PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Is there sufficient evidence for tuning fork tests in diagnosing fractures? A systematic review
AUTHORS	Mugunthan, Kayalvili; Doust, Jenny; Kurz, Bodo; Glasziou, Paul

VERSION 1 - REVIEW

REVIEWER	Dr. Patrick D. Dissmann, MSc PgD FCEM FFSEM Klinikum Lippe GmbH Zentrale Notaufnahme Germany I am the author of one of the reviewed papers (Dissmann 2006). Other than that I have no competing interests.
REVIEW RETURNED	09-Apr-2014

GENERAL COMMENTS	The following study was presented at a College of Emergency Medicine Conference in 2011: Emerg Med J 2011;28:A2-A3 doi:10.1136/emermed-2011-200617.6 CEM abstracts Day 1: CEM Free Paper Session One: Barbour Room West 12:00-13:30 006 Tuning fork testing on ankle injuries: does it improve the accuracy of the Ottawa ankle rules? A Welling, S Cooke, A Dewey, M Archer, B Higgins, G Carss It would be worthwhile contacting the authors about their results, as they have included 1245 patients in their study, which was conducted across several nurse practitioner-led minor injuries units in Hampshire, UK. I have not seen the data published anywhere as yet, but it would be important to check, whether the authors have published their data anywhere, as their sample size could make an important difference to the overall results.

REVIEWER	Zaynab Jawad NW Thames London UK
REVIEW RETURNED	10-Apr-2014

GENERAL COMMENTS	The authors have put together two methods to analyse the accuracy of the tuning fork test. The results are therefore found to be heterogenous. There is no statistical analysis of the significance of this heterogeniety.

	There are too few studies included. Based on this, there is no evidence that the tuning fork test is unreliable as described in the conclusion.

REVIEWER	Mohsen Kazemi, RN, DC, MSc, FRCCSS(C), FCCPOR(C) Canadian Memorial Chiropractic College, Canada
REVIEW RETURNED	18-Jun-2014

GENERAL COMMENTS	The sensitivity and specificity of 128 and 256 Hz tuning forks in detection of simple acute fractures by Kazemi and Roscoe(2000) are reported erroneously. Kazemi and Roscoe reported the sensitivity and specificity as 86.8 and 50% respectively not 89 and 44% reported by the current authors. 89 and 44% are positive and negative predictive values reported by Kazemi and Roscoe (2000). As such other studies statistics should be checked for correct entry. Secondly, as I looked at the data and as the authors correctly stated in their discussion the sensitivity of the tuning fork test is considered high for a clinical test (75-100%). This indicates that the test is clinically useful in picking up fractures when it is positive on average 82% of the time. Hence it will be very useful for quick decision making on the field and remote areas where diagnostic imaging is not available. It is my suggestion that the authors revise the manuscript to reflect this fact. I would like to commend the authors for taking this great task. With minor revisions as I suggested the manuscript would be acceptable for publication.

VERSION 1 – AUTHOR RESPONSE

1)

Reviewer Name Dr. Patrick D. Dissmann, MSc PgD FCEM FFSEM
Regarding the reviewers's comments on the study presented in Emergency Medicine Conference in 2011, the study at the time of presentation was not completed. Our search on major databases have not revealed any publication results of that study. Our initiation of an electronic contact with the primary researcher was not successful.

2)
Reviewer Name Zaynab Jawad
Regarding the reviewers's comments, we have to mention that we have included greater discussion of the heterogeneity of the studies. The number of studies does not allow us to do any form of sub-group analysis.

3)
Reviewer Name Mohsen Kazemi, RN, DC, MSc, FRCCSS(C), FCCPOR(C)
Regarding the reviewer's first comment, During our analysis it was found that there was an error in calculation of sensitivity and specificity (table 3, pg 5)by Kazemi and Roscoe(2000).
As such we have included the corrected values in our analysis. We noted the reviewer is the primary author of this particular study and we will appreciate he will take the note of it.

Regarding the second comment, "I looked at the data and as the authors correctly stated in their discussion the sensitivity of the tuning fork test is considered high for a clinical test (75-100%). This

indicates that the test is clinically useful in picking up fractures when it is positive on average 82% of the time."
We would like to mention that reviewer has misinterpreted the definition of sensitivity. Sensitivity is not a measure to rule in a disease when positive. Highly sensitive test is deemed effective at ruling out fracture when negative. We mentioned in the discussion that the estimated sensitivity (ranging from 75% to 100%) is not sufficient to be relied upon to rule out fractures based on a negative test.

Regarding Editorial comments we would like to correct that author BK has contributed as follows:
1)Substantial contributions to the design of the work and the acquisition, analysis or interpretation of data.
2)Drafting the work or revising it critically for important intellectual content.
3)Final approval of the version getting published.
4)Agreement to be accountable for all aspects of the work