For submission at : https://home.liebertpub.com/publications/telemedicine-and-e-health/54/for-authors

Autonomous Medical Officer Support (AMOS) Software Technology Demonstrations on the International Space Station (ISS)

Douglas Ebert[1], Marlei Walton[2], Vicky E. Byrne[1], Kurt Berens[1], Ariana M. Nelson[3], Ashot Sargsyan[1]

[1] KBR, 2400 NASA Parkway, Houston, TX
[2] NASA-Johnson Space Center
[3] University of Texas Medical Branch, Galveston, TX.

Abstract (150 words) 145 words currently

Performance of medical procedures in spaceflight beyond low Earth orbit (LEO) requires novel solutions to replace real-time ground support because as distance from Earth increases, communication latencies increase, hampering remote guidance. The Autonomous Medical Officer Support Software (AMOS) Technology Demonstrations on the International Space Station (ISS) trialed a novel software tool that shifts the emphasis from preflight training and real-time remote guidance (current ISS paradigm) to a new standard of multi-dimensional in-flight just-in-time (JIT) instruction. The AMOS platform is a skill management tool for all mission phases and currently features comprehensive training and guidance modules for urinary bladder and renal ultrasound examinations. Variability in Subject anatomy, Operator experience, and Operator receptiveness to instruction during autonomous exams are persistent but manageable limitations. Here we report the first successful demonstrations of autonomous imaging activities in the operational setting of spaceflight, validating this autonomous guidance proof of concept.

Manuscript: (2000 word limit for short communication)

Introduction: 1995 words currently (includes acknowledgments)

Current ISS medical operations rely heavily on real-time remote guidance from Earth, which becomes impractical with increased communication latencies during exploration class missions.[1] Successful performance of medical procedures during these missions will require an autonomous solution to replace real-time ground support for Crew Medical Officers (CMOs).[2,3] The Autonomous Medical Officer Support (AMOS) software tool enables independent training and guidance for all crew, shifting from preflight training and remote guidance to in-flight just-in-time (JIT) instruction.[3,4]

The autonomous guidance concepts and software design for AMOS were developed and validated by the multi-disciplinary team for the ground-based Clinical Outcome Metrics for Optimization of Robust Training (COMfORT) study.[5] The COMfORT application was subsequently rebuilt for ISS server compatibility and modular expandability and renamed AMOS. The AMOS tool contains two pilot modules (ultrasound imaging of urinary bladder and kidneys), which, in contrast to the COMfORT modules, guide the user through more comprehensive exams designed to collect clinically useful imaging.[6] The menu-driven software is organized by modules, sections, chapters, and pages, allowing both topic-directed and linear navigation, enabling users to complete complex tasks autonomously with minimal training. The platform includes integrated use tracking and evaluation features to analyze usage patterns and gather feedback.

In contrast to the current practice of generating separate products for 1) preflight training, 2) in-flight JIT training, and 3) procedure execution, the AMOS platform is a single tool for all mission phases of skill management, covering all training aspects and procedural applications (Figure 1). AMOS also provides a streamlined process where multiple levels of autonomy are supported, making it ideal for deployment on the ISS as an exploration analog since it is amenable to incremental integration into operations or research. The modular design enables new content modules to be added without new software certification and is extensible beyond medical content to other applications (*e.g.*, engineering or maintenance).

Operationally beneficial features of the AMOS platform include 1) extending telemedicine capability, 2) enabling autonomous procedure performance, 3) streamlining the process of training and skill retention, 4) reducing preflight training load, and 5) reducing specific medical condition risks through improved monitoring or treatment in repeated operational trials (*e.g.* renal stone and urinary retention risks could be reduced if reliable in-flight autonomous monitoring capability is adopted).[7,8] The AMOS platform has the potential to enable new paradigms for training, skill development and retention, and on-demand performance for medical surveillance or high-priority procedures during exploration missions. This manuscript describes demonstrations that confirm the functionality and value of the AMOS tool in an operational environment relevant to its intended use for progressively Earth-independent medical operations for exploration class spaceflight missions.[9]

Methods: Both the COMfORT and AMOS projects were a collaboration between KBR, National Aeronautics and Space Administration (NASA), and RKT Creative (Detroit, MI). The KBR/NASA team provided the structure, content and user interface requirements, while RKT Creative was responsible for design and coding. These efforts align with NASA standard 3001, which requires that the capabilities for diagnostic and procedural imaging be available for in-mission medical care.[10,11]

AMOS Software

The AMOS software is an extensible markup language (XML)-based training and procedure guidance tool that runs in any web browser; it was installed on the ISS server system in April 2020 and is accessible from any space station computer (SSC). Each SSC was equipped with a link to the software that participants used during study procedures. AMOS was beta tested terrestrially during development with novice ultrasound users to optimize the user experience for astronauts; testing was conducted with NASA Exploration Medical Capability (ExMC) personnel and a physician astronaut, which resulted in improved clarity of instruction and correction of editorial errors. The final version included two comprehensive ultrasound scanning modules: one for the urinary bladder and one for bilateral kidney exam.

Participants

The AMOS technology demonstration (Tech Demo) for ISS was deemed exempt by the NASA Institutional Review Board (IRB). The demonstration required two crew members: an Operator to deploy the hardware and perform the ultrasound exams with guidance of AMOS and a Subject as the volunteer patient. Two sets of ISS crewmembers (Operator and Subject) participated on three separate demonstration occasions. None of the participants had pre-flight training in these procedures or prior experience using AMOS, however future operational use would include pre-flight familiarization with the platform. Results from one pair of crewmembers was not included in the data analysis due to significant protocol deviation in conducting critical initial instruction review.

Protocol

The day before each Tech Demo, the Operator reviewed instructions for the demo activity (10 minutes) and foundational instruction of each of the AMOS modules (30 minutes total), which convey basic principles of ultrasound scanning and anatomy. The Subject reviewed only the instructions (10 minutes). AMOS modules directed the Operator through steps to acquire and save ultrasound imaging scans of the target organs. Ultrasound images were collected using the ISS Ultrasound 2 system (modified GE Vivid q) with a 4C-RS broadband curved array transducer (GE Healthcare, Milwaukee, WI). The Operator collected cine-loop scans of the full and empty urinary bladder in two planes and 2D and color Doppler scans of both kidneys. The ideal scanning order begins with the technically simple bladder exam to build proficiency with basic ultrasound skills before engaging in the more technically complex renal scans, therefore the recommended scan order was: full bladder, empty bladder, right kidney, then left kidney. Setup and configuration were allocated 30 minutes, with all scans scheduled for one session (70 minutes for the bladder and 90 minutes for both kidneys). Participants also completed a 9-item questionnaire within each module. Following the scans, ultrasound data were transferred to the Human Research Facility laptop for later downlink. Operator and Subject astronauts did not have real-time contact with the AMOS Team or Mission Control Center for assistance or clarification on AMOS content for either demo.

Ultrasound and Video Analysis

Each demo was observed via remote video feed and also recorded for later technical and human factors analysis. Images were downlinked then examined by a trained sonographer and rated against a rigorous, multi-component rubric to quantify image quality and clinical utility. Image components were rated from 0 to 3 with 0 being of lowest quality/utility and 3 being the highest. A score of 2 was set as the threshold for clinical adequacy on the scoring rubric (Appendix A).

Click Tracking and Surveys

AMOS tracked Operator usage patterns by logging inputs into an exported .csv file. Surveys regarding software usability and procedure quality were built into the modules and completed during the session to reduce recall bias. Demo 1 crew provided a postflight debrief session 48 days after the scan session; due to crew time constraints, a debrief session was not conducted for Demo 2 or Demo 3.

Results

Ultrasound Imaging

For each demo, one or more clinically adequate images (score of 2 or better) were collected for every view (with one exception, discussed below) and in most cases the average success score met or exceeded the quality threshold (Figure 3). 2D bladder images were rated highest, followed by 2D kidney, and finally color Doppler kidney. This follows the level of complexity for each exam, with color Doppler of the kidney being most complex due to the need to manage Doppler parameters in addition to the 2D image. Quality scores demonstrated success of all imaging studies confirming autonomous training and guidance was achieved using AMOS by crewmembers with no prior training on these exams.

The recommended order for the scans was full bladder, empty bladder, right kidney, then left kidney. Due to the state of the Subject's bladder at the start of demo 1, the crew performed ultrasound scans in the following order: right kidney, left kidney, full bladder, empty bladder, then returned to right kidney. Bilateral kidney imaging was executed in two segments totaling 73 minutes, with improved performance in the final 7-minute repeated right kidney scan which was performed after bladder protocols. The bladder protocol was completed in 25 minutes including ~4 min break for voiding. The altered scan order and repeat of the right kidney exam provided a serendipitous internally controlled opportunity to observe substantial improvement over the course of this short demonstration activity (compare Right and Right Repeat scores in Figure 3B), demonstrating the AMOS tool's capability of supporting a very rapid learning curve.

Human Factors Observations

ISS cabin video analysis identified human factors concerns during one of the demos including sub-optimal positioning and communication lapses. The ultrasound machine and laptop were located on the same side of the cabin, with the Subject sometimes positioned between the Operator and ultrasound controls. This positioning was cumbersome but effective for optimizing some imaging angles. The Subject provided cooperative actions, including repositioning and managing ultrasound controls. The Subject attempted self-scanning, but this proved awkward and uncomfortable, leading the team to conclude that an Operator-Subject team is preferable for kidney scanning procedures, though self-scanning may be feasible in certain situations. One of the demo teams had fewer human factors challenges with a more consistent "floating seated" side-by-side positioning.

During demo 2 the Operator "ended exam" after completing the full bladder series, and the new exam began with the default preset which was incorrect for a bladder scan. Since the Operator did not re-verify settings, the ultrasound was not optimized for bladder, which markedly increased the scan time and reduced the image quality (Figure 3A; Demo 2 Post-Void Horizontal). In another instance, the Operator accidentally hit the invert button, causing a right-left inversion of the image, and increasing the difficulty of image capture. Although this highlights the utility of AMOS even during non-optimal conditions, it also demonstrates a concern for possible error with the human-machine interface, particularly when interfaces are complex.

Software Use Pattern

For all demos, navigation within the AMOS software was primarily through laptop arrow key navigation rather than point-and-click actions. This pattern suggests that first-time users may prefer linear navigation. The software supports various learning and procedure execution styles: in one demo the Operator watched only two video clips, in another demo the Operator only 3, and in one demo, no clips were viewed, indicating Operator preference for studying text and photo material.

Crew In-flight Survey and Postflight Debrief

Operators rated the software high on usability, content organization, and overall satisfaction (Figure 4). Free text responses highlighted the usefulness of the combination of text and videos, and suggested additional instructions for adjusting color Doppler gain and acquiring Doppler images. The Demo 2 Operator also noted the problem regarding ending the exam at the end of the bladder module, suggesting "if-then" wording for repeated scans in the same module. One user also noted that while the software is very useful for nominal scans, more assistance would be needed if autonomous diagnosis was also a goal.

In a postflight debrief, both the Operator and Subject reported no need for ground support during the study procedures, and that the AMOS guide was clear and intuitive. Operator and Subject affirmed that AMOS would be useful for medical procedures in spaceflight. The crew also provided positive feedback and constructive suggestions for future enhancements to support crew autonomy during medical procedures, including integration with portable or wearable screens and incorporation of AI platforms for real-time guidance.

Conclusion

Based on demo data and crew participants' feedback, the AMOS platform's potential to enhance skill proficiency during exploration spaceflight was recognized, with high satisfaction from the crew. As NASA continues to develop progressively Earth-independent medical operations, platforms like AMOS will be required to support CMOs. This proof of concept provides a valuable initiative for future applications of the AMOS platform and the next generation of JIT guidance for exploration class spaceflight.

Acknowledgements and Funding:

Funding for this project was provided by NASA Human Research Program (HRP) via the Exploration Medical Capabilities Element (ExMC). We acknowledge the contributions of: ExMC (Baraquiel Reyna, Lynn Boley, Benjamin Easter, Kris Lehnhardt, Nancy Fleming); the RKT Creative team (Ryan Kinnen, Kelly Comerford, and Tom Smale), the Research Operations and Integration team, Cheryl Flottorp, Gina Vega, Kurt Berens, and the astronaut participants.

References:

1. Parisi M, Panontin T, Wu S-C, et al. Effects of Communication Delay on Human Spaceflight Missions. 2023; doi: 10.54941/ahfe1003920.

2. Pryor M, Ebert D, Byrne V, et al. Diagnosis Behaviors of Physicians and Non-Physicians When Supported by an Electronic Differential Diagnosis Aid. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2019;63(1):68–72; doi: 10.1177/1071181319631420.

3. Russell BK, Burian BK, Hilmers DC, et al. The value of a spaceflight clinical decision support system for earth-independent medical operations. NPJ Microgravity 2023;9(1):46; doi: 10.1038/s41526-023-00284-1.

4. Hurst VW, Peterson S, Garcia K, et al. Concept of Operations Evaluation for Using Remote-Guidance Ultrasound for Exploration Spaceflight. Aerosp Med Hum Perform 2015;86(12):1034–1038; doi: 10.3357/AMHP.3244.2015.

5. Ebert D, Byrne V, Cole R, et al. Clinical Outcome Metrics for Optimization of Robust Training. 2016.

6. Ebert D, Byrne V, Walton M, et al. Autonomous Medical Officer Support (AMOS) ISS Technology Demonstration: Enabling Earth-Independent Procedure Guidance. 2023.

7. Sibonga JD, Pietrzyk R. Evidence Report: Risk of Renal Stone Formation. National Aeronautics and Space Administration; 2017.

8. Siew K, Nestler KA, Nelson C, et al. Cosmic kidney disease: an integrated pan-omic, physiological and morphological study into spaceflight-induced renal dysfunction. Nat Commun 2024;15(1):4923; doi: 10.1038/s41467-024-49212-1.

9. Levin DR, Steller J, Anderson A, et al. Enabling Human Space Exploration Missions Through Progressively Earth Independent Medical Operations (EIMO). IEEE Open Journal of Engineering in Medicine and Biology 2023;4:162–167; doi: 10.1109/OJEMB.2023.3255513.

10. Childress SD, Williams TC, Francisco DR. NASA Space Flight Human-System Standard: enabling human spaceflight missions by supporting astronaut health, safety, and performance. npj Microgravity 2023;9(1):31; doi: 10.1038/s41526-023-00275-2.

11. National Aeronautics and Space Administration. NASA Spaceflight Human-System Standard Volume 1: Crew Health. 2023.

Appendix A: Scoring Rubric

PARAMETER	APPLICABILITY	EXPLANATION	SCORING PRINCIPLE	SPECIFIC SCORING CRITERIA			
				3	2	1	0
PROCEDURE EXECUTION (0-3)	LOOP and GROUP	Grading of compliance with AMOS procedure execution (technical)	Image labeling, system settings and adjustments as prescribed; cine-loop acquisition sequence as prescribed	Full compliance with the procedure prescribed by AMOS software; 3/3	Small deviations from the prescribed procedure	Moderate deviations from the prescribed procedure	Gross deviation from the prescribed procedure
INCLUSION OF TARGET ORGAN VOLUME (0-3)	LOOP	Cine-loops: is the entire volume of the target organ represented in a loop?	Subjective assessment based on loop playback and frame-by-frame review	100% volume of the target organ represented within the sweep	Over 75% of the target organ represented within the sweep	25 - 75% of the target organ volume represented within the sweep	25% or less of the target organ represented within the sweep
	GROUP	Procedure segments: is the entire volume of the target organ represented between multiple loops?	Subjective assessment of all loops in the group. The highest value within group (or higher if instances are mutually	100% volume of the target organ represented among multiple sweeps	Over 75% of the target represented among multiple sweeps	50 - 75% or less of the target organ volume represented among multiple sweeps	50% or less of the target volume represented among multiple sweeps
IMAGING PLANE ORIENTATION (0-3)	LOOP	Alignment of the imaging plane with the prescribed axis of the target organ (adequacy of the rotational /roll position of the probe relative to the target's long axis), or with a prescribed anatomical plane. Mean value used for loop groups	Subjective assessment as described in columns 3,2,1, and 0	Excellent. Target image "stays in place" during the sweep. Linear and volume measurements possible	Good. Target image slightly shifts during the sweep. Measurements possible but may under-estimate some dimensions or volume	Satisfactory. Target image substantially shifts during the sweep. Only limited linear masurements are possible; volume may not be measured	Poor. Target image grossly shifts during the sweep. Misaligment results in gross inadequacy of the imagery. No measurements are
	GROUP		Average of all instances (loops) in the group				
SWEEP SPEED ADEQUACY (0-3)	LOOP	Adequacy of the speed of the sweep for frame-by-frame review	Subjective assessment based on loop playback and frame-by-frame image review	As prescribed; exellent balance of speed and loop size	Acceptable; moderately slower or faster than prescribed	Acceptable: too slow or fast, the number of usable frames too small or image quality diminished	Inadequate: too fast, the number of usable frames too small or image quality diminished
	GROUP		Average of all instances (loops) in the group				
AVERAGE IMAGE QUALITY (0-3)	LOOP and GROUP	Quality of individual images for hypothetical diagnostic evaluation	Subjective assessment based on multiple criteria; amenability to hypothetical clinical inpterpretation	Excellent	Good	Satisfactory	Inadequate
MEASURABILITY	LOOP	Amenability of data for linear or volume measurements	Objective and subjective assessment based on actual measurements, loop playback, and frame-by-frame review	Excellent, reliable results	Acceptable, diminished reliability and possible underestimation	Unreliable	Impossible
EFFECTIVE IMAGING CONTENTS (0-3; derived)	LOOP	Adequacy of imaging epresentation of the target organ within a DICOM instance for hypothetical diagnostic evaluation	Based on columns (I - N) for the given instance (loop)	Adequate imaging representation of target organ as prescribed	Adequate imaging representation of target organ with limitations	Substantially imited imagery of target organ	Grossly inadequate imagery or failure to image the target organ
	GROUP	Adequacy of imaging representation of the target organ between all DICOM instances of the group for hypothetical diagnostic evaluation	The highest value within group (or higher if instances are mutually complementary)				
SUCCESS SCORE	LOOP and GROUP	A composite value (mean) between average quality and effective imaging contents of the instance or group		Satisfies or exceeds AMOS presriptions.			
SUCCESS (Y/N)	LOOP	Overall success of individual cine-loops	Loop successful if the average of IImage Quality and Effective Imaging Content of the loop is > 2.0				
	GROUP	Overall success of the procedure segment	Segment successful if the average of Image Quality and Combined Content of the segment is > 2.0				
COMMENTS		Free-format comments by analysts	Included for reference or to provide clarifications and notes				

Appendix B. Summary of data used in the data review for demo 1

Data	Method of acquisition	Quantity	Format / Quality	Software used for review
Live observation by the team	VPN	Duration of the demonstration	Live video streaming	Browser
ISS Ultrasound II Scanhead video	Digital recording	3 hr 8 min (95 min of active scanning)	Video file	Windows Media Player
Cabin video	Digital recording	3 hr 44 min (overlapping files)	9 separate Video files	Windows Media Player
Cabin photography	Onboard camera	315 photographs	JPG files	Windows Photos
Ultrasound data from Bladder protocol	Download and conversion	17 instances, 4234 images	Standard DICOM	Verification and conversion: EchoPAC™ (GE Healthcare); Verification:

Data	Method of acquisition	Quantity	Format / Quality	Software used for review
Ultrasound data from Kidney protocol	from GE "raw DICOM" format	25 instances, 7014 images	data package	OsiriX MD; Frame-by-frame review and measurements: Onis™ 2.5.1.6 Free (DigitalCore Co. Ltd., Tokyo, Japan)

Amos Paper Figures and Legends

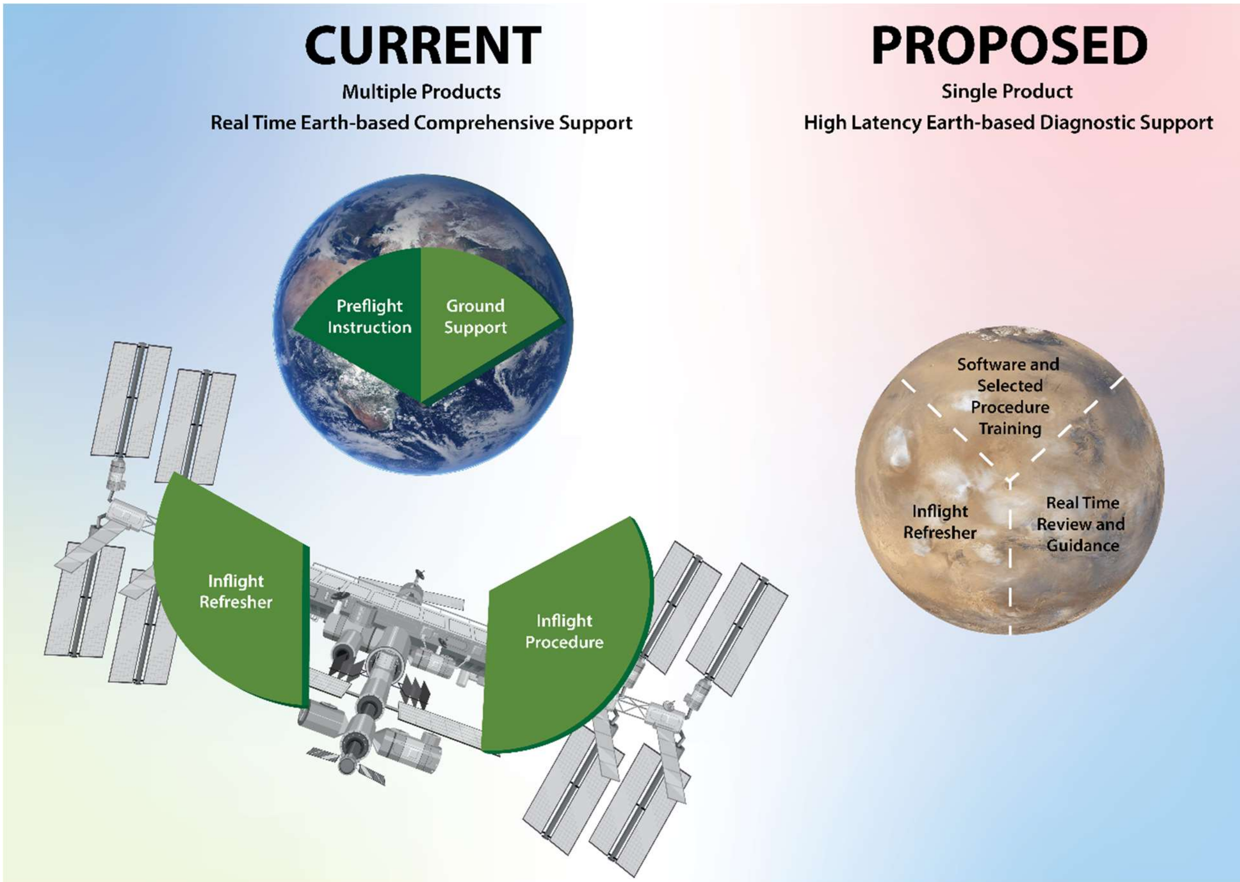

Figure 1: The current ISS training system uses three distinct products during training flow: 1) preflight training materials, 2) infight refresher training, and 3) inflight procedures. Prefight training is 7-16 hours and occurs 6-18 months preflight; crew medical officers (CMOs) receive an additional 4-5 hours of specialized training. The proposed paradigm for exploration uses a single software tool (AMOS or similar) to perform all three functions, streamlining training workflow. With the proposed paradigm, trainees are familiarized with the software and trained only on select procedures that impart highly transferrable skills.

Figure 2: Drs. Jessica Meir (Operator) and Drew Morgan (Subject) conduct an AMOS technical demonstration on the ISS during Expedition 61. (Photo courtesy NASA; ISS062E140375)

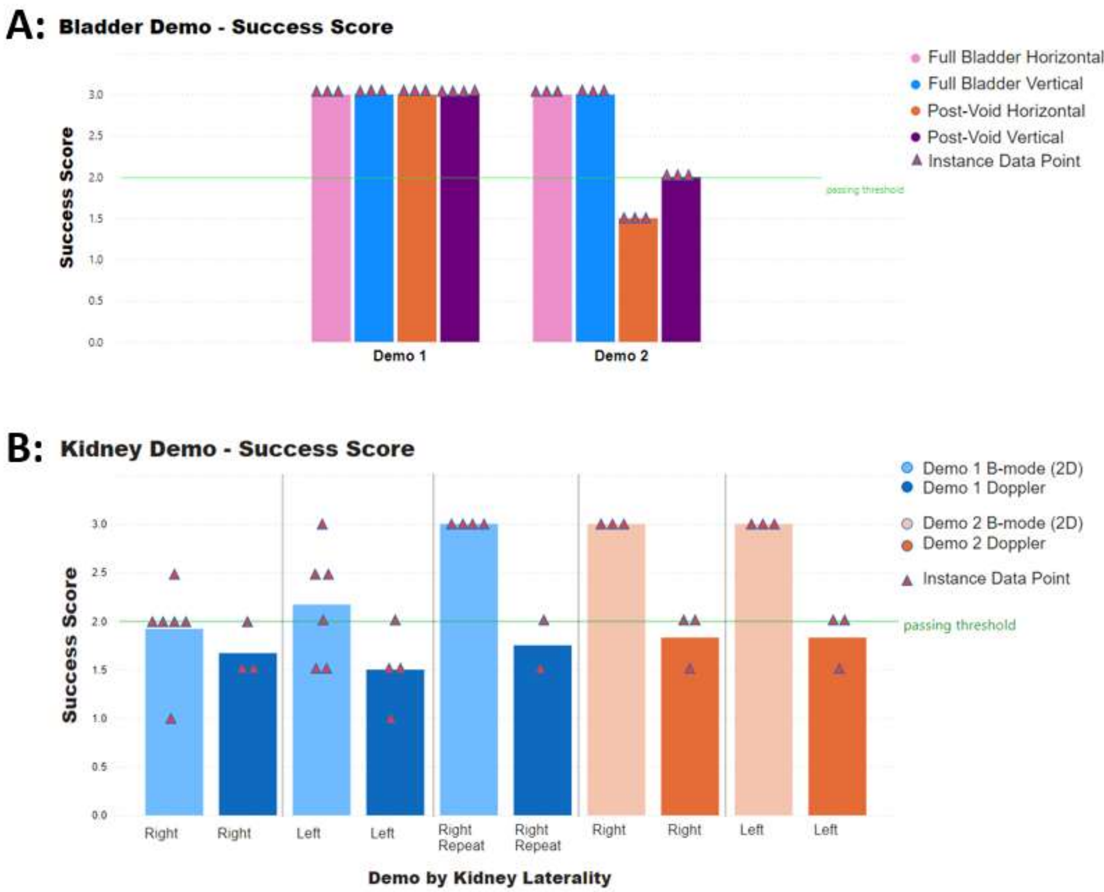

Figure 3: Success scores (0-3) for bladder (Panel A) and kidney (Panel B) exams. Averages of scores for each exam type are indicated by bars and individual imaging instances are shown as triangles. The passing threshold was set at a value of two, which would be considered clinically adequate for diagnosis and measurement; only one imaging instance at two or higher is required to consider the overall exam a success for each Operator.

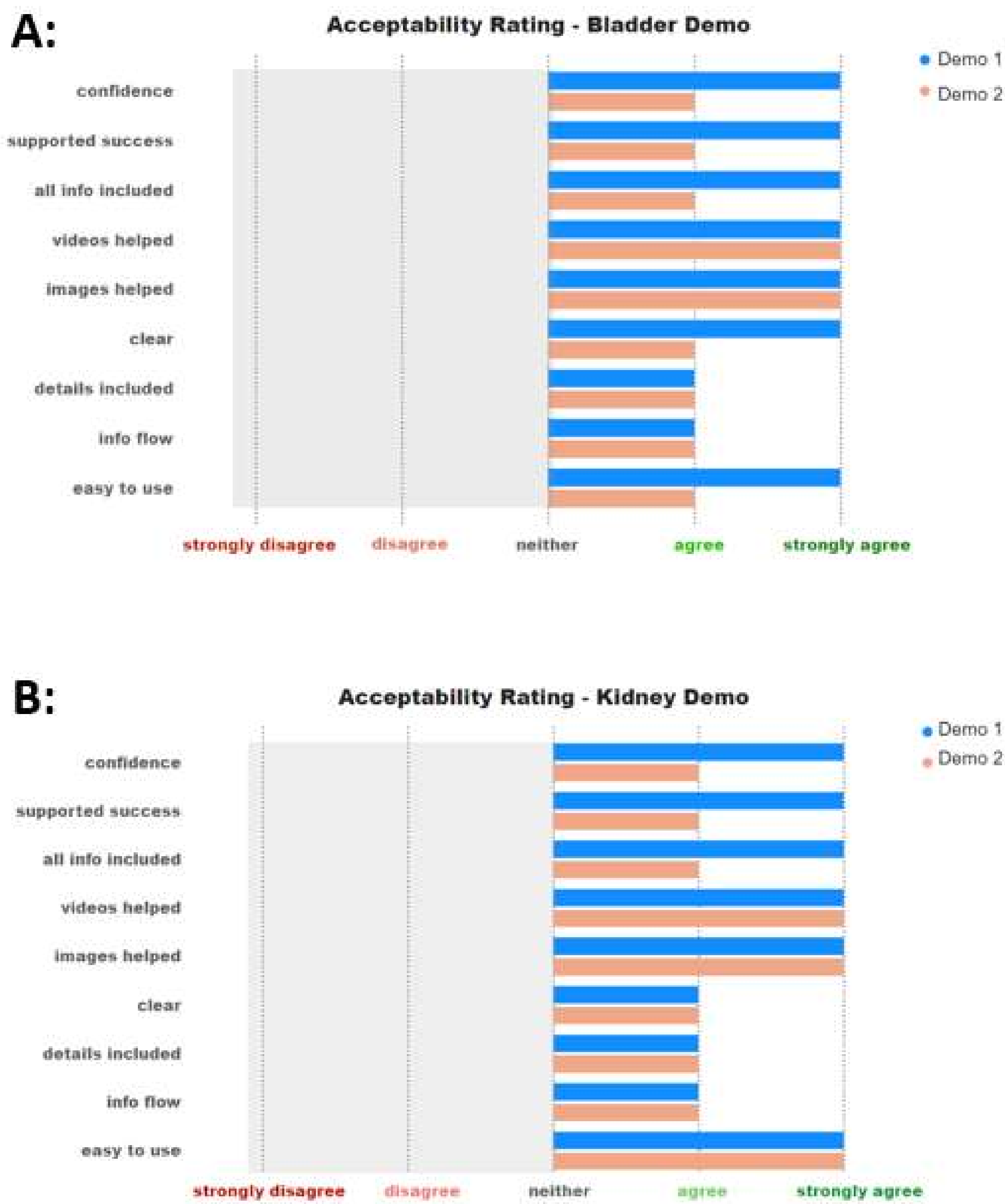

Figure 4: Software acceptability ratings for bladder (Panel A) and kidney (Panel B) AMOS modules; Individual ratings are shown for Operators of both demonstrations.