PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Long-term outcomes of urinary tract infection (UTI) in Childhood (LUCI): protocol for an electronic record-linked cohort study
AUTHORS	Lugg-Widger, Fiona; Angel, Lianna; Cannings-John, Rebecca; Jones, Hywel; Lau, Mandy; Butler, C; Francis, Nick A.; Hay, Alastair; Heginbothom, Margaret; Hood, Kerenza; Paranjothy, Shantini; Vandervoort, Judith; Hughes, Kathryn

VERSION 1 - REVIEW

REVIEWER	Denise Swei Lo Hospital Universitário da Universidade de São Paulo, Brazil
REVIEW RETURNED	17-Jul-2018

GENERAL COMMENTS	Dear authors, I appreciated this important protocol to clarify the long-term outcomes of urinary tract infection. I have a suggestion for improvement in the method section. I suppose that a non-invasive method such as urine collection pads or clean catch urine sample is routinely used as the method for urine collection in accordance with the recommendations in the NICE guideline. However, the Policy of the American Academy of Pediatrics states that: “If a clinician decides that a febrile infant with no apparent source for the fever requires antimicrobial therapy to be administered because of ill appearance or another pressing reason, the clinician should ensure that a urine specimen is obtained for both culture and urinalysis before an antimicrobial is administered; the specimen needs to be obtained through catheterization or suprapubic aspiration (SPA), because the diagnosis of UTI cannot be established reliably through culture of urine collected in a bag (evidence quality: A; strong recommendation).” Therefore, I consider it essential to clarify to the readers what the method for urine sampling is. Reference: AAP SUBCOMMITTEE ON URINARY TRACT INFECTION. Reaffirmation of AAP Clinical Practice Guideline: The Diagnosis and Management of the Initial Urinary Tract Infection in Febrile Infants and Young Children 2–24 Months of Age. Pediatrics. 2016;138(6):e20163026

REVIEWER	Koen Pouwels Public Health England, UK
REVIEW RETURNED	19-Jul-2018

GENERAL COMMENTS	1. Page 4, line 14-106. 'A urine sample is necessary to confirm the presence of UTI in childhood as the presenting symptoms are non-specific and similar to those found in many common childhood illnesses. 106 [6]' Could you please provide some examples of other common childhood illnesses that result in the same symptoms? People without a clinical background, like me, may not know what other illnesses would give similar symptoms (and what the non-specific symptoms are). 2. Page 14, line 296 – 300. As it is currently written, dividing patients into the three groups suggests that the whole time of patients with a mcUTI is assigned to group 1, including the time before acquiring the mcUTI. The next sentence suggests that is not done. I think it would be better to make clear that the groups are based on patient-time (if it indeed is). In addition, it is not entirely clear what is meant with 'will be taken at the point of outcome'. Please clarify what is meant with this. 3. Page 16. 'For the analysis of Research Question 1, using only children whose whole first five years of life were covered by Datastore'. Does this mean that patients with a shorter follow-up are excluded? And doesn't this potentially introduce some bias? If I understand correctly, e.g. patients who die as a consequence of UTI before <5 are excluded, which could introduce some bias. 4. Along the same lines, how is the evaluation for short-term, medium-term and long-term outcomes operationalised? Currently, this is not entirely clear and it would be helpful to clarify whether you're not modelling it in a way that you're building in a selection bias, akin the built-in selection bias of hazard ratios [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653612/]. In addition, it is unclear to me how it is taken into account that after 1 year of follow-up patients are not at risk any more of short-term outcome (<12 months). And if a patient only get's a UTI at 4.8y how is the relatively shorter follow-up while being exposed being taken into account? 5. I do not understand why the authors decided to use multinomial models instead of just providing adjusted survival curves showing the cumulative probabilities over time. That would also be much more intuitive to interpret and you don't have to ignore potential competing risks such as death, or restrict to a population with a minimum amount of follow-up. I'm not yet convinced the proposed statistical analysis using multinomial models is the correct approach. 6. Page 18, line 373-375. 'We will adjust for direct covariates of renal scarring and explore the impact of indirect effects such as mcUTI using causal directed acyclic graph (DAG).' Which covariates are the authors referring to and on what basis are they going to be included in the model? In addition, could it be clarified how indirect effects will be assessed? Just creating a DAG will not provide an answer about indirect effects.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

1. I appreciated this important protocol to clarify the long-term outcomes of urinary tract infection. I have a suggestion for improvement in the method section. I suppose that a non-invasive method such as urine collection pads or clean catch urine sample is routinely used as the method for urine collection in accordance with the recommendations in the NICE guideline. However, the Policy of the American Academy of Pediatrics states that: "If a clinician decides that a febrile infant with no apparent source for the fever requires antimicrobial therapy to be administered because of ill appearance or another pressing reason, the clinician should ensure that a urine specimen is obtained for both culture and urinalysis before an antimicrobial is administered; the specimen needs to be obtained through catheterization or suprapubic aspiration (SPA), because the diagnosis of UTI cannot be established reliably through culture of urine collected in a bag (evidence quality: A; strong recommendation)." Therefore, I consider it essential to clarify to the readers what the method for urine sampling is.

Reference: AAP SUBCOMMITTEE ON URINARY TRACT INFECTION. Reaffirmation of AAP Clinical Practice Guideline: The Diagnosis and Management of the Initial Urinary Tract Infection in Febrile Infants and Young Children 2–24 Months of Age. Pediatrics. 2016;138(6):e20163026

RESPONSE: This has been added in [line 291-298].

Reviewer: 2

1. Page 4, line 14-106. 'A urine sample is necessary to confirm the presence of UTI in childhood as the presenting symptoms are non-specific and similar to those found in many common childhood illnesses. 106 [6]' Could you please provide some examples of other common childhood illnesses that result in the same symptoms? People without a clinical background, like me, may not know what other illnesses would give similar symptoms (and what the non-specific symptoms are).

RESPONSE: This has been addressed [line 106-107].

2. Page 14, line 296 – 300. As it is currently written, dividing patients into the three groups suggests that the whole time of patients with a mcUTI is assigned to group 1, including the time before acquiring the mcUTI. The next sentence suggests that is not done. I think it would be better to make clear that the groups are based on patient-time (if it indeed is). In addition, it is not entirely clear what is meant with 'will be taken at the point of outcome'. Please clarify what is meant with this.

RESPONSE: The exposure period is set at 5 years. What is meant by "taken at point of outcome" is that exposure is a discrete time-varying covariate (group 2/3 (negative or no sample) until first exposure, group 1 (mcUTI) thereafter). For each analysis, the exposure will depend on the timing of the outcome (if applicable). Exposure groups are therefore based on patient-time and we have changed and clarified how the exposure is defined [line 303-313].

3. Page 16. 'For the analysis of Research Question 1, using only children whose whole first five years of life were covered by Datastore'. Does this mean that patients with a shorter follow-up are excluded? And doesn't this potentially introduce some bias? If I understand correctly, e.g. patients who die as a consequence of UTI before <5 are excluded, which could introduce some bias.

Microbiology data (Datastore) is only available to us between 2005 and 2014 and therefore we excluded children whose first 5 years of life were outside this date range or covered partial data. Deaths are still included in this cohort. This has been clarified in lines 342-344.

4. Along the same lines, how is the evaluation for short-term, medium-term and long-term outcomes operationalised? Currently, this is not entirely clear and it would be helpful to clarify whether you're not modelling it in a way that you're building in a selection bias, akin the built-in selection bias of hazard ratios [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3653612/]. In addition, it is unclear to me how it is taken into account that after 1 year of follow-up patients are not at risk any more of short-term outcome (<12 months). And if a patient only get's a UTI at 4.8y how is the relatively shorter follow-up while being exposed being taken into account?

For clarification, for children in Dataset 2 we will follow each child up for 30 days, 1 year, 1-5 years, >5 years post index consultation so all children will have the same period of follow-up. Children included in Dataset 1 will be followed up irrespective of exposure at 5 and 7 years [lines 340-345].

5. I do not understand why the authors decided to use multinomial models instead of just providing adjusted survival curves showing the cumulative probabilities over time. That would also be much more intuitive to interpret and you don't have to ignore potential competing risks such as death, or restrict to a population with a minimum amount of follow-up. I'm not yet convinced the proposed statistical analysis using multinomial models is the correct approach.

RESPONSE: We have decided to review our statistical analysis plan in light of the reviewer's comments. For the primary outcome of renal scarring, it is not the timing between the confirmation of UTI and the renal scarring that is important here but it is presence of renal scarring. Confirmation of an UTI does not impact on identifying renal scarring sooner (in fact renal scarring cannot be diagnosed at a scan very soon after a UTI). Thus we will retain the multinomial analysis and examine outcomes up to the age of 7 years. In addition we shall as suggested, still be running time to event models that will incorporate all children without restriction.

6. Page 18, line 373-375. 'We will adjust for direct covariates of renal scarring and explore the impact of indirect effects such as mcUTI using causal directed acyclic graph (DAG).' Which covariates are the authors referring to and on what basis are they going to be included in the model? In addition, could it be clarified how indirect effects will be assessed? Just creating a DAG will not provide an answer about indirect effects.

RESPONSE: We are mainly adjusting for moderators listed in table 2 (pre exposure variables such as gender, comorbidities, and congenital malformations) that are known or possibly known to be associated with a microbiologically confirmed UTI and outcome. We have clarified the method of assessing indirect effects [line 389-399].

VERSION 2 – REVIEW

REVIEWER	Denise Swei Lo MD, PhD Department of Pediatrics Hospital Universitario da Universidade de Sao Paulo Brazil
REVIEW RETURNED	25-Nov-2018

GENERAL COMMENTS	This study protocol is a valuable contribution to clarify the association of childhood UTI with long-term chronic conditions. It is well written and with a large sample of patients. Unfortunately, it is not clear how urine was sampled. A large number of readers follow the American Academy of Pediatrics recommendations. The diagnosis of UTI may be criticized at the end of the study since the American Academy of Pediatrics recommends that the diagnosis

	of UTI cannot be established reliable through a culture of urine collected in a bag in young children 2-24 months of age. This issue should be discussed as a limitation after completing the study.

REVIEWER	Koen Pouwels Public Health England, United Kingdom
REVIEW RETURNED	12-Nov-2018

GENERAL COMMENTS	The authors have provided helpful clarifications and addressed the comments raised by the reviewers. However, there is one issue that is not fully addressed and one issue that have arised after one of the clarifications. 1. First the issue that is not fully addressed: In response to the comment about the use of multinomial models the authors have reassessed their statistical analysis plan and decided to keep the original analysis because 'it is not the timing between the confirmation of UTI and the renal scarring that is important here but it is presence of renal scarring'. If the follow-up time is the same for everyone and the timing of the event is relevant a time-to-event analysis may indeed not be necessary. However, I don't see how this would address the issue of competing risks (which may cause differences in the time-at-risk). In the hypothetical situation that there would be many deaths, mcUTI could theoretically even be associated with an apparent protective effect against renal scarring if the competing risk (death) is not taken into account as with the current multinomial model. Similarly, it seems that patients who experience renal scarring <5 years are no longer at risk of the outcome 'renal scarring recorded 5-7 years', in other words it is a competing risk for the long-term outcome. By performing a multinomial model one implicitly assumes however that everyone has the same time-at-risk, while the total 'time-at-risk' may actually differ between the two groups for the '5-7 years' outcome (if there is a difference for the <5 years outcome). I would still recommend to use a survival model which takes into account competing risks and differences in time-at-risk. 2. I would like to thank the authors for providing more details about the mediation analysis. The approach using logistic regression models that is described by the authors is only valid for rare outcomes, due to non-collapsibility of the odds ratio. With a common outcome, adding a variable to the logistic regression will even change the odds ratio if it is not a mediator or confounder. From the introduction it appears that renal scarring is actually quite common: 'A systematic review in 2010 found that the prevalence of renal scarring following first childhood UTI was 15%.' Therefore I would recommend to use a log-linear model instead, as this approach avoids the non-collapsibility issue. In addition, I find the following sentence a difficult to read (a bracket is missing): ‘First we will identify the independent variables associated with renal scarring (using an univariable logistic

	regression and identify the mediation variables (mcUTI or not) that are associated with the significant independent variables.' (see for example https://www.annualreviews.org/www.annualreviews.org/ doi/full/10.1146/annurev-publhealth-032315-021402 for a non-technical discussion of mediation analysis and necessary assumptions and issues with common outcomes).

VERSION 2 – AUTHOR RESPONSE

Reviewer: 1

1. This study protocol is a valuable contribution to clarify the association of childhood UTI with long-term chronic conditions. It is well written and with a large sample of patients. Unfortunately, it is not clear how urine was sampled. A large number of readers follow the American Academy of Pediatrics recommendations. The diagnosis of UTI may be criticized at the end of the study since the American Academy of Pediatrics recommends that the diagnosis of UTI cannot be established reliable through a culture of urine collected in a bag in young children 2-24 months of age. This issue should be discussed as a limitation after completing the study.

RESPONSE: We thank reviewer 1 for their comments. We agree that this should be discussed in the main results paper as a limitation.

Reviewer: 2

1. In response to the comment about the use of multinomial models the authors have reassessed their statistical analysis plan and decided to keep the original analysis because 'it is not the timing between the confirmation of UTI and the renal scarring that is important here but it is presence of renal scarring'.

If the follow-up time is the same for everyone and the timing of the event is relevant a time-to-event analysis may indeed not be necessary. However, I don't see how this would address the issue of competing risks (which may cause differences in the time-at-risk). In the hypothetical situation that there would be many deaths, mcUTI could theoretically even be associated with an apparent protective effect against renal scarring if the competing risk (death) is not taken into account as with the current multinomial model.

Similarly, it seems that patients who experience renal scarring <5 years are no longer at risk of the outcome 'renal scarring recorded 5-7 years', in other words it is a competing risk for the long-term outcome. By performing a multinomial model one implicitly assumes however that everyone has the same time-at-risk, while the total 'time-at-risk' may actually differ between the two groups for the '5-7 years' outcome (if there is a difference for the <5 years outcome).

I would still recommend to use a survival model which takes into account competing risks and differences in time-at-risk.

RESPONSE: Thank you for taking the time to comment on this issue again. In the last draft we added in that we would also examine renal scarring using a Cox regression model. We have moved the placement of this detail and added that this analysis would take into account competing risks and differences in time at risk (lines 383-387).

2. I would like to thank the authors for providing more details about the mediation analysis. The approach using logistic regression models that is described by the authors is only valid for rare

outcomes, due to non-collapsibility of the odds ratio. With a common outcome, adding a variable to the logistic regression will even change the odds ratio if it is not a mediator or confounder. From the introduction it appears that renal scarring is actually quite common: 'A systematic review in 2010 found that the prevalence of renal scarring following first childhood UTI was 15%.' Therefore I would recommend to use a log-linear model instead, as this approach avoids the non-collapsibility issue.

(see for example https://emea01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.annualreviews.org%2www.annualreviews.org%2Fdoi%2Ffull%2F10.1146%2Fannurev-publhealth-032315-021402&data=01%7C01%7CLuggFV%40cardiff.ac.uk%7Cb3de7e3872e44c1575c708d665ac0951%7Cbdb74b3095684856bdbf06759778fcbc%7C1&sdata=qZH42imSHqM14QLLlMUmFUBdAxINzIictRiSQfQNeXk%3D&reserved=0 for a non-technical discussion of mediation analysis and necessary assumptions and issues with common outcomes).

RESPONSE: Thank you for providing this very useful reference. We suspect that the prevalence of renal scarring would be lower than the 15% found in the systematic review as our population is less selected than in the studies include in this review. However, as we do not know what the rate of renal scarring will be in our data, we have added using the log-linear model approach in the case where renal scarring is common (lines 394-395).

3. In addition, I find the following sentence a difficult to read (a bracket is missing): ‘First we will identify the independent variables associated with renal scarring (using an univariable logistic regression and identify the mediation variables (mcUTI or not) that are associated with the significant independent variables.’

RESPONSE: We have added the missing bracket into this sentence (line 395).

VERSION 2 – REVIEW

REVIEWER	Koen Pouwels Public Health England, United Kingdom
REVIEW RETURNED	15-Jan-2019

GENERAL COMMENTS	The manuscript has improved and I have no further comments.