PEER REVIEW HISTORY

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

ARTICLE DETAILS

TITLE (PROVISIONAL)	Usefulness of Natriuretic Peptide for the diagnosis of Kawasaki Disease- A Systematic Review and Meta-analysis
AUTHORS	Kuan-Ho Lin, Shy-Shin Chang , Chin-Wei Yu, Shen-Che Lin, Shu-Chun Liu, Hsiao-yun Chao, Meng-tse Gabriel Lee, Jiunn-Yih Wu, Chien-Chang Lee,

VERSION 1 - REVIEW

REVIEWER	Satoru Iwashima Hamamastu University School of Medicine, Departement of Pediatrics. Japan
REVIEW RETURNED	26-Sep-2014

GENERAL COMMENTS	I reviewed the manuscript `` Title: Usefulness of Natriuretic Peptide for the diagnosis of Kawasaki Disease-A Systematic Review and Meta-analysis`` with great pleasure. This is manuscript that the author systematic reviewed to examine the diagnostic value of serum B-type natriuretic peptide (BNP) in acute Kawasaki disease (KD). Serum N-terminal pro-brain natriuretic peptide (NTproBNP) and brain natriuretic peptide (BNP) are often elevated in patients with acute KD, but the NTpro BNP and BNP level in normal children is higher than in adults. Thus, characterization of the normal levels and cut-off values of NTpro BNP and BNP according to age is warranted for proper diagnosis of acute KD. Author concluded NT-proBNP has high diagnostic value for identifying KD in patients with protracted undifferentiated febrile illness and may be used as a diagnostic tool for KD. I agree author`s conclusion, however this manuscript is not including new aspect. This systematic review was not included some reported in this area, such as. Shiraishi et al, reported on Circ J 2013; 77: 2097 – 2101 and Cho SY et al, reported on Ann Clin Lab Sci. 2011 Fall;41(4):360-3. This systematic review was conducted with a relatively small number of reports and not included some reported in this area, such as. Shiraishi et al, reported on Circ J 2013; 77: 2097 – 2101 and Cho SY et al, reported on Ann Clin Lab Sci. 2011 Fall;41(4):360-3. Author should be mention why these two reports were excluded.

REVIEWER	Peter Watson Medical Research Council UK
REVIEW RETURNED	03-Feb-2015

GENERAL COMMENTS	Page 3 lines 10-15 - The positive likelihood ratio is a discriminatory

	criterion on its own so would think the fact this is below the threshold of Jaeschke et al (1994) is sufficient to suggest the BNPs have insufficient discriminatory power to be used in diagnosis. The heterogeneity in the odds ratios between studies is not discussed in the abstract (lines 16-18) but you could say the presence of heterogeneity directly suggests more work is needed to evaluate the relationship between the BNPs and KD. To me the result of this paper is simply that the six odds ratios reported in the studies on page 28 are all greater than one suggesting the presence of KD is related to increasing peptide but that we don't know the magnitude of such a relationship due to limitations in the number and variation of studies. One could, therefore, simply quote the signs of the six odds ratios and say they are all positive hence the direction is the same. I don't think the other analyses in this paper therefore add anything to this simple fact of all the odds ratios being greater than one so I struggle to see the motivation for them. What more information are the ROC curves, likelihood ratios specifically adding to this? I do have doubts about the usefulness of the analyses given the acknowledged limitations of study heterogeneity and low numbers of studies and wonder if these analyses could be removed and a much shorter paper submitted with fewer analyses and the conclusion of the need for further large scale studies? My worry here is: Could people have realised the conclusion of this paper that there is a correlation (of uncertain size) between peptide and KD without reading this paper thereby suggesting this paper is not adding anything new? Usefulness of natriuretic peptide for the diagnosis of Kawasaki Disease - a Systematric Review and Meta-analysis. bmjopen-2014-006703 Page 6. On page 6 prior to any analysis the authors mention that although the relationship between KD and B-type natriuretic peptide (BNP) and its inactive cleave product, NT-pro-BNP, seems tenable (lines 26-40), the strength of such a relationship varies (lines 46-49). The imprecision of the relationship between KD and BNP and NT-pro-BNP is then confirmed following HSROC analyses with this inaccuracy shown by the high heterogeneity mentioned on lines 47-53 on page 12 which naturally yields the conclusion that together with the limited number of studies (page 16, line 50 and elsewhere) and differences between studies which cannot be adjusted for (page 16, line 30), NT-proBNP cannot be used to decide upon the presence or absence of KD and cannot be used as a diagnosis tool (page 16, lines 50-56. The conclusion is that (NT-pro)BNP could only be used as a diagnostic tool if more studies are carried out using a range of thresholds for the diagnosis of KD to see how it relates to both BNP variables and that the small number of studies (five) in this paper precludes an estimate of this relationship. The imprecision of the relationship between BNP and KD is further confirmed by the low value of the positive likelihood ratio on lines 30-33 of page 12 (see later) as the authors acknowledge on line 33 of page 12. There are also issues in comparing studies given the large heterogenity (page 12, line 50) between the odds ratio measuring the (NT-pro)BNP, KT correlation possibly due, as I understand it, to differing concentrations of peptides (page 15, line 54 to page 16, line 7) and the limited number of studies which both suggest more better

	designed sets of larger studies need to be performed to get an idea of the size of this association using, for example, a meta-regression to adjust for confounders. The point here is that the conclusion of this paper, that more work is needed to identify cut-offs for using (NT-pro)BNP to diagnosis KD, is already apparently known prior to this paper from a descriptive look at the literature (of what ends up as only five or so studies in the current paper) as mentioned in the paragraph between lines 20 and 55 on page 6 with lines 46-49 mentioning the heterogeneity of the BNP, KD relationship across studies. In particular the odds ratios in each of these studies (page 28) suggest there is a consistent relationship albeit of an imprecise nature between (NT-pro)BNP and KD which would suggest (NT-pro)BNP is a potentially useful tool to use to predict the presence, or otherwise, of KD but that more work needs to be done to identify the magnitude of its association with KD and, hence, turn it into a diagnostic tool. I, therefore, wonder what the analyses in this paper add to the literature since the conclusion from these analyses that more work is needed to get an accurate idea of the relationship between (NT-pro)BNP and KD is mentioned in the introduction on page 6 between lines 46 and 49. Other comments Page 12, lines 30-42. Using the rules of thumb of Jaeschke et al. (1994) for interpreting sizes of likelihood ratios a positive likelihood ratio of between 2 and 5 suggests, as the authors note, that their model provides only weak diagnostic evidence for the presence of KD (Kawasaki Disease) and is, therefore, not a very useful 'rule-in' test (as is also seen in the low specificity) irrespective of the value of the negative likelihood ratio and, indeed, other criteria such as area under the HSROC curve whose high value of 0.86 (page 12, line 45) is rather irrelevant in that it provides 'good' discrimination yet does not change the key clinical conclusion that given the poor diagnostic size of the positive likelihood ratio the (NT-pro)BNP currently cannot be used as a diagnostic tool because we do not know the precise magnitude of its correlation with KD. (As an aside one could, incidentally, if one was mentioning specificity and sensitivity also have mentioned positive and negative predictive values which are often mentioned in addition to sensitivity and specificity). Page 12, lines 42-48. The HSROC curve in Figure 3 is not commented upon and with just five data points (studies) does not look informative and the area under the curve (page 12, line 45) would therefore appear sufficient to give an idea of diagnostic accuracy. The curve unusually has specificity on the x-axis and would benefit from the more standard plotting of 1-specificity on the x-axis so that the origin is in the bottom left corner of the plot. One could have motivated the use of HSROC curves by saying they take both within and between study variation into account which make them useful diagnostic tools in meta-analysis. Page 12, line 45. Rules of thumb for acceptable values of the area under the curve are given in Hosmer and Lemeshow (2000) for example who suggest an area under the curve of between 0.80 and 0.90 can be regarded as 'good'. Page 12, lines 54-57. Any further details of the lack of publication

	bias such as the results of Eggers test? I would think it would be difficult with just five studies to have a powerful enough test to detect publication bias in any case? Page 15, lines 48-54. The sentence here is grammatically incorrect. References Hosmer DW and Lemeshow SL (2000). Applied Logistic Regression. 2nd Edition. Wiley:New York. Jaeschke R, Guyatt, GH and Saclett DL (1994) Users guides to the medical literature. III. How to use an article abotu a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based medicine Working Group. Journal of the American medical Association, 271(9), 703-707.

VERSION 1 – AUTHOR RESPONSE

Reviewer Name Satoru Iwashima
Institution and Country Hamamastu University School of Medicine, Department of Pediatrics. Japan
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
This systematic review was conducted with a relatively small number of reports and not included some reported in this area, such as. Shiraishi et al, reported on Circ J 2013; 77: 2097 – 2101 and Cho SY et al, reported on Ann Clin Lab Sci. 2011 Fall;41(4):360-3.
Reply:
We thanked the reviewer's for his comments. Cho SY et al's paper is actually included in our analysis (table 2), and we have added Shiraishi et al's paper to our meta-analysis as suggested. Thus, the statistical analysis has been reconducted, and all the tables/figures have been remade. Unfortunately, we still have a relatively small number of reports, as
Kawasaki Disease can be considered a rare disease outside of East Asia. To illustrate, according to Uehara et al, reported on J Epidemiol 2012;22(2):79-85, between 1999-2000 England has an incidence of 8.1 cases (per 100 000), while Japan has an incidence of 218.6 cases (per 100 000) on 2008. The rather low incidence rate makes the collection of KD patients for NT-pro BNP analysis extremely difficult, which might explain why we only find 7 usable reports out of 175 citations.

Reviewer Name Peter Watson
Institution and Country Medical Research Council
UK
Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
Usefulness of natriuretic peptide for the diagnosis of Kawasaki Disease - a Systematic Review and Meta-analysis. bmjopen-2014-006703

Page 6. On page 6 prior to any analysis the authors mention that although the relationship between KD and B-type natriuretic peptide (BNP) and its inactive cleave product, NT-pro-BNP, seems tenable (lines 26-40), the strength of such a relationship varies (lines 46-49). The imprecision of the relationship between KD and BNP and NT-pro-BNP is then confirmed following HSROC analyses with this inaccuracy shown by the high heterogeneity mentioned on lines 47-53 on page 12 which naturally yields the conclusion that together with the limited number of studies (page 16, line 50 and elsewhere) and differences between studies which cannot be adjusted for (page 16, line 30), NT-proBNP cannot

be used to decide upon the presence or absence of KD and cannot be used as a diagnosis tool (page 16, lines 50-56. The conclusion is that (NT-pro)BNP could only be used as a diagnostic tool if more studies are carried out using a range of thresholds for the diagnosis of KD to see how it relates to both BNP variables and that the small number of studies (five) in this paper precludes an estimate of this relationship. The imprecision of the relationship between BNP and KD is further confirmed by the low value of the positive likelihood ratio on lines 30-33 of page 12 (see later) as the authors acknowledge on line 33 of page 12. There are also issues in comparing studies given the large heterogenity (page 12, line 50) between the odds ratio measuring the (NT-pro)BNP, KT correlation possibly due, as I understand it, to differing concentrations of peptides (page 15, line 54 to page 16, line 7) and the limited number of studies which both suggest more better designed sets of larger studies need to be performed to get an idea of the size of this association using, for example, a meta-regression to adjust for confounders. The point here is that the conclusion of this paper, that more work is needed to identify cut-offs for using (NT-pro)BNP to diagnosis KD, is already apparently known prior to this paper from a descriptive look at the literature (of what ends up as only five or so studies in the current paper) as mentioned in the paragraph between lines 20 and 55 on page 6 with lines 46-49 mentioning the heterogeneity of the BNP, KD relationship across studies. In particular the odds ratios in each of these studies (page 28) suggest there is a consistent relationship albeit of an imprecise nature between (NT-pro)BNP and KD which would suggest (NT-pro)BNP is a potentially useful tool to use to predict the presence, or otherwise, of KD but that more work needs to be done to identify the magnitude of its association with KD and, hence, turn it into a diagnostic tool. I, therefore, wonder what the analyses in this paper add to the literature since the conclusion from these analyses that more work is needed to get an accurate idea of the relationship between (NT-pro)BNP and KD is mentioned in the introduction on page 6 between lines 46 and 49.

Reply:

We thanked the reviewer for his comments. As described in the introduction, Kawasaki Disease (KD) is the major cause of cardiac morbidity in pediatric population. However, KD is difficult to diagnose and has a low incidence rate. Uehara et al, reported on J Epidemiol 2012; 22(2):79-85, that England has an incidence of 8.1 cases (per 100 000) between 1999-2000, while Japan has a much higher incidence of 218.6 cases (per 100 000) on 2008. The rather low incidence rate makes the collection of KD patients for NT-pro BNP analysis extremely difficult, which might explain why we only find 7 usable reports out of 175 citations. One of the major advantages of systemic review and meta-analysis is its capability to pool data across different studies, so that the results would not be affected by the chance variability within each individual study. Our manuscript is the first attempt to systematically summarize current evidence on the diagnostic value of NP on Kawasaki disease. After addition of a new study in this revision, the total sample size could reach 1137 with 428 case of KD, which provide adequate sample to evaluate the accuracy of NT-pro BNP in the diagnosis of KD.

Other comments

Page 12, lines 30-42. Using the rules of thumb of Jaeschke et al. (1994) for interpreting sizes of likelihood ratios a positive likelihood ratio of between 2 and 5
suggests, as the authors note, that their model provides only weak diagnostic evidence for the presence of KD (Kawasaki Disease) and is, therefore, not a very
useful 'rule-in' test (as is also seen in the low specificity) irrespective of the value of the negative likelihood ratio and, indeed, other criteria such as area under the HSROC curve whose high value of 0.86 (page 12, line 45) is rather irrelevant in that it provides 'good' discrimination yet does not change the key clinical conclusion that given the poor diagnostic size of the positive likelihood ratio the (NT-pro)BNP currently cannot be used as a diagnostic tool because we do not know the precise magnitude of its correlation with KD.

Reply:

We thanked the reviewer for his insightful comments. In the discussion, on line 295, we have added the sentence “Unfortunately, according to the rules of thumb by Jaeschke in interpreting sizes of

likelihood ratios, our positive likelihood ratio for NT-proBNP provides only weak evidence for KD diagnosis. Even though NT-proBNP is not sufficient for a standalone rule-in test, the widely available blood test may help clinicians decide whether there is a need to arrange confirmatory cardiac sonographic exam for suspected patients. Cardiac sonographic exam is rarely available in the front-end settings such as emergency department or ambulatory clinics. In addition, the current standard for diagnosis of KD is largely based on empirical clinical criteria, and it will be interesting for others to find out whether incorporating NT-proBNP test results would further enhance the accuracy of current criteria in diagnosing atypical KD."

(As an aside one could, incidentally, if one was mentioning specificity and sensitivity also have mentioned positive and negative predictive values which are often mentioned in addition to sensitivity and specificity).
Reply:
We chose not to present positive or negative predictive value because these values are influenced by disease prevalence in the study population, and could not be compared across different studies. Likelihood ratios could provide similar diagnostic information, and is not influenced by the disease prevalence.

Page 12, lines 42-48. The HSROC curve in Figure 3 is not commented upon and with just five data points (studies) does not look informative and the area under the curve (page 12, line 45) would therefore appear sufficient to give an idea of diagnostic accuracy. The curve unusually has specificity on the x-axis and would benefit from the more standard plotting of 1-specificity on the x-axis so that the origin is in the bottom left corner of the plot.
Reply:
After addition of a new study, we now have six points and a sample size of 1,137, which in our opinion could provide useful information for the diagnosis of this rare disease. STATA is the most commonly used software in the meta-analysis of diagnostic test. The HSROC function in the software package does not allow us to change axis.

One could have motivated the use of HSROC curves by saying they take both within and between study variation into account which make them useful diagnostic tools in meta-analysis. Page 12, line 45. Rules of thumb for acceptable values of the area under the curve are given in Hosmer and Lemeshow (2000) for example who suggest an area under the curve of between 0.80 and 0.90 can be regarded as 'good'.
Reply:
In the results, on line 257, we have added the sentence "To take both within and between study variation into account, we constructed HSROC and forest plot for NT-proBNP, which derived an area under the curve (AUC) of 0.87(95% CI: 0.83-0.89) and a summary OR of 21.6(95% CI:8.33-55.97), respectively (figure 3, 4). According to Hosmer and Lemeshow, our AUC of between 0.80 and 0.90 can be regarded as 'good'."

Page 12, lines 54-57. Any further details of the lack of publication bias such as the results of Eggers test? I would think it would be difficult with just five studies to have a powerful enough test to detect publication bias in any case?
Reply:
In the results, on line 287, we have added, "Results of Eggers tests are presented in table 3, but the small number of studies may prevent a meaningful exam of publication bias. "

Page 15, lines 48-54. The sentence here is grammatically incorrect.

Reply:
On line 361, the sentence has been changed to “we do not recommend these cutoff levels for differentiating between KD and other febrile illness.”

References
Hosmer DW and Lemeshow SL (2000). Applied Logistic Regression. 2nd Edition. Wiley:New York.
Jaeschke R, Guyatt, GH and Saclett DL (1994) Users guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based medicine Working Group. Journal of the American medical Association, 271(9), 703-707.
Reply:
Both of these references have been added as suggested.