PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Group Pregnancy Care for refugee background women: a co-designed, multi-method evaluation protocol applying a community engagement framework and an interrupted time series design
AUTHORS	Riggs, Elisha; Yelland, Jane; Mensah, Fiona; Gold, Lisa; szwarc, Josef; Kaplan, Ida; Small, Rhonda; Middleton, Philippa; Krastev, Ann; Mcdonald, E; East, Christine; Homer, Caroline; Nesvadba, Natalija; Biggs, Laura; Braithwaite, Jeffrey; Brown, Stephanie

VERSION 1 – REVIEW

REVIEWER	Merry, Lisa University of Montreal, Faculty of Nursing
REVIEW RETURNED	25-Feb-2021

GENERAL COMMENTS	Thank you for the opportunity to review this paper. The paper describes a complex intervention (group prenatal care for women with a refugee background) including its development and implementation, and presents the plan for its evaluation. I think the paper is an important contribution to the literature, however, for an outside reader the text isn't always easy to follow and understand. There are also a number of flow issues. I appreciate that the authors are trying to present a lot of information with limited space, and that not all details about the program and its evaluation can be presented, but I still think revisions could be made to improve readability and comprehension. Additional tables or appendices could be used to present some information (e.g., definitions and operationalization of key concepts and frameworks). Alternatively, it may be worthwhile to split the paper into more than one publication so that more details can be provided on the various evaluation phases. Abstract There's no mention of where the study is taking place. Also, the way the text is written it's as though GPC is a distinct program. However, from what I understand GPC is a model of care, and the paper is describing how a specific program in Melbourne, Australia, that is based on this model of care, will be implemented and evaluated. Methods: to maintain the tone of how the text is written, I suggest saying something like "The evaluation of the GPC program involves conducting a multi-site, multi-phase, quasi-experimental study using…"

	Regarding the methods, I don't think it's accurate to say it's a "… a multi-site, multi-phase, quasi-experimental study using community-based participatory research methods and an interrupted time series design"- I think it should be stated like it is in the title: "a multi-site, multi-phase, multi- methods study using community-based participatory research methods and an interrupted time series design" (the quasi experimental design is the interrupted time series). The following sentence isn't clear, "Evaluation measures are underpinned by partnerships, community engagement and capacity building" -what is meant by 'underpinned' (I'm not sure this is the right word)? Ethics and dissemination: what is meant by an 'ethics protocol'? I'm not sure about the verb tense in the first sentence, it seems to me that either the protocols were informed, or they will be informed. And I'm not sure if 'informed' is the right word (maybe developed using co-design and participatory principles?). In terms of order, it would be more logical to first say that measures are being piloted… and then say that the dissemination protocol will be or was developed using co-design and participatory principles (it follows the order of the research process). It would be easier to understand the strengths and limitations if they were all written using complete sentences Intro I have a few minor comments on flow and sentence structure: The second sentence of the first paragraph starts with 'these women' implying women from refugee backgrounds in Australia (based on the first sentence), but the sentence is actually referring to refugee women more broadly (i.e, in high income countries). At the end for the first paragraph, the flow between these two sentences is a bit off: "In addition, the psychological and social impacts of torture and other traumatic events can often be experienced intergenerationally." and "This is particularly significant given the accumulating evidence that exposure to stress and trauma preceding, during, and after pregnancy contributes to a range of adverse outcomes (for example infants born preterm, small for gestational age or with low birthweight), with the potential to affect health across the life course." – They seem like two different ideas- one, trauma can be passed on intergenerationally (via nature and nurture) and two, stress and trauma during pregnancy can have adverse effects on infant outcomes and over the long term. I think the link and flow between the two sentences just needs to be clearer.. Third paragraph: "In the state of Victoria, Australia, 40% of all women giving birth WERE born Overseas…" Generally, the order in which ideas are presented in the introduction, doesn't flow:

	Paragraph 1 = Women from a refugee background and their infants are at greater risk (compared to other women) for adverse outcomes Paragraph 2= High quality care is important to ensure the health of moms and babies; the GPC, offered in Melbourne, Australia, aims to offer culturally safe antenatal care for women from a refugee background; this paper describes the program and how it's being evaluated (it seems odd to present the objective of the paper in the middle of the introduction) Paragraph 3= Statistics on refugee women globally and in Australia, particularly in the state of Victoria are presented (there's no mention of Melbourne ?; stats should probably come earlier in the introduction before talking about the health and outcomes of refugee women) Paragraph 4= What good prenatal care should entail, and the benefits of good prenatal care are described (not sure why information on IPV is presented here, it seems very specific for an introduction; is it because you want to make a case for violence informed care? If so, I think the link needs to be clearer) Paragraph 5 = Women from a refugee background experience barriers in accessing pregnancy care and this contributes to the adverse outcomes (would be good to mention explicitly some of the barriers women face); tailored care can improve experiences… Paragraph 3 should probably come first (general stats), and then paragraph 1 followed by paragraph 4, then paragraph 5, and then finish with paragraph 2 (minus the first sentence). The objective of the paper should come at the end of the introduction. Group prenatal Care The first sentence: "The World Health Organization identified Group Pregnancy Care as having the potential to meet the complex needs of populations vulnerable to poor outcomes, with the Australian antenatal care guidelines identifying potential benefits to women from refugee Backgrounds." – what does it mean that the 'guidelines identified benefits'? Minor suggestion, first paragraph: rather than "avenue for sharing information and developing supportive social networks", maybe "forum for sharing…" Paragraph 2: When describing population outcomes, I think it's more accurate to talk about rates or risks. For example in the following sentence: "This included decreased preterm birth for low-income and African American women, increased care attendance for women with opioid addiction, adolescents and low income groups." – I think it would be more accurate to say a 'decreased preterm birth rates among low-income…" and "increased care attendance rates.." – The sentence is also missing the word 'and' between the two outcomes described. The next sentence has the same issue (i.e., rates of low birthweight, caesarean births, etc..). Co-design and implementation

	The first sentence is a bit confusing: "All Australian residents have access to free pregnancy care at public hospitals, free care from public maternal and child health (and other community health) services, and subsidised care from community-based general practitioners and other medical providers through Australia's universal public health insurance scheme (Medicare)"- I don't think 'free care' needs to be said twice; it may be useful to specify what is meant by 'other medical providers' and if I understand correctly, it's only the subsidized care that is covered under Medicare (under what scheme is the 'free care' covered?)? The other sentences in the same paragraph that describe where women access services during pregnancy, seem a bit redundant with the first sentence of the paragraph. What is the "bridging the gap partnership"? It would be informative to have more information on how the 'co-design' took place. The description of the GPC pilot and the 'key elements' table are lacking some details that would help the reader better understand how the program works. For example, what does "culturally appropriate" care mean exactly? Does it mean that care-providers provide care and information according to women's preferences (so women are asked about what they want and care is individualized)? Or does it mean that the group care is organized so that women from the same cultural background attend the same meetings and a priori care and information are adapted on what is felt would be most culturally appropriate? How are different languages accommodated in a given group? Is there a team of bicultural workers that are available to work with different communities? If I understand correctly, there are multiple locations where groups meet (based on the community) but it's the same core team that provides the care?... it sounds like it was piloted with one specific community? When did the pilot happen exactly (2014 to 2016?)? How many women participated in the program? Bottom of page 6 and top of page 7: A framework (trauma and recovery framework) and a number of key concepts are presented: "culturally safe", "access and engagement with prenatal services", "self-efficacy" and "health literacy"- I think there should be more description and definitions provided for these, especially since hypotheses are also being stated in the same paragraph. Likewise, the birth and family health outcomes should be named and explicitly defined. It would be helpful to have a few more details about the qualitative evaluation- how many women participated (was it everyone who was involved in the program)? Did it include feedback from the

	bicultural worker and others involved in the delivery of the program? Partnership and governance Is the team that is responsible for the evaluation the same as the "bridging the gap partnership"? Are the same people who were involved in the initial pilot work also involved in the larger implementation and evaluation? Partners for sustainability Page 7; "The programs are all situated in large public hospitals…" – I thought they were situated in the community? Is it multiple programs, or one program that is available in multiple sites? How many sites/programs are there? Study aims The aims are presented in a section apart from the hypotheses (the hypotheses were presented earlier in the text). General comment: The text goes back and forth between referring to the GPC as a program versus programs. From figure 3, I understand its one program with 4 sites and 7 communities (although it's not clear if this is a selection of sites or all of the sites of the GPC program). The text needs to be consistent throughout. Study design Same comment as the abstract: I don't think it's accurate to say it's a "… a multi-site, multi-phase, quasi-experimental study using community-based participatory research methods and an interrupted time series design"- I think it should be stated like it is in the title: "a multi-site, multi-phase, multi- methods study using community-based participatory research methods and an interrupted time series design" (the quasi experimental design is the interrupted time series). Comparison populations (Aim 1) Page 12: the terminology is a bit confusing- does 'group' mean 'program' (which actually means community)? I don't understand how a similar comparative group will be obtained contemporaneously – won't all women who meet the criteria be invited/referred to participate in the program?

	Program implementation (Aim 2) I know space is limited, but it would be helpful to have some information on 'the Plan Do Study Act method (PDSA)'. I'm assuming based on aim 2, that quantitative data on -intensity, frequency, acceptability, sustainability- will also be gathered in order to measure associations between these variables and outcomes (part a of aim 2)? From the description provided, I don't see how these data are being gathered. Aim 3 (Women's experiences) Is the total sample 490 women ? The verb tense switches between future and past- in the first paragraph, 'women will be invited to complete an interview…' ; in the second paragraph 'standardised measures were used… interviews were audio-recorded.' Is this evaluation phase already done? Health literacy, social connections and experiences of care are key variables and so the measures that were/will be used should be referenced. Since this is a protocol, it would actually be useful to have the tools available in online appendices. Will similar data be collected from a comparison group (like in Aim 1)? I don't really see why it's a nested cohort, when all women in the GPC program will be invited to participate? Aim 4 What information exactly will be collected from the staff and stakeholders in order to evaluate the cost-effectiveness of the program?

REVIEWER	Muscat, Danielle The University of Sydney, Faculty of Medicine and Health
REVIEW RETURNED	03-Mar-2021

GENERAL COMMENTS	This manuscript details the development and planned evaluation of Group Pregnancy Care (GPC) for women and their families from refugee backgrounds. The paper is clear and well-written, and supported by a thorough literature review. The research fills an important gap. However, it would be useful to have further details about the content of GPC. There are also some discrepancies between the impacts and outcomes detailed in the i) program logic, ii) aims and iii) measured outcomes which should be addressed. Finally, the section of the COVID-19 pandemic suggests that much of the planned evaluation detailed in the protocol may not be practically feasible. Further suggestions are provided below.

	1. Figure 1 is very useful to understand the program logic. However, given that the authors state that "We expect that should the program be able to change individual behaviors (e.g. self-efficacy, health literacy) these determinants are on the pathway to improved birth and family health outcomes", I wonder whether a further breakdown of 'impacts' (e.g. self-efficacy; health literacy) and 'outcomes' (e.g. improved birth and family health outcomes) in the Figure would be useful. See here for some examples: https://www.health.nsw.gov.au/research/Publications/developing-program-logic.pdf 2. It may also be useful to provide more detail about "key program elements" in Figure 1. 3. Several of the impacts and outcomes mentioned in the 'Study Aims' are not represented in the program logic, including access and engagement with preventative healthcare, social and emotional wellbeing, and cost-offsets. The authors might consider adding those into a revised figure. 4. It would be useful to get further information about: a) specific content and program elements; b) how health literacy is defined and operationalised by the study authors. 5. How will the interviews of women's experiences "ask about...health literacy"? How will this be operationalised? 6. How were "health literacy, social connections and experiences of care" measured in the structured interview? Were existing measures used? Or were questions purpose-designed for this project?

VERSION 1 – AUTHOR RESPONSE

Reviewer Comments	
Reviewer #1: Thank you for the opportunity to review this paper. The paper describes a complex intervention (group prenatal care for women with a refugee background) including its development and implementation, and presents the plan for its evaluation. I think the paper is an important contribution to the literature, however, for an outside reader the text isn't always easy to follow and understand. There are also a number of flow issues. I appreciate that the authors are trying to present a lot of information with limited space, and that not all details about the program and its evaluation can be presented, but I still think revisions could be made to improve readability and comprehension. Additional tables or appendices	The authorship team have discussed this issue and we have agreed that the information is best presented in one paper. We have revised the paper to improve readability and comprehension.

could be used to present some information (e.g., definitions and operationalization of key concepts and frameworks). Alternatively, it may be worthwhile to split the paper into more than one publication so that more details can be provided on the various evaluation phases.	
Abstract	
--There is no mention of where the study is taking place. Also, the way the text is written, it is as though GPC is a distinct program. However, from what I understand GPC is a model of care, and the paper is describing how a specific program in Melbourne, Australia, that is based on this model of care, will be implemented and evaluated.	We have inserted Melbourne, Australia into the abstract and in the methods section of the text and on Page 10 Whilst other models of Group Antenatal Care exist internationally, this is the first model co-designed with refugee background women implemented by an inter-agency multidisciplinary team. We refer to it as a distinct 'program' as part of this evaluation. The aim is that GPC becomes a model of care that can be implemented with other services/communities where there is an identified need.
--<u>*Methods and analysis:*</u> to maintain the tone of how the text is written, I suggest saying something like "The evaluation of the GPC program involves conducting a multi-site, multi-phase, quasi-experimental study using…"	Change made. Also reflected on Page 9.
--Regarding the methods, I don't think it's accurate to say it's a "… a multi-site, multi-phase, quasi-experimental study using community-based participatory research methods and an interrupted time series design"- I think it should be stated like it is in the title: "a multi-site, multi-phase, multi- methods study using community-based participatory research methods and an interrupted time series design" (the quasi experimental design is the interrupted time series).	Change made. Also reflected on Page 9.
--The following sentence isn't clear, "Evaluation measures are underpinned by partnerships, community engagement and capacity building" - what is meant by 'underpinned' (I'm not sure this is the right word)?	This has been re-phrased.

--<u>*Ethics and dissemination:*</u> What is meant by an 'ethics protocol'? I'm not sure about the verb tense in the first sentence, it seems to me that either the protocols were informed, or they will be informed. And I'm not sure if 'informed' is the right word (maybe developed using co-design and participatory principles?). In terms of order, it would be more logical to first say that measures are being piloted… and then say that the dissemination protocol will be or was developed using co-design and participatory principles (it follows the order of the research process).	This section has been re-phrased incorporating the suggestions made by the reviewer.
-- It would be easier to understand the strengths and limitations if they were all written using complete sentences	These have been amended.
Introduction	
I have a few minor comments on flow and sentence structure: -- The second sentence of the first paragraph starts with 'these women' implying women from refugee backgrounds in Australia (based on the first sentence), but the sentence is actually referring to refugee women more broadly (i.e, in high income countries).	We have removed 'in Australia' as the evidence suggests it is for all women of refugee backgrounds
-- At the end for the first paragraph, the flow between these two sentences is a bit off: "In addition, the psychological and social impacts of torture and other traumatic events can often be experienced intergenerationally." and "This is particularly significant given the accumulating evidence that exposure to stress and trauma preceding, during, and after pregnancy contributes to a range of adverse outcomes (for example infants born preterm, small for gestational age or with low birthweight), with the potential to affect health across the life course." – They seem like two different ideas- one, trauma can be passed on intergenerationally (via nature and nurture) and two, stress and trauma during pregnancy can have adverse effects on infant outcomes and over the long	Thank you for identifying this. We have re-written this and made a new paragraph. Page 5.

term. I think the link and flow between the two sentences just needs to be clearer.	
-- Third paragraph: “In the state of Victoria, Australia, 40% of all women giving birth WERE born Overseas…”	Amended

Generally, the order in which ideas are presented in the introduction, doesn't flow: Paragraph 1 = Women from a refugee background and their infants are at greater risk (compared to other women) for adverse outcomes Paragraph 2 = High quality care is important to ensure the health of moms and babies; the GPC, offered in Melbourne, Australia, aims to offer culturally safe antenatal care for women from a refugee background; this paper describes the program and how it's being evaluated (it seems odd to present the objective of the paper in the middle of the introduction) Paragraph 3 = Statistics on refugee women globally and in Australia, particularly in the state of Victoria are presented (there's no mention of Melbourne ?; stats should probably come earlier in the introduction before talking about the health and outcomes of refugee women) Paragraph 4 = What good prenatal care should entail, and the benefits of good prenatal care are described (not sure why information on IPV is presented here, it seems very specific for an introduction; is it because you want to make a case for violence informed care? If so, I think the link needs to be clearer) Paragraph 5 = Women from a refugee background experience barriers in accessing pregnancy care and this contributes to the adverse outcomes (would be good to mention explicitly some of the barriers women face); tailored care can improve experiences… -- Paragraph 3 should probably come first (general stats), and then paragraph 1 followed by paragraph 4, then paragraph 5, and then finish with paragraph 2 (minus the first sentence). The objective of the paper should come at the end of the introduction.	Thanks for you the suggestion to reorganise this section. We have amended as per the suggestion. Page 5
Group Pregnancy Care for women and their families from a refugee background	
-- The first sentence: "The World Health Organization identified Group Pregnancy Care	This has been edited for clarification. Page 6

as having the potential to meet the complex needs of populations vulnerable to poor outcomes, with the Australian antenatal care guidelines identifying potential benefits to women from refugee Backgrounds." – what does it mean that the 'guidelines identified benefits'?	
-- Minor suggestion, first paragraph: rather than "avenue for sharing information and developing supportive social networks", maybe "forum for sharing…"	Change made. Page 6
--Paragraph 2: When describing population outcomes, I think it's more accurate to talk about rates or risks. For example in the following sentence: "This included decreased preterm birth for low-income and African American women, increased care attendance for women with opioid addiction, adolescents and low income groups." – I think it would be more accurate to say a 'decreased preterm birth rates among low-income…" and "increased care attendance rates.." – The sentence is also missing the word 'and' between the two outcomes described. The next sentence has the same issue (i.e., rates of low birthweight, caesarean births, etc..).	In most cases the studies cited have reported differences in proportions, rather than rates. We have clarified this in the text.
Co-design and implementation of a new model of Group Pregnancy Care	
--The first sentence is a bit confusing: "All Australian residents have access to free pregnancy care at public hospitals, free care from public maternal and child health (and other community health) services, and subsidised care from community-based general practitioners and other medical providers through Australia's universal public health insurance scheme (Medicare)"- I don't think 'free care' needs to be said twice; it may be useful to specify what is meant by 'other medical providers' and if I understand correctly, it's only the subsidized care that is covered under Medicare (under what scheme is the 'free care' covered?)? The other sentences in the same paragraph that describe where women access services during pregnancy, seem a bit	This section has been amended as suggested and a new reference added. Page 6.

redundant with the first sentence of the paragraph.	
--What is the “bridging the gap partnership”?	The partnership has been described on Page 6.
--It would be informative to have more information on how the ‘co-design’ took place.	Given space is limited in the paper, we have referred to the paper where more information can be found on the co-design process that was undertaken. Page 6/7.
--The description of the GPC pilot and the ‘key elements’ table are lacking some details that would help the reader better understand how the program works. For example, what does “culturally appropriate” care mean exactly? Does it mean that care-providers provide care and information according to women’s preferences (so women are asked about what they want and care is individualized)? Or does it mean that the group care is organized so that women from the same cultural background attend the same meetings and a priori care and information are adapted on what is felt would be most culturally appropriate?	Further detail has been provided in Table 1 to better describe the Key elements of GPC. The additions also hope to explain more clearly what is meant by culturally appropriate care in the context of GPC.
How are different languages accommodated in a given group? Is there a team of bicultural workers that are available to work with different communities? If I understand correctly, there are multiple locations where groups meet (based on the community) but it’s the same core team that provides the care?...	Each GPC program is culturally and language specific. One bicultural worker is involved with each program who may speak multiple languages. There is a different core team for each GPC program as it is staffed by the relevant hospital and MCH service in the specific area of Melbourne.
it sounds like it was piloted with one specific community? When did the pilot happen exactly	GPC was piloted with the one community, the Karen community (from Burma) in 2016. The

(2014 to 2016?)? How many women participated in the program?	qualitive evaluation was completed with 19 women who participated in GPC.
--Bottom of page 6 and top of page 7: A framework (trauma and recovery framework) and a number of key concepts are presented: "culturally safe", "access and engagement with prenatal services", "self-efficacy" and "health literacy"- I think there should be more description and definitions provided for these, especially since hypotheses are also being stated in the same paragraph. Likewise, the birth and family health outcomes should be named and explicitly defined.	We have inserted some definitions of key study terms. Page 7. We are monitoring a broad range of perinatal outcomes in the evaluation. Specifically, we have conceptualised 'family outcomes' as part of evaluation measures to assess 'stressful life events' and 'intimate partner abuse', for example.
--It would be helpful to have a few more details about the qualitative evaluation- how many women participated (was it everyone who was involved in the program)? Did it include feedback from the bicultural worker and others involved in the delivery of the program?	The number of women (n=19) who participated in the evaluation of the pilot study has been inserted. Page 8. Due to space limitations we have included brief details, and have provided the reference for the published paper should readers wish to read about it in detail. Feedback from staff were not included in this evaluation/paper – there is a focus on exploring these views with the evaluation described in this protocol paper. Page 14
Partnership and Governance	
--Is the team that is responsible for the evaluation the same as the "bridging the gap partnership"? Are the same people who were involved in the initial pilot work also involved in the larger implementation and evaluation?	Clarification has been provided to explain that the GPC partnership and team build upon the Bridging the Gap partnership. The funding provided for this evaluation supported an expansion of the partnership to include additional hospitals and maternal and child health services.
Partners for sustainability	
	GPC is situated in the community. Thank you for identifying this, we have amended. Page 8

--Page 7; "The programs are all situated in large public hospitals…" – I thought they were situated in the community? Is it multiple programs, or one program that is available in multiple sites? How many sites/programs are there?	Figure 3 provides an overview of the GPC sites.
Study Aims	
--The aims are presented in a section apart from the hypotheses (the hypotheses were presented earlier in the text).	We have moved the hypotheses to accompany the aims. Page 10.
--General comment: The text goes back and forth between referring to the GPC as a program versus programs. From figure 3, I understand its one program with 4 sites and 7 communities (although it's not clear if this is a selection of sites or all of the sites of the GPC program). The text needs to be consistent throughout.	Thank you for identifying this. We have thoroughly checked for consistency throughout the paper.
Study design	
--Same comment as the abstract: I don't think it's accurate to say it's a "… a multi-site, multi-phase, quasi-experimental study using community-based participatory research methods and an interrupted time series design"- I think it should be stated like it is in the title: "a multi-site, multi-phase, multi- methods study using community-based participatory research methods and an interrupted time series design" (the quasi experimental design is the interrupted time series).	This has been amended. Page 11
<u>Comparison populations (Aim 1):</u> --Page 12: the terminology is a bit confusing- does 'group' mean 'program' (which actually means community)?	We agree the terminology is confusing and have removed the work program where possible. GPC is referred to as such and 'community' refers to the cultural group.
I don't understand how a similar comparative group will be obtained contemporaneously –	Yes, this is correct. All eligible women will be invited to attend GPC. However, we anticipate

won't all women who meet the criteria be invited/referred to participate in the program?	that not all women will be able to attend (due to unavailability on the selected day, work commitments, don't feel they need the support of the group etc). This group of women will be considered the comparison.
Program implementation (Aim 2): --I know space is limited, but it would be helpful to have some information on 'the Plan Do Study Act method (PDSA)'	As requested, we have included some detail about PDSA. Page 14.
--I'm assuming based on aim 2, that quantitative data on -intensity, frequency, acceptability, sustainability- will also be gathered in order to measure associations between these variables and outcomes (part a of aim 2)? From the description provided, I don't see how these data are being gathered.	Thank you for identifying this. We have edited Aim 2 to provide clarity about what will be assessed and collected.
Aim 3 (Women's experiences) --Is the total sample 490 women ?	Yes. This has been made clear on Page 13.
-- The verb tense switches between future and past- in the first paragraph, 'women will be invited to complete an interview…' ; in the second paragraph 'standardised measures were used… interviews were audio-recorded.' Is this evaluation phase already done?	We have edited to ensure consistent tense. Thank you for identifying this.
-- Health literacy, social connections and experiences of care are key variables and so the measures that were/will be used should be referenced. Since this is a protocol, it would actually be useful to have the tools available in online appendices.	There were no existing standard measures that were deemed to be appropriate for health literacy or social connections, thus, these are study designed measures drawing upon the existing literature and what the partners/services wanted to know. We have drawn upon previous research of the investigators to include questions/measures used in previous studies. We are happy to provide the Interview schedule upon request.

-- Will similar data be collected from a comparison group (like in Aim 1)?	Due to funding and time constraints, there is no comparison group for the data collected through interviews.
-- I don't really see why it's a nested cohort, when all women in the GPC program will be invited to participate?	Thank you for identifying this, it has been removed.
Aim 4 --What information exactly will be collected from the staff and stakeholders in order to evaluate the cost-effectiveness of the program?	Information collected related to GPC implementation costs: staff salaries, venue, professional development, and consumables e.g. refreshments). These total costs are then used to estimate cost-savings in-terms of reducing pre-term births and low birth weight infants.
Reviewer#2: This manuscript details the development and planned evaluation of Group Pregnancy Care (GPC) for women and their families from refugee backgrounds. The paper is clear and well-written, and supported by a thorough literature review. The research fills an important gap. However, it would be useful to have further details about the content of GPC. There are also some discrepancies between the impacts and outcomes detailed in the i) program logic, ii) aims and iii) measured outcomes which should be addressed. Finally, the section of the COVID-19 pandemic suggests that much of the planned evaluation detailed in the protocol may not be practically feasible. Further suggestions are provided below. -- Figure 1 is very useful to understand the program logic. However, given that the authors state that "We expect that should the program be able to change individual behaviors (e.g. self-efficacy, health literacy) these determinants are on the pathway to improved birth and family health outcomes", I wonder whether a further breakdown of 'impacts' (e.g. self-efficacy; health	Thank you for identifying these points for us to consider and for the reference to developing program logics. There are many ways that a program logic can be designed. Our diagram has been designed with the team of Investigators and has been used to inform all aspects of the evaluation so far, therefore we feel we cannot change it, though we acknowledge it may have some limitations.

literacy) and 'outcomes' (e.g. improved birth and family health outcomes) in the Figure would be useful. See here for some examples: https://www.health.nsw.gov.au/research/Publications/developing-program-logic.pdf	
-- It may also be useful to provide more detail about "key program elements" in Figure 1.	We have added detail to Figure 1 in response to the comments from reviewer 1, which hopefully provide more detail. We have also inserted a reference to the study website which presents further details.
-- Several of the impacts and outcomes mentioned in the 'Study Aims' are not represented in the program logic, including access and engagement with preventative healthcare, social and emotional wellbeing, and cost-offsets. The authors might consider adding those into a revised figure.	The program logic has been agreed to and approved by the study Investigators. We consider the additional figures that have been provided as part of this paper (Figures 2 and 3) to provide sufficient detail of the evaluation aims. These are separate to the conceptual framework of the 'GPC model'.
-- It would be useful to get further information about: a) specific content and program elements; b) how health literacy is defined and operationalised by the study authors.	a) Further detail has been provided in table 1 and a reference to the study website provided to refer to further information. b) some key definitions have been included on Page 7.
-- How will the interviews of women's experiences "ask about…health literacy"? How will this be operationalised?	No standard measure application to refugee maternal health literacy was identified. We developed our own set of questions that can be categorised within the 8 domains of the Health Literacy Questionnaire (Osborne 2013).
-- How were "health literacy, social connections and experiences of care" measured in the structured interview? Were existing measures used? Or were questions purpose-designed for this project?	There were no existing standard measures that were deemed to be appropriate for health literacy or social connections, thus, these are study designed measures drawing upon the existing literature and what the partners/services wanted to know. We also drew upon the research of the study investigators who have previously applied interview questions on these topics that have been well received by women.