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. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Prospective analysis of the accuracy of diagnosis of carpal tunnel syndrome using a web based questionnaire
AUTHORS	Bland, Jeremy; Rudolpher, Stephan; Weller, Peter

VERSION 1 - REVIEW

REVIEWER	Einar Wilder-Smith Neurology, Yong Loo Lin School of Medicine National University Singapore
REVIEW RETURNED	17-Apr-2014

GENERAL COMMENTS	This is a well written clear and concise paper.

REVIEWER	Professor David Warwick University of Southampton]UK
REVIEW RETURNED	08-Jun-2014

GENERAL COMMENTS	**Strong Points** • As listed by the Authors, Prospective design, Large numbers, Objective confirmation of diagnosis using best available current methods, Unselected patient population • It is prudent to examine the accuracy of a questionnaire when presented in different formats • The Authors properly used the patient rather than the hand as the object of study, important when there is a bilateral disease (p5 Line 35) • The ability to tune the questionnaire to sensitivity makes it a more versatile and safer tool **Points to Improve** • Add the correlation statistic to the abstract (line 30)Add the sample size to the abstract so the power of the study is apparent • Rephrase p3, line 20. • Line 57 p4 should read *ethics* rather than *ethical* • The description of Figure 2 (p6, lines 52-57) is not clear to me. There is no vertical axis legend. I do not understand the distribution.

	• Figure 3 is not intuitively clear either. How do the shades on the right match the distribution of CTS severity on the left? If there is a 30+probability of CTS, what does a distribution of 50% severity mean; why does the shading change? I am afraid that for me I do not see the meaning of the data. • The STARD flow diagram has been inserted, as I understand from the Author's disclaimer, as a requirement of the BMJ. The diagram is inappropriate, as the authors properly plea (Page 12, lines 50-56). The BMJ as an academic journal, should lift this requirement which forces an arbitrary and misleading cut off for data which are necessarily continuous rather than categorical. • In order to know whether or not the Reader might subject his patient to the questionnaire, it would be helpful for the paper to give an indication of how long it takes to complete, or at least how many questions there are. **Points that cannot be improved** • It is assumed that neurophysiology studies are a gold standard. A proportion of patients can have CTS but normal studies. The Authors acknowledge this (p 9, lines 48-50). A reference should be inserted here. • Some patients chose not to use the web form; the authors have not shown whether these non-responders may have skewed the utility of the web based version (whereas they have analysed those who filled the questionnaire but did not attend for NCS) • An alternative, and probably more suitable methodology, would have been for each patient to fill both the web based and the paper questionnaire with a consequent an performing an intra-observer correlation and possibly qualitative analysis (which did they find easier etc) , rather than the historical comparison in Figure 1. • The Authors use a historical comparison for Figure 1- how large was the control sample size, how were the data was this validated and where were they published? • Why is the correlation between the website score and the severity od CTS so weak (Figure 3) • The patients had neurophysiology prior to clinical examination. Might it would be more cost effective to use clinical examination and then not perform NCS in those with a very high diagnostic probability of CTS on clinical grounds (a validated approach) rather than perform NCS even in those with a low diagnostic probability, and indeed in those with a very high diagnostic probability). The Authors acknowledge this (3053).p9, lines. • The patients have not had a clinical examination prior to the diagnosis of CTS. This is an inversion of usual clinical medicine in which history and examination precede the diagnostic test. • There are other CTS scores (eg the Boston score). It would

	be helpful to have some indication in the paper to help the Reader understand that there are other schemes and why the the Authors' believe theirs is better. **CONCLUSION** The Authors have presented an interesting web based tool to calculate the probability of neurophysiologically-demonstrable impaired median nerve conduction. The underlying flaw is that carpal tunnel syndrome is, by definition a conglomeration of symptoms and signs; yet there is no clinical examination within this model. NCS are not universally accurate and arguably the gold standard is response to carpal tunnel release. The other issue is that the primary question in the study was to compare a web based version with a paper version (page 2, line 7-8) yet the rest of the paper concentrates on a different outcome, namely a prospective correlation of the web questionnaire and NCS. The only part which reflects the primary objective is Figure 1 which uses an unquantified historical control. The Authors need to refocus the objectives on the methodology.

REVIEWER	David Ring Massachusetts General Hospital
REVIEW RETURNED	11-Jun-2014

GENERAL COMMENTS	1. Methods 56-57: Ethical approval is not needed to collect the data as part of patient care, but it is required to publish it. 2. I don't see diagnostic performance characteristics for the web-based questionnaires. I'd be concerned about and overly sensitive, nonspecific diagnostic instrument. That might lead to unnecessary surgery. 3. I did the questionnaire and it is very long. Also I often wondered what scientific data supported including some of the questions. Many of them seem to relate to risk factors that are at best debatable.

REVIEWER	Linda Chesterton Keele University England UK
REVIEW RETURNED	13-Jun-2014

GENERAL COMMENTS	This is an interesting and large prospective study investigating the

	accuracy of diagnosis of the common condition of CTS via a questionnaire on a publicly available website. It has potential to offer primary care clinicians a useful clinical tool to aid certainty of diagnosis and identify the severity of condition and thus appropriate treatment choices The introduction is focused and clearly identifies current issues for clinicians with the diagnosis of CTS including the potential issues of a lack of agreed gold standard and the use NCS as a comparator in the study. The major concern with reporting this study would be the adequacy of ethical approval. The implications of a lack of ethical submission for this study need further clarification especially in relation to use of patient data and data storage, it is not entirely clear what data is held on the website and what explanations are given to the patients nor are aspects of uncontrolled third party access that are referred to. The authors state that approval was felt unnecessary, however patient data is used for research purposes rather than for clinical benefit to the individual patient. Patient consent has not apparently been provided. Other issues that the authors should address are as follows: • The paper (and especially the abstract) would benefit from greater consistency in reference to the primary objective; i.e. of comparing the probability of CTS when diagnosed via a questionnaire on the web site with results of NC studies. The comparison of results to a previous paper version of the questionnaire are secondary but are (for example) stated as the first result. This is somewhat confusing in the current presentation and distracts from the main theme of the paper. Perhaps consider always referring to this as a secondary issue. • No exclusions due to co-morbidity give potentially generalisable results to the population referred to the unit. Generalising to populations beyond this may not be possible however only 58% completed the online questionnaire. Was the reason for non-completion (of the attenders) determined? Is the web accessible to all? issues of health and IT literacy may be relevant and should be addressed especially in terms of the referral (v study) population demographics – this should be addressed in the discussion. • The authors acknowledge that patients will already have been assessed by their GP prior to referral. In addition to the limitations identified, there may therefore also be an element of recall bias introduced when patients are completing the questionnaire. This factor has not been discussed. • The demographic characteristics of patients are not described • The figures for the results may benefit from greater clarification in terms of how they are constructed and interpreted as they display complex and multiple outcomes. Figure 2 needs a legend • The issue of the weak correlation Pearson r=.03) between the questionniare and severity shown on NCS is not really addressed The authors suggest that patients presenting in general practice with likely CTS should be directed to complete the questionnaire and managed depending on the results. Could the authors suggest at what level they would recommend referral for NCS, or if treatments or referrals to surgery should be targeted to patients with particular

	scores? • A description of how the authors foresee the questionnaire being used within clinical pathways for investigation and management would be interesting. Confirming the requirement for ethical approval for this study is essential

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name Einar Wilder-Smith
Institution and Country Neurology, Yong Loo Lin School of Medicine
National University Singapore
Please state any competing interests or state 'None declared': None declared

This is a well written clear and concise paper.

Reviewer: 2
Reviewer Name Professor David Warwick

Points to Improve
· Add the correlation statistic to the abstract (line 30)Add the sample size to the abstract so the power of the study is apparent (Added)
· Rephrase p3, line 20. (Rephrased)
· Line 57 p4 should read ethics rather than ethical (Changed)
· The description of Figure 2 (p6, lines 52-57) is not clear to me. There is no vertical axis legend. I do not understand the distribution. (Converted to table 1 - hopefully clear)
· Figure 3 is not intuitively clear either. How do the shades on the right match the distribution of CTS severity on the left? If there is a 30+probability of CTS, what does a distribution of 50% severity mean; why does the shading change? I am afraid that for me I do not see the meaning of the data. (Now figure 2 - a detailed explanation of this complex figure has been added to the text)
· The STARD flow diagram has been inserted, as I understand from the Author's disclaimer, as a requirement of the BMJ. The diagram is inappropriate, as the authors properly plea (Page 12, lines 50-56). The BMJ as an academic journal, should lift this requirement which forces an arbitrary and misleading cut off for data which are necessarily continuous rather than categorical. (We all agree about this I think but the diagram is still there for an example cutoff figure)
· In order to know whether or not the Reader might subject his patient to the questionnaire, it would be helpful for the paper to give an indication of how long it takes to complete, or at least how many questions there are. (added to text)

Points that cannot be improved
· It is assumed that neurophysiology studies are a gold standard. A proportion of patients can have CTS but normal studies. The Authors acknowledge this (p 9, lines 48-50). A reference should be inserted here. (reference inserted - we would contend that, though not a gold standard, they are probably no more fallible than ordinary clinical diagnosis)
· Some patients chose not to use the web form; the authors have not shown whether these non-responders may have skewed the utility of the web based version (whereas they have analysed those who filled the questionnaire but did not attend for NCS) (we have also now looked at the characteristics of the patients who did not fill in the form online, apart from being slightly older they are similar in all respects, demographically and clinically)
· An alternative, and probably more suitable methodology, would have been for each patient to fill

both the web based and the paper questionnaire with a consequent an performing an intra-observer correlation and possibly qualitative analysis (which did they find easier etc) , rather than the historical comparison in Figure 1. (See the notes in track changes - this approach too would have had its problems)
· The Authors use a historical comparison for Figure 1- how large was the control sample size, how were the data was this validated and where were they published? (Data and reference added)
· Why is the correlation between the website score and the severity od CTS so weak (Figure 3) (Essentially because the website algorithms are tuned to predict the likelihood of CTS, not it's severity. The fact that there does turn out to be a significant correlation with neurophysiological severity is an incidental finding)
· The patients had neurophysiology prior to clinical examination. Might it would be more cost effective to use clinical examination and then not perform NCS in those with a very high diagnostic probability of CTS on clinical grounds (a validated approach) rather than perform NCS even in those with a low diagnostic probability, and indeed in those with a very high diagnostic probability). The Authors acknowledge this (3053).p9, lines. (This is an interesting topic for debate but would be out of place in this paper)
· The patients have not had a clinical examination prior to the diagnosis of CTS. This is an inversion of usual clinical medicine in which history and examination precede the diagnostic test. (The website questionnaire is the equivalent logically of history taking. Physical examination is rarely helpful in CTS until very late stages or when it indicates the presence of other pathology)
· There are other CTS scores (eg the Boston score). It would be helpful to have some indication in the paper to help the Reader understand that there are other schemes and why the the Authors' believe theirs is better. (We have mentioned some of this and made previous comparisons - see notes in track changes)

CONCLUSION

The Authors have presented an interesting web based tool to calculate the probability of neurophysiologically-demonstrable impaired median nerve conduction.

The underlying flaw is that carpal tunnel syndrome is, by definition a conglomeration of symptoms and signs; yet there is no clinical examination within this model. NCS are not universally accurate and arguably the gold standard is response to carpal tunnel release.

The other issue is that the primary question in the study was to compare a web based version with a paper version (page 2, line 7-8) yet the rest of the paper concentrates on a different outcome, namely a prospective correlation of the web questionnaire and NCS. The only part which reflects the primary objective is Figure 1 which uses an unquantified historical control. The Authors need to refocus the objectives on the methodology. (we tend to see this as only one question - we already knew the extent to which the paper form of the questionnaire was able to predict the diagnosis of CTS. The current study simply repeats the same exercise for the web version to check that the questionnaire does not perform differently when presented in this way. In both papers the analysis is the same, to compare the questionnaire prediction of CTS with the NCS results.)

Reviewer: 3
Reviewer Name David Ring
Institution and Country Massachusetts General Hospital
Please state any competing interests or state 'None declared': None declared.

1. Methods 56-57: Ethical approval is not needed to collect the data as part of patient care, but it is required to publish it. (Now obtained)

2. I don't see diagnostic performance characteristics for the web-based questionnaires. I'd be concerned about and overly sensitive, nonspecific diagnostic instrument. That might lead to unnecessary surgery. (The diagnostic performance characteristics can be adjusted to favour either sensitivity or specificity - an example version is shown in the STARD diagram and other possibilities

can be derived from the ROC curve)

3. I did the questionnaire and it is very long. Also I often wondered what scientific data supported including some of the questions. Many of them seem to relate to risk factors that are at best debatable. (It is long but hard to remove some of the 'low value' questions because of the way the neural network analysis works. There are also some questions in the current online version which are not used in the diagnostic algorithm but are there to serve clinical purposes or to support other studies - remember that this online tool is now a major way in which we collect information from patients attending our carpal tunnel clinic)

Reviewer: 4
Reviewer Name Linda Chesterton
Institution and Country Keele University England UK
Please state any competing interests or state ‘None declared’: None

This is an interesting and large prospective study investigating the accuracy of diagnosis of the common condition of CTS via a questionnaire on a publicly available website. It has potential to offer primary care clinicians a useful clinical tool to aid certainty of diagnosis and identify the severity of condition and thus appropriate treatment choices The introduction is focused and clearly identifies current issues for clinicians with the diagnosis of CTS including the potential issues of a lack of agreed gold standard and the use NCS as a comparator in the study.

The major concern with reporting this study would be the adequacy of ethical approval. The implications of a lack of ethical submission for this study need further clarification especially in relation to use of patient data and data storage, it is not entirely clear what data is held on the website and what explanations are given to the patients nor are aspects of uncontrolled third party access that are referred to. The authors state that approval was felt unnecessary, however patient data is used for research purposes rather than for clinical benefit to the individual patient. Patient consent has not apparently been provided.

(Ethical approval now sought and gained. The mechanism by which data is held on the website anonymously is discussed in the main text)

Other issues that the authors should address are as follows:
• The paper (and especially the abstract) would benefit from greater consistency in reference to the primary objective; i.e. of comparing the probability of CTS when diagnosed via a questionnaire on the web site with results of NC studies. The comparison of results to a previous paper version of the questionnaire are secondary but are (for example) stated as the first result. This is somewhat confusing in the current presentation and distracts from the main theme of the paper. Perhaps consider always referring to this as a secondary issue. (We have re-ordered the presentation in the abstract, but as pointed out above we do not really appreciate the distinction being made here)

• No exclusions due to co-morbidity give potentially generalisable results to the population referred to the unit. Generalising to populations beyond this may not be possible however only 58% completed the online questionnaire. Was the reason for non-completion (of the attenders) determined? Is the web accessible to all? issues of health and IT literacy may be relevant and should be addressed especially in terms of the referral (v study) population demographics – this should be addressed in the discussion. (we have added some demographics and also made some comparisons with patients who did not complete the questionnaire online but came for NCS and completed the paper version during the same period. The only patients now not represented in the analysis are the very small number who were referred by their GP, did not visit the website, and did not turn up for testing either. The patients who completed the questionnaire on peaper rather than online were a little older but the same sex distribution, had an exactly similar rate of CTS diagnosis and similar distribution of NCS severity grade when CTS was present and similar questionnaire diagnostic scores, Boston severity and Boston functional status scores to those who completed the questionnaire online.)

• The authors acknowledge that patients will already have been assessed by their GP prior to referral. In addition to the limitations identified, there may therefore also be an element of recall bias introduced when patients are completing the questionnaire. This factor has not been discussed. (As we have no way of knowing what occurs in a 5 minute GP consultation prior to referral we are not sure that anything can useful be said about this)

• The demographic characteristics of patients are not described (now included)

• The figures for the results may benefit from greater clarification in terms of how they are constructed and interpreted as they display complex and multiple outcomes. Figure 2 needs a legend (Figure altered and amended - see responses to reviewer 2)

• The issue of the weak correlation Pearson r=.03) between the questionniare and severity shown on NCS is not really addressed (See response to reviewer 2)

The authors suggest that patients presenting in general practice with likely CTS should be directed to complete the questionnaire and managed depending on the results. Could the authors suggest at what level they would recommend referral for NCS, or if treatments or referrals to surgery should be targeted to patients with particular scores? (A variety of policies could be adopted but discussing them at length in this paper would make it very long. Locally we would choose to divert patients with scores <20% away from the GP direct access neurophysiology/CTS clinic but would continue to test all the high scoring patients for a mixture of local service, research and clinical reasons, others might choose to take patients with a score of >70% as almost certain CTS with a significant probability of severe disease requiring surgery and fast track these patients to a hand surgery service)

• A description of how the authors foresee the questionnaire being used within clinical pathways for investigation and management would be interesting. (See answer to last point)

Confirming the requirement for ethical approval for this study is essential (Done)