], Lynn Saile[1], Lynn Boley[1], D ug Butler[1], Jerry Myers[3]

1. Wyle Science, Technology & Engineering Group
2. University of Texas Medical Branch, Galveston, TX
3. NASA Glenn Research Center

Presentation Overview

- Important thoughts before model development
- Beginning the verification, validation, and credibility (VV&C) journey
- Defining Integrated Medical Model (IMM) framework
- IMM Compliance Matrix
- IMM Verification and Validation (V&V)
- IMM Credibility Scoring
- Summary

Thoughts Before Model Development

"Essentially, all models are wrong, but some are useful."

George Box (1987); Professor Emeritus of Statistics at the University of Wisconsin

- **Establishing model and simulation (M&S) credibility starts before model development**
- **M&S credibility includes modeling team with end user and/or customer**
- **Successful M&S will have ongoing credibility assessments throughout life of model**

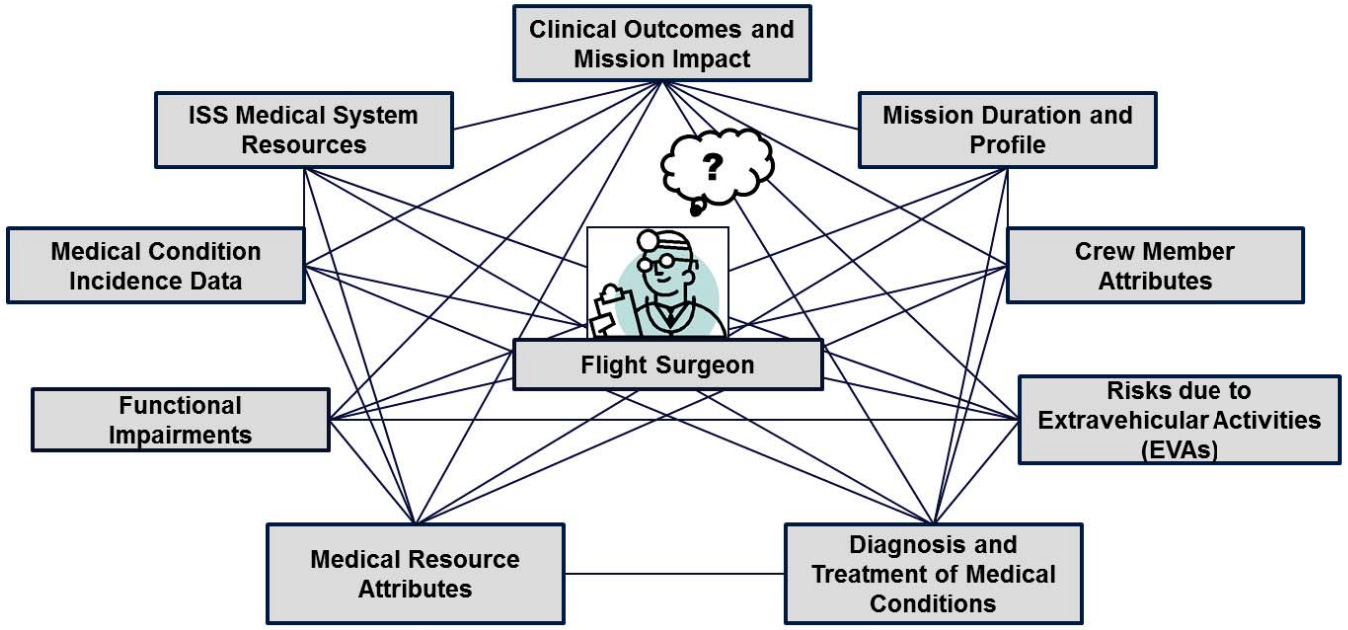
Guiding VV&C Document Chronology

- Recommended practices references
- IMM Verification and Validation Plan (2008)
- NASA-STD-7009 Standard for Models and Simulations (2008/2013)
- NASA-HDBK-7009 Implementation Guide for NASA-STD-7009 (2013)
- NASA-STD-7009 Guidance Document for Human Health and Systems Models and Simulations (estimated 2014)

Beginning the VV&C Journey

Risk of unacceptable health and mission outcomes due to limitations of in-flight medical capabilities.

What is the likelihood of a medical evacuation?	What is the risk of Loss of Crew Life due to illness on ISS?	What medical devices should we have on ISS?	What should be in the Exploration Medical Kit?

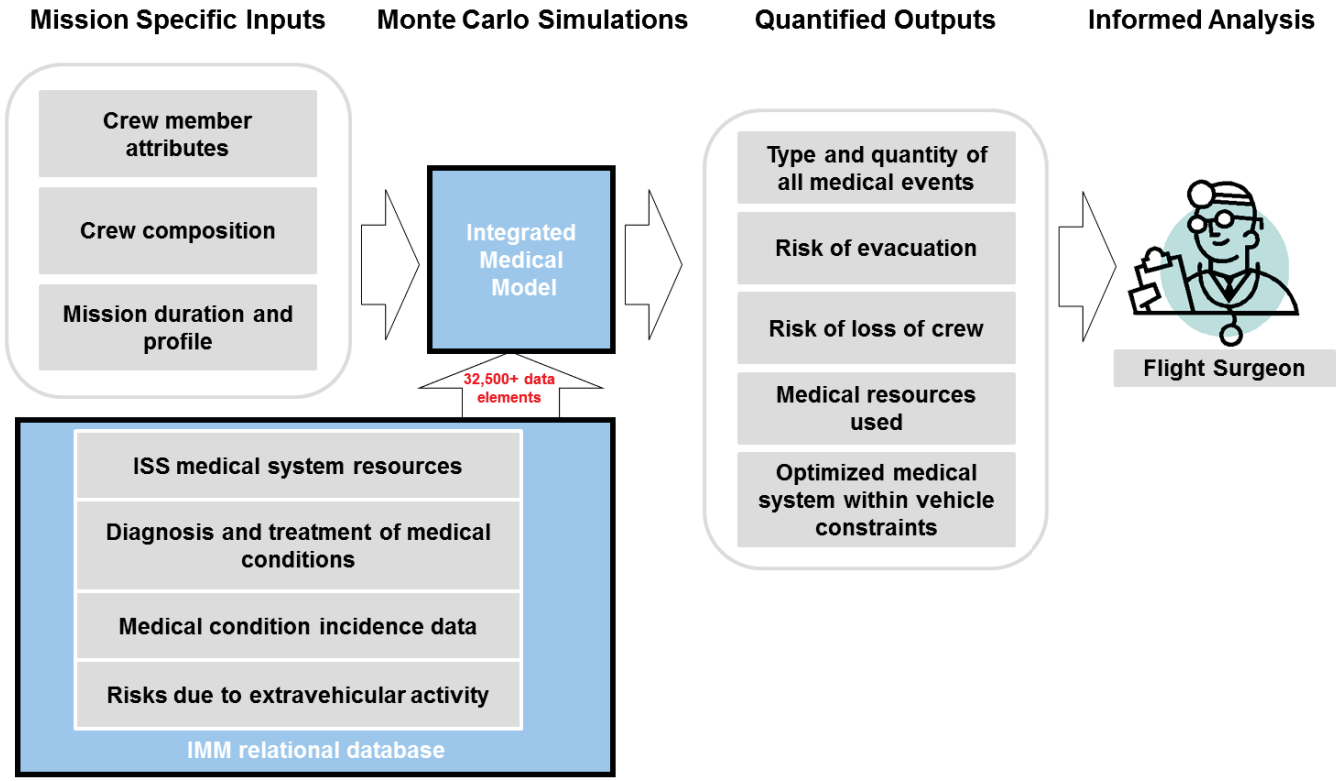


Goal: Develop an integrated, quantified, evidence-based probabilistic risk forecasting model to be used as a decision support tool for NASA crew health and mission planners that can help guide exploration planning, requirements development, and R&D technology investment activities.

Defining IMM Framework

Real world system (RWS): the most likely health state of the astronaut crew with a defined set of medical resources used to diagnose and treat medical events, and most likely mission impacts due to medical causes

Approach: Employ best-evidence clinical research methods, probabilistic risk assessment (PRA) techniques using ISS as a precursor to exploration missions



ISS baseline: RWS in the context of a microgravity environment in low earth orbit

IMM 7009 Compliance Matrix Summary (IMM V3 SAS)

49 Requirements

Compliant	Compliance Incomplete	Not Applicable
37	9	3

Compliant (7009 requirements section)	**Areas of incomplete compliance** (7009 requirements (reqs) section- # of reqs in given section out of total # of compliance incomplete reqs)
Programmatics (4.1)	Verification, validation, and uncertainty (4.4- 7 of 9 reqs)
Models (4.2)	Training (4.6- 2 of 9 reqs)
Simulations and analyses (4.3)	
Assessing the credibility of M&S results (4.7)	
Reporting results to decision makers (4.8)	

IMM V&V Plan Summary (IMM V3 SAS)

IMM Element	Internal Review	% Complete	External Review	% Complete
IMM list of medical conditions	verification validation	100 100	NASA non-NASA	60 0
CliFF V&V process	verification validation	80 80	NASA or non-NASA	2
CliFF content	verification validation	98 98	NASA non-NASA	3* 0
External modules (conditional)	verification validation	50 50	NASA or non-NASA	50
IMM overall approach	verification validation	80 80	NASA non-NASA	80 0
Model programming	verification validation	100 95	non-NASA	0
Software and code documentation	verification validation	90 80	non-NASA	0
IMM database process	verification validation	20 20	non-NASA	0
IMM output components	verification validation	94 26	NASA	5
IMM output results	validation	20	NASA	5

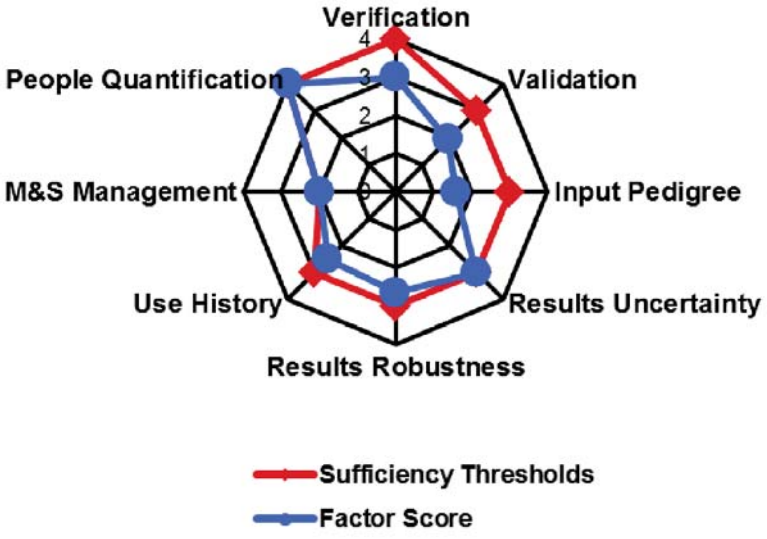
*based on completion of 100 CliFFs

IMM Credibility Scoring (IMM V3 SAS)

	Credibility Assessment Factors	Evidence			Technical Review		Factor Score	Weighted Subfactor Score	Overall Score	Sufficiency Threshold
		Score*	Weight+	Threshold*	Score*	Threshold*				
1	Verification	3	0.075	4	3	3	3	0.225	2.47	3.075
2	Validation	1.5	0.15	3	3	3	1.95	0.2925		
3	Input Pedigree	1	0.175	3	3	3	1.6	0.28		
4	Results Uncertainty	3	0.2	3	3	3	3	0.6		
5	Results Robustness	2.5	0.15	3	3	3	2.65	0.3975		
6	Use History	2.5	0.15	3	N/A	N/A	2.5	0.375		
7	M&S Management	2	0.05	2	N/A	N/A	2	0.1		
8	People Qualifications	4	0.05	4	N/A	N/A	4	0.2		

* 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



Subfactors	Weight
Evidence	0.7
Technical Review	0.3

Legend	
	CAS Score > Threshold *Exceeds credibility requirements*
	Threshold ≥ CAS Score ≥ (Threshold-0.5) *Ready for use*
	(Threshold-0.5) > CAS Score ≥ (Threshold-1.0) *Use with caution*
	CAS Score < (Threshold-1.0) *Use not recommended or to be used with EXTREME CAUTION by subject matter experts only*

Summary

- **Early integration of credibility assessment as part of the IMM development and implementation process yielded:**
  - Confidence in the IMM for customer and end user decision makers
  - Significant risk forecast contributions to operations, science and technology planning, and exploration research
    - ISS Program acceptance
    - Transition to operational model and simulation tool
    - Completion of service requests from broad end user consortium
- **Reporting tools have evolved over the lifetime of the IMM project to better communicate VV&C status**
  - Comprehensive VV&C approach combining IMM V&V Plan with NASA-STD-7009 requirements
  - Important to appropriately weight the different credibility assessment factors for the problem of interest
- **The greater medical community recognizes the importance of rigorously vetting computational models and looks to NASA for leadership**

Getting It Right: Better Validation Key to Progress in Biomedical Computing - Bringing models closer to reality - *10/19/12*

Questions?

For more information about how to apply the NASA-STD-7009 to models and simulations: “How to Develop and Interpret Credibility Assessments of Numerical Models for Human Research: NASA-STD-7009 Demystified” Thursday, February 13 2:20pm in Iris/Tulip

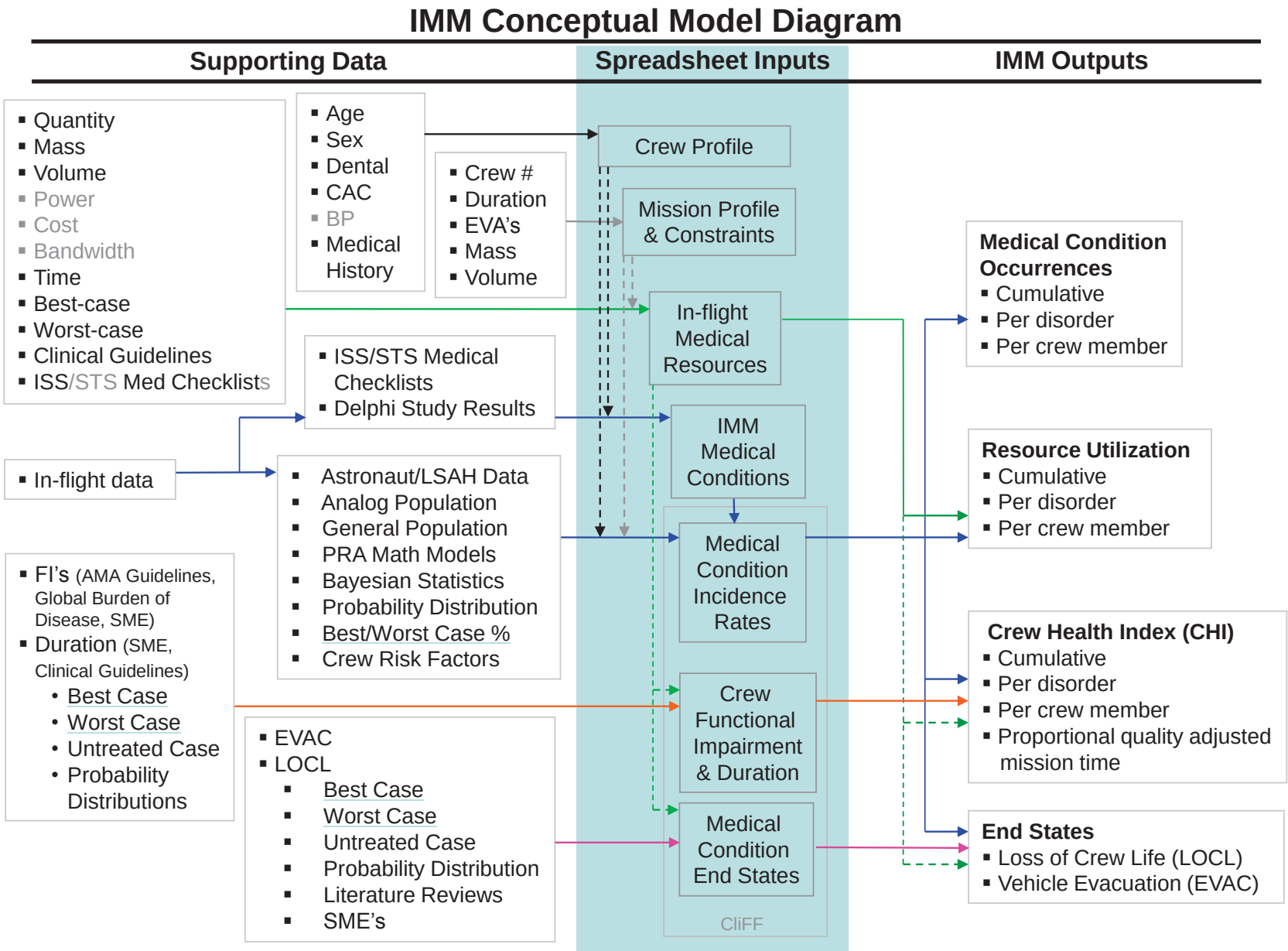
Back-up Slides

Comparison of IMM v3S and IMM v3M: ISS 6 Mission

ISS6 Mission – Comparison of main IMM output parameters

Primary Output Component	IMM V3S Mean/ Probability	IMM V3S 95% Confidence Interval Lower Bound	IMM V3S 95% Confidence Interval Upper Bound	IMM V3M Mean/ Probability	IMM V3M 95% Confidence Interval Lower Bound	IMM V3M 95% Confidence Interval Upper Bound
Total Medical Events	105.63	87	125	105.66	87	125
CHI*	91.01	76.12	97.44	91.55*	75.81	98.07
EVAC Probability	0.1123	0.1092	0.1154	0.1136	0.1107	0.1167
LOCL Probability	0.0059	0.0051	0.0066	0.0058	0.0051	0.0066
Treated CHI*	92.03	79.63	97.52	92.86*	80.25	98.15
Untreated CHI*	16.02	2.44	32.51	17.44*	3.39	34.32
Treated EVAC Probability	0.0473	0.0452	0.0494	0.0470	0.0450	0.0490
Treated LOCL Probability	0.0053	0.0046	0.0061	0.0048	0.0041	0.0054

IMM Conceptual Model (IMM V3 SAS)



IMM Background

- 100 specific medical conditions, including 10 space adaptation conditions
- Developed using International Space Station (ISS) based medical capability
- Assumes ISS operational environment
- Bounds clinical outcome uncertainty
  - Best-case scenario
  - Worst-case scenario
  - Untreated-case scenarios

IMM NASA-STD-7009 Risk Assessment

- IMM within scope of NASA-Standard-7009
- IMM is subject to 7009 compliance

Risk and Criticality Assessment Matrix for IMM

Modeling & Simulation Results Influence	5: Controlling				
	4: Significant			IMM	
	3: Moderate				
	2: Minor				
	1: Negligible				
		IV: Negligible	III: Moderate	II: Critical	I: Catastrophic
		Decision Consequence			

The broad applications of IMM contribute to the risk and criticality assessment of IMM v3.0.

IMM Core Questions

- **For a specified mission scenario:**
  - What medical conditions are most likely to occur?
  - What medical resources are most likely to be utilized?
  - What is the probability of evacuation or loss of crew life due to medical events?
  - What is the expected crew functional impairment and what is the uncertainty range of that impairment due to medical events?
  - What are the optimal medical resources to minimize crew functional impairment and the probability of evacuation or loss of crew life?

IMM Practical Application

- **Medical systems design & optimization (current)**
  - In-flight medical systems and electronic health records
  - Develop design handbooks relating risk-mass-volume per DRM
- **Managing science and technology portfolios (current)**
  - Use sensitivity analyses to align budgets with most influential risk drivers for an established set of design reference missions (DRM)
- **Communicating health risks to programs, international partners (current)**
  - Risk of evacuation, risk of loss of crew life, crew certification
- **Establish training priorities respective to mission (future)**
  - Use sensitivity analyses to align limited training hours with most influential risk drivers for a particular mission and crew
- **Requirements development (current)**
  - Use quantifiable IMM risk analyses as rationale to requirements
- **Knowledge management (current)**
  - Use IMM database and central library to manage (review, vet, update) organizational memory for every medical condition of interest

IMM Customer-defined 7009 Sufficiency Thresholds

Sufficiency Thresholds

Level	Verification	Validation	Input Pedigree	Results Uncertainty	Results Robustness	Use History	M&S Management	People Qualifications
4	Numerical errors small for all important features.	Results agree with real-world data.	Input data agree with real-world data.	Non-deterministic & numerical analysis.	Sensitivity known for most parameters; key sensitivities identified.	De facto standard.	Continual process improvement.	Extensive experience in and use of recommended practices for this particular M&S.
3	Formal numerical error estimation.	Results agree with experimental data for problems of interest.	Input data agree with experimental data for problems of interest.	Non-deterministic analysis.	Sensitivity known for many parameters.	Previous predictions were later validated by mission data.	Predictable process.	Advanced degree or extensive M&S experience, and recommended practice knowledge.
2	Unit and regression testing of key features.	Results agree with experimental data or other M&S on unit problems.	Input data traceable to formal documentation.	Deterministic analysis or expert opinion.	Sensitivity known for a few parameters.	Used before for critical decisions.	Established process.	Formal M&S training and experience, and recommended practice training.
1	Conceptual and mathematical models verified.	Conceptual and mathematical models agree with simple referents.	Input data traceable to informal documentation.	Qualitative estimates.	Qualitative estimates.	Passes simple tests.	Managed process.	Engineering or science degree.
0	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.
	M&S Development		M&S Operations			Supporting Evidence		

Medical Conditions in IMM by Category

Medical Illness

Acute Chest Pain/Angina
Acute Prostatitis
Allergic Reaction
Anaphylaxis
Anxiety
Appendicitis
Atrial Fibrillation
Back Pain (SAS)
Behavioral Emergency
Cardiogenic Shock
Choking (foreign body inhalation)
Constipation (SAS)
Cough (URI/Pneumonia)
Dental Abscess
Dental Crown Replacement
Dental Temporary Filling
Dental Toothache
Dental Tooth Loss
Depression
Diarrhea
Eye Abrasion
Eye Corneal Ulcer
Eye Infection
Gastroenteritis
Glaucoma
Headache (SAS)
Hemorrhoid
Indigestion
Insomnia (SAS)
Kidney Stones
Late Insomnia
Medication OD/Misuse
Mouth Ulcer/Cold Sore
Nasal Congestion (SAS)
Nosebleed (SAS)
Otitis Externa
Otitis Media
Pharyngitis
Seizures
Sepsis
Sinus Infection
Skin Infection
Skin Rash
Space Motion Sickness (SAS)
Stroke
Sudden Cardiac Arrest
Urinary Incontinence (SAS)
Urinary Retention (SAS)
Urinary Tract Infection
Vaginal Yeast Infection

SAS = space adaptation syndrome

Medical Conditions in IMM by Category

Injury/Trauma:

- Abdominal Injury
- Back Injury
- Chest Injury/Pneumothorax
- Eye Abrasion
- Eye Penetration
- Elbow Dislocation
- Finger Dislocation
- Fingernail Delamination (EVA)
- Head Injury (TBI)
- Hip/Proximal Femur Fracture
- Hypovolemic Shock
- Lumbar Spine Fracture
- Neck Injury
- Neurogenic Shock
- Paresthesias/Hot Spots (EVA)
- Shoulder Dislocation
- Skin Abrasion/Laceration
- Sprain/Strain
- Wrist Fracture

Environmental:

- Acute Radiation Sickness
- Altitude Sickness
- Barotrauma (ear/sinus block)
- Burns
- Decompression Sickness (EVA)
- Eye Chemical Burn
- Headache (CO_2 induced)
- Smoke Inhalation
- Toxic Exposure

Likelihood: IMM Evidence Base

- **Only U.S. medical in-flight data used**
- **Astronaut health information**
  - published reference sources
  - lifetime surveillance of astronaut health (LSAH) database
  - medical records including flight surgeon interviews
- **ISS Expeditions 1 thru 13 (2006)**
- **STS-01 through STS-114 (2005)**
- **Apollo, Skylab, Mir (U.S. crew only)**
- **Review of crew medical charts**
- **Flight surgeon subject matter expertise**
- **Analog, terrestrial data**

Plan for LSAH Validation Data Set