PEER REVIEW HISTORY

BMJ Paediatrics Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Paediatrics Open. The paper was subsequently accepted for publication at BMJ Paediatrics Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Improving case ascertainment of congenital anomalies: findings from a prospective birth cohort with detailed primary care record linkage.
AUTHORS	Bishop, Chrissy; Small, Neil; Mason, Dan; Corry, Peter; Wright, Prof John; PARSLOW, ROGER; Bittles, Alan; Sheridan, Eamonn

VERSION 1 - REVIEW

REVIEWER	Miller, Nicola National Congenital Anomaly and Rare Disease Registration Service (NCARDRS), Public Health England Competing interests: None
REVIEW RETURNED	19-Aug-2017

GENERAL COMMENTS	This is an interesting and useful contribution in the developing context of Congenital anomaly ascertatinment and registration. Whilst i have recommended mionor revisions some of these are significant to the inferences made in the paper. My comments are as follows: - the introduction should include an acknowledgement that 61% of reported congenital anomalies are detected in the antenatal period - the date of diagnosis for congenital anomalies is missing in 14% of live births. As the paper acknowledges the recording of date of diagnosis in later diagnoses of CA is not always accurate (or present). This may impact on the claim of a 30% increase in ascertainment using this method of data linkage - risk factors for CAs are being collected in the national system NCARDRS - The national register would not tolerate a 17% discrepancy rate in accuracy of a data source without requiring subsequent validation of that data. This is very important to maintain the high quality of the data. The implication of this is that whilst this may serve well as an indicative source of ascertainment the cases would need further investigation to be accceptable as registerable cases. This may not be viable, nationally. It may be worth doing some further analysis of which groups of cases have a) a high level of accuracy with the hospital diagnoses and b) are most likely missing from the otherwise ascertained cases. - page 14 NCARDRS; line 18 ?prognosis - the dramatic increase in skeletal dysplasias is very interesting (and perhaps emphasised because of the demographics in the study cohort). I am very curious about the increased nervous system diagnoses and would have liked more detail around this - were specific conditions identified? A detailed breakdown would be

	interesting. - the cardiac conditions identified later tend to be less serious conditions (small VSD, ASD) these might be more properly (and more accurrately) be identified using tertiary cardiac database data which NCARDRS is trying to procure. There is no discussion in the paper about whether there are better ways of improving the data collection for paediatric cases, using surgical data or other disease specific data for example? it might be useful to consider whether there are specific conditions being picked up in this data linkage that would not be picked up by otherwise improving paediatric data collection. - is this generalisable to all primary care data systems?

REVIEWER	Salemi, Jason Baylor College of Medicine, Houston, TX, United States Competing interests: None
REVIEW RETURNED	23-Aug-2017

GENERAL COMMENTS	The authors set the stage for an important investigation as to the extent to which primary care records can improve the completeness of ascertainment of birth defects. This supposition was made primarily because traditional registries only surveil for birth defects through age 1 whereas primary care records can be used to follow later into childhood when many defects may manifest, or to overcome cases missed in the first year of life by existing registry mechanisms. From that standpoint, there is clear value to the idea behind the paper. However, I feel the paper does not provide a clear message because in some ways I find the paper could do a better job of describing the ways in which primary care records are improving surveillance and describing the accuracy of primary care records for identifying birth defects. In other ways, the addition of measures of association, to me, muddies the water and detracts from the paper's intended message. In fact, although the authors describe an objective of the paper to "determine whether magnitudes of association for risk factors persisted", there is no mention of this in the Discussion. Below, I pose some questions and suggestions to hopefully make the paper more useful to its intended audience. 1. In my opinion, I think this paper should focus on the added contribution of primary care records to the surveillance of birth defects, above and beyond what is currently being done. To that end, it would be valuable to explicitly describe in the paper, the current data sources for the surveillance system to which the primary care records database is being compared. Is it that current surveillance efforts include only inpatient hospital data? Only inpatient, ambulatory, specialty care, emergency department? Do they include primary care records, but simply don't extend past age one? Furthermore, the measures of association do nothing for me in this paper, especially since the paper prepares me to learn about improvement in surveillance, not an investigation of risk factors. I genuinely feel that there is plenty for the surveillance aspects for this paper, and the measures of association can become another paper. 2. The paper really hones in on the utility of primary care records because of this deficiency in capturing infants with birth defects AFTER their first birthday (referencing the fact that only 2% of all registrants are captured after one year of age). However, if I am

	understanding Table 1 correctly, if I restrict only to the first year of life, primary care records DOUBLE the rate of all birth defects and with the exception of abdominal defects, increases the rate of defect subgroups. So it's not just following past the first birthday. There seems to be a fundamental deficiency in the BINOCAR system that fails to identify children with birth defects. Do primary care records really add that many new cases? If so, explore the why behind this, and not just the intense focusing on following past the first birthday – I feel this discrepancy in rates in infants less than one is even more profound. 3. There is a comment in the paper regarding differences in BINOCAR and BiB, namely the inclusion/exclusion of infants with metabolic disorders depending on the presence of associated defects. I do not understand why the BiB data cannot be modified to agree with the BINOCAR so that we can begin to isolate the independent effect of adding primary care data instead of wondering how much of the difference is due to other differences in the methodologies, like what to do with children with a metabolic disorder but no other birth defects. 4. I am not that familiar with the BiB. Is the cohort representative of the general population or might the sample include children who may be at higher risk of birth defects than the general population, due to participation bias? Again, just trying to think of other reasons the rate is so much higher in BiB than in BINOCAR. 5. I apologize if I missed it, but the authors made it seem as though they were going to "determine the accuracy of the primary care diagnoses information" by comparing them to medical records (gold standard), but I saw no tables, figures, or mention of positive predictive value or accuracy of the defects identified by primary care records, neither overall nor by defect subset. Was this actually an intended aim of the study? 6. Figure 1 was a little hard to follow. First, if you agree with my earlier sentiments, I would consider removing the regression analyses as the currently exist. Second, I would add that 733 infants with anomalies from the linked BiB-GP data are carried down after 1618 are excluded (so it is easy to tell that the 733 + 127 from phase 1 equal the 860. It is also important to note that if BINOCAR represents the "existing surveillance system" to which you are comparing your new approach, then there are 127 of your 860 cases in BiB that were not captured by primary care records, right? I just need to understand the precise differences in the registries being compared – BiB which is basically a hospital notification system and primary care records…to BINOCAR, which includes? Is this an apples to apples comparison in which the only real difference is the inclusion of primary care records from birth to age 5 or are there other important differences that could partly or mostly explain the differences we are observing? 7. In Table 1, I am confused as to how chromosomal anomalies are excluded from BINOCAR, yet there is a rate of chromosomal anomalies of 43.0? 8. For Table 2, what would be more valuable and coincide more with the study aims, would be to compare characteristics of infants diagnosed with a birth defect through primary care records only versus those that were identified by another source (e.g., hospital

	notification system). That would facilitate an understanding of the unique contribution of primary care records and in which population subgroups are primary care records making a difference in improving surveillance. 9. For Figure 2, I believe you currently have the 860 records identified by the BiB, both phases. Why not compare the proportion captured by phase 1 (first line), and then add a line for the additional proportion captured by phase 2? Again, your goal does not seem to be how many were captured by BiB in each age group, but instead what the added value of primary care records is, right? Also, in the Figure, is "age 0" the time period between birth and the child's first birthday or something different? I ask because if that is the case, it seems highly unlikely that less than 10% of all birth defects are captured after the child turns 1. If I am misunderstanding the axis, please revised to make clearer. 10. An appendix would be helpful to describe the mapping of CTV3 Read codes to ICD-10 codes or at least which fall under major defect categories. Ultimately, as someone who is always investigating the impact of new data sources (once they are deemed sufficiently accurate in their diagnoses) on the completeness of birth defects surveillance/registries, I really love the idea and the effort put forth. I do feel the paper is currently too multifaceted, and tough to follow in terms of really honing in on the independent additional contribution of primary care records, and how much of that additional contribution is because we are following kids past age 1 versus the ability to primary care records to identify birth defects missed by other data sources prior to the child's first birthday. There are several papers by Salemi et al, Tanner et al, and Rutkowski et al investigating the relative contribution of various data sources, the accuracy of hospital discharge diagnosis codes for birth defects, which might give some additional framing ideas for presenting your information. Not necessary, just may be useful.

VERSION 1 – AUTHOR RESPONSE

Reviewers comments (Italics) and our response	Changes made in the manuscript
<u>Reviewer: 1</u> *Comments to the Author* *This is an interesting and useful contribution in the developing context of Congenital anomaly ascertainment and registration. Whilst i have recommended minor revisions some of these are significant to the inferences made in the paper. My comments are as follows:* *- the introduction should include an acknowledgement that 61% of reported congenital anomalies are detected in the antenatal period.*	
1. Thank you for pointing out the statistic on antenatal detection. However we found the	Added the following response to the

NCARDS report (2015) states that 71% of CA are detected prenatally. Therefore we added this statistic to the limitations section where we had already mentioned the pitfall of primary care data vis a vis showing terminations of pregnancy for CA.	manuscript "NCARDRS highlights 71% of CA are detected antenatally, and 42% of CA diagnosed antenatally resulted in termination."
- the date of diagnosis for congenital anomalies is missing in 14% of live births. As the paper acknowledges the recording of date of diagnosis in later diagnoses of CA is not always accurate (or present). This may impact on the claim of a 30% increase in ascertainment using this method of data linkage	
2. This is a very good point although we can't find where the 14% figure came from without a reference. If referring to the 2015 NCARDRS report, it states just under a third of CA were diagnosed in the postnatal period in a live birth and time of diagnosis was known for 46% of these? So this indicates 54% had missing dates? We need clarification of the figures you provide. However in the primary care data there were no missing dates of diagnosis. All CA diagnoses were associated with a primary care appointment, which was associated with a date. One study also extracting CA from primary care data found the dates of diagnosis in the primary care record were not always the same as CA diagnoses made by the diagnosing consultant when checked with medical records, but the dates were found to be within a 30-day discrepancy window. We referenced the study sighting this point in the limitations (reference 17). We hope the reviewers feel this is adequate.	Rephrased the comment and highlighted the reference in the paper.
- risk factors for CAs are being collected in the national system NCARDRS	
We have added this	Wording to this effect added.
- The national register would not tolerate a 17% discrepancy rate in accuracy of a data source without requiring subsequent validation of that data. This is very important to maintain the high quality of the data. The implication of this is that whilst this may serve well as an indicative source of ascertainment the cases would need further investigation to be acceptable as registerable cases.	

This may not be viable, nationally. It may be worth doing some further analysis of which groups of cases have a) a high level of accuracy with the hospital diagnoses and b) are most likely missing from the otherwise ascertained cases.	
3. We appreciate the comment about national registers not tolerating a 17% discrepancy rate, and thank the reviewer for mentioning this. The 17% discrepancy rate was investigated in detail in our study, and the main reason for this discrepancy rate was found to be clinicians using ICD-10 codes outside of the CA ICD-10 chapter in the first Born in Bradford congenital anomaly study (Sheridan et al 2013), used for validating the primary care CA diagnoses. We spoke to the clinicians diagnosing children from the Sheridan et al (2013) study who explained the conditions some of the children presented with were so rare they could not find the appropriate ICD-10 code to match within the Q chapter of the ICD-10. The Algorithm written to extract Read-codes from primary care data in the present study, only included conditions in the ICD-10 Q chapter, as advised by EUROCAT therefore would not have selected conditions outside of the ICD-10 Q chapter. There were a few other reasons for the 17% discrepancy rate, that were the child had died, or they had changed GP practice more than once. We have added a sentence to the paper to clarify this and hope this satisfies your query.	Added the following: We also found 127(17%), which did not match between phase 1 and 2 methodologies. On further inspection these cases had ICD-10 codes outside of the CA chapter as recommended by EUROCAT. The clinicians responsible for the phase 1 study explained this was due to some conditions being so rare they could not find an appropriate code within the recommended CA ICD-10 chapter. A small number of the 127 children had died or moved primary care practice more than once, causing potential errors in their diagnoses records.
- page 14 NCARDRS; line 18 ?prognosis	
Removed, thank you	
- the dramatic increase in skeletal dysplasias is very interesting (and perhaps emphasised because of the demographics in the study cohort). I am very curious about the increased nervous system diagnoses and would have liked more detail around this - were specific conditions identified? A detailed breakdown would be interesting.	
4. The increase in nervous system disorders was due to hearing loss. A geneticist was consulted regarding the categorization of hearing loss and agreed it was best to place these CA in neurological disorders. Skeletal dysplasia's were increased primarily due to short stature. But there were only a few cases.	Modified what was there already and included an additional reference for the increase in Asthma prevalence: Considering the percentage increase by bodily system group, skeletal dysplasias increased considerably from age one to age five (210%),

	primarily due to diagnoses of short stature, followed by nervous system (77%), due to an inflation of hearing loss in Bradford, and respiratory (44%) disorders, of which are confirmed to be high in Bradford (ref 27). Reference added - Bradford Joint Strategic Needs Assessment. Long Term Conditions. 2014.
- the cardiac conditions identified later tend to be less serious conditions (small VSD, ASD) these might be more properly (and more accurately) be identified using tertiary cardiac database data which NCARDRS is trying to procure. There is no discussion in the paper about whether there are better ways of improving the data collection for paediatric cases, using surgical data or other disease specific data for example? it might be useful to consider whether there are specific conditions being picked up in this data linkage that would not be picked up by otherwise improving paediatric data collection.	
5. The reviewer is correct to point out some heart conditions such as VSD and ASD are sometimes less severe, and perhaps could be more accurately identified using cardiac databases. However we feel our data may reflects a true increase in cardiac conditions. This is due to some evidence to suggest the late detection is caused by some cardiac diagnoses being missed at antenatal screening, due to difficulties in diagnosis. We have added an explanatory sentence on this and added a reference.	Late detection of heart CA could be attributable to some cases being missed at antenatal screening, due to detection being difficult.[25]
- is this generalisable to all primary care data systems?	
Yes we believe so as long as the data quality is good enough.	
I hope these are helpful *kind regards* **Thank you, we did find these comments helpful**	

Reviewer: 2 *Comments to the Author* *Thank you for the opportunity to review this paper. The authors set the stage for an important investigation as to the extent to which primary care records can improve the completeness of ascertainment of birth defects. This supposition was made primarily because traditional registries only surveil for birth defects through age 1 whereas primary care records can be used to follow later into childhood when many defects may manifest, or to overcome cases missed in the first year of life by existing registry mechanisms. From that standpoint, there is clear value to the idea behind the paper. However, I feel the paper does not provide a clear message because in some ways I find the paper could do a better job of describing the ways in which primary care records are improving surveillance and describing the accuracy of primary care records for identifying birth defects. In other ways, the addition of measures of association, to me, muddies the water and detracts from the paper's intended message. In fact, although the authors describe an objective of the paper to "determine whether magnitudes of association for risk factors persisted", there is no mention of this in the Discussion. Below, I pose some questions and suggestions to hopefully make the paper more useful to its intended audience.*	
Thank you for the comment regarding one of this papers original aims to determine if the magnitudes of association for risk factors persist. Agreed this was not mentioned in the discussion. We did however mention this in the key messages. We continue to focus this paper on risk factors and we have included a comment on this in the discussion. This does however go against your advice to separate the risk factors from the use of primary care data for case ascertainment into two studies. But we will explain why we have taken this line as we respond to the rest of your comments.	We also explored whether magnitudes of association for risk factors of CA persisted with this increased study population. Words added. We found no substantial change in risk factors, even with a slightly difference CA profile. Changes to statistical significance of risk factors would have had implications for comparative analyses between registries with different ascertainment methods.
1. In my opinion, I think this paper should focus on the added contribution of primary care records to the surveillance of birth defects, above and beyond what is currently being done. To that end, it would be	

valuable to explicitly describe in the paper, the current data sources for the surveillance system to which the primary care records database is being compared. Is it that current surveillance efforts include only inpatient hospital data? Only inpatient, ambulatory, specialty care, emergency department? Do they include primary care records, but simply don't extend past age one? Furthermore, the measures of association do nothing for me in this paper, especially since the paper prepares me to learn about improvement in surveillance, not an investigation of risk factors. I genuinely feel that there is plenty for the surveillance aspects for this paper, and the measures of association can become another paper.	
Thank you to the reviewer for suggesting this however, we feel that the unique contribution of this paper, is adding to the body of evidence on validating primary care data as a source of CA case ascertainment, due to its longitudinal nature and its ability to capture diagnoses beyond age one, and by linking to multi-ethnic cohort data from the born in Bradford study, and comparing to a previous highly established CA study also using born in Bradford data (published in the Lancet) was able to a) compare diagnoses to this previous Bradford study as a validation exercise, b) re-affirm risk factors persist in this large population sample. The results therefore have clinical implications for Bradford, and national implications for the use of primary care data for CA case ascertainment. We have added a section to the discussion to explain this.	We have rewritten this by adding this section to the discussion, "We also assessed the effect of improved ascertainment on the point estimates and statistical significance of the risk factors for CA. We found no substantial change in these risk factors, even with a slightly different CA profile. Changes to statistical significance of risk factors would have had implications for comparative analyses between registries with different ascertainment methods."
2. The paper really hones in on the utility of primary care records because of this deficiency in capturing infants with birth defects AFTER their first birthday (referencing the fact that only 2% of all registrants are captured after one year of age). However, if I am understanding Table 1 correctly, if I restrict only to the first year of life, primary care records DOUBLE the rate of all birth defects and with the exception of abdominal defects, increases the rate of defect subgroups. So it's not just following past the first birthday. There seems to be a fundamental deficiency in the BINOCAR system that fails to identify children with birth defects. Do primary care records really add that many new cases? If so, explore the why behind this, and not just the intense focusing on following past the first birthday – I feel this discrepancy in rates in infants less than one is even more profound.	

Secondly, there was a study done by EUROCAT looking into late diagnosis, and the potential deficiencies in the BINOCAR and EUROCAT systems regarding late diagnosis. This study is available on the EUROCAT website, was published in 2009 by the EUROCAT team, and found that they agreed there are some late conditions which were mainly in subgroups eye, microcephaly, teratogeneic syndromes and fetal alcohol syndrome. Because the late conditions were exclusive to these few subgroups, EUROCAT decided they were not important enough to warrant detection after the age of 1, unless specifically investigating conditions in the aforementioned subgroups, and was referenced in the present study.	Added this sentence: Despite the large difference in BINOCAR rates at age 1, and BiB rates at age 1, the phase 1 study[10] found a similar 305.74 per 10,000 live births, helping to explain the influence of the Bradford demographics on the high numbers, before additional cases are added post age 1 using primary care records.
3. There is a comment in the paper regarding differences in BINOCAR and BiB, namely the inclusion/exclusion of infants with metabolic disorders depending on the presence of associated defects. I do not understand why the BiB data cannot be modified to agree with the BINOCAR so that we can begin to isolate the independent effect of adding primary care data instead of wondering how much of the difference is due to other differences in the methodologies, like what to do with children with a metabolic disorder but no other birth defects.	
This is a really useful comment and we are happy to explain our approach as it helps consolidate the argument of this paper and underlines methodological flaws in BINOCAR/EUROCAT. Firstly in response to your comment, the BiB data cannot be modified as BiB does not collect terminations or detection of congenital anomalies in the antenatal period, as recruitment for the study is at 26-28 weeks gestation. This is mentioned in the paper at the end of the discussion section on page 15. We have modified the BiB data to agree with BINOCAR in the only way possible, which was to exclude metabolic and chromosomal disorders and calculate the rates from this sample. We could not exclude chromosomal conditions alone, because BINOCAR includes metabolic disorders only if they have a structural anomaly. In Bradford hospital and the BiB original congenital anomaly study, metabolic conditions were registered if they had a structural anomaly or not. This is mentioned on page 7 of the paper. It was therefore not possible from the data in this study to determine which metabolic conditions from the BiB data also had a structural anomaly. Metabolic conditions were therefore removed, which ensured that the rates were not over estimated, and are in fact more likely to be underestimated with all metabolic conditions removed.	

4. I am not that familiar with the BiB. Is the cohort representative of the general population or might the sample include children who may be at higher risk of birth defects than the general population, due to participation bias? Again, just trying to think of other reasons the rate is so much higher in BiB than in BINOCAR.	
We feel this comment additionally reflects why it is important to retain the risk factors analysis. As described in the introduction of this study, BiB and Bradford are not representative of the general population. There is a high level of ethnic diversity, almost half of the population are of Pakistani heritage, and there is approximately 60% of consanguineous marriage in this group. As the referenced Sheridan et al (2013) paper discovered, consanguinity is a major risk factor for congenital anomalies in Pakistani communities. If the risk factor analysis using the increased population captured by interrogation of the primary care database, was not repeated in this paper, the large increase in numbers would appear, as you put it, to be subject to participation bias. Given the results of the risk factor analysis remain the same as the previous Bradford congenital anomaly study (Sheridan et al 2013), this validates primary care data as a source of gaining more congenital anomalies from the same at risk groups.	
5. I apologize if I missed it, but the authors made it seem as though they were going to "determine the accuracy of the primary care diagnoses information" by comparing them to medical records (gold standard), but I saw no tables, figures, or mention of positive predictive value or accuracy of the defects identified by primary care records, neither overall nor by defect subset. Was this actually an intended aim of the study?	
We thank the reviewer for the comment regarding clarification to determine the accuracy of the primary care diagnoses information by comparing to medical records. As the process was, using a population of children identified by clinician diagnosis (Sheridan et al 2013), in Bradford, this paper compares children identified by searching the primary care database for read medical codes, to those identified by clinician diagnoses, of which there were 296 identified by both methodologies. There were an additional 437 identified by primary care data and 127 that did not match. On closer inspection of the 127 that did not match, this was primarily due to clinicians diagnosing conditions and using read-codes that were not in the ICD-10 chapter. The main reason for the 127 discrepancy rate was found to be clinicians using ICD-10 codes outside of the congenital anomaly chapter in the first Born in Bradford congenital anomaly study (Sheridan et al 2013)	This is the section also requested by the first reviewer and was added to the manuscript. We also found 127(17%), which did not match between phase 1 and 2 methodologies. On further inspection these cases had ICD-10 codes outside of the CA chapter as recommended by EUROCAT. The clinicians responsible for the phase 1 study explained this was due to some conditions being so rare they could not find an appropriate code within the recommended CA ICD-10 chapter. A small number of the 127 children had died or moved primary care practice more than once, causing potential

which was used for validating the primary care data. We personally spoke to the clinicians diagnosing children from the Sheridan et al (2013) study and they explained the conditions some of the children presented with were so rare they could not find the appropriate ICD-10 code to match within the Q chapter of the ICD-10. The Algorithm written to extract Read-codes from primary care data in the present study, only included conditions in the ICD-10 Q chapter, as advised by EUROCAT therefore would not have selected conditions outside of the ICD-10 Q chapter. There were a few other reasons for the 17% discrepancy rate, that were the child had died, or they had changed GP practice more than once. This issue was picked up by the first reviewer also so we have added a sentence to the paper to explain this and hope this satisfies your query. Furthermore, evidence also using primary care data to select populations of children with congenital anomalies, suggests that as primary care data has been validated using medical records, and as reviewers point out this is the gold standard, at least an 80% agreement rate was found in multiple studies, and 83% in this study, thus there may not be a requirement to continue validating primary care data.	errors in their diagnoses records.
6. Figure 1 was a little hard to follow. First, if you agree with my earlier sentiments, I would consider removing the regression analyses as the currently exist. Second, I would add that 733 infants with anomalies from the linked BiB-GP data are carried down after 1618 are excluded (so it is easy to tell that the 733 + 127 from phase 1 equal the 860. It is also important to note that if BINOCAR represents the "existing surveillance system" to which you are comparing your new approach, then there are 127 of your 860 cases in BiB that were not captured by primary care records, right? I just need to understand the precise differences in the registries being compared – BiB which is basically a hospital notification system and primary care records…to BINOCAR, which includes? Is this an apples to apples comparison in which the only real difference is the inclusion of primary care records from birth to age 5 or are there other important differences that could partly or mostly explain the differences we are observing?	
Born in Bradford (BiB) is a cohort study collecting information from mothers and children in Bradford, and provides information which generated the risk factors. It is not a congenital anomaly register, and was used in this study to ascertain a study population, utilize the detailed information collected to compensate for the missing data often present in primary care studies, and to create variables for risk factors. For example ethnicity and socioeconomic status are not recorded well in primary care data,	

but were in BiB, and BiB asks mothers in the cohort detailed information about their health during pregnancy. The BiB study was referenced in the introduction on page 3, and was not explained in full to reduce the word count to ensure this paper met the journals word count regulations. BINOCAR is the national congenital anomaly register, and the present study introduces primary care data as a more comprehensive source of congenital anomaly case ascertainment.	
7. In Table 1, I am confused as to how chromosomal anomalies are excluded from BINOCAR, yet there is a rate of chromosomal anomalies of 43.0?	
We thank the reviewer for pointing out their confusion regarding excluding chromosomal anomalies from BINOCAR, yet there is a rate of chromosomal anomalies of 43.0. This is a typo and has now been removed from the table.	
8. For Table 2, what would be more valuable and coincide more with the study aims, would be to compare characteristics of infants diagnosed with a birth defect through primary care records only versus those that were identified by another source (e.g., hospital notification system). That would facilitate an understanding of the unique contribution of primary care records and in which population subgroups are primary care records making a difference in improving surveillance.	
Again we thank the reviewer for the suggestion to compare children diagnosed with congenital anomalies from the primary care records to children diagnosed by another source. However this is exactly what is done in this study. The dataset of children with congenital anomalies used for comparison (Sheridan et al 2013), was a hospital notification system in Bradford hospital, and diagnoses were confirmed by clinicians, the present study compares congenital anomaly diagnosis from interrogation of the primary care database for all GP practices in Bradford. The study further explains both in table 2 and in the discussion (page 14), that the subgroups for nervous system and skeletal dysplasia's increased the most when using the primary care data. We have now added specific diagnoses, which caused the inflation of these bodily system subgroups as per the advice of the first reviewer.	This comment (which was also added as a response to reviewer 1.) and reference added. Considering the percentage increase by bodily system group, skeletal dysplasias increased considerably from age one to age five (210%), primarily due to diagnoses of short stature, followed by nervous system (77%), due to an inflation of hearing loss in Bradford, and respiratory (44%) disorders, of which are confirmed to be high in Bradford (Bradford Joint Strategic Needs Assessment. Long Term Conditions. 2014).
9. For Figure 2, I believe you currently have the 860 records identified by the BiB, both phases. Why not compare the proportion captured by phase 1 (first line), and then add a line for the additional	

proportion captured by phase 2? Again, your goal does not seem to be how many were captured by BiB in each age group, but instead what the added value of primary care records is, right? Also, in the Figure, is "age 0" the time period between birth and the child's first birthday or something different? I ask because if that is the case, it seems highly unlikely that less than 10% of all birth defects are captured after the child turns 1. If I am misunderstanding the axis, please revised to make clearer.	
Thank you for the useful comment about possible amendments to figure 2. However as the aim of this graph is to show the additional diagnoses made at each year of age, given the data collection methods of phase 1, which were up to the child's first birthday only, this graph does show what you are suggesting, as if data collection were to be cut off at age one, the line would not extend to age 5. We have however changed the explanation of this plot in the text of the paper to add clarification. Age 0 refers to the number of diagnoses that were made as the child was born. Age 1 refers to the period between birth and the child's first birthday, during which 528 congenital anomalies were found (61%).	We have amended the text to read "Figure 2 demonstrates the age of the child when they received their first CA diagnosis. Without the additional cases from primary care data, this plot would only show diagnosis up to age one, a total of 600 children. Primary care data adds a further 260(30%) of cases".
10. An appendix would be helpful to describe the mapping of CTV3 Read codes to ICD-10 codes or at least which fall under major defect categories.	
We thank the reviewers for the advice to contain an appendix indicating the mapping of CTV3 to ICD-10, or at least which fall under major defect categories. The reason this was not included in the paper is because the mapping is an online available resource which was referenced in the paper, and the classification of ICD-10 codes into major congenital anomalies, and removal of minor anomalies, is also an online available resource, also referenced in the methods section of the paper on page 5. We were also wary that there may be issues about rare conditions and identifiability	
Ultimately, as someone who is always investigating the impact of new data sources (once they are deemed sufficiently accurate in their diagnoses) on the completeness of birth defects surveillance/registries, I really love the idea and the effort put forth. I do feel the paper is currently too multifaceted, and tough to follow in terms of really honing in on the independent additional contribution of primary care records, and how much of that additional contribution is because we are following kids past age 1 versus the ability to primary care records to identify birth defects missed by other data sources prior to the child's first birthday. There are several papers by Salemi et al, Tanner et al, and	

Rutkowski et al investigating the relative contribution of various data sources, the accuracy of hospital discharge diagnosis codes for birth defects, which might give some additional framing ideas for presenting your information. Not necessary, just may be useful.	
We are grateful for the supportive comments and excellent advice offered here.	