Developing and implementing novel techniques during primary space medicine

data systematic reviews.

Andrew Winnard PhD, Nick Caplan PhD, Claire Bruce-Martin MSc, Patrick Swain MSc, Rochelle Velho MBChB, Roberto Meroni PhD, Virginia Wotring PhD, Volker Damann MD, Tobias Weber PhD, Simon Evetts PhD, Joanthan Laws MRes

Word count for abstract: 242

Word count for narrative: 2906

Number of references: 34

Number of tables: 1

Number of figures: 3
Abstract:

Background

The Aerospace Medicine Systematic Review Group was set up in 2016 to facilitate high quality and transparent synthesis of primary data to enable evidence-based practice. The group identified many research methods specific to space medicine that need consideration for systematic review methods. The group has developed space medicine specific methods to address this and trialed usage of these methods across seven published systematic reviews. This paper outlines evolution of space medicine synthesis methods and discussion of their initial application.

Methods

Space medicine systematic review guidance has been developed for protocol planning, quantitative and qualitative synthesis, sourcing grey data, and assessing quality and transferability of space medicine human spaceflight simulation study environments.

Results

Decision algorithms for guidance and tool usage were created based on usage. Six reviews used quantitative methods in which no meta analyses were possible due to lack of controlled trials or reporting issues. All reviews scored the quality and transferability of space simulation environments. One review was qualitative. Several research gaps were identified.

Conclusion
Successful use of the developed methods demonstrates usability and initial validity. The current space medicine evidence base resulting in no meta analyses to be possible shows the need for standardized guidance on how to synthesize data in this field. It also provides evidence to call for increasing use of controlled trials, standardizing outcome measures and improving minimum reporting standards. Space medicine is a unique field of medical research that requires specific systematic review methods.
Introduction

The Aerospace Medicine Systematic Review Group (AMSRG) was set up in 2016 to facilitate high quality and transparent synthesis of primary data to enable evidence based medical practice. The benefits of the group were outlined in the May 2017 Aerospace Medicine and Human Performance journal President’s Page, as developing the knowledge base and improving the quality and value of research while highlighting research gaps to strengthen arguments for funding by demonstrating research needs.

International gold standards for synthesizing traditional medical data are outlined generally in the Cochrane handbook and should be reported to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Individual specialist topic areas tend to establish groups to address their specific research intricacies and practices. The most thorough example of specialist groups are those within Cochrane, however Cochrane groups require substantial recurring funding to setup and maintain databases of centrally quality scored trials, up to date methods, protocol registrations and systematic review publication databases staffed by information specialists. The AMSRG is a good initial step towards high quality regulation whilst funding is limited. All good quality systematic reviews should also conform to PRISMA guidelines. Aerospace is one of the final remaining fields of medicine to do so, a gap the AMSRG is addressing.

AMSRG has identified that space medicine has many research methods and limitations that are specific to the field and impact the systematic review process, such as: small sample sizes; a lack of controlled trials; and common usage of ground-based spaceflight analogues. These aspects necessitate adaptation of systematic review
synthesis methods and often require consideration regarding how findings from different research settings may safely transfer to the operational space medicine environment\textsuperscript{18}. To date, space medicine has relied on individual operational expert opinions or non-standardized, evidence books\textsuperscript{19} that do not adhere to internationally recognized systematic review standards. In addition, the sample sizes of spaceflight and many analog studies are also small compared to those used in terrestrial clinical medicine which can lead to more individual observations. This prevents standardized, transparent, repeatable, and easily updatable syntheses within which risk of bias, certainty and transferability can be addressed. The AMSRG has published seven systematic reviews as a group, five done in collaboration with the European Astronaut Centre on operationally driven topics (highlighted in Table I). During these reviews, space medicine specific methods and decision algorithms have been developed, and trialed, by the AMSRG. This paper provides a summary of these methods, trial usage, discussion for ongoing usage, and development from research, operational and political perspectives.

**Methods**

*Protocol planning and pre-scoping*

The AMSRG follows Cochrane guidance that requires initial planning and documenting methods decisions in a written protocol. Protocols state a clear question, scope, search strategy, inclusion criteria and analysis decisions, such as sub groupings and statistical choices. Importantly, protocols record these decisions before results are available to
reduce potential reviewer bias around methods decisions after synthesis results are available. The protocol stage also helps ensure a manageable scope likely to return the required data to run the statistical methods selected. While Cochrane requires a fully published protocol\(^3\), the AMSRG strongly recommends a protocol and pre-scoping step\(^28\), in which decisions are made on search terms, scope, quality scoring and synthesis methods and quick, pre-scoping searches are performed to check existence of relevant data. For a review to be considered systematic, the protocol and pre-scoping must result in a systematic search strategy using Boolean logic and inclusion criteria that is detailed in the final publication alongside valid quality scoring and synthesis methods to enable the review to be transparent and repeatable. An example search strategy table, to find studies detailing human biomechanical and cardiopulmonary changes due to partial gravity, from a AMSRG and EAC collaborative review is presented in Figure 1.
As space medicine has many sources of data that sit outside common medical journal databases, the AMSRG provides a list of search locations to help locate grey literature, such as technical reports, for inclusion in reviews\textsuperscript{30}. With regards to protocol planning, guides to help with both qualitative and quantitative methods have been made available. These guides have evolved over the years and have been modified by the AMSRG group\textsuperscript{12, 29}.

Deciding main synthesis method
To date, AMSRG has followed a decision algorithm that primarily recommends full quantitative meta-analysis and preference for reviews to be based on controlled clinical trials where possible, following Cochrane quantitative methods. However, based on the initial review questions tackled, it is apparent that controlled trials, repeated studies and use of standardized outcome measures across studies often do not exist or have been poorly reported, making Cochrane based meta-analyses not possible for some research questions. Therefore, alternate acceptable analyses were performed in a hierarchy of effect size analysis, qualitative analyses^{20, 26} (individually or combined with quantitative analysis within an integrated approach)^{9} and finally, if no other method is possible and there is a justified reason for continuing, reporting of significance testing results from included papers^{27}. In extreme cases if no published data exists on a key topic, an empty review could be published to provide comprehensive gap analysis and used to stimulate primary research. However, care should be taken to ensure an empty review is not solely due to an overly specific question^{34}. The methods decision algorithm is presented in Figure 2.
Figure 2 - Methods decision algorithm, with the decisions running down the middle starting from the top, dashed lines show alternate or supplementary decision options.

Effect size analysis converts all reported data to standardized units that can be reported in a single unit with a confidence interval to enable comparison across studies and outcomes to identify overall trends. The AMSRG recommends considering the use of effect size bias correction using the Hedges method that corrects for small sample sizes common in space medicine\textsuperscript{11}. The AMSRG provides a spreadsheet to manage data extraction and calculates both basic and Hedges corrected effect sizes\textsuperscript{33}. Results are still displayed on Forest plots, but effect size analysis does not complete the final meta-
analysis step, so there is no overall synthesis statistic, heterogeneity step or diamond to represent pooled effect on the Forest plots. The Cochrane handbook refers to this as “summarizing effect estimates”\(^\text{16}\).

Qualitative methods are useful when data, technical reports, and discussions clearly exist in published materials on an important topic, but not within controlled trials. These methods are also useful to combine with quantitative analysis to capture a more holistic multi-nodal dialogic explanation of a research topic\(^\text{15}\), this is known as an integrated approach\(^\text{4,21}\). The AMSRG qualitative guidance centers around thematic analysis, identifying common themes across the evidence base before structuring them into thematic maps that explore the relationships and potential hierarchies that exist within the data and potentially integrating quantitative and qualitative data together to provide insights toward answering research questions\(^\text{1,26}\).

Significance test reporting has not been formally supported by the AMSRG quantitative methods guides to date as it is limited compared to meta and effect analysis. However, as it has been used by a small number of initial reviews due to lack of effect size data it has now been added to the decision algorithm. The method involves reporting the P values of studies when there is no other data available to calculate effect sizes but there are strong justifications for still acknowledging the results of some studies.

**Developing options for quality scoring**

Risk of bias has been assessed using Cochrane risk of bias tools\(^\text{6,8}\) for any controlled or within participant trials. However, it is very specific to randomized controlled trials and so has often not been fully applicable to space medicine studies. In addition, the
Cochrane tool is not suited to scoring many grey literature sources such as technical reports. Several tools have been developed and trialed for scoring quality, usage and transferability of human spaceflight terrestrial based analogues. A bed rest quality score reports both greater quality and transferability as a higher score on an eight point scale that considers various factors considered important to the quality of bed rest as an astronaut microgravity simulation. A rank has been compiled that lists how well various ground simulations are likely to accurately model astronaut biomechanical, cardiovascular and metabolic changes, this gives an indication of how safe it is to transfer findings from the models to astronaut medicine. For reviews including grey literature there is a tool for scoring the quality of non-empirical data sources such as technical documents. This tool is based on scoring documents greater if they are well sourced, clearly written and based upon previous research as opposed to documents lacking citations and not having clearly explained methods for how findings or conclusions were reached. An algorithm to help decide which tools might be used for various types of reviews is shown in Figure 3.
Figure 3 - decision algorithm for considering use of various tools

Results

As of August 2020, seven space medicine systematic reviews had been completed and published in peer reviewed scientific journals (see Table I). Six of the reviews used quantitative analysis and one used qualitative. All the reviews conformed to PRISMA.
standards and followed Cochrane guidelines wherever they were not using AMSRG specific methods, guides and tools based on the decision algorithms. AMSRG specific methods were used in every instance to assess quality of space medicine specific data including space simulation analogue research environments and technical reports. All of the reviews were human space medicine topics. In all quantitative reviews, a full meta-analysis was not possible as the outcomes were highly heterogeneous across studies and there was poor reporting of space medicine studies. However, all quantitative studies were able to perform effect size analysis and only two\textsuperscript{10, 27} had to rely on significance result reporting. One review\textsuperscript{13} used qualitative methods to present a thematic analysis of technical documents from grey literature sources required to answer the technical nature of the research question. Three reviews\textsuperscript{25, 31, 32} included bed rest spaceflight simulation studies and reported bed rest quality scores. One review\textsuperscript{24} used the quality of partial gravity simulations to list how well various ground simulations are likely to accurately model astronaut biomechanical, cardiovascular and metabolic changes. A summary of all seven reviews, the research questions each posed and the AMSRG methods used within the methods is presented in Table I.

<table>
<thead>
<tr>
<th>Review</th>
<th>Aim</th>
<th>AMSRG methods used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valayer, Kim \textsuperscript{27A}</td>
<td>Evaluate if caloric restriction and dietary fasting can mitigate the adverse effects of ionizing radiation for deep space exploration</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Significance result reporting.</td>
</tr>
<tr>
<td>Sandal, Kim \textsuperscript{25A}</td>
<td>Evaluate the effectiveness of nutritional countermeasures, as a standalone intervention, to ameliorate musculoskeletal and cardiopulmonary deconditioning in gravitationally unloaded humans</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bed rest quality score.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Summary</td>
<td>Protocol &amp; pre-scoping</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Laws, Caplan</td>
<td>Identify the technical constraints of the Orion MPCV or transferable spacecraft that impact on the capability of astronauts to exercise effectively</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td>Winnard, Scott</td>
<td>Determine the time when humans exposed to simulated microgravity while not performing CM reach a moderate or large effect on muscle health outcomes, linked to if astronauts could have periods of no exercise on Moon or Mars missions.</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td>Konda, Karri</td>
<td>Investigate exercise countermeasures for attenuating musculoskeletal deconditioning during long duration bed rest.</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td>Winnard, Nasser</td>
<td>Assess interventions for counteracting changes, and reducing injury risks, to the lumbopelvic region, during microgravity exposure</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td>Richter, Braunstein</td>
<td>Determine the human cardiopulmonary and biomechanical changes expected to occur in partial gravity to inform future Moon and Mars mission operations.</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
</tbody>
</table>

* Pre-publication development version. ^ EAC collaborative review

Several gaps have been identified from the published reviews. Gaps are explained within each review and this is recommended by the AMSRG for all reviews. To provide an easily accessible summary overview of the gaps, they are also listed on the AMSRG gap analysis web page [http://aerospacedmed.rehab/gap-analysis](http://aerospacedmed.rehab/gap-analysis). A very brief summary of the gaps listed on the webpage as of August 2020 is as follows:

- Space agencies are advised to make information available about medically relevant constraints of spacecraft as accessible publications rather than solely in grey literature.
• Additional systematic reviews should carry on from the initial muscle review to establish the time for onset of musculoskeletal and cardiovascular effects without countermeasures to inform potential no-exercise periods during space missions.

• More primary data are needed on expected muscle changes in microgravity when not taking exercise countermeasures.

• Primary research data linked to muscle changes in spaceflight need to standardize outcome measures used and consider including patient reported measures within the standardized list.

• More studies are required to investigate the countermeasures for minimizing risk to the lumbopelvic region during spaceflight.

• More data on the expected human physiological effects of various g-loading environments are needed to inform medical operations for Lunar and Martian missions.

Discussion

Seven AMSRG methods specific to space medicine systematic reviews, which cover both quantitative and qualitative analysis, have been developed to address the unique limitations, research gaps and challenges of space medical research. Decision algorithms to guide researchers in which AMSRG methods to use have been trialed. Having AMSRG as a central group to standardize and publish review methods, along with a summary of any identified gaps, is beneficial to managing information in a single place, providing sector oversight, and improving research quality in space medicine.
Developing an algorithm to decide the best review method to use

Space medical research presents challenges that have prevented the use of meta-analysis in the AMSRG’s current reviews. These challenges include: the use of different spaceflight simulation models; a lack of controlled clinical trials; a large number of heterogeneous outcome measures; a lack of standardization; and, in some cases, poor reporting of study data. The identification of these challenges by AMSRG, and the development or adaptation of methods to address them, is evidence of the need for a centralized group to provide this guidance and gap analysis for the sector. The AMSRG decision algorithm, supported by the Cochrane handbook, prioritizes analyses in the order of: meta-analysis; effect size analysis; and then significance test reporting. The algorithm provides the option of considering qualitative analysis at all levels, either to supplement quantitative data in an integrative approach or as an alternative when data is available but in a format that does not support quantitative methods. As significance testing is now a formal accepted step in ASMRG methods, the AMSRG quantitative guidance will be updated to include additional significance testing options such as combining P values and vote counting, to bring the AMSRG guidance fully in line with all valid non meta-analysis options suggested by Cochrane. It should be noted that there are limitations to the non-meta-analysis methods detailed in the Cochrane handbook and so caution must be taken when using them. For example, care must still be taken when using effect size and/or P value-based analyses to recognize the original units of included measures and not make unreasonable comparisons such as comparing wildly differing outcomes. In addition, the reasons why space medicine reviews are forced to use alternate methods should be identified and addressed and the AMSRG has
recommended increasing performing of controlled clinical trials, determining and then using standardized outcome measures relevant to space operations and ensuring reporting standards of space medicine research supports systematic reviews. Reporting all experimental group means, standard deviations and group sample sizes is required for meta-analysis and should be set as the minimum journal publishing requirements in addition to reporting any statistical results such as P values. Furthermore, basic data on medical requirements/constraints of human spaceflight environments such as the internal volume of spacecraft that is available for operational essential activities including exercise countermeasures should be made easily available and accessible to the research community.

Quality scoring tools for space medicine

The wide variety of ground-based space simulations used in research need to be assessed for risk of bias, quality and transferability of studies using specifically tailored methods as opposed to generic quality tools. To address this for bed rest, the most commonly encountered simulation within the published reviews to date, a specific quality scoring tool was developed. This was developed by communicating with a team of experts to establish and then agree the final criteria and was done as part of one the first completed reviews\textsuperscript{31}. This has since been supplemented with an AMSRG ranking system that indicates how well parabolic flight, bed rest, isolation and suspension study’s findings can be transferred to actual astronaut settings during spaceflight. However, additional detailed quality scoring tools for each environment would be beneficial to develop. To date, where spaceflight data was not existing, AMSRG reviews
have remained within the scope of human simulation research only, as there were concerns that animal models might be too severely limited for transferability. In some cases, non-empirical sources may also provide useful insights at a human level and the AMSRG has provided a quality assessment tool for such documents. The use of the AMSRG developed tools in successfully published reviews, several of which have been done in collaboration with operational space agency medical staff from the European Space Agency, shows an initial level of validity of the tools. Going forwards, it would be useful to also test both the inter- and intra-rater reliability of the tools. Both reliabilities are important as best practice is to use agreement of multiple reviewers when scoring papers. While the methods for scoring space medicine studies reported in this document consider quality and transferability, it has been mostly possible to use Cochrane’s risk of bias tool for assessing bias. Where study designs made the Cochrane risk of bias inappropriate, alternative validated tools were available such as the Physiotherapy Evidence Database scale\textsuperscript{22}, and Quality in Prognostic Studies\textsuperscript{5}. It is likely that review teams will be able to use existing tools such as these to score bias and use AMSRG tools to further assess the quality and transferability of the various ground-based simulations.

In conclusion, space medicine is a unique field of medical research that requires specific systematic review methods to be developed to enable safe, transparent, reproducible synthesis of primary data to develop a robust evidence base that underpins space medical operations. After performing seven systematic reviews in aerospace medicine, adopting traditional systematic review tools by the Cochrane group has been challenging and has required modification to capture the full breadth of
primary sources available in aerospace medicine. The AMSRG group has built and will continue to build on the relevant foundations required to curate a central repository of educational resources, which are required to perform systematic reviews in aerospace medicine using evidence-based methodology without compromising scientific rigor.

Limitations

The limitations of the methods covered here are in the developmental nature of them. As already mentioned, the tools have been developed and then trialed within the initial reviews to establish them and test their validity and usability in the field. It would be useful to also establish the reliability of the tools that involve author scoring of studies.
References


### Table I - AMSRG reviews and methods used

<table>
<thead>
<tr>
<th>Review</th>
<th>Aim</th>
<th>AMSRG methods used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valayer, Kim</td>
<td>Evaluate if caloric restriction and dietary fasting can mitigate the adverse effects of ionizing radiation for deep space exploration</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Significance result reporting.</td>
</tr>
<tr>
<td>Sandal, Kim</td>
<td>Evaluate the effectiveness of nutritional countermeasures, as a standalone intervention, to ameliorate musculoskeletal and cardiopulmonary deconditioning in gravitationally unloaded humans</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bed rest quality score.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis</td>
</tr>
<tr>
<td>Laws, Caplan</td>
<td>Identify the technical constraints of the Orion MPCV or transferable spacecraft that impact on the capability of astronauts to exercise effectively</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of non-empirical data sources score</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Qualitative methods guide to conduct a thematic analysis.</td>
</tr>
<tr>
<td>Winnard, Scott</td>
<td>Determine the time when humans exposed to simulated microgravity while not performing CM reach a moderate or large effect on muscle health outcomes, linked to if astronauts could have periods of no exercise on Moon or Mars missions.</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bed rest quality score.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis.</td>
</tr>
<tr>
<td>Konda, Karri</td>
<td>Investigate exercise countermeasures for attenuating musculoskeletal deconditioning during long duration bed rest.</td>
<td>Protocol &amp; pre-scoping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Significance result reporting.</td>
</tr>
<tr>
<td>Winnard, Nasser</td>
<td>Assess interventions for counteracting changes, and reducing injury risks, to the lumbopelvic region, during microgravity exposure</td>
<td>Protocol &amp; pre-scoping. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bed rest quality score. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis. *</td>
</tr>
<tr>
<td>Richter, Braunstein</td>
<td>Determine the human cardiopulmonary and biomechanical changes expected to occur in partial gravity to inform future Moon and Mars mission operations.</td>
<td>Protocol &amp; pre-scoping. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of partial gravity simulations tool. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quantitative methods guide. *</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Effect size analysis *</td>
</tr>
</tbody>
</table>

* Pre-publication development version. ^ EAC collaborative review
Captions for Figures

Figure 1 - example search strategy table from Richter et al. 2017

Figure 2 - Methods decision algorithm, with the decisions running down the middle starting from the top, dashed lines show alternate or supplementary decision options.

Figure 3 - decision algorithm for considering use of various tools