Validation of the NASA Integrated Medical Model: A Space Flight Medical Risk Prediction Tool

Jerry Myers[a], Yamil Garcia[b], John Arellano[c], Lynn Boley[b], Debra Goodenow[a], Eric Kerstman[d], Matthew Koslovsky[b], David Reyes[d], Lynn Saile[b], Wafa Taiym[b], Millennia Young[e]

[a] National Aeronautics and Space Administration, Glenn Research Center, Cleveland, OH, USA
[b] KBRwyle, Houston, TX, USA
[c] MEIT, Houston, TX, USA
[d] University of Texas Medical Branch, Galveston, TX, USA
[e] National Aeronautics and Space Administration, Johnson Space Center, Houston, TX, USA

Abstract: The Human Research Program funded the development of the Integrated Medical Model (IMM) to quantify the medical component of overall mission risk. The IMM uses Monte Carlo simulation methodology, incorporating space flight and ground medical data, to estimate the probability of mission medical outcomes and resource utilization. To determine the credibility of IMM output, the IMM project team completed two validation studies that compared IMM predicted output to observed medical events from a selection of Shuttle Transportation System (STS) and International Space Station (ISS) missions. The validation study results showed that the IMM under-predicted the occurrence of ~10% of the modeled medical conditions for the STS missions and over-predicted ~20% of the modeled medical conditions for the ISS missions. These findings imply that the strength of IMM predictions to inform decisions depends on simulated mission specifications including length. This discrepancy could result from medical recording differences between ISS and STS that possibly influence observed incidence rates, IMM combining all "mission type" data as constant occurrence rate or fixed proportion across both mission types, misspecification of symptoms to conditions, and gaps in the literature informing the model. Some of these issues will be alleviated by updating the IMM source data through incorporation of the observed validation data.

Keywords: PRA, aerospace medicine, validation, credibility, simulation, NASA

1. INTRODUCTION

The Integrated Medical Model (IMM) represents an aspect of the NASA Human Research Program's (HRP) effort to quantitatively estimate medical risks to astronauts for existing operational missions. The IMM was developed to join medical and human health information acquired over years of crewed spaceflight to inform current mission medical risks, future space flight vehicle design, mission resource requirements' specifications, and mission requirements associated with commercial space flight ventures.

Historically, medical environment design and operation uses both qualitative and quantitative assessment of risk to optimize clinical outcomes and resource utilization. In July 2001, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) implemented the requirement that accredited hospital and treatment settings must conduct at least one proactive risk analysis annually. This requirement serves to achieve clinical outcome optimization and maintain accreditation [1]. Specific implementations of risk assessments vary widely but generally fall into the following programs: failure mode and effects analysis (FMEA), fault tree analysis, and quality management programs[2]; with FMEA and its derivatives historically being the most commonly used. The common use of FMEA in clinical operations risk assessment likely stemmed from its acceptance in other operational environments like NASA who, prior to 1986, depended on FMEA and hazard analysis (HA) as the

means to assess mission risk [3]. FMEA is similar to multidisciplinary root cause analysis, but is prospective rather than retrospective when applied to healthcare [1]. It relies on the calculation of a risk priority number, which combines a 10-point scale of severity, occurrence, and detectability assessed by multidisciplinary teams at the target institutions. Due to this focused assessment, which is based on the local institution's employee experience, FMEA often does not consider population and multi-institutional information and lacks the ability to identify complex system, combinational effects. This reduces its ability to support planning and new technology development. Efforts to use Delphi studies [4] or otherwise modify FMEA [5] to improve its applicability to healthcare risk assessment and prevent predictable failure modes have been proposed in recent years. This continued until the general acceptance of other, better quantitative methods based on probabilistic analysis.

The acceptance of quantitative risk analysis approaches has led to more acceptance of data-driven healthcare risk assessment processes, such as those based on fault tree and probabilistic risk analysis (PRA) approaches. PRA techniques relate a set of potential outcomes of interest to critical events representing the operational environment, typically implemented via event tree and fault tree analyses. By parameterizing these event trees with representative probabilities and uncertainties of the events, a quantitative assessment of the risk of the defined outcomes can be performed [3]. In addition to healthcare, other technology-driven industries, such as nuclear, space, food safety, and environmental protection are using these techniques to prospectively evaluate existing risks and the cost-benefit of new technologies, processes, and the optimization of resources [6].

The healthcare industry has moved to adopt PRA for the additional benefit that it quantitatively supports cost utility estimates and medical decision support [7]–[11]. Particularly, recent healthcare focus on informed decision-making has benefitted from quantitative risk modeling by improving the evidence supporting design and funding capture in the development of new healthcare technologies [12]. Resource allocation in the planning for natural disaster response and disease outbreaks have benefitted from such evidence modeling [13], [14]. PRA derived techniques, such as Sociotechnical PRA (ST-PRA) have proven to be important risk vs cost vs outcomes utility estimate tools for medical staff, hospital administrators, and government decision makers, when compared to qualitative techniques [1], [9], [15]. Hospital admittance practices and resource planning have utilized PRA type methods, such as probabilistic mortality models, to improve other risk-scoring admittance techniques, and as a means to stratify treatment allocations [7], [11], [16]–[18]. Further application in these areas has led to implementation of optimization techniques to refine resource allocation and placement in general healthcare and disaster settings [14], [19], [20]. The literature is brimming with Markov probabilistic models related to the risk of specific applications or treatment processes. Predicting falls, caries, stroke outcomes, hospital (discharge) re-admittance after cardiac event, and the impacts of diabetes treatment are just a sampling of the myriad applications to which probabilistic techniques have been used to evaluate healthcare treatment and technology [8], [10], [21]–[25]. Similarly, NASA recently adopted PRA techniques in the assessment of specific medical conditions which require additional insight due to the unique environment of space flight and the lack of observable events thus far such as in bone fracture[26], [27], head injury [28] and decompression sickness [29].

The NASA-HRP intends for the IMM to provide a more global means to quantify the medical component of total mission risk in a manner comparable to space flight system risk estimates performed within engineering and mission PRAs. The IMM utilizes PRA techniques to simultaneously incorporate space flight and ground medical data to assess the need for particular medical resources and capabilities across various mission scenarios. The IMM approach simulates the occurrence and resolution of predicted medical events along a planned mission timeline to estimate the probability of mission medical outcomes such as medical impairment, loss of life, and resource utilization.

The NASA-HRP requires all models and simulations that can have moderate to high impact on crew health or mission success to be vetted in accordance to NASA Standards for Models and Simulations, NASA-STD-7009a [30]. This standard focuses on establishing the credibility, defined as the belief that model output is representative of how the real world system will perform, by assessing eight credibility factors: Verification, Validation, Development Data Pedigree, Input Data Pedigree, Uncertainty

Characterization, Results Robustness, Model Use History, and Model Management. Since 7009a [30] focuses on engineering systems, the IMM adapted the processes established so that they can be readily applied to predictive models that are more prevalent in health care and biomedical research. To determine the credibility of IMM in support of mission planning decision making, NASA undertook a validation study that compared IMM predicted output to directly observed medical events and outcomes from a selection of Shuttle Transportation System (STS) and International Space Station (ISS) missions.

2. METHODS

2.1 IMM Implementation

2.1.1. IMM Concept

Keenan et al. 2015 [31] describe the underlying concept and overall implementation of the IMM. Briefly, the IMM architecture follows the practices of probabilistic risk assessment as outlined in the NASA PRA implementation guidance [3]. However, the implementation diverges from strict PRA implementation to accommodate the broad assumptions required for medical treatment and outcome simulations. As illustrated in Figure 1, the IMM takes as user-specified input mission characteristics including mission length, number of EVAs, and certain crewmember characteristics including sex and medical factors. Currently, 100 medical conditions are modeled for applicable space flight medical conditions, and incident rates are set based on crew characteristics. For example, the incidence rate used to simulate a certain condition may differ depending on if the crewmember has had surgery in the past or not.

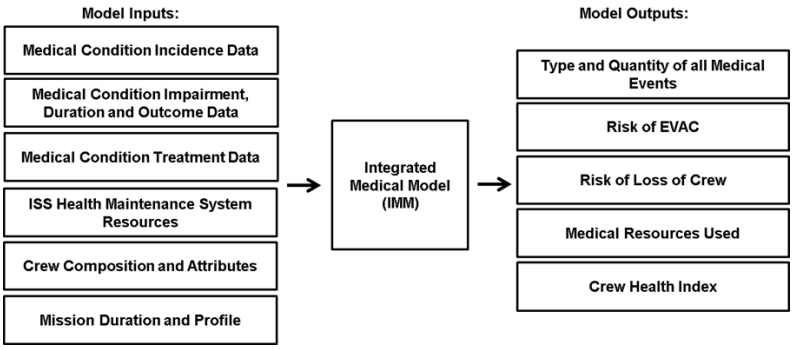


Figure 1. IMM input and output parameters. Note that Crew Health Index is a normalized measure of Available Mission Time – Quality Time Lost. Reproduced from [31]**.**

2.1.2 Medical Condition Occurrence, Treatment and Outcomes

The IMM assumes each medical condition occurs and is addressed independently of the occurrence of other conditions throughout the planned mission timeline. Generated incidence rates IMM (described in section 2.1.3) for rate-dependent conditions are assumed constant for the duration of a simulated mission and event occurrences are governed by a Poisson Process (exponential waiting times between events). The IMM assumes conditions associated with specific mission events, such as during adaptation to the spaceflight environment, extravehicular activity (EVA) or following solar particle events, follow a binomial distribution. The IMM captures the severity of a simulated medical condition by generating a best- or worst-case event scenario; each scenario is associated with separate medical event outcome distributions (Figure 2). Outcomes associated with these two event paths represent a continuum of possible outcomes for the affected crewmember given defined resource, treatment, and environmental constraints. Resource types and quantities, used to model medical risk mitigation in the IMM, are derived from the International Space Station (ISS) Health Maintenance System [32]. Treatments, specified for each medical condition/scenario path, define required quantities of medical resources, the per-day dosage, and a resource category. The pharmaceutical category allows for the model to consider suitable alternates from the same category when primary resource is depleted in the treatment of a simulated condition. The IMM generates outcomes for a condition based on the proportion of treatment available allowing for partial credit in having some but not all of the resources

required for treatment of a simulated event. The IMM implements this partial credit by defining outcome distributions for the fully-treated and untreated distributions as the extremes, and using the proportion of treatment available to shift continuously between them. The IMM allocates the medical resources assigned to treat each medical event from the medical kit in the order of medical event occurrence (at the time of simulated onset). The IMM allows for treatment modification within the simulation to account for remaining mission time relative to the end of the mission or to account for concurrent condition treatments from the same crewmember.

Primary outcomes quantifying the impact of medical events on the mission are measured by the quality time lost (QTL) and Crew Health Index (CHI), probability of evacuation (pEVAC), probability of loss of crew life (pLOCL), and total number of medical events (TME).

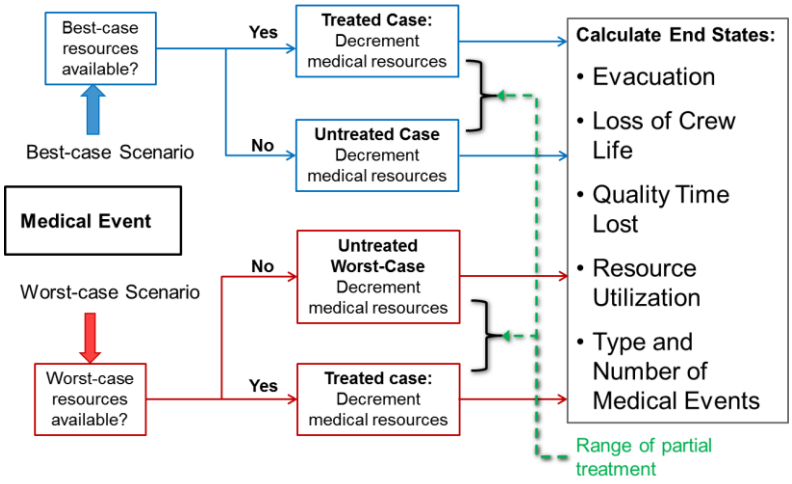


Figure 2. Illustration of IMM implementation of condition treatment and outcomes. 4 paths bound the limits of available treatment and outcome processes. When limited or incomplete treatment is predicted, IMM weights outcomes on the remaining availability of critical treatment components.

2.1.3 Data Informing the IMM

A SQL database, the integrated Medical Evidence Database (iMED), houses the medical-condition-model inputs. Subject matter experts assess the quality of the medical data, and the iMED management enforces a strict configuration management process to maintain medical data consistency. The iMED includes 100 medical conditions considered to be of concern by the space flight medical community. Whenever possible, space flight observed medical conditions, i.e. in-flight data, informs the incidence data for the medical conditions simulated in the IMM. The NASA Lifetime Surveillance of Astronaut Health (LSAH) and published literature provides the IMM with in-flight incidence data estimates [33]–[35]. The current version of the model uses in-flight data from shuttle missions STS 1-114, except STS-51-L (Challenger) and STS-107 (Columbia), ISS expeditions 1-13, Apollo, Skylab, and Shuttle/Mir. Data from some later flights inform medical condition inputs related to Spaceflight Associated Neuro-ocular Syndrome (SANS, formally known as visual impairment and intracranial pressure or VIIP syndrome). Where observational data are insufficient to adequately define the in-flight medical risk, the IMM uses terrestrial analog and general population data including Bayesian analysis incorporating pre- and post-flight astronaut data and terrestrial data [36], analog condition terrestrial data, and external probabilistic modules, to estimate medical-event incident likelihoods [37].

2.2 IMM Validation

The validation of IMM used both qualitative and quantitative techniques in order to best utilize in-flight observed data for comparison to model predictions. Typically, to achieve HRP-defined credibility levels requires the validation to take a strictly quantitative approach, which is not always possible with this type of predictive model. The IMM approach attempts to address the complexities of the medical system data limitations, uncertainties associated with clinical interpretation of historical data, and the data-limited scope of predictive space flight medical modeling (simply not enough observed time to obtain precise estimates of incidence).

2.2.1 Referent Data

To evaluate the IMM model, the referent data for validation consisted of observed medical incidence not previously incorporated into the primary iMED data repository. Specifically, medical observations, mission lengths and crew profiles from ISS Expeditions (Exp) 14 through 39/40 and, and STS 115 through STS 135 composed the referent real world system (RWS) dataset. This RWS medical data included the information from the medical record that could include information such as type of medical condition or symptoms, whether the condition occurred during the initial physiological adaptation to space or later in a mission, or if the condition occurred during extravehicular activity (EVA). If recorded during the mission, the medical capabilities used to evaluate and treat each condition were also included.

For a select number of conditions, where iMED incidence data included inflight experience for some of the missions included in this study, both observed and predicted counts of these conditions were set to zero for the overlapping missions. This was done to ensure that the validation data were completely independent of the model predicted data, validating on newly observed data only. The choice was made to use the most conservative estimate in assessing the time sequence to ensure that none of the observed comparison data for this study included any prior iMED incidence data.

2.2.2 Validation Simulations and Comparisons

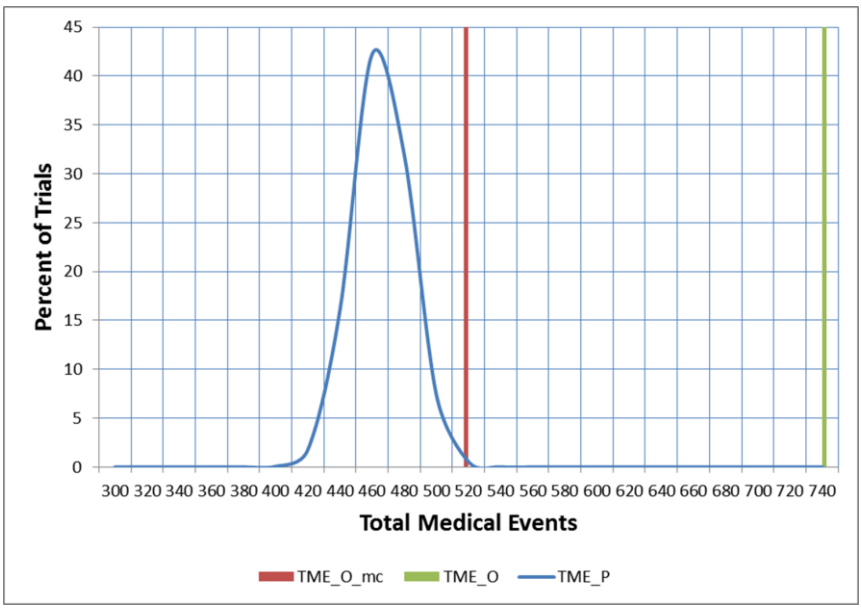
The validation study utilized a separate set of IMM mission simulations corresponding to each of the 31 ISS and 21 STS missions in the referent set. In each simulation, iMED incidence using available US space flight data and the appropriate subject matter expert identified terrestrial and space analog data, not in the referent data set, informed the IMM. Each simulation assessed the impact of 100 medical conditions that NASA medical operations have observed during spaceflight, or believe could have a high potential to occur, or could have a significant mission impact. One hundred thousand trials (simulations of that particular mission) were generated for each mission. Adequate model convergence was assessed by confirming that the main outputs exhibited a less than 5% change in their calculated standard deviation over the last two 1,000 trial increments.

The validation comparison focused on the RWS observed and IMM predicted number of total medical events (TME - combined RWS observations or IMM predictions across the entire set of missions in the RWS dataset), medical consumable utilization and predictions of LOCL and EVAC. Note that QTL and CHI cannot be used in validation comparisons as currently there exists no direct means to acquire these as observable outcomes on US space missions.

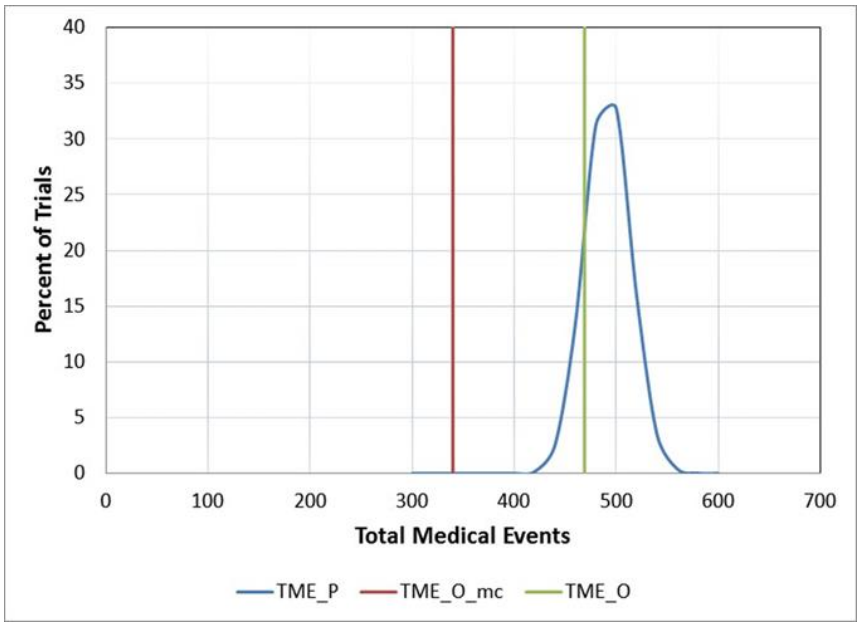
3. RESULTS

Figure 3 illustrates the comparison of the observed and predicted medical events combined across all RWS missions. It should be noted that the RWS referent data contained additional observed conditions that are not within the 100 conditions modeled by the IMM. The total number of observed events, including those not modeled by IMM (TME_O), are included in Figure 3 for completeness. When considering only the 100 medical conditions in the validation study (TME_O_mc in Figure 3), referent observations of total medical events generally under-predicted STS observations and over-predicted the number of the ISS observations. More specifically, as illustrated in Figure 4, the IMM predicted within a 90% CI in 13 of the 21 STS missions and 15 of the 31 ISS missions. When observations existed

outside the prediction CI, IMM tended to under-predict the TME for STS missions (5 of 21 missions) and over-predict the TME for ISS missions (15 out of 31 missions).



(a) STS

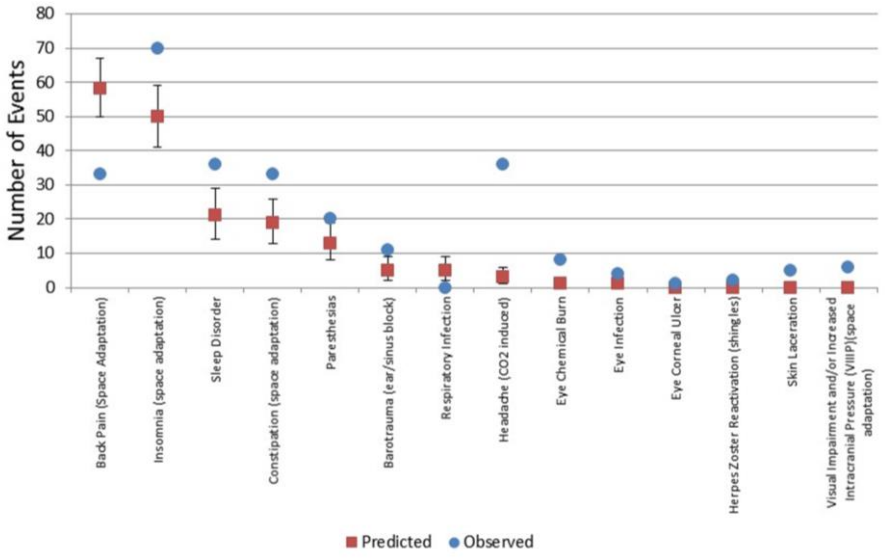


(b) ISS

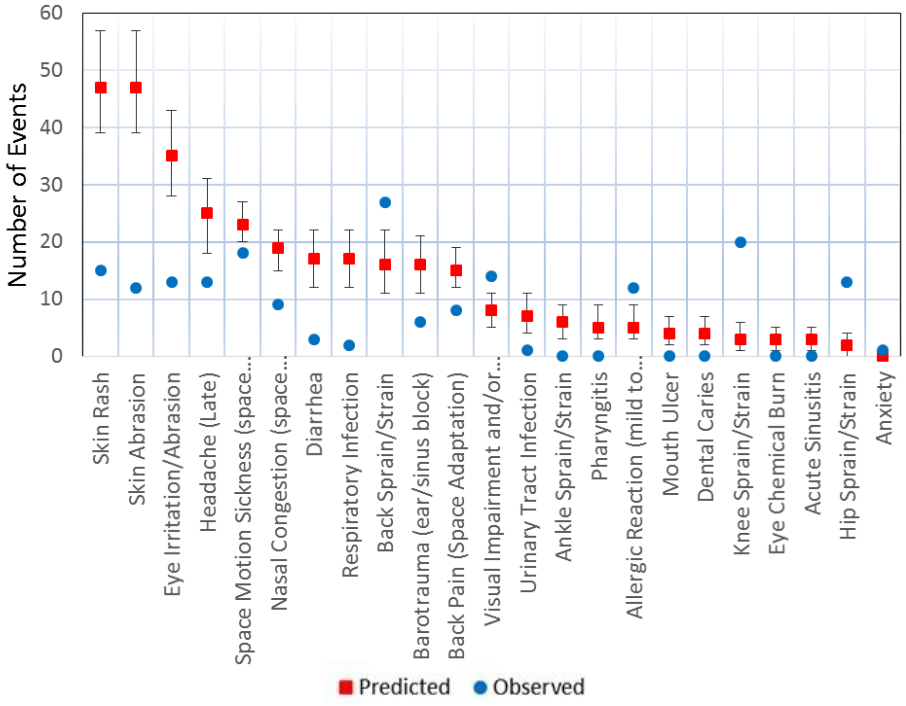
Figure 3. Distribution of total medical events predicted (_P) across trials vs. observed (_O_mc, red line) for RWS (a) STS and (b) ISS missions. The subscript (O_, green line) refers to observed medical conditions that include additional reported medical conditions than are modelled in the IMM.

On a condition basis, IMM predicted 20% of the STS and 15% of the ISS individual medical conditions events within prediction uncertainty defined by confidence limits estimated by the simulation percentiles. As shown in Figure 4, 14% of STS and 24% of ISS individual medical condition predictions fell outside of the prediction uncertainty. Of note, all but two of the STS out of range conditions were

under-predicted. The remainder of the individual medical condition events exhibited indeterminate comparison due to no observed incidences in the referent data set as there is not enough resolution to get a stable estimate of the incidence rate. In the indeterminate case, the model is not inconsistent with the zero observed events, but more observed missions are needed to get stable estimates of inflight incidence for comparison to predicted incidence.



(a) STS



(b) ISS

Figure 4. Out of range predictions for the per condition cumulative comparisons for RWS (a) STS missions and (b) ISS missions. For out of range predictions, STS mission predictions under-predicted the number of events for all but two conditions, while ISS mission predictions over-predicted the number of events for all but 6 conditions

Each rank comparison of observed and predicted medical consumables was considered an excellent match if within <=2 rank difference between observed and predicted, a fair match if difference >=3 but <=5, and a poor match if >5 rank difference. These qualitative assessments of resource utilization represent a potential threshold where IMM predictions would affect decision-making. Qualitatively, the IMM predictions of medication utilization showed either fair or excellent correspondence (Table 1) with the observed RWS for all medication categories for STS. In addition, the IMM predictions of medication utilization showed either fair or excellent correspondence with the observed RWS for all medication categories for ISS, with the exception of steroids. The IMM tended to under-predict the use of steroids on ISS. This discrepancy may be related to IMM resource table inputs and may present an opportunity to improve the IMM input data. Additionally, we estimated the correlation between the rankings of medical categories in terms of required resources in the observed RWS and IMM predictions (STS and ISS). For both scenarios, we estimated a positive correlation between the IMM predictions for STS and ISS with the observed RWS (Kendall Tau-b = 0.76 and Kendall Tau-b = 0.57, respectively) indicating not disparate orderings of categories.

Table 1. Rank comparison of predicted and observed medical consumable utilization by resource category.

	STS			ISS		
Medical Resource Category	**Observed**	**Predicted**	**Match**	**Observed**	**Predicted**	**Match**
Antacids	10	13	Fair	10	12	Excellent
Antibiotics	7	8	Excellent	7	3	Fair
Antidiarrheals	11	7	Fair	11	8	Fair
Antiemetics	3	1	Excellent	3	6	Fair
Antifungals	9	10	Excellent	9	9	Excellent
Antihistamines	4	3	Excellent	4	4	Excellent
Antivirals	13	12	Excellent	13	14	Excellent
Decongestants	6	5	Excellent	6	7	Excellent
Hypnotics	2	2	Excellent	2	2	Excellent
Laxatives	12	11	Excellent	12	10	Excellent
Non-opioid Analgesics	1	4	Fair	1	1	Excellent
Ophthalmic Lubricants	8	9	Excellent	8	5	Fair
Opioid Analgesics	14	14	Excellent	14	11	Fair
Steroids	5	6	Excellent	5	13	Poor

The RWS did not report any instances of medically induced considerations of EVAC or observations of LOCL. Consistent with this observation, the IMM estimated low probabilities for LOCL and EVAC (not shown). Comparisons of observed and predicted EVAC and LOCL counts, illustrated in Table 2, illustrate IMM's consistency with zero observed LOCL and EVAC events. However, as with indeterminate condition comparisons, without some observed events, it's impossible to determine if the IMM predicted rate is accurate.

Table 2. LOCL and EVAC predictions compared to RWS observations. Predicted counts are estimated using the median of the simulated distribution. Confidence intervals are estimated by the 5th and 95th percentiles of the simulation distribution. A confidence limit of (0, 0) indicates that IMM predicted a 0 LOCL or EVAC count in more than 95% of the generated trials, as estimated by the 5th and 95th percentiles of the simulation distribution.

STS	Predicted Number	90% Confidence Interval
EVAC RWS = 0	0	(0, 1)
LOCL RWS = 0	0	(0, 0)

ISS	Predicted Number	90% Confidence Interval
EVAC RWS = 0	0	(0, 1)
LOCL RWS = 0	0	(0, 0)

4. CONCLUSIONS

With respect to the 100 medical conditions included in the IMM, IMM predictions represent a reasonable first estimate of the medical risk for both STS and ISS type missions, but care must be taken when utilizing the output for decision-making purposes. These findings show that IMM exhibits variations in strength to inform decisions as mission length varies, with shorter missions having the tendency to under predict total medical events and longer missions the tendency to over predict total medical events. However, clinical evaluation of resource utilization predictions infers that the predicted required medical resources are representative of resource utilization on the ranked scale. There wasn't enough data to determine accuracy in quantity required. The full difference in the STS and ISS IMM-modeled predictions compared to the reference RWS observations cannot be fully determined within the scope of this analysis. Differences may be due to relative proportions of space adaptation conditions, or issues with estimates of incidence rates made under the different ISS and STS medical reporting conditions. There may also be underlying differences that are not captured within the IMM approach of combining all "mission type" data and assuming a constant occurrence rate over a mission. We conjecture that the predictive performance of the IMM will improve as the iMED is updated with reference RWS data.

The success and generalization of using the NASA model and simulation credibility methods to support biomedical and health care modeling has also generated substantial interest by the broader medical community. Institutions like the National Institutes of Health (NIH) Interagency Modeling and Analysis Groups – Committee for Credible Practices in Modeling and Simulation have adopted aspects of this approach to develop similar standards and rules for health care modeling. (https://www.imagwiki.nibib.nih.gov/content/committee-credible-practice-modeling-simulation-healthcare-msm-2014).

Acknowledgements

We would like to acknowledge Marlei Walton and Shannon Melton of KBRwyle, DeVon Griffin and Kelly Gilkey of NASA GRC for their contributions to the development and design of the IMM validation activity. We would also like to acknowledge the NASA Human Research Program, Exploration Medical Capability Element, and Lifetime Survey of Astronaut Health for their support of this project.

References

[1] D. a Marx and a D. Slonim, "Assessing patient safety risk before the injury occurs: an introduction to sociotechnical probabilistic risk modelling in health care.," *Qual. Saf. Health Care*, vol. 12 Suppl 2, p. ii33-i38, 2003.

[2] J.-E. Rah, R. P. Manger, A. D. Yock, and G.-Y. Kim, "A comparison of two prospective risk analysis methods: Traditional FMEA and a modified healthcare FMEA," *Med. Phys.*, vol. 43, no. 12, pp. 6347–6353, Dec. 2016.

[3] M. Stamatelatos and H. Dezfuli, "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners," *NASA Tech. Rep.*, vol. SP-2011-34, no. December, p. 323, 2011.

[4] P. B. Southard, S. Kumar, and C. A. Southard, "A modified Delphi methodology to conduct a failure modes effects analysis: a patient-centric effort in a clinical medical laboratory.," *Qual. Manag. Health Care*, vol. 20, no. 2, pp. 131–51, 2011.

[5] H. W. W. Potts, J. E. Anderson, L. Colligan, P. Leach, S. Davis, and J. Berman, "Assessing the validity of prospective hazard analysis methods: a comparison of two techniques.," *BMC Health Serv. Res.*, vol. 14, no. 1, p. 41, Jan. 2014.

[6] K. M. Thompson, "Variability and uncertainty meet risk management and risk communication.," *Risk Anal.*, vol. 22, no. 3, pp. 647–54, Jun. 2002.

[7] A. Gandjour and E.-J. Weyler, "Cost-effectiveness of referrals to high-volume hospitals: an analysis based on a probabilistic Markov model for hip fracture surgeries.," *Health Care Manag. Sci.*, vol. 9, no. 4, pp. 359–69, Nov. 2006.

[8] M. P. M. H. Rutten-van Mölken, J. B. Oostenbrink, M. Miravitlles, and B. U. Monz, "Modelling the 5-year cost effectiveness of tiotropium, salmeterol and ipratropium for the treatment of chronic obstructive pulmonary disease in Spain.," *Eur. J. Health Econ.*, vol. 8, no. 2, pp. 123–35, Jun. 2007.

[9] S. C. Comden, D. Marx, M. Murphy-Carley, and M. Hale, *Using Probabilistic Risk Assessment to Model Medication System Failures in Long-term Care Facilities*. 2005.

[10] J. B. Oostenbrink, M. P. M. H. Rutten-van Mölken, B. U. Monz, and J. M. FitzGerald, "Probabilistic Markov model to assess the cost-effectiveness of bronchodilator therapy in COPD patients in different countries," *Value Heal.*, vol. 8, no. 1, pp. 32–46, 2005.

[11] K. H. Lich *et al.*, "Strategic planning to reduce the burden of stroke among veterans: using simulation modeling to inform decision making.," *Stroke*, vol. 45, no. 7, pp. 2078–84, Jul. 2014.

[12] A. Briggs, M. Sculpher, J. Dawson, R. Fitzpatrick, D. Murray, and H. Malchau, "The use of probabilistic decision models in technology assessment : the case of total hip replacement.," *Appl. Health Econ. Health Policy*, vol. 3, no. 2, pp. 79–89, 2004.

[13] J. A. Sobieraj *et al.*, "Modeling hospital response to mild and severe influenza pandemic scenarios under normal and expanded capacities.," *Mil. Med.*, vol. 172, no. 5, pp. 486–90, May 2007.

[14] M. R. Zolfaghari and E. Peyghaleh, "Implementation of equity in resource allocation for regional earthquake risk mitigation using two-stage stochastic programming.," *Risk Anal.*, vol. 35, no. 3, pp. 434–58, Mar. 2015.

[15] R. Garside, M. Pitt, R. Anderson, S. Mealing, R. D'Souza, and K. Stein, "The cost-utility of cinacalcet in addition to standard care compared to standard care alone for secondary hyperparathyroidism in end-stage renal disease: a UK perspective.," *Nephrol. Dial. Transplant*, vol. 22, no. 5, pp. 1428–36, May 2007.

[16] L. I. Iezzoni, M. Shwartz, A. S. Ash, and Y. D. Mackiernan, "Predicting in-hospital mortality for stroke patients: results differ across severity-measurement methods.," *Med. Decis. Making*, vol. 16, no. 4, pp. 348–56.

[17] D. Kansagara *et al.*, "Risk Prediction Models for Hospital Readmission," *JAMA*, vol. 306, no. 15, p. 1688, Oct. 2011.

[18] J. Hippisley-Cox and C. Coupland, "Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score.," *BMJ Open*, vol. 3, no. 8, p. e003482, Aug. 2013.

[19] J. A. Moore, J. J. Gordon, M. Anscher, J. Silva, and J. V Siebers, “Comparisons of treatment optimization directly incorporating systematic patient setup uncertainty with a margin-based approach.,” *Med. Phys.*, vol. 39, no. 2, pp. 1102–11, Feb. 2012.
[20] E. C. Parker, S. S. Survanshi, P. B. Massell, and P. K. Weathersby, “Probabilistic models of the role of oxygen in human decompression sickness.,” *J. Appl. Physiol.*, vol. 84, no. 3, pp. 1096–102, Mar. 1998.
[21] S. Singh, H. Sun, and A. H. Anis, “Cost-effectiveness of hip protectors in the prevention of osteoporosis related hip fractures in elderly nursing home residents.,” *J. Rheumatol.*, vol. 31, no. 8, pp. 1607–13, Aug. 2004.
[22] H. P. Selker *et al.*, “Patient-specific predictions of outcomes in myocardial infarction for real-time emergency use: a thrombolytic predictive instrument.,” *Ann. Intern. Med.*, vol. 127, no. 7, pp. 538–56, Oct. 1997.
[23] K. N. Page, A. L. Barker, and J. Kamar, “Development and validation of a pressure ulcer risk assessment tool for acute hospital patients.,” *Wound Repair Regen.*, vol. 19, no. 1, pp. 31–7, Jan. 2011.
[24] A. J. Palmer *et al.*, “Computer modeling of diabetes and its complications: a report on the Fifth Mount Hood challenge meeting.,” *Value Health*, vol. 16, no. 4, pp. 670–85, Jun. 2013.
[25] M. E. Moss and D. T. Zero, “An overview of caries risk assessment, and its potential utility.,” *J. Dent. Educ.*, vol. 59, no. 10, pp. 932–40, Oct. 1995.
[26] E. S. Nelson, B. Lewandowski, A. Licata, and J. G. Myers, “Development and Validation of a Predictive Bone Fracture Risk Model for Astronauts,” *Ann. Biomed. Eng.*, vol. 37, no. 11, pp. 2337–2359, Nov. 2009.
[27] C. M. Sulkowski, K. M. Gilkey, B. E. Lewandowski, S. Samorezov, and J. G. Myers, “An extravehicular suit impact load attenuation study to improve astronaut bone fracture prediction.,” *Aviat. Space. Environ. Med.*, vol. 82, no. 4, pp. 455–62, Apr. 2011.
[28] A. S. Weaver, A. D. Zakrajsek, B. E. Lewandowski, J. E. Brooker, and J. G. Myers, “Predicting head injury risk during International Space Station increments.,” *Aviat. Space. Environ. Med.*, vol. 84, no. 1, pp. 38–46, Jan. 2013.
[29] J. Conkin, K. V Kumar, M. R. Powell, P. P. Foster, and J. M. Waligora, “A probabilistic model of hypobaric decompression sickness based on 66 chamber tests.,” *Aviat. Space. Environ. Med.*, vol. 67, no. 2, pp. 176–83, Feb. 1996.
[30] NASA, “NASA-STD-7009A: Standard for Models and Simulations,” *Nasa Technical Standard*, vol. 7009A, no. I. pp. 07–11, 2008.
[31] A. Keenan *et al.*, “The Integrated Medical Model: A Probabilistic Simulation Model Predicting In-Flight Medical Risks,” 2015.
[32] NASA, “International Space Station Integrated Medical Group Medical Checklist.” NASA Mission Operations Directorate, Houston, TX, 2011.
[33] P. B. Hamm, A. E. Nicogossian, S. L. Pool, M. L. Wear, and R. D. Billica, “Design and current status of the longitudinal study of astronaut health.,” *Aviat. Space. Environ. Med.*, vol. 71, no. 6, pp. 564–70, Jun. 2000.
[34] R. A. Scheuring, C. H. Mathers, J. A. Jones, and M. L. Wear, “Musculoskeletal injuries and minor trauma in space: incidence and injury mechanisms in U.S. astronauts.,” *Aviat. Space. Environ. Med.*, vol. 80, no. 2, pp. 117–24, Feb. 2009.
[35] E. L. Kerstman, R. A. Scheuring, M. G. Barnes, T. B. DeKorse, and L. G. Saile, “Space adaptation back pain: A retrospective study,” *Aviat. Sp. Environ. Med.*, vol. 83, no. 1, pp. 2–7, 2012.
[36] K. M. Gilkey, M. P. Mcrae, E. A. Grif, A. S. Kalluri, and J. G. Myers, “Bayesian Analysis for Risk Assessment of Selected Medical Events in Support of the Integrated Medical Model Effort,” no. July, pp. 1–50, 2012.
[37] L. Boley, “IMM-Plan-101 Rev2, Integrated Medical Model Comprehensive Configuration Management Plan.” KBRWyle Science Technology and Engineering, Houston, TX.