

## 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

<b>TITLE (PROVISIONAL)</b>	Healthcare use for children with complex needs: using routine health data linked to a multiethnic, ongoing birth cohort
<b>AUTHORS</b>	Bishop, Chrissy; Small, Neil; Parslow, Roger; Kelly, Brian

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Rosario Caruso, PhD, RN Head of Health Professions Research and Development Unit IRCCS Policlinico San Donato 20097 San Donato Milanese (Mi) Italy
<b>REVIEW RETURNED</b>	24-Jul-2017

<b>GENERAL COMMENTS</b>	This study explore the healthcare services demand of children with CA, introducing new knowledge at this regard. The authors are aware about the study limits and the methodology is clear and correct.
-------------------------	---

<b>REVIEWER</b>	Karrie Downing National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention
<b>REVIEW RETURNED</b>	14-Aug-2017

<b>GENERAL COMMENTS</b>	<p>I read with interest the article “Healthcare use for children with complex needs: using routine health data linked to an ongoing birth cohort.” Linking data from the large prospective Born in Bradford cohort to medical records, the authors compared primary care and hospital service utilization among children with and without congenital anomalies. The research question is important, and the data allows for a uniquely comprehensive and informative analysis on the healthcare utilization of children with congenital anomalies. While the manuscript has the potential to offer valuable information on this currently under-studied issue, especially in relation to primary care utilization, the analysis, writing, and presentation of the results require too great a revision.</p> <p>The text is often difficult to follow because sentences get wordy (many are run-ons) and contain grammatical errors. The authors should review the paper to simplify text and clarify language. Furthermore, the results in the text do not accurately describe the findings portrayed in the tables, and, at times, the comparisons are unclear and inappropriate.</p>
-------------------------	--

In Table 2, it is unnecessary to show the coefficient values for the confounders, which are largely driven by children without CA—this table should be focused on the coefficient value for CA only. If the authors are interested in examining factors associated with healthcare use among children with CA, the model should be limited to children with CA; other associations among the general population distract from the main point of the paper. In addition, Figure 1 is confusing and the numbers do not add up as the reader might expect. Overall, the authors should update their results to focus on healthcare utilization among children with congenital anomalies, ensure results are accurate based on the models and samples used, and improve the overall clarity of this manuscript.

Examples of suggested revisions can be found below:

#### Abstract

A suggested edit for the aim sentence in Objectives (page 1 line 30): “Our aim was to examine healthcare use of children with CA and consequential complex needs compared to those without CA.” Assessing their need and demand seems redundant.

The design/method section is lacking description. Include the study location, the names of data sources (Born in Bradford, SystemOne, etc.) and a brief description of the performed statistical analyses in the abstract.

In the design (page 1, line 37), it should say the authors “conducted” a sub analysis on.... unless this sub-analysis has been performed before. If these 400 children are part of the original cohort, make that clear.

Are 5 years olds included in the analysis? The 0<5 notation through the abstract/manuscript is confusing.

In the conclusions (page 2 line 20), the statement regarding hospital service use among children with CA “especially in economically deprived and ethnically diverse neighborhoods” can’t be made because your analysis is describing individuals, not neighborhoods, and you didn’t limit your table 2 analysis exclusively to children with CA.

In the first strengths and limitations bullet, remove the comma  
In the last bullet, “we” performed...

#### Manuscript

Intro, page 2

On line 46, 93% survive to what? Adulthood?

On line 53, “it has been suggested”- this sentence makes an important point but is difficult to follow and needs to be rephrased.

On line 56, clarify that this sentence refers to primary care in England (since primary care may not be ideal for this role in other countries like the US) and that primary care may be ideal because of the NHS health system where GPs can act as gatekeepers.

Intro, Page 3

On lines 6-10, the sentence beginning with "These complexities" is a run-on. Please simplify.

On line 12, what does "interchangeable service requirements" mean?

On line 18, the terms in this sentence are not easily understood. What do "aggravated," "patient complexity," "levels of deprivation," and "primary care practice provision" mean in this context?

On line 28, the authors should state they selected data on children.... linked to primary care records...

On line 36, remove the word "addressing"

Methods, page 3

On line 43, is BiB still recruiting the 12450 pregnant women or just following up?

Figure 1 is confusing and the numbers don't add up. Include all exclusionary criteria, including incomplete questionnaires, children without one hospital episode, etc. If the authors choose to keep Figure 1, it should be cited somewhere in the manuscript.

Methods, page 4

Why was having a non-birth hospital event part of the eligibility criteria?

On line 20, edit to "including for those"

On line 38, was it convenience sample or were the selected at random from the cohort?

Statistical analysis, page 4

On line 54, the authors should use a non-poisson negative binomial regression model to account for over-dispersion.

Statistical analysis, page 5

Does Model 3 control for the measures of ill health? Otherwise how are they "added"?

On line 17, change "measure if" to "measure of"

Confounders, page 5

Define the abbreviation GCSE. Is it possible have exactly 5 GCSE equivalents? If so, what category of education would they belong to?

Define economically deprived.

How far out in relation degree does other blood go?

Results, page 6

Tables

Change the name of the first row in Table 1 to Total.

It is unclear why table 1 is stratified by race/ethnicity, since this was not an objective of the study and other tables/figures are not stratified by race/ethnicity.

If this manuscript is trying to describe children with congenital anomalies, why isn't table 2 limited to only those with congenital anomalies?

Ensure that you are using consistent terminology to describe all characteristics and confounders throughout the text and tables (e.g. primary care consultations and GP appointments). Additionally, all acronyms should be defined in a footnote.

Table 3 is mentioned in the text before Table 2.

All tables should stand alone (i.e. with descriptive titles including population, age, dataset, years; acronyms spelled out (Table 3: spell out A&E); and footnotes, when appropriate).

#### Text

The text describing the results needs to be re-visited to ensure the language is accurately describing what is reported in the tables.

Some examples follow:

On lines 14-15, these reported percentages should reflect what is shown in the table (so either the table needs to be rearranged to show 53%, 35% and 13% respectively or the text should state 7.2% 5.4% and 5.2%).

On lines 17-19, this statement is not accurate. Based on the way this is written the percentage would be  $(183/706) * 100$  or 26%. Forty-nine percent represents the number of Pakistani children with CA from first cousin unions.

#### Discussion

The authors state "When accounting for confounders, children with CA were predicted to require an increase in primary care consultations ( $\beta 0.24$ , 95% CI 1.18-0.30)..."; the authors should check the beta coefficient and confidence interval here and in Table 2, since they do not make sense.

Throughout the manuscript, it is unclear if the results on Pakistani children are among children with CA or all Pakistani children. The author should clarify if assessing interactions between CA status and healthcare utilization by race/ethnicity is an objective of the paper.

The description of the overall findings are incorrect based on the models the authors ran. For example, the authors state "Although we find children from Pakistani heritage are predicted to use more primary care consultations (0.40, 95% CI 0.36-0.44) than children without CA..." In this sentence, the authors are comparing Pakistani children with children without CA, which makes no sense.

In the following sentence, the authors state "We also find children with CA from economically deprived neighborhoods have an increased risk of using hospital services, but not primary care consultations."

	<p>However, the model the authors ran includes all children, not just those with CA and the findings on “economic depravity” is not specific for children with CA.</p> <p>The authors should better describe the implications of their results.</p>
--	---

<b>REVIEWER</b>	Elaine Meehan Murdoch Children's Research Institute, Melbourne, Australia
<b>REVIEW RETURNED</b>	25-Sep-2017

<b>GENERAL COMMENTS</b>	<p>Thank you for the opportunity to review this manuscript by Bishop and colleagues regarding healthcare use for children with congenital anomalies. Overall, this is a well-designed research study, and of strong methodological quality. The authors should be commended for their use of rich datasets and data linkage techniques to explore this important topic. In recent years, interest in the patterns of medical service use in children with complex conditions has grown, and the use of data linkage to capture data from multiple sources has become popular. This manuscript contributes to that growing body of literature, and provides information on health service use among children with congenital anomalies in particular. I do have some concerns, however, in a number of areas. I have described these in the comments below, in the order in which they come up in the manuscript. I think, given its innovative design and wealth of data, this study would be a good contribution to the research literature; however, I recommend a number of changes/additions to the manuscript.</p> <p><b>Abstract</b></p> <ul style="list-style-type: none"> <li>• Objectives - The aim specified in the abstract (“to examine healthcare use of children with CA and consequential complex needs, assessing their need and demand”) does not completely match with the aims/objectives of the study that are described in the last paragraph of the introduction. The study investigates actually health service utilisation in this group of children, which is very different to assessing their “need and demand”. I recommended this aim be re-worded in the abstract to reflect what is in the body of the paper.</li> <li>• Setting – As per BMJ Open’s instructions for authors, the following information should come under ‘Setting’: “level of care e.g. primary, secondary; number of participating centres. Generalise; don’t use the name of a specific centre, but give geographical location if important”. At present, there is none of this information included - geographical location, in particular, should be included. The information that is currently included in this section (about the data collected) would be more suited under one of the other Abstract sub-headings.</li> <li>• Primary and secondary outcome measures – minor point: 0&lt;5 years should read 0-5 years? (throughout the manuscript)</li> </ul> <p><b>Introduction</b></p> <ul style="list-style-type: none"> <li>• The introduction is lacking a description, or at the very least, a definition of congenital anomalies. I feel this should be placed before the current introductory statement regarding survival. Including a short description of the main differences between minor and major CA, as well as a prevalence estimate, would be beneficial.</li> <li>• “Around 93% of children with a congenital anomaly (CA) survive...” – Survive what? Birth?</li> </ul>
-------------------------	---

Beyond the neonatal period? It's not clear (I did check reference 1 but I wasn't able to ascertain exactly what you are referring to in that document either). Please clarify this by specifying the period of survival you are referring to.

- "with the bulk of evidence being American studies..." – I suggest replacing "being" with "coming from".
- "We compare longitudinal healthcare use for children ages 0<5....." – I suggest replacing "0<5" with "0-5" and specifying "years" after 0-5. This terminology (0<5) is used throughout the manuscript; I recommend changing to "0 – 5 years" throughout.

#### Methods

- Under case ascertainment and coding methods – "...hospital episode statistics, which are split into elective, accident and emergency and other emergency admissions...." – do the hospital episode statistics pertain to hospital admissions only? What about A&E presentations (that do not lead to an admission to hospital), and outpatient appointments? This is not very clear I recommend including a better description of what the "hospital episode statistics" capture. (I note that this is touched on in paragraph 2 of the discussion, but suggest it would be more appropriate to have a description of this upfront in the methods section).
- The terms "diagnostic CTV3 Read medical codes" and ICD-10 codes should be defined/described.
- I recommend defining major and minor CA (also see first comment above, under Introduction), and providing justification for excluding minor CA from the analysis.
- Statistical analysis seems appropriate and is well described (though I am not a statistical expert. The manuscript may benefit from having a statistical review).

#### Results

- Sentence 3 ("Children of Pakistani heritage were born...") is a bit misleading. It suggests that 53% of children of Pakistani heritage were born with a CA (when it is really only 7%). I recommend re-wording this sentence to reflect what you actually mean which is: of all children with a CA, 53% were of Pakistani heritage, 35% were of white British heritage, and 13% were of an "other" ethnic group.
- Sentence 5 ("Children of Pakistani heritage with CA had on average 1.61 more primary care consultations, and 0.39 more hospital admissions per year than children without CA") – I'm not sure how meaningful it is to present these data in this way. Was the difference in the number of primary consultations and hospital admissions between the two groups actually statistically significant? This information would assist the reader in understanding how big a difference this actually is.
- The main reasons for hospital admissions are discussed under broad categories (respiratory conditions etc.). How were these categories defined? And what data were used to determine the main reasons for admissions? I am assuming that discharge diagnoses were used to ascertain the main reasons for admissions and that the ICD-10 groupings were used to categorise them; however, these are not explained/well defined.
- How is "economic deprivation" defined? This is probably described well in the original study; however, because it is such an important focus of this study, it should be defined in the methods section of this paper.
- Is it possible to present the data based on age, and to look at how healthcare utilisation changed over the first five years of life?

	<p>For example, did children have more hospital encounters early in life (soon after diagnosis), and did these then drop off as they got older? Did they use primary care services more as they got older? To me, these are important questions.</p> <p>Discussion</p> <ul style="list-style-type: none"> <li>• Summarizing the main findings of the study at the start of the discussion is useful; however, the repetition of <math>\beta</math> values and confidence intervals in the discussion section of a paper is generally not advisable (these should be included in the results section only, and the discussion section used to actually interpret the results and put them in context).</li> <li>• Good point about care co-ordination and the potential for this to impact on the patterns of health service utilisation in this group.</li> <li>• A discussion of the implications of the study findings is lacking. Do any of the findings have the potential to inform/change current practice? What should the next steps be?</li> </ul> <p>Minor points</p> <ul style="list-style-type: none"> <li>• There are some grammatical errors throughout the manuscript. I recommend the author read through thoroughly and amend where appropriate.</li> <li>• I think rounding the percentages in the tables to one decimal place would suffice and make the table easier to read. In addition, some of the percentage values in the tables run over two lines – the table should be amended to ensure they are on one line only.</li> <li>• I suggest splitting paragraph 1 of the results section into 2 or 3 separate paragraphs.</li> <li>• Table 2: Under Economic Deprivation, should “Not Economically deprived” be above “Economically deprived”?</li> </ul>
--	---

### VERSION 1 – AUTHOR RESPONSE

#### Reviewer 1

I read with interest the article “Healthcare use for children with complex needs: using routine health data linked to an on-going birth cohort.” Linking data from the large prospective Born in Bradford cohort to medical records, the authors compared primary care and hospital service utilization among children with and without congenital anomalies. The research question is important, and the data allows for a uniquely comprehensive and informative analysis on the healthcare utilization of children with congenital anomalies. While the manuscript has the potential to offer valuable information on this currently under-studied issue, especially in relation to primary care utilization, the analysis, writing, and presentation of the results require too great a revision.

The text is often difficult to follow because sentences get wordy (many are run-ons) and contain grammatical errors. The authors should review the paper to simplify text and clarify language. Furthermore, the results in the text do not accurately describe the findings portrayed in the tables, and, at times, the comparisons are unclear and inappropriate. In Table 2, it is unnecessary to show the coefficient values for the confounders, which are largely driven by children without CA—this table should be focused on the coefficient value for CA only. If the authors are interested in examining factors associated with healthcare use among children with CA, the model should be limited to children with CA; other associations among the general population distract from the main point of the paper.

In addition, Figure 1 is confusing and the numbers do not add up as the reader might expect. Overall, the authors should update their results to focus on healthcare utilization among children with congenital anomalies, ensure results are accurate based on the models and samples used, and improve the overall clarity of this manuscript.

#### Abstract

A suggested edit for the aim sentence in Objectives (page 1 line 30): “Our aim was to examine healthcare use of children with CA and consequential complex needs compared to those without CA.” Assessing their need and demand seems redundant.

Thank you removed.

The design/method section is lacking description. Include the study location, the names of data sources (Born in Bradford, SystemOne, etc.) and a brief description of the performed statistical analyses in the abstract.

Thanks added the following sentence:

Observational secondary analysis of data from the Born in Bradford study, a large prospective birth cohort, linked to primary care data held in SystemOne and hospital episode statistics. Negative binomial regression with 95% confidence intervals was performed to predict healthcare use.

In the design (page 1, line 37), it should say the authors “conducted” a sub analysis on.... unless this sub-analysis has been performed before. If these 400 children are part of the original cohort, make that clear.

Thank you added words to that affect.

Are 5 years olds included in the analysis? The 0<5 notation through the abstract/manuscript is confusing.

Thank you for pointing this out. I have changed to 0–<5 years as children were included from birth up until their 5th birthday and not beyond. If I put 0–5 this would infer I had included children that were older than 5 but not yet 6, which I did not do. Specified this the first time I mention 0–<5. I hope this clarifies the point.

In the conclusions (page 2 line 20), the statement regarding hospital service use among children with CA “especially in economically deprived and ethnically diverse neighborhoods” can’t be made because your analysis is describing individuals, not neighborhoods, and you didn’t limit your table 2 analysis exclusively to children with CA.

Thank you. Added the following sentence. “hospital services are most in demand for children with CA, but are also in high demand for children who were economically deprived whether they had a CA or not.”

In the first strengths and limitations bullet, remove the comma  
In the last bullet, “we” performed...

Thank you made the suggested corrections.

Manuscript



Intro, page 2

On line 46, 93% survive to what? Adulthood?

Yes to adulthood added this to the manuscript.

On line 53, "it has been suggested"- this sentence makes an important point but is difficult to follow and needs to be rephrased.

Thank you added the following sentence:

This may be due in part to a lack of longitudinal data related to the nature of multi-disciplinary involvement required by children with complex needs

Thank you for pointing this out. Added "in the UK"

On line 56, clarify that this sentence refers to primary care in England (since primary care may not be ideal for this role in other countries like the US) and that primary care may be ideal because of the NHS health system where GPs can act as gatekeepers.

Thank you added

Intro, Page 3

On lines 6-10, the sentence beginning with "These complexities" is a run-on. Please simplify.

Thank you, changed to "multi-disciplinary care"

On line 12, what does "interchangeable service requirements" mean?

Thank you, changed to "Multidisciplinary care needs"

On line 18, the terms in this sentence are not easily understood. What do "aggravated," "patient complexity," "levels of deprivation," and "primary care practice provision" mean in this context?

Agreed. Changed to:

"The need and demand for primary care services in particular are intensified by patient complexity, levels of deprivation, and primary care practice provision".

On line 28, the authors should state they selected data on children.... linked to primary care records...

Added thank you

On line 36, remove the word "addressing"

Removed thank you

Methods, page 3

On line 43, is BiB still recruiting the 12450 pregnant women or just following up?

Thank you changed to recruited.

Figure 1 is confusing and the numbers don't add up. Include all exclusionary criteria, including incomplete questionnaires, children without one hospital episode, etc. If the authors choose to keep Figure 1, it should be cited somewhere in the manuscript.

Agreed. I have removed it and added extra explanation of the figures and number of CA retrieved to the text on page 4.

Methods, page 4

Why was having a non-birth hospital event part of the eligibility criteria?

We already had all the non-birth events as every child included was registered to the BiB cohort at birth. So we had hospital data for those children who also had either an elective, emergency or other emergency admission after birth.

On line 20, edit to "including for those"

Included thank you.

On line 38, was it convenience sample or were they selected at random from the cohort?

They were selected at random from the cohort. Thanks for pointing this out. Removed convenience sample.

Statistical analysis, page 4

On line 54, the authors should use a non-poisson negative binomial regression model to account for over-dispersion.

Thank you for pointing out this typo. The model used was negative binomial regression. Or nbreg command in stata. Removed the term poisson.

Statistical analysis, page 5

Does Model 3 control for the measures of ill health? Otherwise how are they "added"?

Yes thank you for this grammatical suggestion removed the word added.

On line 17, change "measure if" to "measure of"

Can't find this typo.

Confounders, page 5

Define the abbreviation GCSE. Is it possible to have exactly 5 GCSE equivalents? If so, what category of education would they belong to?

A GCSE is a standard level of education in the UK. I have added a reference to their description. I hope this clarifies the point for you.

Define economically deprived.

Thank you for pointing this out. Added the following sentence and referenced to define: Economic deprivation was measured using means-tested benefit status.

In the UK, being in receipt of means-tested benefits is recognised as measure of income poverty, as these benefits are frequently the only source of income and are paid at rates that put individuals below standard poverty lines.

How far out in relation degree does other blood go?

It was any relation. Added this explanation to the text.

Results, page 6

Tables

Change the name of the first row in Table 1 to Total. It is unclear why table 1 is stratified by race/ethnicity, since this was not an objective of the study and other tables/figures are not stratified by race/ethnicity.

I have changed the title to include "multi-ethnic ongoing birth cohort." It is necessary to stratify by ethnicity due to the nature of the population being studied.

If this manuscript is trying to describe children with congenital anomalies, why isn't table 2 limited to only those with congenital anomalies?

Because we are also trying to show the affects of other factors on health care use, not just CA. I have re-written the aims at the end of the introduction. I also did limit an additional analysis to children with CA, but this data was not included in the paper. I have made reference to this analyses on page 5 in statistical analysis, but it can be offered as an appendix if required.

Ensure that you are using consistent terminology to describe all characteristics and confounders throughout the text and tables (e.g. primary care consultations and GP appointments). Additionally, all acronyms should be defined in a footnote.

Thank you changed any occurrences of GP appointments to PC.

Table 3 is mentioned in the text before Table 2.

Thanks for pointing this out. Changed the organization of paragraphs.

All tables should stand alone (i.e. with descriptive titles including population, age, dataset, years; acronyms spelled out (Table 3: spell out A&E); and footnotes, when appropriate).

Thank you made some adjustments to the tables. Changed A & E to accident and emergency thank you.

I hope the corrections suffice.

Text

The text describing the results needs to be re-visited to ensure the language is accurately describing what is reported in the tables. Some examples follow:

On lines 14-15, these reported percentages should reflect what is shown in the table (so either the table needs to be rearranged to show 53%, 35% and 13% respectively or the text should state 7.2% 5.4% and 5.2%).

We appreciate this suggestion, however the table does not show the ethnicity split. We have included the ethnicity split as a result and not referenced the table.

On lines 17-19, this statement is not accurate. Based on the way this is written the percentage would be  $(183/706)*100$  or 26%. Forty-nine percent represents the number of Pakistani children with CA from first cousin unions.

Ok thanks for pointing this out. Changed.  
Discussion

The authors state “When accounting for confounders, children with CA were predicted to require an increase in primary care consultations ( $\beta 0.24$ , 95% CI 1.18-0.30)...”; the authors should check the beta coefficient and confidence interval here and in Table 2, since they do not make sense.

We thank the reviewer for pointing this out. This was a typing error and has now been changed.

Throughout the manuscript, it is unclear if the results on Pakistani children are among children with CA or all Pakistani children. The author should clarify if assessing interactions between CA status and healthcare utilization by race/ethnicity is an objective of the paper.

Agreed this was not well defined. We have rewritten the aims as follows:

Our aims were therefore to explore the longitudinal healthcare use for children from birth up to their fifth birthday (0–<5) with CA and without. We do this by linking demographic and socioeconomic data from a large prospective birth cohort covering a deprived and ethnically diverse population, to children’s primary care records, hospital episodes statistics and referral information. In doing so this study examines the effects of having a CA, and consequential ill health, on primary care use, use of hospital services and referrals to multidisciplinary specialists. We also investigate the influence of demographic and socioeconomic factors on health care use.

The description of the overall findings are incorrect based on the models the authors ran. For example, the authors state “Although we find children from Pakistani heritage are predicted to use more primary care consultations (0.40, 95% CI 0.36-0.44) than children without CA...” In this sentence, the authors are comparing Pakistani children with children without CA, which makes no sense.

Yes I agree see below comment.

In the following sentence, the authors state “We also find children with CA from economically deprived neighborhoods have an increased risk of using hospital services, but not primary care consultations.” However, the model the authors ran includes all children, not just those with CA and the findings on “economic depravity” is not specific for children with CA.

Thank you for pointing this out. We did stratify the analyses for all three outcomes by CA, except these results (and table) were not shown to fit the maximum table requirements for the journal. We overlooked however including a “results not shown” comment, which I have now added in the methods. There is another option of including as an ancillary table in the appendix. Happy to include this if you feel its necessary.

Included the following: We also stratified the analyses for all three outcomes by CA (results not shown).

The authors should better describe the implications of their results.

Agreed and this was also a comment from the second reviewer. Added some sentences to the discussion.

#### Reviewer 2

Thank you for the opportunity to review this manuscript by Bishop and colleagues regarding healthcare use for children with congenital anomalies. Overall, this is a well-designed research study, and of strong methodological quality. The authors should be commended for their use of rich datasets and data linkage techniques to explore this important topic. In recent years, interest in the patterns of medical service use in children with complex conditions has grown, and the use of data linkage to capture data from multiple sources has become popular. This manuscript contributes to that growing body of literature, and provides information on health service use among children with congenital anomalies in particular. I do have some concerns, however, in a number of areas. I have described these in the comments below, in the order in which they come up in the manuscript. I think, given its innovative design and wealth of data, this study would be a good contribution to the research literature; however, I recommend a number of changes/additions to the manuscript.

#### Abstract

- Objectives - The aim specified in the abstract (“to examine healthcare use of children with CA and consequential complex needs, assessing their need and demand”) does not completely match with the aims/objectives of the study that are described in the last paragraph of the introduction. The study investigates actual health service utilisation in this group of children, which is very different to assessing their “need and demand”. I recommended this aim be re-worded in the abstract to reflect what is in the body of the paper.

Agreed and removed the need and demand comment from the abstract.

- Setting – As per BMJ Open’s instructions for authors, the following information should come under ‘Setting’: “level of care e.g. primary, secondary; number of participating centres. Generalise; don’t use the name of a specific centre, but give geographical location if important”. At present, there is none of this information included - geographical location, in particular, should be included. The information that is currently included in this section (about the data collected) would be more suited under one of the other Abstract sub-headings.

Thank you for pointing this out. Added the following sentence and hope this clarifies:

Bradford is a large city in the north of England. This study investigates children’s interactions with primary, secondary and tertiary healthcare services.

- Primary and secondary outcome measures – minor point: 0<5 years should read 0-5 years? (throughout the manuscript)

#### Introduction

I have changed to 0–<5 years as children were included from birth up until their 5th birthday and not beyond. Specified this the first time I mention 0–<5. I hope this clarifies the point.

- The introduction is lacking a description, or at the very least, a definition of congenital anomalies. I feel this should be placed before the current introductory statement regarding survival. Including a short description of the main differences between minor and major CA, as well as a prevalence estimate, would be beneficial.

Added definition and prevalence estimate. I feel the differences between major and minor CA are addressed on page 4 of the methods.

- “Around 93% of children with a congenital anomaly (CA) survive...” – Survive what? Birth? Beyond the neonatal period? It’s not clear (I did check reference 1 but I wasn’t able to ascertain exactly what you are referring to in that document either). Please clarify this by specifying the period of survival you are referring to.

Survive to adulthood. Also clarified for reviewer one. Changed and highlighted.

- “with the bulk of evidence being American studies...” – I suggest replacing “being” with “coming from”.

Thank you. Added

- “We compare longitudinal healthcare use for children ages 0<5.....” – I suggest replacing “0<5” with “0-5” and specifying “years” after 0-5. This terminology (0<5) is used throughout the manuscript; I recommend changing to “0 – 5 years” throughout.

As above I have changed to 0–<5 years as children were included from birth up until their 5th birthday and not beyond. Specified this the first time I mention 0–<5. I hope this clarifies the point.

#### Methods

- Under case ascertainment and coding methods – “...hospital episode statistics, which are split into elective, accident and emergency and other emergency admissions...” – do the hospital episode statistics pertain to hospital admissions only? What about A&E presentations (that do not lead to an admission to hospital), and outpatient appointments? This is not very clear I recommend including a better description of what the “hospital episode statistics” capture. (I note that this is touched on in paragraph 2 of the discussion, but suggest it would be more appropriate to have a description of this upfront in the methods section).

Thank you for asking this question. Hospital episode statistics include A and E presentations also. Added the following sentence:

Hospital events included admissions for elective procedures, other emergencies, and A and E presentations.

- The terms “diagnostic CTV3 Read medical codes” and ICD-10 codes should be defined/described.

Thank you added on page 4.

- I recommend defining major and minor CA (also see first comment above, under Introduction), and providing justification for excluding minor CA from the analysis.

Added a sentence to explain this a little further on page 4, although the justification is that EUROCAT advises this.

- Statistical analysis seems appropriate and is well described (though I am not a statistical expert. The manuscript may benefit from having a statistical review).

#### Results

- Sentence 3 (“Children of Pakistani heritage were born...”) is a bit misleading. It suggests that 53% of children of Pakistani heritage were born with a CA (when it is really only 7%). I recommend re-wording this sentence to reflect what you actually mean which is: of all children with a CA, 53% were of Pakistani heritage, 35% were of white British heritage, and 13% were of an “other” ethnic group.

Thank you reworded and highlighted on page 6.

- Sentence 5 (“Children of Pakistani heritage with CA had on average 1.61 more primary care consultations, and 0.39 more hospital admissions per year than children without CA”) – I’m not sure how meaningful it is to present these data in this way. Was the difference in the number of primary consultations and hospital admissions between the two groups actually statistically significant? This information would assist the reader in understanding how big a difference this actually is.

We felt it was important to report the absolute numbers, as it was meaningful to clinicians. We later show in the statistical analysis that hospital admissions were in more demand than primary care consultations.

- The main reasons for hospital admissions are discussed under broad categories (respiratory conditions etc.). How were these categories defined? And what data were used to determine the main reasons for admissions? I am assuming that discharge diagnoses were used to ascertain the main reasons for admissions and that the ICD-10 groupings were used to categorise them; however, these are not explained/well defined.

Thank you for this useful point. I have added clarification to the footnote of table 3.

- How is “economic deprivation” defined? This is probably described well in the original study; however, because it is such an important focus of this study, it should be defined in the methods section of this paper.

Thank you for pointing this out. I have added the following sentence and referenced:

Economic deprivation was measured using means-tested benefit status. In the UK, being in receipt of means-tested benefits is recognised as measure of income poverty, as these benefits are frequently the only source of income and are paid at rates that put individuals below standard poverty lines.

- Is it possible to present the data based on age, and to look at how healthcare utilisation changed over the first five years of life? For example, did children have more hospital encounters early in life (soon after diagnosis), and did these then drop off as they got older? Did they use primary care services more as they got older? To me, these are important questions.

I agree these are important questions. And you are correct there is more healthcare use for children in Bradford in general in the first two years of life, which was discovered in some preliminary (unpublished) research related to this project. But these patterns do not stay the same for children with CA, because of the severity of their conditions and ongoing healthcare needs. This study aimed to firstly quantify that children with CA in fact do use increases of primary, secondary and tertiary care, and to investigate some of the socio demographic reasons for healthcare use in this diverse, deprived population. Perhaps another study can assess changes over the first few years of life, but like I say preliminary investigations showed us that this was not so clear cut for these children with very complex needs.

Discussion

- Summarizing the main findings of the study at the start of the discussion is useful; however, the repetition of  $\beta$  values and confidence intervals in the discussion section of a paper is generally not advisable (these should be included in the results section only, and the discussion section used to actually interpret the results and put them in context).

Yes thank you for pointing this out I have removed the repetition of  $\beta$  values.

- Good point about care co-ordination and the potential for this to impact on the patterns of health service utilisation in this group.

Thanks!

- A discussion of the implications of the study findings is lacking. Do any of the findings have the potential to inform/change current practice? What should the next steps be?

Yes indeed they do. I agree this point has not been labored appropriately. Added some sentences on clinical implications to the discussion.

Minor points

- There are some grammatical errors throughout the manuscript. I recommend the author read through thoroughly and amend where appropriate.

Thank you hopefully now resolved.

- I think rounding the percentages in the tables to one decimal place would suffice and make the table easier to read. In addition, some of the percentage values in the tables run over two lines – the table should be amended to ensure they are on one line only.

Thank you for pointing this out. Changes have been made.

- I suggest splitting paragraph 1 of the results section into 2 or 3 separate paragraphs.

Thank you done this.

- Table 2: Under Economic Deprivation, should “Not Economically deprived” be above “Economically deprived”?

I have changed the order thank you.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Dr Elaine Meehan, Research Officer Murdoch Childrens Research Institute, Melbourne, Australia
<b>REVIEW RETURNED</b>	29-Nov-2017

<b>GENERAL COMMENTS</b>	<p>I feel that the authors have adequately addressed my concerns outlined in the initial review, and I recommend this manuscript be accepted. I just picked up a few very minor things in the text:</p> <ol style="list-style-type: none"> <li>1. The references for the first two sentences of the Introduction have been inserted incorrectly (they are currently formatted as author, year rather than a superscript number).</li> <li>2. ...‘contaminant increases in costs to the healthcare services’ – I’m not sure if the word ‘contaminant’ is used correctly here?</li> <li>3. Regarding the use of the term 0-&lt;5 throughout (which the authors did address in their response), it may be more appropriate to just use the term 0-4 throughout. This isn’t a huge issue though, and I will leave to the authors’ discretion.</li> </ol>
-------------------------	--



	<p>4. Regarding the last sentence of the statistical analysis 'We also stratified the analyses for all three outcomes by CA (results not shown)' - If you're not providing these data in the main manuscript or as an appendix, or at the very least summarizing the results of this analysis in the text, I would suggest leaving this sentence out.</p> <p>5. Fifth sentence of the results section – 13% is specified twice for 'other' ethnicities – one should be deleted.</p>
--	---

## VERSION 2 – AUTHOR RESPONSE

Dear reviewers,

Thank you again for your helpful comments. I have addressed as follows:

1. The references for the first two sentences of the Introduction have been inserted incorrectly (they are currently formatted as author, year rather than a superscript number) - Thank you, changed and now referenced correctly.
2. ...'contaminant increases in costs to the healthcare services' – I'm not sure if the word 'contaminant' is used correctly here? - Thank you changed the wording.
3. Regarding the use of the term 0-<5 throughout (which the authors did address in their response), it may be more appropriate to just use the term 0-4 throughout. This isn't a huge issue though, and I will leave to the authors' discretion - Thank you. We have left it as 0-<5, with an explanation in the last paragraph of the introduction.
4. Regarding the last sentence of the statistical analysis 'We also stratified the analyses for all three outcomes by CA (results not shown)' - If you're not providing these data in the main manuscript or as an appendix, or at the very least summarizing the results of this analysis in the text, I would suggest leaving this sentence out - Thank you. Point taken and removed.
5. Fifth sentence of the results section – 13% is specified twice for 'other' ethnicities – one should be deleted - Thank you. Removed.

We do hope the manuscript now meets recommended standards and thanks again for all the helpful pointers. We feel the manuscript is a lot better for it.