PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Adaptation of potentially preventable medication-related hospitalisation indicators for Indigenous populations in Australia using a modified Delphi technique
AUTHORS	Spinks, Jean; Kalisch Ellett, Lisa; Spurling, Geoffrey; Theodoros, Theo; Williamson, Daniel; Wheeler, Amanda

VERSION 1 – REVIEW

REVIEWER	Donna Xu Purdue University, USA
REVIEW RETURNED	17-Aug-2019

GENERAL COMMENTS	1. In the introduction section, please provide the background of PPMRH indicators used in Australia and other countries. Please add the objectives of this study. 2. The indicators are developed for Indigenous populations. Why chose this population? In Table 3, how many indicators are specific to this population? 3. In Page 10 Lines 6-9: “In addition to the original 45 indicators,[11] panellists identified a further 56 new indicators. Hence, the Master List of indicators at the start of Round 1 rating comprised 102 indicators." 45+56=101, not 102. 4. Please describe the column of clinical presentation in Table 2. 5. It is better to discuss the similarities and differences between the existing indicators and the newly developed indicators. Moreover, compare the newly developed indicators specific for the Indigenous population with indicators for other populations.

REVIEWER	Cheryl Barnabe University of Calgary, Canada
REVIEW RETURNED	23-Sep-2019

GENERAL COMMENTS	The authors aimed to develop potentially preventable medication-related hospitalization indicators for Indigenous populations in Australia, stated in the introduction as 'developing and testing the feasibility of a culturally appropriate, strengths-based, medication review service', although it is not clear to me that this aim was achieved. The work expanded an existing list of indicators - from an original 45 indicators and adding another 57 for consideration in a Delphi process. From there, 41 were accepted as unchanged, 7 were rejected and the remainder were either modified or merged to leave a final list of 81 indicators. The participants in this process

	were either physicians or pharmacists working in an Indigenous health setting, or from public health. My concerns are 1) considering the development of the original 45 indicators, did your list of indicators simply expand back to a larger list that had already gone through a Delphi process to contract; 2) it is not clear in the results which indicators are original, modified, or from the expanded list; 3) are the selected indicators particularly relevant to Indigenous populations only, and do they reflect the culturally appropriate, strengths-based target you had selected? None of the participants are Indigenous, and it is not clear to me what aspects are really specific to Indigenous contexts. In essence, you are describing that a small group of experienced clinicians working in a particular setting expanded and modified an existing list, but did not take cultural aspects into consideration.

REVIEWER	Christian Dayé Science, Technology and Society (STS) Unit, Graz University of Technology, Austria
REVIEW RETURNED	07-Oct-2019

GENERAL COMMENTS	The present article describes the use of a Delphi procedure to revise and develop a list of indicators of potentially preventable medication-related hospitalisations (PPMRHs). Since my expertise is in social research methodology and specifically in Delphi designs, I was asked by the editors to focus my review on methodological issues. I want to emphasize that I consider it beyond my competence to assess how this study relates to other work done in the field itself. The Delphi design used in the study is well described and the points made to justify the approach are convincing. The modifications made in contrast to what appears to be (or have been) the "standard" form of Delphi—especially the alteration between phases of quantitative assessment and qualitative discussion—have been wise and obviously fruitful. The fact that only a small number of experts participated in the study is a more general problem and not singular to this specific study, although the results might be more convincing if the numbers were higher (both for the quantitative and the qualitative phases). While I thus have virtually no concerns with the choice of method and the study design, two points remained open that relate to the interpretation of the results. Both are basic problems with Delphi studies, so it might be valuable if the authors could address them in a paragraph or two. The first concerns the term “consensus.” On the empirical level, what we observe—especially in the quantitative Delphi phases—is a convergence of expert opinions. To interpret this convergence as “consensus” is non-trivial and requires considerable either specific additional features in the study methodology or background knowledge of the rationales motivating the experts’ assessments. There is reason to believe that the interpretation of convergence as consensus is justified in the study under scrutiny, yet the evidence is scattered and could be assembled at some point in the text. Among the issues corroborating the interpretation of convergence as consensus are: (1) the feeding back of reasons

	for non-affirmative answers; (2) the additional qualitative phases, where the reasoning behind the assessments has partly been discussed; (3) and the wording of the Delphi questionnaire, which made clear to the participants that they were involved in a decision process, a measure counteracting fatigue as the real cause of convergence. It should be pointed out, however, that the decision by the study authors not to collect reasons for affirmative answers ("Accept indicator unchanged") potentially obfuscates dissension, to wit in situations where substantially different reasons lead experts effectively to select the same answer. The second point concerns another fundamental problem with Delphi, namely interdependencies. Delphi is strong when it comes to creating long lists of events or indicators, but it essentially treats them as independent and equally important. This has been a major problem in Delphi's field of origin, future studies, where it quickly became clear that there existed considerable path dependencies between the various events: If event A occurred, this could considerable change the likelihood of events B, C, and D. Now, I lack the medical competence to be ultimately certain about that, but given that paths dependencies exist also in human diseases, it seems to me that even if treating the proposed PPMRH indicators as independent and equally important was unproblematic from a medical standpoint, the authors might wish to note that they are aware of this methodological implication of their method of choice.

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer: 1
Reviewer Name: Donna Xu
Institution and Country: Purdue University, USA
Please state any competing interests or state 'None declared': None declared

1. In the introduction section, please provide the background of PPMRH indicators used in Australia and other countries.

This has been added:

"Clinical indicators have been developed and used in a number of countries to measure PPMRHs which link sub-optimal care involving medication use with subsequent hospitalisation [9-11]. However, differences have been found, for example, between the UK and USA in terms of the inclusion of particular indicators, presumably guided by the prevalence of different health conditions in different population groups and health system differences.[12] Thus, although a set of PPMRH indicators have been developed for use in the general Australian population [13,14], it cannot be assumed that this is a robust measure for specific subsets of the Australian population with distinct healthcare needs, like Indigenous people."

Please add the objectives of this study.

This has been added:

"The objective of this study was to develop a meaningful and clinically relevant outcome measure for use in the Indigenous Medication Review Service pilot study (IMeRSe)[15], which is trialling the feasibility of a culturally appropriate, strengths-based, medication review service."

2. The indicators are developed for Indigenous populations. Why chose this population?

The target population for the IMeRSe feasibility study is Indigenous people. This issue has been addressed in the manuscript by the response to point 1. above:

"Clinical indicators have been developed and used in a number of countries to measure PPMRHs which link sub-optimal care involving medication use with subsequent hospitalisation [9-11]. However, differences have been found, for example, between the UK and USA in terms of the inclusion of particular indicators, presumably guided by the prevalence of different health conditions in different population groups and health system differences.[12] Thus, although a set of PPMRH indicators have been developed for use in the general Australian population [13,14], it cannot be assumed that this is a robust measure for specific subsets of the Australian population with distinct healthcare needs, like Indigenous people."

…alongside the additional text provided in the introduction:

"Previous research has shown that Indigenous people encounter barriers to accessing medication review services[16, 17], thus the aim of IMeRSe is to overcome these barriers and meet the health needs of the population.[15"

…..and the Objective statement at the end of the introduction:

"The objective of this study is to develop a meaningful and clinically relevant outcome measure for use in the Indigenous Medication Review Service pilot study (IMeRSe)[15], which is trialling the feasibility of a culturally appropriate, strengths-based, medication review service."

In Table 3, how many indicators are specific to this population?

Table 3 outlines all included indicators. We have added an additional column to specify the source of the indicator as either:

(i) original (Kalish et.al 2012)

(ii) original (denoting that the original has been adapted in light of updated clinical guidelines); or*

(iii) new.

3. In Page 10 Lines 6-9: “In addition to the original 45 indicators,[11] panellists identified a further 56 new indicators. Hence, the Master List of indicators at the start of Round 1 rating comprised 102 indicators." 45+56=101, not 102.

Thank you for noting this error, it has been corrected.

4. Please describe the column of clinical presentation in Table 2.

The Table 2 header now reads:

“Table 2: Number of clinical indicators, grouped by clinical presentation and round.”

5. It is better to discuss the similarities and differences between the existing indicators and the newly developed indicators. Moreover, compare the newly developed indicators specific for the Indigenous population with indicators for other populations.

Additional text has been added to the Discussion section which reads:

“In comparison to the general Australian population list, the new list contains more neurological indicators (expanded from 7 to 14), vaccine preventable diseases (expanded from 0 to 12) and “other” indicators (expanded from 0 to 9), which better reflects the health burden of the Indigenous population. For example, trachoma and rheumatic heart disease are health issues seen in the Indigenous population, but very rarely in the general Australian population.”

Reviewer: 2
Reviewer Name: Cheryl Barnabe
Institution and Country: University of Calgary, Canada
Please state any competing interests or state ‘None declared’: None declared

The authors aimed to develop potentially preventable medication-related hospitalization indicators for Indigenous populations in Australia, stated in the introduction as 'developing and testing the feasibility of a culturally appropriate, strengths-based, medication review service', although it is not clear to me that this aim was achieved.

The indicator list was developed as one of the main outcome measures for the IMerSe feasibility study. The IMeRSe study is 'developing and testing the feasibility of a culturally appropriate, strengths-based, medication review service'. This has been clarified by the addition of the objective statement at the end of the Introduction section which reads:

“The objective of this study was to develop a meaningful and clinically relevant outcome measure for use in the Indigenous Medication Review Service pilot study (IMeRSe)[15], which is trialling the feasibility of a culturally appropriate, strengths-based, medication review service.”

The work expanded an existing list of indicators - from an original 45 indicators and adding another 57 for consideration in a Delphi process. From there, 41 were accepted as unchanged, 7 were rejected and the remainder were either modified or merged to leave a final list of 81 indicators. The participants in this process were either physicians or pharmacists working in an Indigenous health setting, or from public health.

My concerns are: 1) considering the development of the original 45 indicators, did your list of indicators simply expand back to a larger list that had already gone through a Delphi process to contract;

This has been clarified by additional text added in the Discussion section:

“The development of this updated clinical indicator list is an important step in addressing MRPs for Indigenous Australians. This study builds on earlier work which identified a set of clinical indicators for use in a general Australian population.[130, 141] The new list includes 81 indicators, sourced from 34 of the original 45 indicators and 47 newly identified indicators. In comparison to the list for the general Australian population, the new list contains more neurological indicators (expanded from 7 to 14), vaccine preventable diseases (expanded from 0 to 12) and “other” indicators (expanded from 0 to 9), which better reflects the health burden of the Indigenous population. For example, trachoma and rheumatic heart disease are health issues seen in the Indigenous population, but rarely in the general Australian population.”

2) it is not clear in the results which indicators are original, modified, or from the expanded list;

This has been clarified by the addition of a column in Table 3 in response to Reviewer 2:

Table 3 outlines all included indicators. We have added an additional column to specify the source of the indicator as either:

(i) original (Kalish et.al 2012)

(ii) original (denoting that the original has been adapted in light of updated clinical guidelines); or*

(iii) new.

3) are the selected indicators particularly relevant to Indigenous populations only, and do they reflect the culturally appropriate, strengths-based target you had selected? None of the participants are Indigenous, and it is not clear to me what aspects are really specific to Indigenous contexts. In

essence, you are describing that a small group of experienced clinicians working in a particular setting expanded and modified an existing list, but did not take cultural aspects into consideration.

The Indigenous status of panellists was not reported in the manuscript for concerns around anonymity. The response below is provided by one of the Indigenous panellists/authors:

"A number of key features characterise Aboriginal and Torres Strait Islander Health in contemporary Australia.

1) The first is the ongoing impacts of colonisation, dispossession of land, language and culture and the subsequent impacts of these on both the social determinant and epigenetic determinants of chronic disease in Aboriginal and Torres Strait Islander Australians.

2) The second is the challenge of providing Aboriginal and Torres Strait Islander led care across both the system of care type and incident and prevalent diseases which contextualises the past, present and potential future for Aboriginal and Torres Strait Islander Australians when colonial impacts still pervade contemporary Indigenous populations and communities and act as significant barriers to participation at all levels of society and community.

3) The concept of reconciliation and a shared journey that brings both Indigenous and non-Indigenous populations together in a way that acknowledges and understands the past, and seeks ways to remove the stain of colonisation into the future "shared journey, shared future" should be central to all health research undertaken in Australia.

This study and the IMeRSe program is predicated on the reality of the circumstances above. It looks to address clear gaps in the continuum of care that have immediate and potentially fatal impacts on the population and looks to integrate a model of care that allows for access to services that the vast bulk of other Australians take for granted, but are not available for Indigenous Australians, despite the excess disease burden and barriers to access health services which people face on a day to day basis.

The study is a collaboration between Aboriginal and Torres Strait Islander and other Australian researchers and clinicians and is stronger for this collaboration - the mere fact of lack of clear identification in author citation is in no way a measure of the cultural capability of the team or the processes undertaken by the researchers in conjunction with services and community members.

It is always challenging to contextualise our own knowledge regarding definitions of cultural capability and safety into a completely different setting e.g. Canada First Nations, Inuit and Métis people and Aboriginal and Torres Strait Islander Australians. Given this there perhaps should be a re-focussed

approach from Reviewer 2 to examine the elements of the model and outcomes from the study rather than the cultural veracity/authenticity of the researchers and panellists".

Further, a statement has been added in the Methods section that states:

"Ideally, Indigenous clinicians and researchers would constitute the whole of the CVG, however whilst the CVG did have Indigenous representation and attempts were made to include more, we were not able to convene an entirely Indigenous CVG."

The indicator list developed relates to clinical scenarios encountered in Indigenous health settings in Australia, rather than the cultural aspects of primary care. This concern appears to be related to issue 1. for which additional text was added in the Discussion section for clarification:

"The development of this updated clinical indicator list is an important step in addressing MRPs for Indigenous Australians. This study builds on earlier work which identified a set of clinical indicators for use in a general Australian population.[130, 141] The new list includes 81 indicators, sourced from 34 of the original 45 indicators and 47 newly identified indicators. In comparison to the for the general Australian population list, the new list contains more neurological indicators (expanded from 7 to 14), vaccine preventable diseases (expanded from 0 to 12) and "other" indicators (expanded from 0 to 9), which better reflects the health burden of the Indigenous population. For example, trachoma and rheumatic heart disease are health issues seen in the Indigenous population, but very rarely in the general Australian population."

Reviewer: 3
Reviewer Name: Christian Dayé
Institution and Country: Science, Technology and Society (STS) Unit, Graz University of Technology, Austria
Please state any competing interests or state 'None declared': None declared

The present article describes the use of a Delphi procedure to revise and develop a list of indicators of potentially preventable medication-related hospitalisations (PPMRHs). Since my expertise is in social research methodology and specifically in Delphi designs, I was asked by the editors to focus my review on methodological issues. I want to emphasize that I consider it beyond my competence to assess how this study relates to other work done in the field itself.

The Delphi design used in the study is well described and the points made to justify the approach are convincing. The modifications made in contrast to what appears to be (or have been) the "standard" form of Delphi—especially the alteration between phases of quantitative assessment and qualitative discussion—have been wise and obviously fruitful. The fact that only a small number of experts participated in the study is a more general problem and not singular to this specific study, although the results might be more convincing if the numbers were higher (both for the quantitative and the qualitative phases).

While I thus have virtually no concerns with the choice of method and the study design, two points remained open that relate to the interpretation of the results. Both are basic problems with Delphi studies, so it might be valuable if the authors could address them in a paragraph or two.

The first concerns the term “consensus.” On the empirical level, what we observe—especially in the quantitative Delphi phases—is a convergence of expert opinions. To interpret this convergence as “consensus” is non-trivial and requires considerable either specific additional features in the study methodology or background knowledge of the rationales motivating the experts’ assessments. There is reason to believe that the interpretation of convergence as consensus is justified in the study under scrutiny, yet the evidence is scattered and could be assembled at some point in the text. Among the issues corroborating the interpretation of convergence as consensus are: (1) the feeding back of reasons for non-affirmative answers; (2) the additional qualitative phases, where the reasoning behind the assessments has partly been discussed; (3) and the wording of the Delphi questionnaire, which made clear to the participants that they were involved in a decision process, a measure counteracting fatigue as the real cause of convergence. It should be pointed out, however, that the decision by the study authors not to collect reasons for affirmative answers (“Accept indicator unchanged”) potentially obfuscates dissension, to wit in situations where substantially different reasons lead experts effectively to select the same answer.

We thank the Reviewer for this eloquently stated point of clarification. The following text has been added to the discussion section:

“It must be noted that use of the term “consensus” here, especially in the early phases of the Delphi process, is in fact “convergence” of expert opinion, however, as: (i) panellists were made aware that they were involved in a decision-making process at the start; (ii) justification for non-acceptance was fed-back to the group between rounds; and (iii) face-to-face discussions were held to reach agreement in Round 3, consensus has been assumed.”

The second point concerns another fundamental problem with Delphi, namely interdependencies. Delphi is strong when it comes to creating long lists of events or indicators, but it essentially treats them as independent and equally important. This has been a major problem in Delphi’s field of origin, future studies, where it quickly became clear that there existed considerable path dependencies between the various events: If event A occurred, this could considerable change the likelihood of events B, C, and D. Now, I lack the medical competence to be ultimately certain about that, but given that paths dependencies exist also in human diseases, it seems to me that even if treating the proposed PPMRH indicators as independent and equally important was unproblematic from a medical standpoint, the authors might wish to note that they are aware of this methodological implication of their method of choice.

Again, we than the Reviewer for their insight. The following text has been added as a limitation in the Discussion section:

“Finally, the authors note that the final list of clinical indicators developed here are not necessarily independent of each other, nor are they of equal weighting of clinical seriousness. Thus, this issue will need to be accounted for in the data analysis of the PPMRHs for the IMeRSe study.”

VERSION 2 – REVIEW

REVIEWER	Christian Dayé Science, Technology and Society (STS) Unit ISDS - Institute of Interactive Systems and Data Science Graz University of Technology Austria
REVIEW RETURNED	04-Nov-2019

GENERAL COMMENTS	The two issues I raised in my first review were addressed in the revised version.