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)	Complex Care for Kids Ontario: Protocol for a mixed-methods randomized controlled trial of a population-level care coordination initiative for children with medical complexity
AUTHORS	Orkin, Julia; Chan, Carol; Fayed, Nora; Lin, Jia Lu Lilian; Major, Nathalie; Lim, Audrey; Peebles, Erin; Moretti, Myla; Soscia, Joanna; Sultan, Roxana; Willan, Andrew; Offringa, Martin; Guttmann, Astrid; Bartlett, Leah; Kanani, Ronik; Culbert, Erin; Hardy-Brown, Karolyn; Gordon, Michelle; Perlmutar, Marty; Cohen, Eyal

VERSION 1 - REVIEW

REVIEWER	Rita Mangione-Smith University of Washington, USA
REVIEW RETURNED	07-Dec-2018

GENERAL COMMENTS	This study protocol regarding a randomized controlled trial of an intervention focused on improving care coordination for children with medical complexity (CMC) is well written and the study (which is already underway) addresses a critically important topic area related to effective models of care for this vulnerable population of children. Issues to consider: 1) The study cited as reference #28, is not accurately described. The protocol should be corrected to reflect that this was not a randomized controlled trial. It was a longitudinal cohort study that examined FECC quality measure sensitivity to change over-time. The patients included in the cohort were CMC who were enrolled in a regional care coordination program and then followed over an 18-month period with FECC measures assessed at baseline, 6 months, 12 months, and 18 months. This study did demonstrate improvement in several of the FECC measures over-time, but there was no control comparison group and study participants were not randomized. 2) Is reference #33 referring to this website: https://www.seattlechildrens.org/research/centers-programs/child-health-behavior-and-development/labs/mangione-smith-lab/measurement-tools/ ? If so, the investigators may want to consider also citing the following publications related to FECC development, field testing, and validation that were published after their current study began in January 2017: Gidengil C, Parast L, Burkhart Q, Brown J, Elliott MN, Lion KC, McGlynn EA, Schneider EC, Mangione-Smith R. Development and

	Implementation of the Family Experiences with Coordination of Care Survey Quality Measures. Acad Pediatr. 2017; pii: S1876-2859(17)30118-3. doi: 10.1016/j.acap.2017.03.012. [Epub ahead of print]. PMID: 28373108 Parast L, Burkhart Q, Gidengil C, Schneider EC, Mangione-Smith R, Lion KC, McGlynn EA, Carle A, Britto MT, Elliott MN. Validation of New Care Coordination Quality Measures for Children with Medical Complexity. Acad Pediatr. 2018 Mar 14. pii: S1876-2859(18)30126-8. doi: 10.1016/j.acap.2018.03.006. [Epub ahead of print] PMID: 29550397 3) For the planned cost analysis, it appears that the investigators do not intend to include intervention costs, e.g., the costs associated with the "key workers." If they do intend to include intervention costs, a description of how they plan to do so should be included in the protocol. If not, this is problematic as others who may want to replicate the care coordination services they are providing through CCKO and testing with this intervention trial, will need to know the costs of doing so.

REVIEWER	Damhnat McCann University of Tasmania, Australia
REVIEW RETURNED	26-Feb-2019

GENERAL COMMENTS	Thank you for the opportunity to review this well written paper which outlines the protocol for an evaluation of a care coordination program for children with medical complexity. The fact that it is a population level evaluation that includes multiple sites, with focused outcomes developed in conjunction with parents, are key strengths of the study. The authors rightly assert that population level evaluations using RCT are needed to inform the evidence base in this area. I wish them well in the conduct of their study and look forward to the results. This is the protocol for an ongoing study, but I have some comments relating to the study background and the level of detail provided in the methods that hopefully will be helpful. Background Page 5 Line 47: I would avoid making a generalised assertion that marital discord is a direct consequence of being a parental caregiver and suggest the term 'marital strain' if discussing this area. Some of the more recent research has shown only a small effect size or no increased risk of relationship breakdown (eg. Namkung et al, 2015; Risdal & Singer, 2004; Tossebro & Wendelborg, 2017. Page 7 beginning line 8: The background information relating to previously published RCTs would be strengthened by including a wider range of studies that have used RCT to evaluate care coordination programs (eg. Farmer et al, 2011; Looman et al., 2015; McKissick et al, 2017). Methods/Design Page 9 line 54: How will exclusion criteria (g) be assessed? Greater detail is needed regarding the circumstances in which this criteria is anticipated to occur. Does this relate for example to children who are expected to be placed in out of home care during the study period?

	Page 11 line 54 and reference list: References 31 and 32 appear to be the same reference. Page 13 line 3: What is the justification for using only one subscale from KIDSCREEN-52 rather than the entire tool? Has the use of the single subscale to assess quality of life and emotional health been validated in studies with similar populations? Line 7: Is the VAS only administered once at each of the four data collection timepoints? If so, how will the analysis of this data contribute to the study outcomes, considering that the single time points may not represent the child's pain experience more generally during the study period? Line 41: An adapted version of the KIDSCREEN-52 subscale is being used with the parents – has this been validated for use in an adult population? Page 11 line 40: Children who have a change in their clinical status and meet exclusion criteria related to increasing severity of their condition will be removed from the waitlist and treated in the intervention arm. How will these children be accounted for in the analysis? While the principle of intention to treat is being applied to the study, these children would otherwise have been excluded and therefore represent a significantly different population from the other study participants. Page 15 line 5: The sample size is based on a loss to follow up of only 10%. This seems low for a study conducted over two years and requiring parents to complete quite detailed questionnaires. Is this calculation based on previous research with this population? Line 24: It would be good to have a little more detail regarding the recruitment of participants. Line 42: Will children be included as participants in the qualitative interviews? Page 16 line 26: It would be helpful to have a copy of the demographic questionnaire included as an appendix. Is a functional assessment of the child included in the demographic questionnaire? References Farmer, J. E., Clark, M. J., Drewel, E. H., Swenson, T. M., & Ge, B. (2011). Consultative care coordination through the medical home for CSHCN: A randomized controlled trial. Maternal and Child Health Journal, 15(7), 1110-1118. doi:10.1007/s10995-010-0658-8 Looman, W. S., Antolick, M., Cady, R. G., Lunos, S. A., Garwick, A. E., & Finkelstein, S. M. (2015). Effects of a Telehealth Care Coordination Intervention on Perceptions of Health Care by Caregivers of Children With Medical Complexity: A Randomized Controlled Trial. J Pediatr Health Care, 29(4), 352-363. doi:10.1016/j.pedhc.2015.01.007 McKissick, H. D., Cady, R. G., Looman, W. S., & Finkelstein, S. M. (2017). The Impact of Telehealth and Care Coordination on the Number and Type of Clinical Visits for Children With Medical Complexity. Journal of Pediatric Health Care, 31(4), 452-458. doi:10.1016/j.pedhc.2016.11.006

	Namkung, E. H., Song, J., Greenberg, J. S., Mailick, M. R., & Floyd, F. J. (2015). The Relative Risk of Divorce in Parents of Children With Developmental Disabilities: Impacts of Lifelong Parenting. Am J Intellect Dev Disabil, 120(6), 514-526. doi:10.1352/1944-7558-120.6.514 Risdal, D., & Singer, G. H. S. (2004). Marital Adjustment in Parents of Children with Disabilities: A Historical Review and Meta-Analysis. Research and Practice for Persons with Severe Disabilities, 29(2), 95-103. doi:10.2511/rpsd.29.2.95 Tossebro, J., & Wendelborg, C. (2017). Marriage, Separation and Beyond: A Longitudinal Study of Families of Children with Intellectual and Developmental Disabilities in a Norwegian Context. J Appl Res Intellect Disabil, 30(1), 121-132. doi:10.1111/jar.12225

VERSION 1 – AUTHOR RESPONSE

1. The study cited as reference #28, is not accurately described. The protocol should be corrected to reflect that this was not a randomized controlled trial. It was a longitudinal cohort study that examined FECC quality measure sensitivity to change over-time. The patients included in the cohort were CMC who were enrolled in a regional care coordination program and then followed over an 18-month period with FECC measures assessed at baseline, 6 months, 12 months, and 18 months. This study did demonstrate improvement in several of the FECC measures overtime, but there was no control comparison group and study participants were not randomized. We have amended the description of this study in the manuscript to reflect the fact that the study was a longitudinal cohort study instead of a randomized controlled trial. Please refer to page 12 for the amendment.

2. Is reference #33 referring to this website:

https://www.seattlechildrens.org/research/centersprograms/child-health-behavior-anddevelopment/labs/mangione-smith-lab/measurement-tools

If so, the investigators may want to consider also citing the following publications related to FECC development, field testing, and validation that were published after their current study began in January 2017:

Gidengil C, Para st L, Burkhart Q, Brown J, Elliott MN, Lion KC, McGlynn EA, Schneider EC,

Mangione-Smith R. Development and Implementation of the Family Experiences with

Coordination of Care Survey Quality Measures. Acad Pediatr. 2017; pii: S1876-2859(17)30118-3. doi:10.1016/j.acap.2017.03.012. [Epub ahead of print]. PMID: 28373108

Parast L, Burkhart Q, Gidengil C, Schneider EC, Mangione-Smith R, Lion KC, McGlynn EA, Carle A, Britto MT, Elliott MN. Validation of New Care Coordination Quality Measures for Children with Medical Complexity. Acad Pediatr. 2018 Mar 14. pii: S1876-2859(18)30126-8.

doi:10.1016/j.acap.2018.03.006. [Epub ahead of print] PMID: 29550397

We have amended the reference to include the website link. Please refer to the references list to see this change.

We have also included the references you provided; thank you very much for providing this list of references to us.

3. For the planned cost analysis, it appears that the investigators do not intend to include intervention costs, e.g., the costs associated with the “key workers.” If they do intend to include intervention costs, a description of how they plan to do so should be included in the protocol. If not, this is problematic as others who may want to replicate the care coordination services they are providing through CCKO and testing with this intervention trial, will need to know the costs of doing so.

Yes we plan on doing an analysis of the intervention costs: a cost-effectiveness analysis will be performed alongside this clinical trial to estimate the incremental costs (or savings) of the CCKO initiative compared to standard care in reducing hospitalization. Both a health care system and societal perspective will be used with a time horizon of 12 months. Please see pages 13-14 for more details.

4. Page 5 Line 47: I would avoid making a generalised assertion that marital discord is a direct consequence of being a parental caregiver and suggest the term ‘marital strain’ if discussing this area. Some of the more recent research has shown only a small effect size or no increased risk of relationship breakdown (eg. Namkung et al, 2015; Risdal & Singer, 2004; Tossebro & Wendelborg, 2017.

We have amended that statement to say “marital strain” instead. Please refer to page 5 for this amendment.

5. Page 7 beginning line 8: The background information relating to previously published RCTs would be strengthened by including a wider range of studies that have used RCT to evaluate care coordination programs (eg. Farmer et al, 2011; Looman et al., 2015; McKissick et al, 2017).

We have added the above-mentioned studies to strengthen the background information. Please refer to pages 5-6 for this additional information.

6. Page 9 line 54: How will exclusion criteria (g) be assessed? Greater detail is needed regarding the circumstances in which this criteria is anticipated to occur. Does this relate for example to children who are expected to be placed in out of home care during the study period? When the triaging team receives the referrals, a social history for the child is provided; within the social history it will reflect the child’s home situation. Exclusion criteria (g) is typically applied to children who are placed in out of home care with no set date of returning to the care of the biological parent(s). It would not be feasible to enroll these children into the research study because it is unknown if their current caregiver will be the same caregiver over the duration of the study (2 years). Please refer to page 9 for more details regarding this process.

7. Page 11 line 54 and reference list: References 31 and 32 appear to be the same reference. We have consolidated those 2 references into one reference. Please refer to the references list for this change.

8. Page 13 line 3: What is the justification for using only one subscale from KIDSCREEN-52 rather than the entire tool? Has the use of the single subscale to assess quality of life and emotional health been validated in studies with similar populations?

Given that quality of life is a secondary outcome, a 52 item scale was inappropriately long to administer. The 52 version has subscales that each have their own respective psychometric properties as reported in the manual. The “feelings’ subscale are the general overall questions about a child’s life. Therefore it was appropriate to use only the subscale that had content validity for assessing QOL in this population.

It has been used in this way in the PICU Weecover* cohort study, which is a similar population. Also, the use of single subscales of the KIDSCREEN-52 is valid when there is content and theoretical reasons for looking at one component of QOL as is the case in this group. *Reference: https://www.ncbi.nlm.nih.gov/pubmed/29394221

9. Line 7: Is the VAS only administered once at each of the four data collection timepoints? If so, how will the analysis of this data contribute to the study outcomes, considering that the single time points may not represent the child’s pain experience more generally during the study period?

Yes the VAS is administered once at each of the four data collection timepoints.

We agree that timing and frequency of pain measurement is always a challenge. The elicitation of VAS in this study is non-specific to location and in type because in this non-categorical group the location and types of pain vary considerably (e.g., pain due to contractures, GI pain, constipation, medical procedures, etc). Based on previous reviews of pain scales in this situation asking about general pain, cues proxies to respond about their impressions of their child’s pain in a recent window or since the last turning point in their health*. We felt the limitations were acceptable for the 6 month to 1 year window between measurement points in a betweengroups study, recognizing this method would inappropriate for clinical evaluation and tracking. *Reference: https://www.ncbi.nlm.nih.gov/pubmed/16996689

10. Line 41: An adapted version of the KIDSCREEN-52 subscale is being used with the parents – has this been validated for use in an adult population?

No this has not been validated for use in the adult population, but there is face and content validity of the items for adults.

To measure adult quality of life we are administering the Diener SWLS.

The KIDSSCREEN “Feelings” subscale is being used to perform cross validation on the veracity of the parent’s proxy of the child’s quality of life using identical items. This is a methodological analysis to control for parental depression effects in proxy reporting of child quality of life.

11. Page 11 line 40: Children who have a change in their clinical status and meet exclusion criteria related to increasing severity of their condition will be removed from the waitlist and treated in the intervention arm. How will these children be accounted for in the analysis? While the principle of intention to treat is being applied to the study, these children would otherwise have been excluded and therefore represent a significantly different population from the other study participants.

At the time of adjudication these children did not meet any exclusion criteria; post priori data would not influence this adjudication. Per Dentry et al's JAMA article*, with the intention to treat (ITT) principle, we will analyze these children according to the group they were assigned to at the time of randomization – regardless of whether or not they received the intended treatment. In order to obtain an unbiased estimate of the effect of the different treatment groups we must analyze all patients per their original treatment group. *Reference: https://jamanetwork.com/journals/jama/fullarticle/1884555

12. Page 15 line 5: The sample size is based on a loss to follow up of only 10%. This seems low for a study conducted over two years and requiring parents to complete quite detailed questionnaires. Is this calculation based on previous research with this population?

In our experiences with this population we have overserved similar losses. Moreover, we have accounted for this loss to follow up by adding 10% in our sample size calculations.

13. Line 24: It would be good to have a little more detail regarding the recruitment of participants. Please see pages 14-15 for a little more detail regarding the recruitment of participants for the study.

14. Line 42: Will children be included as participants in the qualitative interviews?

No, as the majority of children within complex care are very young and/or non-verbal, children are not included as participants in the qualitative interviews. Please see page 15 for the clarification that only parents will be participants for the qualitative interviews.

15. Page 16 line 26: It would be helpful to have a copy of the demographic questionnaire included as an appendix. Is a functional assessment of the child included in the demographic questionnaire?

No, a functional assessment of the child is not included in the demographic questionnaire (please see Appendix 5 for a copy of the demographic questionnaire). However, we collect functional information about each child in a clinical information form (please see Appendix 6 for a copy of this clinical information form). Moreover, the triaging team completes a functional assessment of each referral using the standard operational definition for children with medical complexity who are medically fragile and/or technology dependent (please refer to figure 2 of the appendix).

VERSION 2 – REVIEW

REVIEWER	Rita Mangione-Smith Professor and Chief, Division of General Pediatrics & Hospital Medicine University of Washington, Department of Pediatrics Seattle, WA, USA I was involved in the development of the Family Experiences with Coordination of Care (FECC) measure set being utilized in this RCT as the primary outcome measure.
REVIEW RETURNED	01-Apr-2019

GENERAL COMMENTS	The authors have addressed all of the concerns I raised on initial review of this paper. I have one new issue the authors should address to strengthen the methods section of the paper. The FECC survey consists of 20 individual care coordination quality measures. The 20 measures do not function as a composite score and there are no FECC survey domains. It would be helpful to know if the authors are collecting data on all 20 FECC measures included in the survey or a subset of these measures for their primary outcome assessment. If they are using a sub-set of the measures, they should indicate which measures they intend to report. It would also be important to know that the FECC SD calculated from the prior validation sample and used for power calculations was based on only those FECC measures that will be assessed. Once these clarifications are added, I have no further suggestions for improving this manuscript. The authors have been responsive to the reviewer comments they received and the paper is well written.

REVIEWER	Damhnat McCann University of Tasmania, Australia
REVIEW RETURNED	15-Apr-2019

GENERAL COMMENTS	Thanks for the opportunity to review the revised version of this paper which outlines the protocol for an evaluation of a care coordination program for children with medical complexity. The authors have made minor adjustments to the paper which provides additional detail to assist reader understanding. I also appreciate the clarification provided with regard to the data collection tools. I wish the researchers well with the study and look forward to seeing the findings.

Correction: *Complex care for kids Ontario: protocol for a mixed-methods randomised controlled trial of a population-level care coordination initiative for children with medical complexity*

Orkin J, Chan CY, Fayed N, *et al*. Complex care for kids Ontario: protocol for a mixed-methods randomised controlled trial of a population-level care coordination initiative for children with medical complexity. *BMJ Open* 2019;9:e028121. doi: 10.1136/bmjopen-2018-028121.

This article was previously published with errors in data.

- In the 'Primary outcome' section, there's a typo: it says "as they hadve content relation to:"; it should instead read "as they have content relation to:"
- In the 'Sample size' section,
  - The last sentence in the first paragraph is missing the second part. It should read: 'projected smallest clinically important difference of 0.5 of the within-patient SD, which is recommended by the developer as a moderate effect size.'
  - The second paragraph should start with 'The required sample size is considered feasible as it is estimated that a pool of about 250 patients are readily identifiable for recruitment at CCKO sites.'

BMJ Open 2019;**9**:e028121corr1. doi:10.1136/bmjopen-2018-028121corr1