Extended Reality Applications for Space Health

Mahdi Ebnali; Phani Paladugu; Christian Miccile; Sandra Hyunsoo Park; Barbara Burian; Steven Yule; Roger D. Dias

INTRODUCTION: Spaceflight has detrimental effects on human health, imposing significant and unique risks to crewmembers due to physiological adaptations, exposure to physical and psychological stressors, and limited capabilities to provide medical care. Previous research has proposed and evaluated several strategies to support and mitigate the risks related to astronauts' health and medical exploration capabilities. Among these, extended reality (XR) technologies, including augmented reality (AR), virtual reality (VR), and mixed reality (MR) have increasingly been adopted for training, real-time clinical, and operational support in both terrestrial and aerospace settings, and only a few studies have reported research results on the applications of XR technologies for improving space health. This study aims to systematically review the scientific literature that has explored the application of XR technologies in the space health field. We also discuss the methodological and design characteristics of the existing studies in this realm, informing future research and development efforts on applying XR technologies to improve space health and enhance crew safety and performance.

KEYWORDS: space health, space medicine, extended reality, augmented reality, virtual reality, mixed reality, immersive technology.

Ebnali M, Paladugu P, Miccile C, Park SH, Burian B, Yule S, Dias RD. *Extended reality applications for space health*. Aerosp Med Hum Perform. 2023; 94(3):122–130.

For over 60 yr of human space exploration, the National Aeronautics and Space Administration (NASA) and other space agencies have shown a growing interest and dedicated tremendous effort to advancing the space exploration field. Keeping astronauts safe and healthy throughout missions has been a major area of focus since the earliest days of space exploration.[29] These goals are even more critical when considering the recent global efforts of governmental agencies and private companies to enable commercial spaceflights for non-astronaut travelers. Understanding the effects of spaceflight on human physiology and developing mitigation strategies to protect space travelers' health is crucial as future missions move beyond low Earth orbit to lunar explorations and into deep space destinations.[19]

The space environment imposes unique physical and psychological challenges for crewmembers, affecting their performance during routine activities in space, and creating unprecedented levels of health-threatening hazards and potential accidents.[11] An extensive body of literature has already shown that long-term exposure to microgravity conditions significantly affects spatial orientation, sensorimotor coordination, and neurophysiological adaptive responses.[7,45] Previous studies have also reported deleterious psychological effects of isolation and confinement during space missions. Furthermore, delays and disruptions in the communication between mission control and the spacecraft can create additional challenges to crewmembers,[26,48] especially during medical events, compromising crew health and capacity and posing a significant risk to mission success.[6,42]

Previous research has proposed and evaluated several strategies to support and mitigate the inherent risks related to space health and medical exploration capabilities. Effective clinical tools and medical training have been identified as critical requirements for space explorations,[21] especially for long-duration missions. The NASA human research roadmap identifies critical gaps in current knowledge in the areas of medical decision-making and crew clinical skills required to enable extended missions and/or autonomous operations.[1,31] Several researchers are investigating different types of interventions, including clinical decision support systems and

From the STRATUS Center for Medical Simulation, Boston, MA, USA.

This manuscript was received for review in June 2022. It was accepted for publication in December 2022.

Address correspondence to: Mahdi Ebnali, Ph.D., 10 Vining St., Boston, MA 02115, USA; mebnali-heidari@bwh.harvard.edu.

DOI: https://doi.org/10.3357/AMHP.6131.2023

medical training to support astronauts during medical event management.[42] High-fidelity simulation, for example, has emerged as an effective methodology for educating, practicing, and evaluating team performance in aviation, spaceflight, and medicine.[25] However, the physical nature of simulation and associated expertise and equipment required to run real-time training programs limits scalability and portability. More deployable and scalable solutions are required to train and maintain clinical competency in operational settings, particularly in remote and resource-constrained environments such as space.

Extended reality (XR) technologies have increasingly been adopted for training and operational support in both terrestrial[30,33] and space settings,[35,44] and a growing number of studies have shown the potential for space crews to enhance their operational and behavioral skills using these immersive technologies.[24,35,44] XR refers to a wide range of technology that blends the physical and the digital worlds in a reality-virtuality continuum. Experiences in which graphics are overlaid onto video streams of the physical world are defined as augmented reality (AR), and experiences that present a fully digital experience are known as virtual reality (VR). Mixed reality (MR) covers experiences between these two extremes.[2]

Research addressing the use of XR for space health applications is relatively scarce; however, a few promising preliminary findings have been reported in the past decades on the utility and effectiveness of XR technologies for promoting space health. The aim of the present study was to systematically review the scientific literature and describe how XR technologies, including AR, VR, and MR, have been applied to the space health field. Additionally, we will discuss the methodological and design characteristics of the existing studies in this realm, informing future research and development efforts on applying XR technologies to improve space health and enhance crew safety from low Earth orbit, through lunar explorations, to deep space missions such as Mars.

METHODS

This study was carried out according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines.[38]

Data Source

In May 2021, a literature search was conducted within the following databases for studies published between January 1900 and June 2021: MEDLINE (PubMed), the IEEE Xplore Digital Library, the Association for Computing Machinery (ACM) Digital Library, PsycINFO (Ovid), EMBASE, and the Web of Science Core Collection.

Search Strategy

We developed a comprehensive search strategy to retrieve all studies with a focus on XR technologies in the space health domain. We adapted the Medical Subject Headings (MeSH) terms and keywords from our MEDLINE search strategy to the other databases according to the specific syntax required for each database. No database filters were used and the search query was applied to both title and abstract. To ensure literature saturation, we also hand-searched the reference list of all included articles for other potential inclusions. We used the following search query: (astronauts OR "space flight" OR "space research" OR spacecraft OR "extraterrestrial environment" OR "extravehicular activity" OR hypogravity OR moon OR mars OR astronaut* OR cosmonaut* OR orbit* OR gravity* OR microgravity* OR "space mission" OR "space exploration") AND ("virtual reality" OR "augmented reality" OR "extended reality" OR "mixed reality" OR immersive).

Eligibility Criteria and Selection Process

Only peer-reviewed original studies were included. We excluded studies that were not peer-reviewed original full manuscripts, not space-related, not healthcare-related, not XR-related, and not in English. All studies retrieved by the search strategy were imported into a web-based systematic review management platform (Covidence). Using the Covidence platform and in pairs, four researchers independently screened titles and abstracts based on the eligibility criteria. In case of disagreement, another researcher made the final decision. Subsequently, the same four researchers read and screened full-text articles for inclusion/exclusion, and another researcher solved disagreements. The PRISMA flow diagram showing screening and selection results is shown in **Fig. 1**.

Data Extraction

A data extraction form was designed in the Covidence platform and three independent researchers extracted data, in pairs, from eligible studies. A fourth researcher compared the data extracted from each pair and solved disagreements. If necessary, a fifth researcher was consulted. The Covidence dataset was extracted in CSV format to allow descriptive data analysis. The following fields were extracted from all the studies: title, methods, study design, sex and age of participants, the total number of participants, type of participants, medical specialty, clinical condition, study setting, XR modality, primary purpose, design features, addressed issues and limitations of XR tools, usefulness and usability measures assessment, type of XR device, technology acceptance assessments, objective, and subjective assessment metrics.

Data Synthesis and Quality Assessment

We conducted a qualitative narrative synthesis of all included studies, providing a descriptive analysis based on study design and setting, population, type of medical condition and specialty, type of XR technology, and measurements. Four independent researchers, in pairs, evaluated the methodological quality of all the studies using the validated Medical Education Research Study Quality Instrument (MERSQI),[10] which is a 10-item instrument that assesses 6 domains of research quality (study design, sampling, data type, validity of assessments, data analysis, and outcomes). Each domain

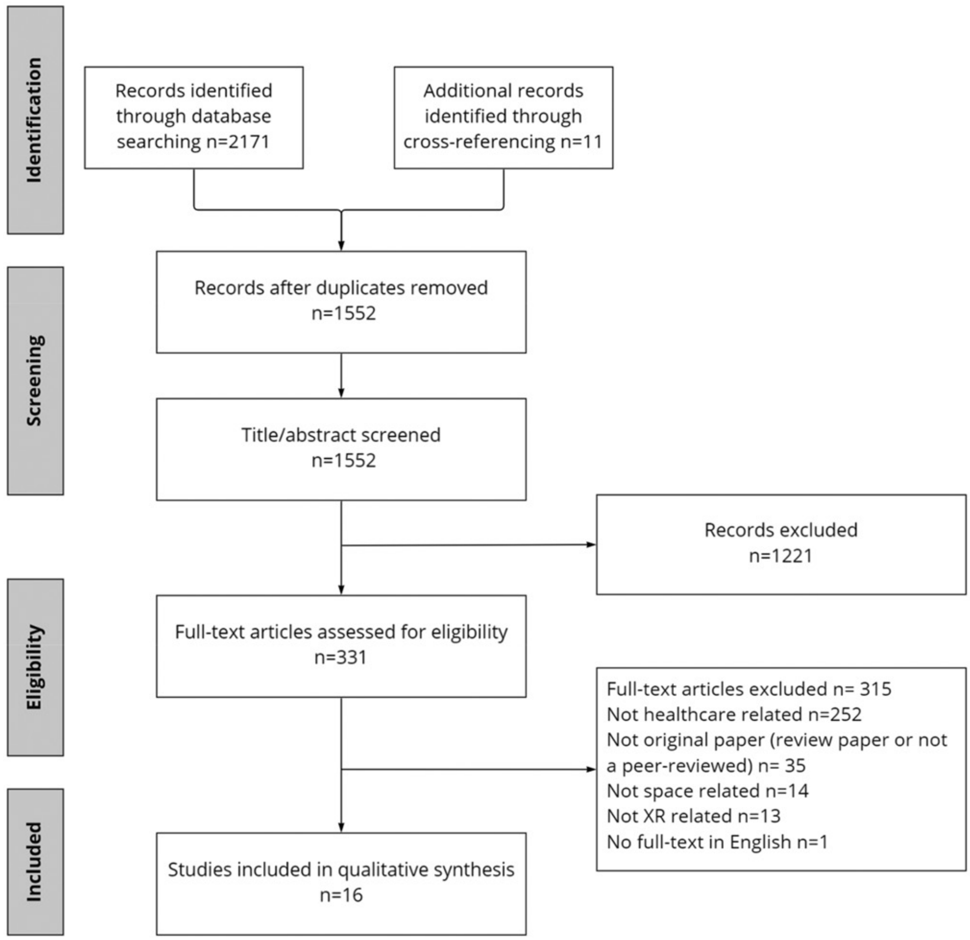

Fig. 1. PRISMA flow diagram.

receives a score from 0 to 3 for a maximum total score of 18. We calculated and reported the mean MERSQI score, based on the two authors' individual ratings, for each included study, as well as the mean and standard deviation MERSQI score among the studies.

RESULTS

A total of 2182 studies were imported for screening and 630 duplicates were removed. We screened 1552 studies based on title and abstract, from which 331 were assessed for eligibility based on the full text. After the full-text screening, 315 studies were excluded and a total of 16 studies were included in the final systematic review. **Fig. 2** shows the distribution of the studies based on the date of publication.

Study Design and Setting

Four studies (25%) were pre/post-test, four (25%) case report or descriptive, three (19%) cohort, two (13%) randomized control trials, one (6%) nonrandomized control trial, one (6%) cross-sectional, and one (6%) qualitative study. All studies (N = 16, 100%) were conducted in simulated settings and no study was conducted in a real space environment. Among the simulation-based studies, 10 studies (62.5%) performed their experiments in an intravehicle setting, 4 studies (25%) explored an extravehicle setting, and 6 studies (37.5%) included microgravity simulations.

Population

While three studies did not mention participants, the remaining 13 studies included in aggregation a total of 777 participants, representing astronauts, clinicians, engineers, and students. More specifically, the participants were: physicians in five studies (31%), engineers in four studies (25%), astronauts in two studies (12.5%), flight surgeons in two studies (12.5%), other healthcare professionals in two studies (12.5%), other healthcare students in two studies (12.5%), and medical students in one study (6.2%). Moreover, most of the studies (N = 11, 68.8%) included other types of participants such as analog astronauts, family members of astronauts, master students, and nonspecified students.

Medical Conditions

Medical specialties were classified into two main groups: surgical and nonsurgical. The majority of studies developed and/or evaluated XR technologies involving nonsurgical specialties (N = 12, 79%) and only 4 studies (25%) evaluated XR

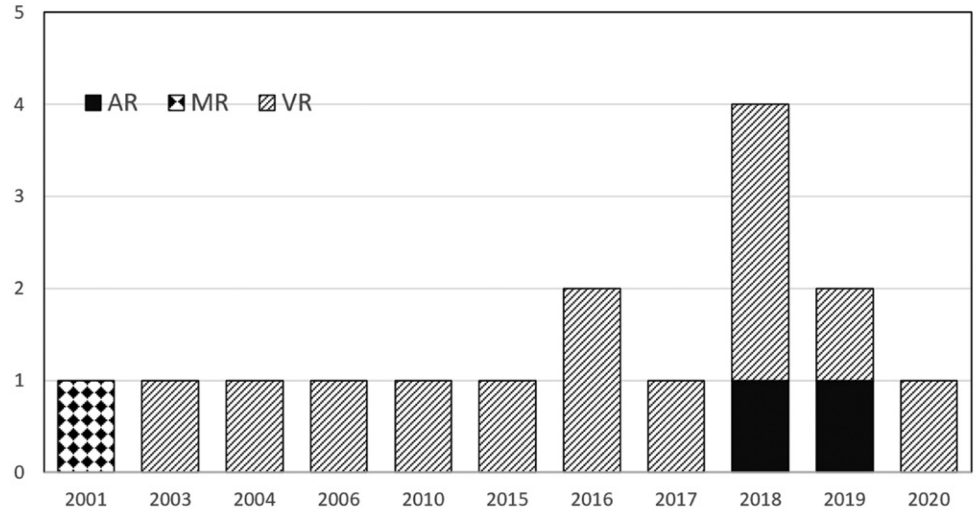

Fig. 2. Number of studies by year of publication, grouped by type of XR tools (VR, AR, and MR).

applications in surgical specialties. The nonsurgical specialties were: three (25%) in Neurology, three (25%) in Rehabilitation Medicine, two (15.6%) in Radiology, two (15.6%) in Emergency Medicine, and two (15.6%) in Psychiatry. **Table I** and **Table II** also show the details of the studies separated based on surgical and nonsurgical applications.

XR Technology

Among all included papers, 13 studies (81.2%) used VR technology, two studies (12.5%) used AR, and one study (6.2%) reported an MR application.

Hardware. Based on the technological embodiment continuum proposed by Flavián et al.,[16] as shown in **Table III**, we categorized XR hardware used in these studies into four groups: stationary external devices (e.g., desktops, fixed displays), portable external devices (e.g., smartphones), wearable devices (e.g., HDM), and implanted devices (e.g., microchips or smart contact lenses).

Table IV also shows the types of XR devices used in the studies. Most studies used either PC/projector tools or custom VR environments. Oculus Rift and Microsoft HoloLens were the only two commercially available tools used in these studies.

Measurements

We extracted the type of measures these studies used to evaluate XR applications. The measures were categorized in the following domains:[5]

1. Self-report: Out of 16 studies, 9 (56.2%) used various types of self-reported tools, including the orientation preference questionnaire,[8,9] system usability score (SUS),[28] simulator sickness questionnaire,[8,47] usability interview,[28] usability structured questionnaire,[29] trainer's rating of instructional design method,[28] NASA Task Load Index (TLX),[15,17] Short Stress State Questionnaire,[15] social connectedness and satisfaction questionnaire,[17,46] and expert scoring questionnaire.[29]
2. Observation-based: Task performance also was measured in two studies[15,34] using observational methods. They focused on performance accuracy and time until completion of a task. In one study also, the reviewers measured the number of VR-based laparoscopic tasks that were successfully completed and the percentage of task attempts.[39]
3. Knowledge tests: Only two studies (12.5%) used knowledge tests as a metric to evaluate the effectiveness of an intervention. Finseth et al. used a written quiz for wayfinding and emergency response abilities of users as a result of a VR-based intervention for spaceflight hazard training during a graduated stress exposure condition.[15] Limbu et al. also evaluated users' knowledge after practicing with a sensor-based AR system designed for aircraft maintenance, medical imaging, and astronaut training.[29]
4. Motion-Kinematic: a wide range of metrics associated with human motion and kinematics was used to evaluate the effectiveness of XR applications or to modify and customize the applications: hand movement,[3,27] waist trajectories,[27] torso alignment and force transducer data,[20] path length,[39] knee extension and joint motion velocity,[47] stride length and step length symmetry,[14] center of pressure, body angle, and orientation.[9] For example, the hands' position and motion were measured using magnetic trackers and sensor gloves to evaluate a surgical simulator developed for astronaut training.[3]
5. Physiological data: In general, only three studies (18.8%) used physiological metrics, including EEG,[27,47] heart rate variability, and blood pressure.[15]
6. System-related data: Although we only included studies that collected data from human participants, a few studies also reported on metrics associated with system performance, which can directly or indirectly affect the human experience in using XR applications. Barnes et al. used software features such as integration accuracy and update rate to discuss the feasibility of implementation of a surgical training system for astronauts.[3] Other studies[12,20] involved measurement of

Table I. Summary of the Studies Included in This Review With a Focus on XR-Based Surgical Applications.

REF (YEAR)	NO. PARTICIPANTS	STUDY DESIGN	XR TECHNOLOGY	MEDICAL SPECIALTY	CLINICAL CONDITION	SETTING	ASSESSMENT METRICS	MERSQI	PRIMARY PURPOSE
Ross[43] (2001)	Not reported	Case report—descriptive	MR/PC-Monitor	Surgical	Elective/Chronic	Spacecraft	Not reported	9.5	Telemedicine
Barnes et al.[3] (2003)	Not reported	Case report—descriptive	VR/PC-Monitor	Surgical	Elective/Chronic	Other	Kinematics/Others	6.6	Clinical training/telemedicine
Harnett et al.[20] (2004)	20	Case report—descriptive	VR/VR laparoscopy simulator	Surgical	Not specified	Microgravity	Observation based/Kinematics/Others	8	Clinical training/telemedicine
Panait et al.[39] (2008)	14	Pre-test/post-test	VR/VR laparoscopy simulator	Surgical	Elective/Chronic	Microgravity	Observation-based/Kinematics/Others	10.75	Usability/human factors

system execution time (surgery time), video streaming rate, video recording time, and system latency to evaluate an XR-based data collection system for human performance.

Primary Purpose of Space Health XR Applications

1. Usability: The primary purpose of most of the studies ($N = 10$, 62.5%) was usability evaluation of XR applications in which they tested usability aspects of nine VR interventions and one AR intervention.
2. Virtual diagnosis/therapeutics: Seven studies (43.8%) also evaluated XR applications with a focus on virtual diagnosis/therapeutic aspects such as motion sickness and physiological adaptations to microgravity[9,12,14] and stress/isolation management.[15,17,46] All of the virtual diagnosis/therapeutic applications were evaluated in VR setups.
3. Medical training: Six studies (37.5%) performed research on evaluating XR applications as clinical training tools using VR[3,20,37] and AR.[9,12,29]
4. Telemedicine: Telemedicine was the main purpose of four studies (25%); one of these studies used MR[43] and three studies used VR[3,17,37] in their experiments.

Methodological Quality

The mean (SD) total MERSQI score of the studies was 9.36 (2.7) out of 18, ranging from 3.5 to 13.2. The average MERSQI scores for individual studies are listed in Table I and Table II.

DISCUSSION

This study provides a systematic review of the literature regarding the application of XR-based space health technologies, distilling evidence from a wide range of interventions, type of XR technology, medical conditions, application purposes, experimental setups, and assessment metrics. We identified 16 articles reporting the use of XR technologies in space-health applications which were all conducted in simulated scenarios. Practical methodological limitations were found in the studies which relied on small sample size and weak study designs, adversely influencing the consistency and quality of the studies' conclusions. Additionally, the low MERSQI scores across all studies demonstrate the overall poor quality of the evaluation methods used. In the following sections, we discuss the main findings, strengths, and limitations of the studies and outline suggested directions for future research and development of XR applications for space health.

Participants

Various types of participants, including physicians, engineers, astronauts, flight surgeons, and students, were recruited in the studies. However, only two studies included astronauts as research subjects. The findings from these studies are still useful since more non-astronaut travelers are expected to participate in space explorations. Future studies should investigate how design considerations of space health XR applications need to be tailored for specific types of space travelers.

Table II. Summary of the Studies Included in This Review With a Focus on XR-Based Nonsurgical Applications.

REF (YEAR)	NO. PARTICIPANTS	STUDY DESIGN	XR TECHNOLOGY	MEDICAL SPECIALTY	CLINICAL CONDITION	SETTING	ASSESSMENT METRICS	MERSQI SCORE	PRIMARY PURPOSE
Yoshimitsu[47] (2010)	9	Pre-test/ post-test	VR/Not reported	Rehabilitation	Other	Microgravity	Physiological/Kinematics	12	Usability/human factors
Chen et al.[8] (2015)	32	Pre-test/ post-test	VR/PC-Monitor	Neurology	Emergency/ Acute	Other	Self-report/Kinematics	9	Usability/human factors
Wu et al.[46] (2016)	53	Cross sectional	VR/Custom VR	Psychology	Elective/ Chronic	Other	Self-report	8	Usability/human factors
Eikema et al.[14] (2016)	20	Non-RCT	VR/Custom VR	Rehabilitation	Other	Other	Kinematics	11.5	Usability/human factors
Chen et al.[9] (2017)	34	RCT	VR/Oculus Rift	Neurology	Elective/ Chronic	Other	Self-report/Kinematics	11	Usability/human factors
Limbu et al.[28] (2018)	142	Cohort	AR/HoloLens	Radiology	Elective/Chronic	Spacecraft	Self-report	9.5	Clinical Training/Telemedicine
Galunder et al.[17] (2018)	12	Cohort	VR/Not reported	Psychology	Emergency/Acute	Other	Self-report/Others	9.25	Stress Management
Finseth et al.[15] (2018)	20	RCT	VR/Custom VR	Emergency Medicine	Elective/Chronic	Microgravity	Self-report/Knowledge/ Observation/Physiological	13.2	Usability/human factors
Del Mastro et al.[12] (2018)	3	Case report	VR/Oculus Rift	Rehabilitation	Elective/Chronic	Other	Self-report/Others	4.5	Clinical Training/Telemedicine
Limbu et al.[29] (2019)	398	Cohort study	AR/HoloLens	Radiology	Elective/Chronic	Spacecraft	Self-report/Knowledge Tests	12	Clinical Training/Telemedicine
Li et al.[27] (2019)	20	Pre-test/ post-test	VR/PC-Monitor	Neurology	Emergency/Acute	Other	Self-report/Physiological/ Kinematics	11.5	Usability/human factors
Nasser et al.[37] (2020)	Not Reported	Qualitative	VR/Custom VR	Emergency Medicine	Emergency/Acute	Spacecraft	Others	3.5	Clinical Training/Telemedicine

Study Designs

The majority of the studies lack rigor in the design. More than 80% of the studies did not include control groups and only two studies (13%)[9,15] used randomized controlled experiments. Furthermore, no study followed up participants to measure outcomes longitudinally and none of the training research assessed long-term knowledge and/or skills retention.

Most of the XR tools (75%) used in the studies were non-immersive (2D displays and nonwearable devices). Therefore, it can be difficult to generalize findings from these studies regarding the actual effectiveness of immersive XR technologies in space health. Moreover, as no full validation strategies were outlined, it is unclear whether the proposed XR applications published to date are of true value in improving astronauts' health and space mission safety. Particularly for studies focused on training aspects of XR applications in space health, it is critical to investigate the transfer of learning into real-life settings and the progression of knowledge and skills over time through proper longitudinal studies. Future investigations of XR-based educational applications for astronauts' training in medical event management could result in significant advances in our understanding of the suitability of this technology for supporting medical learning and clinical practice in space.

Primary Purpose of XR Applications

The majority of the studies aimed to investigate the usability of XR applications for space health, where virtual therapeutic techniques, clinical training, and telemedicine were the three main focuses of these applications. Surprisingly, XR-based decision support and procedural clinical guidance was not investigated in any of the studies. XR technologies, particularly AR, have been widely used as a clinical support/guidance tool for terrestrial applications.[4,22,40] Due to minimal interaction with ground controllers, communication limitations, and prolonged nature of the mission, just-in-time (JIT) training and real-time clinical guidance systems are critical for long deep-space missions. NASA and the Translational Research Institute for Space Health (TRISH) have highlighted the importance of the development of augmented clinical tools (ACTs) to support astronauts by providing planned composite medical education and real-time care delivery guidance systems during spaceflights.[34] XR could be used to seamlessly deliver ACTs, helping astronauts to effectively manage medical events in space. The Augmented Reality Coach (AR-Coach) is an example of a clinical guidance application in which a virtual coach system guides the crew in real time on how to perform point-of-care ultrasound during medical emergencies in space.[32]

Medical Specialty

Most of the studies focused on nonsurgical conditions. Although the frequency of nonsurgical medical conditions is higher than surgical events in space travel, surgical care should be an essential component of space health solutions for developing effective and comprehensive medical exploration capabilities to ensure the health and safety of space travelers. Although the feasibility of performing certain surgical

Table III. Group of XR Hardware Used in the Studies (Based on Human Technological Embodiment Continuum[37]).

XR HARDWARE GROUPS	NUMBER OF STUDIES (%)
Stationary external devices[3,8,14,15,17,20,27,39,43,46,47]	11 (68.8)
Wearable devices[9,12,15,28,29]	5 (31.2)
Portable external devices[37]	1 (6.2)
Implantable devices	0

procedures in space has already been determined,[13] only a few relevant investigations have taken place due to challenges associated with the availability of clinical expertise and both diagnostic and operational resources during space missions. Considering the promising application of XR technologies in surgery in terrestrial settings,[40] it is critical to investigate potential opportunities and challenges of immersive technologies to ensure safe and effective surgical care in space. For instance, AR can be used to provide real-time or near-real-time support to flight surgeons from Earth or other space stations, enhancing collaboration in telesurgery. Providing a unified view of the surgical field by superimposing digital information onto AR glasses helps flight surgeons and the medical expert on Earth or another space station effectively communicate in managing medical emergencies. AR information also reduces looking back and forth between different sources of information. Virtual projections of subsurface anatomy allow flight surgeons to identify, anticipate, and avoid critical structures before they are exposed. Future studies need to establish baseline concepts of functionalities and design features of XR applications that could provide surgical support during spaceflight, from the preoperative to the postoperative period. Surgical support can be provided as a real-time guide or surgical training program which prepares space travelers for various surgical procedures.

XR-Based Medical Training

Most of the studies did not clearly describe the learning theories they used to guide the design or application of XR in space health. Generalizability is not only reduced by the difference in XR settings, it has also been affected by divergent underlying learning theories. To date, it is unclear if the use of XR technology in training astronauts is likely to contribute to astronauts' safety and mission success. Although ground-based evidence supports that as training tools and content become more engaging and reliable, more learning outcomes may be expected, and patients will ultimately benefit,[36] there is still a lack of comprehensive theoretical guidance for developing XR-based medical training curricula. Most of the investigations on Earth or in space related to XR medical training have been focused on the acquisition of technical skills. However, it is crucial to incorporate nontechnical skills into XR-based medical training curricula, particularly for space applications where errors can pose a significant risk to mission success. For instance, training situational awareness in high-risk environments is reported as a critical component of operation safety, but is lacking in XR-based medical educational curricula for space.[4]

There is not enough evidence to inform the design of suitable learning activities with XR systems, where knowledge and skill development could be integrated into the astronauts' capabilities during space missions. Therefore, further research should illuminate minimum requirements of XR systems' and models' designs, features, and functionalities, as well as how to effectively use them for healthcare education.

Table IV. Type of XR Devices Used in the Studies.

XR DEVICE	NUMBER OF STUDIES (%)
Oculus Rift[9,12]	2 (12.5)
HoloLens[28,29]	2 (12.5)
Custom VR environments[14,15,37]	4 (25)
PC/monitor or projector[3,8,27,43]	4 (25)
VR laparoscopy simulators[20,39]	2 (12.5)
Not reported[17,47]	2 (12.5)

XR Hardware, Devices, and Modalities

A high variability was observed in XR tools used in the studies; however, most of them did not use portable or wearable tools such as head-mounted displays. Considering the publication year of the papers, it is not surprising that only four studies used HMDs available in the market. Wearable VR devices emerged with the introduction of the PC-connected Oculus Rift prototype in 2010 and progressed rapidly over the course of the last decade. The cost of VR headsets has dropped dramatically and computer hardware capable of running these headsets is virtually mainstream. XR systems are now part of affordable standalone AR and VR headsets, which are expected to be cheaper and lighter in upcoming years. Future studies need to investigate XR applications deployed in these new hardware technologies.

Evaluation Measures

Alongside the poor methodological quality of the studies included in this review, there was a large variability in assessment metrics, which compromises the generalizability and reliability of results obtained from these studies. Methodological quality assessed by the MERSQI score was low for most studies, indicating that several aspects of research in this field can be improved. Although some studies employed validated tools for assessing the effectiveness of XR applications, most did not follow a systematic evaluation approach. Further studies to strengthen the existing evidence would require assessing behavioral and patient-related outcomes through more rigorous and systematic approaches. Future studies should include larger sample sizes and standard validated measurement methods.

In addition to space, the findings from this systematic review can be used to support the design and development of XR applications for terrestrial settings such as telemedicine, medical training, and clinical decision support in remote and austere environments. The paradigm of health care delivery has shifted dramatically from hospital-based to homecare and telehealth. XR technology in these contexts has the potential to provide useful resources. This study may also help researchers and developers to gain a better understanding of the challenges in developing XR-based solutions for improving patient care, particularly in low-resource environments.

Limitations

Several limitations should be considered when interpreting the results of this systematic review. Given the rapid growth of XR technology in recent years, it is likely that research involving certain XR applications in space health has not yet been published or is under patent/copyright restrictions, precluding their inclusion in this review. Clinical utility and validity of XR applications included in this review were limited by high variability in sample size, design of the study, medical conditions, and type of XR tools. Included studies also covered a wide range of main purposes, which may limit the specific scope of findings considering high variability in XR applications' primary purposes. Future review studies should target a narrower concentration of main purposes for XR applications. This relatively new research field must build on more validated metrics to investigate the impacts of XR tools on astronauts' health and performance. In addition, further research should include the analysis of other moderating variables in order to provide a better understanding of the impact of XR on space health.

One of the other major drawbacks of most of the included papers is that they were not clear about the design and prototyping stage of the XR systems and how they incorporated astronauts' needs and feedback into the design process of the XR applications. Human-centered design has emerged as a promising and versatile approach to engage users in the design and adaptation of healthcare digital systems to better meet clinicians' and patients' needs, resulting in fewer usability issues and human errors, plus a higher adoption rate of technologies.[18,41] In terrestrial medicine, for example, unsatisfactory clinicians' perceptions of a system's content and design are associated with less successful technology implementations.[23]

Conclusion

To our knowledge, this is the first study that systematically reviews the existing literature on XR applications for space health. We reviewed applications of XR technologies that focused on space health and encompassed a broad range of experimental design, XR tools, medical specialties, clinical conditions, space-related experiment setups, and assessment metrics. The limited number of the studies and wide variation in the design of the studies, medical conditions being studied, and primary purpose of the XR applications pose substantial challenges to reporting compelling evidence in support of successful implementations of XR in space health. There was a lack of consistently positive outcomes and high-quality studies for all XR modalities.

XR technology is in the early stages of application within space health, but it has enormous potential for supporting astronauts and non-astronaut space travelers during medical event management. Real-world applications of XR in space health not only require designing pertinent functionality and features, but also identifying appropriate clinical guidelines and training methodologies to better address the needs of astronauts and other space travelers during medical event management in space.

ACKNOWLEDGMENT

Financial Disclosure Statement: This work was funded by the Translational Research Institute for Space Health (TRISH) through NASA Cooperative Agreement NNX16AO69A (Grant # T0506). The authors declare no conflict of interest.

Authors and Affiliations: Mahdi Ebnali, Ph.D., Department of Emergency Medicine, Mass General Brigham, Harvard Medical School, Boston, MA; Phani Paladugu, Christian Miccile, Sandra Hyunsoo Park, M.D., and Roger D. Dias, M.D., Ph.D., Human Factors and Cognitive Engineering Lab, STRATUS Center for Medical Simulation, Brigham and Women's Hospital, Boston, MA, USA; Sandra Hyunsoo Park and Roger D. Dias, Department of Emergency Medicine, Harvard Medical School, Boston, MA, USA; Barbara Burian, Ph.D., Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA, USA; and Steven Yule, Ph.D., Department of Clinical Surgery, University of Edinburgh, Edinburgh, Scotland, United Kingdom.

REFERENCES

1. Alwood JS, Ronca AE, Mains RC, Shelhamer MJ, Smith JD, Goodwin TJ. From the bench to exploration medicine: NASA life sciences translational research for human exploration and habitation missions. NPJ Microgravity. 2017; 3(1):5.
2. Andrews C, Southworth MK, Silva JNA, Silva JR. Extended reality in medical practice. Curr Treat Options Cardiovasc Med. 2019; 21(4):18.
3. Barnes B, Menon AS, Mills R, Bruyns CD, Twombly A, et al. Virtual reality extensions into surgical training and teleoperation. In: 4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine; 24-26 April 2003; Birmingham, UK. New York: IEEE; 2003:142–145.
4. Barsom EZ, Graafland M, Schijven MP. Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc. 2016; 30(10):4174–4183.
5. Barteit S, Lanfermann L, Bärnighausen T, Neuhann F, Beiersmann C. Augmented, mixed, and virtual reality-based head-mounted devices for medical education: systematic review. JMIR Serious Games. 2021; 9(3):e29080.
6. Billica RD, Simmons SC, Mathes KL, McKinley BA, Chuang CC, et al. Perception of the medical risk of spaceflight. Aviat Space Environ Med. 1996; 67(5):467–473.
7. Blaber E, Marçal H, Burns BP. Bioastronautics: the influence of microgravity on astronaut health. Astrobiology. 2010; 10(5):463–473.
8. Chen W, Chao JG, Chen XW, Wang JK, Tan C. Quantitative orientation preference and susceptibility to space motion sickness simulated in a virtual reality environment. Brain Res Bull. 2015; 113:17–26.
9. Chen W, Chao JG, Zhang Y, Wang JK, Chen XW, Tan C. Orientation preferences and motion sickness induced in a virtual reality environment. Aerosp Med Hum Perform. 2017; 88(10):903–910.
10. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle–Ottawa Scale-Education. Acad Med. 2015; 90(8):1067–1076.
11. Cucinotta FA. Space radiation risks for astronauts on multiple International Space Station missions. PLoS One. 2014; 9(4):e96099.
12. Del Mastro A, Schlacht IL, Benyoucef Y, Groemer G, Nazir S. Motigravity: a new VR system to increase performance and safety in space operations simulation and rehabilitation medicine. In: Arezes P, editor.

Advances in safety management and human factors. New York: Springer International Publishing; 2018:207–217.

13. Drudi L, Ball CG, Kirkpatrick AW, Saary J, Grenon SM. Surgery in space: where are we at now? Acta Astronaut. 2012; 79:61–66. Erratum in: Acta Astronautica 2014; 93:129 .
14. Eikema DJA, Chien JH, Stergiou N, Myers SA, Scott-Pandorf MM, et al. Optic flow improves adaptability of spatiotemporal characteristics during split-belt locomotor adaptation with tactile stimulation. Exp Brain Res. 2016; 234(2):511–522.
15. Finseth TT, Keren N, Dorneich MC, Franke WD, Anderson CC, Shelley MC. Evaluating the effectiveness of graduated stress exposure in virtual spaceflight hazard training. J Cogn Eng Decis Mak. 2018; 12(4):248–268.
16. Flavián C, Ibáñez-Sánchez S, Orús C. The impact of virtual, augmented and mixed reality technologies on the customer experience. J Bus Res. 2019; 100:547–560.
17. Galunder SS, Gottlieb JF, Ladwig J, Hamell J, Keller PK, Wu P. A VR ecosystem for telemedicine and non-intrusive cognitive and affective assessment. In: 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH); 16-18 May 2018; Vienna, Austria. New York: IEEE; 2018:1–6.
18. Garcia LJ, Pichler RF, Seitz EM, Merino GSA, Do Amaral Gontijo L, Merino EAD. Diagnosis and identification of key issues of usability for reducing medication errors. Strategic Design Research Journal; 2017; 10(1):67–78.
19. Garcia M. International Space Station. 2015 Jan. 12. [Accessed 2021 Dec. 16]. Available from https://www.nasa.gov/mission_pages/station/main/index.html.
20. Harnett BM, Broderick T, Doarn CR, Rafiq A, Muth T, Merrell RC. Dynamic automated data collection for human performance. Journal of Information Technology in Healthcare. 2004; 2(3):175–186.
21. HRR - Gap - TRAIN-02. We need to identify effective methods and tools that can be used to train for long-duration, long-distance space missions. (Previous title: SHFE-TRAIN-02). [Accessed 2021 Dec. 17]. Available from https://humanresearchroadmap.nasa.gov/gaps/gap.aspx?i=310.
22. Hu XS, Nascimento TD, Bender MC, Hall T, Petty S, et al. Feasibility of a real-time clinical augmented reality and artificial intelligence framework for pain detection and localization from the brain. J Med Internet Res. 2019; 21(6):e13594.
23. Jha AK, DesRoches CM, Campbell EG, Donelan K, Rao SR, et al. Use of electronic health records in U.S. hospitals. N Engl J Med. 2009; 360(16):1628–1638.
24. Karasinski JA, Joyce R, Carroll C, Gale J, Hillenius S. An augmented reality/internet of things prototype for just-in-time astronaut training. In: Lackey S, Chen J, editors. Virtual, augmented and mixed reality. New York: Springer International Publishing; 2017:248–260.
25. Kuypers MI. Emergency and wilderness medicine training for physician astronauts on exploration class missions. Wilderness Environ Med. 2013; 24(4):445–449.
26. Landon LB, Slack KJ, Barrett JD. Teamwork and collaboration in long-duration space missions: going to extremes. Am Psychol. 2018; 73(4): 563–575.
27. Li Y, Liu A, Ding L. Machine learning assessment of visually induced motion sickness levels based on multiple biosignals. Biomed Signal Process Control. 2019; 49:202–211.
28. Limbu B, Jarodzka H, Klemke R, Wild F, Specht M. From AR to expertise: a user study of an augmented reality training to support expertise development. J Univers Comput Sci. 2018; 24(2):108–128.
29. Limbu B, Vovk A, Jarodzka H, Klemke R, Wild F, Specht M. WEKIT. One: a sensor-based augmented reality system for experience capture and re-enactment. In: Scheffel M, Broisin J, Pammer-Schindler V, Ioannou A, Schneider J, editors. Transforming learning with meaningful technologies. EC-TEL 2019. Lecture notes in computer science. Springer Cham; 2019; 11722:158–171.
30. Lin JC, Yu Z, Scott IU, Greenberg PB. Virtual reality training for cataract surgery operating performance in ophthalmology trainees. Cochrane Database Syst Rev. 2021; 12:CD014953.
31. Lyndon B. Johnson Space Center. Human health and performance risks of space exploration missions: evidence reviewed by the NASA Human Research Program. Houston (TX): NASA Lyndon B. Johnson Space Center; 2009.
32. Mahdi Ebnali. AR-coach: using augmented reality (AR) for real-time clinical guidance during medical emergencies on deep space exploration missions. In: AHFES; 2022. [Accessed 2022 May 18].
33. Mantovani F, Castelnuovo G, Gaggioli A, Riva G. Virtual reality training for health-care professionals. Cyberpsychol Behav. 2003; 6(4):389–395.
34. Mars K. The Translational Research Institute for Space Health (TRISH). 2017 Nov. 22. [Accessed 2022 May 16]. Available from https://www.nasa.gov/hrp/tri.
35. Montgomery K, Thonier G, Stephanides M, Schendel S. Virtual reality based surgical assistance and training system for long duration space missions. Stud Health Technol Inform. 2001; 81:315–321.
36. Naik VN, Brien SE. Review article: simulation: a means to address and improve patient safety. Can J Anaesth. 2013; 60(2):192–200.
37. Nasser M, Peres N, Knight J, Haines A, Young C, et al. Designing clinical trials for future space missions as a pathway to changing how clinical trials are conducted on Earth. J Evid Based Med. 2020; 13(2): 153–160.
38. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021; 372:n160.
39. Panait L, Merrell RC, Rafiq A, Dudrick SJ, Broderick TJ. Virtual reality laparoscopic skill assessment in microgravity. J Surg Res. 2006; 136(2):198–203.
40. Rahul K, Raj VPMD, Srinivasan K, Deepa N, Kumar NS. A Study on virtual and augmented reality in real-time surgery. In: 2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW). New York: IEEE; 2019:1–2.
41. Ratwani RM, Savage E, Will A, Arnold R, Khairat S, et al. A usability and safety analysis of electronic health records: a multi-center study. J Am Med Inform Assoc. 2018; 25(9):1197–1201.
42. Robertson JM, Dias RD, Gupta A, Marshburn T, Lipsitz SR, et al. Medical event management for future deep space exploration missions to Mars. J Surg Res. 2020; 246:305–314.
43. Ross MD. Medicine in long duration space exploration: the role of virtual reality and broad bandwidth telecommunications networks. Acta Astronaut. 2001; 49(3–10):441–445.
44. Salamon N, Grimm JM, Horack JM, Newton EK. Application of virtual reality for crew mental health in extended-duration space missions. Acta Astronaut. 2018; 146:117–122.
45. Scott JM, Warburton DER, Williams D, Whelan S, Krassioukov A. Challenges, concerns and common problems: physiological consequences of spinal cord injury and microgravity. Spinal Cord. 2011; 49(1):4–16.
46. Wu P, Morie J, Wall P, Ott T, Binsted K. ANSIBLE: virtual reality for behavioral health. Procedia Eng. 2016; 159:108–111.
47. Yoshimitsu K, Shiba N, Matsuse H, Takano Y, Matsugaki T, et al. Development of a training method for weightless environment using both electrical stimulation and voluntary muscle contraction. Tohoku J Exp Med. 2010; 220(1):83–93.
48. Yule S, Robertson JM, Mormann B, Smink DS, Lipsitz S, et al. Crew autonomy during simulated medical event management on long duration space exploration missions. Hum Factors. 2022; 2022: 187208211067575.