PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Study Protocol: A Survey Exploring Patients' and Healthcare Professionals' Expectations, Attitudes, and Ethical Acceptability Regarding the Integration of Socially Assistive Humanoid Robots in Nursing
AUTHORS	Mlakar, Izidor; Kampič, Tadej; Flis, Vojko; Kobilica, Nina; Molan, Maja; Smrke, Urška; Plohl, Nejc; Bergauer, Andrej

VERSION 1 – REVIEW

REVIEWER	Nuria Valles-Peris Universitat Autonoma de Barcelona
REVIEW RETURNED	27-Jul-2021

GENERAL COMMENTS	Dear Authors, Thank you very much for giving me the opportunity to know and comment on your interesting work on the study of expectations and attitudes towards social robots in healthcare. Your study is interesting, and I have no doubt that you will get valuable results in the future. However, I consider that there are some questions in the design of the study on which it would be necessary to deepen and reflect, to introduce some concretions and/or modifications. In the following paragraphs I make some suggestions for improving your protocol: General issues: In the title it refers to humanoid robots and in the text to social robotic systems (SRS). This issue is relevant, because most of the robots currently being introduced in hospitals are not humanoid. On the other hand, the reticence and fears related to humanoid robots may be different from the considerations related to other SRS or task automation. SRS is a too generic term that may refer to different types of robots. As explained in the article, very diverse SRS performing different tasks may be used in healthcare. For the data to have more consistency, it is important to specify what type of SRS we are talking about and adapt the survey to that robot (health-care telepresence, disinfection of the hospital or clinic; prescription and meal dispensing, telepresence robots for communicating, pet robots for company...). Or have the hospital staff and patients surveyed already interacted with a particular robot and attitudes about that robot are analyzed? If that is the case, it is necessary to specify. Objectives and hypotheses

	The article states that the objective is "investigating the feasibility and technical and nontechnical considerations of integration of social robotic systems as perceived by healthcare professionals and patients in hospitals in the region". However, the feasibility and technical considerations cannot be analyzed on the basis of surveys on attitudes and expectations of healthcare professionals and patients. In this case, I think that this issue could be solved changing the writing, making clear that the acceptability, attitudes and expectations about SRS can be analyzed and from there recommendations can be made on how to design robots according to the needs and concerns that arise. At the end of the article, picking up this question, it is said that the results will be used to model the design of the social robotics systems, in relation to what? To the appearance? To the functionality of the robot? The distribution of tasks between the medical staff and the robot? How it will be introduced in the hospital? Without knowing what type of SRS is to be introduced and for what purpose, the expectations and attitudes of patients in the abstract have the risk of reinforcing an idea of polarization that is very common in social robotics, organized on the basis of utopian or dystopian imaginaries. This type of approach does not give us valuable information for designing robots for hospitals, while it can reinforce a series of negative attitudes of healthcare personnel and patients towards robots. On the other hand, “attitudes” and “ethical reservations” play a very important role in the research questions. However, it is not explained how attitudes and ethical reservations are conceptualized and measured in the questionnaires used. It is necessary to justify why the TSES, EAS and NARS were selected and what these questionnaires contribute to the study of attitudes and ethical issues. Have they been validated by other studies? What are the previously identified strengths and risks of these questionnaires? What kind of conclusions have they reached? These questions would be very valuable in the introduction of the study and in the discussion. Methods It is explained that different questionnaires will be used: for healthcare professionals the Technology-Specific Expectation Scale (TSES) and Ethical Acceptability Scale (EAS); and for patients, the Technology-Specific Expectation Scale (TSES), and The Negative Attitudes towards Robots Scale (NARS). Why different questionnaires were given to the two groups? This could led you to don’t have comparable results. The title of the tools alone (NARS) presupposes a negative bias towards the attitude of the patients. It is necessary to justify the choice of different questionnaires for the different collectives or to assess the adequacy of this choice. Ethical, legal and regulatory aspects Beauchamp and Childress' Principles of Biomedical Ethics do not make much sense in this protocol if they are not developed or specified how they are taken into account. The same is true for the generic standards of ethics and integrity of the research that are included. It needs to be explicitly stated that all participants will sign an

	informed consent. Data storage and privacy The paper states that an “Anonymous survey regarding patient attitudes towards the robot in a hospital environment for the needs of the international project Horizon 2020 will be carried out for the preliminary needs of the project. For the most part, questionnaires with the Likert scale will be used, and among personal data we will only collect information on gender, age and level of education of patients, and for employees, an additional information on their occupation.” I don't understand this question, will be distributed previously to the study a“survey regarding patient attitudes towards the robot in a hospital environment for the needs of the international project Horizon 2020 will be carried out for the preliminary needs of the project”? Is it a different survey? How has it been designed? Is it to adapt the questions of the questionnaires to be conducted? Please, clarify this question. Patient and Public involvement The statement that if it is a survey, patients cannot be involved refers to a choice made by the investigators, but is not inherent to a survey, as the principles of RRI highlight. A prospective pilot phase can be carried out before distributing the questionnaires, in order to find out what are patients and public concerns, and to introduce some questions on the subject in the questionnaire. Another question regarding this issue: who distributes the questionnaire to patients? Outcomes It would be useful to include among the health care workers expected differences in relation to the type of profession and tasks they perform in the hospital (if they are doctors or nurses or other types of workers). Among both groups, healthcare professionals and patients, differences are to be expected depending on the context in which the robot is used and the specific uses of the robot.

REVIEWER	runa lazzarino Middlesex University
REVIEW RETURNED	03-Dec-2021

GENERAL COMMENTS	“The cohort will include subjects from the region rather than country.” : aren’t country hospitals going to be involved too? “Technological developments in the fields of robotics and AI can provide significant cost Savings” : This argument is valid, but at date it refers to a situation happening in the distant future, surely in the field of social robotics, and AI devices assisting the nursing workforce. I would not over-emphasise this argument, as there are indeed also great concerns about the deepening of health access disparities due to the introduction of advanced technologies in the HC sector. “Although the systems exhibit robust, autonomous capabilities and initial concerns regarding physical safety around people have been

	at least partially addressed, the uptake of the technology is arguably slow" : These are very important concerns, and the AI autonomy is at an embryonic stage not robust at all, as well as is the addressing of ethical (beyond social contact reduction) and safety concerns. The second paragraph of the background should be re-written more carefully, also by providing definition, separating of which barriers we are talking about, and distinguish between workforce and patients. "To sum up, the SRS, if designed correctly, can have a significant impact especially on 'other nursing services" ; This sentence is vague and inaccurate "However, a more anticipatory and contemporary position towards technology in nursing must be established with both healthcare professionals and patients." : the background for this strong affirmation had been sufficiently well built. "Building on this baseline the proposed study implements three questionnaires focusing on: i) general acceptance of robots in a setting of nursing (NARS [23]), ii) ethical and professional reservations (EAS [22]) and iii) functional (technological) expectations (TSES [21]). The tools chosen are widely used questionnaires for evaluating" : this should go in the methods, I suggest instead strengthening your background. "…recognizing sociocultural bias as a benefit rather than limitation. Since psychometrically evaluated tools will be used, cross-cultural differences will be evaluated against relevant studies in similar settings" : the cultural variant has not been introduced and explained before, it is not even in the research questions. "The aim of this study is to determine professionals' and patients' attitudes and expectations towards social robotic system (SRS)" OR "The main objective is to assess applicability, ethical and technical considerations regarding integration of SRS" : please ensure consistency in aim, objectives and research questions, and terminology…. Below you have also: "The study is a cross-sectional survey investigating the feasibility and technical and nontechnical considerations of integration of social robotic systems as perceived by healthcare professionals and patients in hospitals in the region." "The cross-sectional study will include at least 1,000 healthcare professionals and at least 500 inpatients answering to two questionnaires." : the total questionnaires are three, is it? Please, it reads confusing which questionnaires are completed and how by which kind of cohort and from which setting, and how the age, gender etc. are collected and for which purpose. I think that this is a protocol, and methods should be absolutely very clear. How do you plan to recruit participants from other hospitals for example? Why at your selected hospital the survey will be paper and pencil? "The exclusion criteria reflect the primary domains of possible integration, thus patients hospitalized at the pediatric clinic, department of psychiatry and the

	clinic for gynecology and perinatology will not be included." I suppose you mean 'integration' of the kind of robots/AI device (still unclear, as not defined – please define what you exactly mean with social robotic system in your study). This is an interesting point, and I would expand more on this. "Most recently, COVID-19 has shown us the need for understanding and coordination in dealing with new outbreaks" : COVID has not been dealt with in the introduction, maybe worth doing and again I would over-emphasise the reduction of costs, currently AI, especially autonomous tech in HC is very costy and underdeveloped, social robots are barely implemented in the real world. Why is the project of reference mentioned in the discussion? Explain acronyms is full first, please. Please proofread article

REVIEWER	Hojjat Zeraati Tehran University of Medical Sciences, Biostatistics
REVIEW RETURNED	08-Dec-2021

GENERAL COMMENTS	• Page 3 Line 30: The word" Cohort" needs to be replaced with "study sample". • Page 6-7: Is there any relationship between attitude of the patients and their gender? • Page 6-7: Is there any relationship between attitude of the patients and their disease? • Page 6-7: Is there any relationship between attitude of the patients and their hospital part and reason of hospitalization (Surgery or internal or ICU…)? • Study design, page 7-8: In a cross-sectional study random selection samples must be done. The method(s) of random sampling needs (need) to be clearly stated. • Study design, page 7-8: The method of sample size computation (1000 and 500) needs to be clearly stated. • Study design, page 7-8: The format of the questioning needs to be cleared (paper-pencil in the abstract, and electronically in page 7 line 52). • Page 7 Line 58: What is the relation between patient-care and supporting staff? It needs to be omitted. • Page 8 Line 8: Why will select two hospital? Are there only this two hospital in region? • Sample size, page 13: The method of sample size computation (1000 and 500) needs to be clearly stated. • Analysis, page 13-14: The method for testing the normality of distributions need to be mentioned. • Analysis, page 13-14: The first assumption for using Anova, Mancova, t-test and…, is a random sampling, so the method of randomization need clearly stated. • It is recommended that the questionnaires need to be appended.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Nuria Valles-Peris, Universitat Autonoma de Barcelona Comments to the Author:

Dear Authors,

Thank you very much for giving me the opportunity to know and comment on your interesting work on the study of expectations and attitudes towards social robots in healthcare. Your study is interesting, and I have no doubt that you will get valuable results in the future. However, I consider that there are some questions in the design of the study on which it would be necessary to deepen and reflect, to introduce some concretions and/or modifications.

In the following paragraphs I make some suggestions for improving your protocol:

General issues:

1. **In the title it refers to humanoid robots and in the text to social robotic systems (SRS). This issue is relevant, because most of the robots currently being introduced in hospitals are not humanoid. On the other hand, the reticence and fears related to humanoid robots may be different from the considerations related to other SRS or task automation. SRS is a too generic term that may refer to different types of robots. As explained in the article, very diverse SRS performing different tasks may be used in healthcare. For the data to have more consistency, it is important to specify what type of SRS we are talking about and adapt the survey to that robot (health-care telepresence, disinfection of the hospital or clinic; prescription and meal dispensing, telepresence robots for communicating, pet robots for company...). Or have the hospital staff and patients surveyed already interacted with a particular robot and attitudes about that robot are analyzed? If that is the case, it is necessary to specify.**

Thank you for the comment. We have replaced the SRS with a less generic term, socially assistive humanoid robots (SAHR).

Objectives and hypotheses

2. **The article states that the objective is "investigating the feasibility and technical and nontechnical considerations of integration of social robotic systems as perceived by healthcare professionals and patients in hospitals in the region". However, the feasibility and technical considerations cannot be analyzed on the basis of surveys on attitudes and expectations of healthcare professionals and patients. In this case, I think that this issue could be solved changing the writing, making clear that the acceptability, attitudes and expectations about SRS can be analyzed and from there recommendations can be made on how to design robots according to the needs and concerns that arise.**

We have modified the paper to consistently express the main objective, the investigation of attitudes, expectations, and ethical considerations regarding the integration of socially assistive humanoid robots into nursing and care workflow. We have also explained that such investigation will provide critical knowledge on how to design and implement robots.

3. **At the end of the article, picking up this question, it is said that the results will be used to model the design of the social robotics systems, in relation to what? To the appearance? To the functionality of the robot? The distribution of tasks between the medical staff and the robot? How it will be introduced in the hospital? Without knowing what type of SRS is to be introduced and for what purpose, the expectations sand attitudes of patients in the abstract have the risk of reinforcing an idea of polarization that is very common in social robotics, organized on the basis of utopian or dystopian imaginaries. This type of approach does not give us valuable information for designing robots for hospitals, while it can reinforce a series of negative attitudes of healthcare personnel and patients towards robots.**

The robots we target will be deployed into clinical settings as social, digital assistants engaging with patients and participating as a digital support tool during grand rounds. Within the context of patient engagement, the robots will lead the domain-centric discourse in order to collect patient-reported outcomes and patients' health quality measures. The units will further contribute with gamification (i.e.

simple breathing exercises, Q&A sessions, etc.) to complement the regular nursing routine. In the context of decision support tools, the robots will offer person-centric access to electronic health records and a rule-based alerting system. Thus, this study aims to prevent oversimplification of the context of targeted environment and identify practical barriers and reservations that must be considered when designing the behavior of the units before placing them into the clinical routine. Of course, we expect that new, practical issues will arise during the real-life operations. Moreover, the study could reveal potential misconceptions about SAHRs and point to specific myths or fears that should be addressed with future educational programs. Lastly, the results of the proposed study will also reveal which subpopulations of providers and patients may need additional information regarding the safety and potential benefits of SAHRs. Thus

4. **On the other hand, "attitudes" and "ethical reservations" play a very important role in the research questions. However, it is not explained how attitudes and ethical reservations are conceptualized and measured in the questionnaires used. It is necessary to justify why the TSES, EAS and NARS were selected and what these questionnaires contribute to the study of attitudes and ethical issues. Have they been validated by other studies? What are the previously identified strengths and risks of these questionnaires? What kind of conclusions have they reached? These questions would be very valuable in the introduction of the study and in the discussion.**

Due to another reviewer's suggestion, we moved the part about questionnaires to the Methods section (Design and setting). However, we definitely agree with you that the choice of these questionnaires and their description have to be much more detailed. As such, each questionnaire that will be used in the study is now properly described (what they measure, how many items, previous validations, some results, and identified strengths).

Methods

5. **It is explained that different questionnaires will be used: for healthcare professionals the Technology-Specific Expectation Scale (TSES) and Ethical Acceptability Scale (EAS); and for patients, the Technology-Specific Expectation Scale (TSES), and The Negative Attitudes towards Robots Scale (NARS). Why different questionnaires were given to the two groups? This could led you to don't have comparable results. The title of the tools alone (NARS) presupposes a negative bias towards the attitude of the patients. It is necessary to justify the choice of different questionnaires for the different collectives or to assess the adequacy of this choice.**

Thank you for your comment. Although we have put considerable thought into the selection of questionnaires, our thought process was not properly explained in the manuscript. While both groups will fill out the questionnaire that measures expectations, and we will be able to compare this variable between the two samples, the employees will additionally fill out the ethical acceptability scale, while patients will fill out the negative attitudes scale. This is due to the fact that employees are the ones who are the primary decision-makers in this context and have to weigh the pros/cons and ethical issues related to the integration of social robots. We argue that patients are less likely to consider the ethical applicability of social robots, but certainly have attitudes towards their use (and can thus answer NARS).

Also – ideally both groups would answer all three questionnaires, but the study had to be as concise and easy for participants as possible (to minimize the needed effort on the participant side and gather a large sample of participants). As such, we only used the two most relevant questionnaires in each sample. This is now also explained in the manuscript.

Ethical, legal and regulatory aspects

6. **Beauchamp and Childress' Principles of Biomedical Ethics do not make much sense in this protocol if they are not developed or specified how they are taken into account. The same is true for the generic standards of ethics and integrity of the research that are included.**
   It needs to be explicitly stated that all participants will sign an informed consent.

Thank you for this comment. We omitted the Beauchamp and Childress' Principles of Biomedical Ethics and other generic standards of ethic and integrity. We also stated that all participants will sign an informed consent. In the reformulation of the paragraph, we followed other study protocols already published by BMJ Open (e.g., Yash et al., 2020; Borg et al., 2021).

Borg D, Rae K, Fiveash C Queensland Family Cohort Research Collaborative, et al. Queensland Family Cohort: a study protocol. BMJ Open 2021;11:e044463. doi: 10.1136/bmjopen-2020-044463

Yash Pal R, Kuan WS, Tiah L, et al. End-of-life management protocol offered within emergency room (EMPOWER): study protocol for a multicentre study. BMJ Open 2020;10:e036598. doi: 10.1136/bmjopen-2019-036598

Data storage and privacy

7. **The paper states that an "Anonymous survey regarding patient attitudes towards the robot in a hospital environment for the needs of the international project Horizon 2020 will be carried out for the preliminary needs of the project. For the most part, questionnaires with the Likert scale will be used, and among personal data we will only collect information on gender, age and level of education of patients, and for employees, an additional information on their occupation." I don't understand this question, will be distributed previously to the study a"survey regarding patient attitudes towards the robot in a hospital environment for the needs of the international project Horizon 2020 will be carried out for the preliminary needs of the project"? Is it a different survey? How has it been designed? Is it to adapt the questions of the questionnaires to be conducted? Please, clarify this question.**

We have removed the part related to HosmartAI to avoid confusion. We wanted to highlight that this study is funded and a part of a European Project. Within the umbrella of HosmartAI project, the integration of robots to clinical workflow will be carried out in two subsequent clinical studies.

Patient and Public involvement

8. **The statement that if it is a survey, patients cannot be involved refers to a choice made by the investigators, but is not inherent to a survey, as the principles of RRI highlight. A prospective pilot phase can be carried out before distributing the questionnaires, in order to find out what are patients and public concerns, and to introduce some questions on the subject in the questionnaire.**

We have modified the statement to reflect the fact that investigators did not involve patients in study design.

9. **Another question regarding this issue: who distributes the questionnaire to patients?**

The questions to the patients are distributed by the staff. This is now elaborated under section: Methods - Design and setting.

Outcomes

10. **It would be useful to include among the health care workers expected differences in relation to the type of profession and tasks they perform in the hospital (if they are doctors or nurses or other types of workers).**
    Among both groups, healthcare professionals and patients, differences are to be expected depending on the context in which the robot is used and the specific uses of the robot.

As the current literature has not yet studied the differences between different occupations in expectations regarding social robots and ethical acceptability, we sadly cannot form very specific

predictions. However, we intend to explore different occupations' expectations and ethical acceptability, which will contribute to more deliberate investigation of such differences in the future.

Reviewer: 2
Dr. runa lazzarino, Middlesex University Comments to the Author:

1. **"The cohort will include subjects from the region rather than country." : aren't country hospitals going to be involved too?**

We have modified the statement: The cohort will include only subjects from Slovenia.

2. **"Technological developments in the fields of robotics and AI can provide significant cost Savings" : This argument is valid, but at date it refers to a situation happening in the distant future, surely in the field of social robotics, and AI devices assisting the nursing workforce. I would not over-emphasise this argument, as there are indeed also great concerns about the deepening of health access disparities due to the introduction of advanced technologies in the HC sector.**

We have modified the sentence and added the barriers and issue related to ethics and justice in the 'Summary' section of the introduction. We added two relevant references.

3. **"Although the systems exhibit robust, autonomous capabilities and initial concerns regarding physical safety around people have been at least partially addressed, the uptake of the technology is arguably slow": These are very important concerns, and the AI autonomy is at an embryonic stage not robust at all, as well as is the addressing of ethical (beyond social contact reduction) and safety concerns. The second paragraph of the background should be re-written more carefully, also by providing definition, separating of which barriers we are talking about, and distinguish between workforce and patients.**

We agree with your comment and have carefully refined the background to clearly highlight the main barriers related to both patients and professionals.

4. **"To sum up, the SRS, if designed correctly, can have a significant impact especially on 'other nursing services"; This sentence is vague and inaccurate**

After the revision of the section, the sentence was modified.

5. **"However, a more anticipatory and contemporary position towards technology in nursing must be established with both healthcare professionals and patients." : the background for this strong affirmation had been sufficiently well built.**

We have modified the introduction section to further emphasize the importance of co-creation. And the goal of the proposed study is exactly that co-design of the robot's behavior thorough understanding of expectations and reservations of patients and staff alike.

6. **"Building on this baseline the proposed study implements three questionnaires focusing on: i) general acceptance of robots in a setting of nursing (NARS [23]), ii) ethical and professional reservations (EAS [22]) and iii) functional (technological) expectations (TSES [21]). The tools chosen are widely used questionnaires for evaluating": this should go in the methods, I suggest instead strengthening your background.**

We strongly agree with the comment and have thus deleted all mentions of questionnaires in the background section. We have moved a more detailed description of the questionnaires to the 'Methods' section. We have further elaborated on the background.

7. **"…recognizing sociocultural bias as a benefit rather than limitation. Since psychometrically evaluated tools will be used, cross-cultural differences will be**

evaluated against relevant studies in similar settings": the cultural variant has not been introduced and explained before, it is not even in the research questions.

We understand that this bit was written in a rather confusing manner. As such, it was deleted.

8. **"The aim of this study is to determine professionals' and patients' attitudes and expectations towards social robotic system (SRS)" OR "The main objective is to assess applicability, ethical and technical considerations regarding integration of SRS": please ensure consistency in aim, objectives and research questions, and terminology…. Below you have also: "The study is a cross-sectional survey investigating the feasibility and technical and nontechnical considerations of integration of social robotic systems as perceived by healthcare professionals and patients in hospitals in the region."**

We have modified the paper to consistently express the main objective which is applicability, ethical and functional considerations regarding integration of SHAR into nursing and care workflow.

9. **"The cross-sectional study will include at least 1,000 healthcare professionals and at least 500 inpatients answering to two questionnaires." : the total questionnaires are three, is it? Please, it reads confusing which questionnaires are completed and how by which kind of cohort and from which setting, and how the age, gender etc. are collected and for which purpose. I think that this is a protocol, and methods should be absolutely very clear. How do you plan to recruit participants from other hospitals for example? Why at your selected hospital the survey will be paper and pencil?**

We have modified this section to clearly state which questionnaires will be completed by which study cohort and in which setting, and how information on gender, age etc. will be collected. The purpose of collecting the specific data is elaborated on in the Objectives and hypotheses as well as the Outcomes section of the protocol. The data will be collected in digital form, only letter of informed consent will be signed in paper format (we have now clarified this in the manuscript as well). Other hospitals were reached directly through the research and administration departments of the hospitals.

10. **"The exclusion criteria reflect the primary domains of possible integration, thus patients hospitalized at the pediatric clinic, department of psychiatry and the clinic for gynecology and perinatology will not be included." I suppose you mean 'integration' of the kind of robots/AI device (still unclear, as not defined – please define what you exactly mean with social robotic system in your study). This is an interesting point, and I would expand more on this.**

Thank you for pointing out this issue. We have now clarified what we mean by social robot systems; we actually refer to socially assistive humanoid robots. We have also rewritten this sentence to make it clearer.

11. **"Most recently, COVID-19 has shown us the need for understanding and coordination in dealing with new outbreaks" : COVID has not been dealt with in the introduction, maybe worth doing and again I would over-emphasise the reduction of costs, currently AI, especially autonomous tech in HC is very costy and underdeveloped, social robots are barely implemented in the real world.**

We have revised the introduction, to highlight the main barriers of the SAHR in hospital settings, focusing on oversimplification. The research itself is not focused COVID. Arguably during the pandemic individuals became more open to the idea of autonomous technology in HC. However, the main barriers driven by oversimplification and failure to reflect the complexity of the setting still remain and will return after the pandemic.

12. **Why is the project of reference mentioned in the discussion?**

Based on editor's recommendation, we have removed the discussion section.

13. **Explain acronyms is full first, please.**

We have revised the paper to ensure the first mentioning of acronyms is accompanied by the full explanation.

14. **Please proofread article**

We have revised the paper using Grammarly and with professional language editing service.

Reviewer: 3
Prof. Hojjat Zeraati, Tehran University of Medical Sciences

Comments to the Author:

1. **Page 3 Line 30: The word" Cohort" needs to be replaced with "study sample".**

Word "Cohort" was replaced by "study sample".

2. **Page 6-7: Is there any relationship between attitude of the patients and their gender?**

Based on previous literature, we cannot form specific hypotheses regarding the comparison between men/women (most studies did not find any differences). However, this will be explored. This is now clarified in the manuscript as well.

3. **Page 6-7: Is there any relationship between attitude of the patients and their disease?**

Data on disease will not be collected. Based on previous literature, however, it is hard to predict whether this could be an important factor.

4. **Page 6-7: Is there any relationship between attitude of the patients and their hospital part and reason of hospitalization (Surgery or internal or ICU…)?**

Similarly to our answer above, data on hospital part/reason for hospitalization will not be collected. Based on previous literature, however, it is hard to predict whether this could be an important factor.

5. **Study design, page 7-8: In a cross-sectional study random selection samples must be done. The method(s) of random sampling needs (need) to be clearly stated.**

Method of sampling is now clearly stated. Even though we agree that random sampling would be ideal, the setting of our study does not enable us to employ this method, as not random members of the population can be accessed (e.g., all potential inpatients). Therefore, we will follow a non-probability method of sampling while being aware of the limitations it brings (e.g., limited generalizability), as was already done in several papers published by BMJ Open (e.g., Choi et al., 2020; Wheelock et al., 2017).

Choi EY, Pyo J, Lee W, et al. Nurses' experiences of patient safety incidents in Korea: a cross-sectional study. BMJ Open 2020;10:e037741. doi: 10.1136/bmjopen-2020-037741

Wheelock A, Miraldo M, Thomson A, et al. Evaluating the importance of policy amenable factors in explaining influenza vaccination: a cross-sectional multinational study. BMJ Open 2017;7:e014668. doi: 10.1136/bmjopen-2016-014668

6. **Study design, page 7-8: The method of sample size computation (1000 and 500) needs to be clearly stated.**

Data on sample sizes was removed from pages 7-8. However, it is now explained in much more detail later on in the manuscript, especially on page 13 (Sample size determination section; explained in our response to your comment #10).

7. **Study design, page 7-8: The format of the questioning needs to be cleared (paper-pencil in the abstract, and electronically in page 7 line 52).**

We corrected this discrepancy; the format of data collection will be digital in both study samples.

8. **Page 7 Line 58: What is the relation between patient-care and supporting staff? It needs to be omitted.**

In the new version of the text, this has been omitted.

9. **Page 8 Line 8: Why will select two hospital? Are there only this two hospital in region?**

This was misarticulated. We have contacted hospitals in the region.

10. **Sample size, page 13: The method of sample size computation (1000 and 500) needs to be clearly stated.**

Thank you for pointing out this issue. The manuscript was a bit confusing in stating how many participants will be invited and how many participants we actually plan to recruit. Our justification for the target sample sizes were also not elaborated enough. In the new version of the manuscript, we fixed this issue and thoroughly explained the factors we took into account.

11. **Analysis, page 13-14: The method for testing the normality of distributions need to be mentioned.**

Thank you for your suggestion. We have now explained how we will test the normality of distributions in our study.

12. **Analysis, page 13-14: The first assumption for using Anova, Mancova, t-test and…, is a random sampling, so the method of randomization need clearly stated.**

We agree that it makes sense to clarify the nature of our data – for this please see the section on Methods – Design and setting, 4th paragraph of the manuscript. However, very rarely samples in similar studies are in fact obtained by random sampling, as this is not a prerequisite per se for the application of methods such as *t*-test, ANOVA, or MANOVA (they do assume, however, the homogeneity of variance, independence in observations, and for *F* statistics the dependent variable to be measured at least at interval level). But, the application of these methods does require the sample to be representative of the population (which is indeed the easiest to achieve by random sampling) when operating under the population model approach in order to be able to generalize the inferences made by specific research. As in practice this doesn't always hold true or it is impossible to assess, several methods have been developed to address this, such as operating within randomization model through permutations of group membership in order to assess the probability of non-arbitrairty in group membership (which is actually considered as superior, e.g., Ludbrook & Dudley, 1998; and Ernst, 2004), or employing bootstrap methods, which will be used if any indications for them will be detected.

Ernst, M. D. Permutation methods: A basis for exact inference. Statistical Science 2004; 19(4), 676-685.

Ludbrook, J., & Dudley, H. Why permutation tests are superior to t and F tests in biomedical research. American Statistician, 1998; 52(2), 127-132.

13. It is recommended that the questionnaires need to be appended.

We understand that appending the questionnaires would make it easier for readers to gain an insight into the measured constructs, but this is not possible due to complex copyright reasons in academia. As the questionnaires are not our property (but, instead, the property of the original author and/or the journal that published the original validation), appending them could lead to negative consequences. We would like to note, however, that questionnaires are available in their entirety in research articles that are cited in our manuscript.

VERSION 2 – REVIEW

REVIEWER	Nuria Valles-Peris Universitat Autonoma de Barcelona
REVIEW RETURNED	13-Jan-2022

GENERAL COMMENTS	Dear authors, Thank you very much for the new version of the article, and congratulations for the work done, because I consider that it has been substantially improved since the last version and many issues have been clarified. I just want to point out some small issues, so that you can assess whether it is appropriate to adjust some minor issues. Prior to this, however, I would like to point out that I am not an expert in quantitative data analysis and, therefore, I am not going to make any assessment of the section "Data analysis and statistics". Objectives, p.7: It is stated that: “R3: What are the main ethical reservations that may impact the acceptance of robotic units in hospital care in healthcare professionals?” However, a questionnaire cannot determine “what are” the ethical reservations about the introduction of a robot in a hospital; this requires a different kind of bioethical study. With the tools proposed in the article, it is possible to know the opinion of professionals and patients on whether they consider its use to be ethically acceptable or not. Therefore, I would change the sentence or remove this objective. Design and methods, p.8: It is said that “The study is a cross-sectional survey evaluating expectations, attitudes, and ethical considerations “, but nowhere is it explained what aspects of expectations and attitudes are measured. On ethical considerations, I have already explained that in order to be rigorous, it is necessary to change the way of expressing it. It appears by first time “TSES” and “EAS”. Please, include the complete name of the questionnaires. The EAS designed to assess the appropriateness of the use of robots for therapy with children with autism is now used for a generalist assessment. This is a problem, because we know that expectations and attitudes change depending on whether they refer to a particular technologies and context, or whether they refer to a general context (where stereotypes, etc. are much more involved). I propose, since particular applications are already designed into the project, to include these specifications in the questions. Or, if it is already done, explain it in the paper. Patient and Public Involvement, p.13:

	It is stated that “Healthcare professionals of the participating medical institution were involved in the study design.” Please, explain briefly how it has been done. Outcomes, p.14: In Table 2, I think that instead of “expected cohorts” would be better “expected results”. Yours sincerely, Núria

REVIEWER	runa lazzarino Middlesex University
REVIEW RETURNED	12-Jan-2022

GENERAL COMMENTS	- in the title, pleas add that you 'exploring' (more than assessing probably) the ethical considerations. This may be considered part of the attitudes, however, as you spelled it out in the introduction, maybe this third component has a particularly relevant role within the economy of the study. - similarly: why don't we have in the title 'regarding the integration...' rather than 'towards the use...'? - in last highlight, do not add new aims/elements: I refer to 'perceived barriers'. Find a formula and stick to it throughout please. Also, sometimes in text you again refer to barriers: if barriers belong to the core aims, it must be made clear since from onset please and included in research questions. - nothing is said about the Slovenian context. Please, add something in relation to the HC sector, or technology in it etc. - R5-R7 also come as a surprise, and what do you mean with 'characteristics'? Not technical term, anticipate what you have in the following part (ie, demographic variables). I feel that your hypotheses were not paved enough in the background. Usually, you do not introduce new background literature after the introduction... "Moreover, based on previous literature [27,28], which investigated the role of age, education...". Readers should not be waiting to read the outcomes to have the full picture of what you intend to explore and find. - Similar call for consistency: they way you refer to robots is too varied. You now have SAHR in abstract and title. You use the acronym on page 6 without explaining it in full or defining it, and in the previous pages you use different terms to refer to the robots. Please, rectify, define and be consistent. - I find the 'design and setting' section too unclear and confusing, despite the provision of the table which is meant to help clarity. I think both design and way of presenting it are to be simplified. What is the setting (Maribor or not, and which more: sometimes you have the plural)? Who and how will be recruited within it? What questionnaires will they complete?... the questionnaires and their scales etc can be uploaded separately as a table. What is this and how will it be used and justified?: 'Additionally, a few additional dichotomous questions will be posed to participants as well (e.g., “Do you think the robot could answer patient’s questions about treatment?). how do you assess 'highly relevant data'? - change date of start of the study please - give example of sensitive data you are not collecting

REVIEWER	Hojjat Zeraati Tehran University of Medical Sciences, Biostatistics

REVIEW RETURNED	23-Feb-2022

GENERAL COMMENTS	The reviewer completed the checklist but made no further comments.

VERSION 2 – AUTHOR RESPONSE

REVIEWER 2

- in the title, pleas add that you 'exploring' (more than assessing probably) the ethical considerations. This may be considered part of the attitudes, however, as you spelled it out in the introduction, maybe this third component has a particularly relevant role within the economy of the study.

- similarly: why don't we have in the title 'regarding the integration...' rather than 'towards the use...'?

Yes, thank you for your suggestion. We have now replaced “towards the use” with “regarding the integration”. We have also emphasized ethical acceptability in the title.

- in last highlight, do not add new aims/elements: I refer to 'perceived barriers'. Find a formula and stick to it throughout please. Also, sometimes in text you again refer to barriers: if barriers belong to the core aims, it must be made clear since from onset please and included in research questions.

Thank you for pointing this out. According to your suggestion and the suggestion of the other reviewer, we now use “ethical acceptability” instead of “perceived barriers” in the highlights section as well as throughout the manuscript.

- nothing is said about the Slovenian context. Please, add something in relation to the HC sector, or technology in it etc.

We have added a sentence to the introduction that provides brief information on the Slovenian HC sector and the technology in it. We have also cited an article on digital health in Slovenia, where readers can find additional information.

- R5-R7 also come as a surprise, and what do you mean with 'characteristics'? Not technical term, anticipate what you have in the following part (ie, demographic variables). I feel that your hypotheses were not paved enough in the background. Usually, you do not introduce new background literature after the introduction... "Moreover, based on previous literature [27,28], which investigated the role of age, education...". Readers should not be waiting to read the outcomes to have the full picture of what you intend to explore and find.

Thank you for your comment. We use “demographic characteristics” instead of simply “characteristics” in the revised manuscript. We also specify the variables in brackets. We also understand your point regarding the role of age and education. The new version of the manuscript now mentions age and education earlier in the text.

- Similar call for consistency: they way you refer to robots is too varied. You now have SAHR in abstract and title. You use the acronym on page 6 without explaining it in full or defining it, and in the previous pages you use different terms to refer to the robots. Please, rectify, define and be consistent.

Thanks! We now use “socially assistive humanoid robots” (or SAHR) throughout the text when referring to our study. Besides defining the abbreviation in the abstract, we have also spelled it out in the introduction (when it is first mentioned).

- I find the 'design and setting' section too unclear and confusing, despite the provision of the table which is meant to help clarity. I think both design and way of presenting it are to be simplified. What is the setting (Maribor or not, and which more: sometimes you have the plural)? Who and how will be recruited within it? What questionnaires will they complete?... the questionnaires and their scales etc can be uploaded separately as a table. What is this and how will it be used and justified?: 'Additionally, a few additional dichotomous questions will be posed to participants as well (e.g., “Do you think the robot could answer patient’s questions about treatment?). how do you assess 'highly relevant data'?

We have modified the section to clear up inconsistencies and answer questions that you posed in your comment. However, not a lot can be deleted, as the information provided is necessary to explain the study design (and was also explicitly demanded by the reviewers in previous stages of review). Since we are aware that a lot of information is provided, the tables offer a much shorter overview of the study design. We will recruit healthcare professionals and patients. Healthcare professionals will fill out TSES and EAS, while patients will fill out TSES and NARS.

Only patients and healthcare professionals from the University Clinical Centre Maribor will be invited to participate in the study. We deleted plural in this context. The part about “highly relevant data” was deleted.

- change date of start of the study please

Since the article was submitted in June 2021, and we have to comply with the timeline proposed in our project application, the study has already started in December 2021 (immediately after the first round of review). As such, we have changed the estimated start of the study from September 2021 to December 2021 in the manuscript.

- give example of sensitive data you are not collecting

We added a few examples of sensitive data we are not collecting to the manuscript. Examples include (but are not limited to) data revealing racial or ethnic origin, sexual orientation, religious beliefs, … We will also not collect names, diagnosis data, etc.

REVIEWER 1

Thank you very much for the new version of the article, and congratulations for the work done, because I consider that it has been substantially improved since the last version and many issues have been clarified.

I just want to point out some small issues, so that you can assess whether it is appropriate to adjust some minor issues. Prior to this, however, I would like to point out that I am not an expert in quantitative data analysis and, therefore, I am not going to make any assessment of the section "Data analysis and statistics".

Thank you for the positive feedback! We also believe that the manuscript has become much clearer thanks to the constructive comments we received during the review process.

Objectives, p.7:

It is stated that: “R3: What are the main ethical reservations that may impact the acceptance of robotic units in hospital care in healthcare professionals?”

However, a questionnaire cannot determine “what are” the ethical reservations about the introduction of a robot in a hospital; this requires a different kind of bioethical study. With the tools proposed in the article, it is possible to know the opinion of professionals and patients on whether they consider its use to be ethically acceptable or not. Therefore, I would change the sentence or remove this objective.

Thank you very much for your comment. You are absolutely right. As such, we have modified R3. Instead of “what are the main ethical reservations”, we now use “To what extent do healthcare professionals” (…). The focus is hence shifted from identifying the main ethical reservations to exploring the extent of ethical acceptability.

Design and methods, p.8:

It is said that “The study is a cross-sectional survey evaluating expectations, attitudes, and ethical considerations “, but nowhere is it explained what aspects of expectations and attitudes are

measured. On ethical considerations, I have already explained that in order to be rigorous, it is necessary to change the way of expressing it.

The specific instruments that measure expectations, attitudes, and ethical considerations related to the integration of SAHRs are described in the abstract, on page 9, and then again in Table 1. We have also cleared up the terminology used. We now use “ethical acceptability” throughout the text.

It appears by first time “TSES” and “EAS”. Please, include the complete name of the questionnaires.

The EAS designed to assess the appropriateness of the use of robots for therapy with children with autism is now used for a generalist assessment. This is a problem, because we know that expectations and attitudes change depending on whether they refer to a particular technologies and context, or whether they refer to a general context (where stereotypes, etc. are much more involved). I propose, since particular applications are already designed into the project, to include these specifications in the questions. Or, if it is already done, explain it in the paper.

In the previous version, TSES, EAS, and NARS were only fully named in the abstract. In the new version, they are also defined in text (when mentioned for the first time). We cannot change the questionnaires at this point. However, we have added a short explanation in the manuscript on why we believe that the EAS can also be used in a general context.

Patient and Public Involvement, p.13:

It is stated that “Healthcare professionals of the participating medical institution were involved in the study design.” Please, explain briefly how it has been done.

We elaborated on this in the manuscript. Healthcare professionals are among the co-authors of the paper. They helped with the study design (selection of variables and questionnaires, relevant demographic variables) and provided advice on how to carry the study out (e.g., regarding the data collection). This was done in the form of workshops and electronic communication.

Outcomes, p.14:

In Table 2, I think that instead of “expected cohorts” would be better “expected results”.

Yes, thank you. We fixed this.