

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/262189211>

Poor multi-rater reliability in TCM pattern diagnoses and variation in the use of symptoms to obtain a diagnosis

Article in *Acupuncture in Medicine* · May 2014

DOI: 10.1136/acupmed-2013-010473 · Source: PubMed

CITATIONS

9

READS

98

3 authors:



Oddveig Birkeflet

University of Oslo

14 PUBLICATIONS 65 CITATIONS

[SEE PROFILE](#)



Petter Laake

University of Oslo

153 PUBLICATIONS 5,118 CITATIONS

[SEE PROFILE](#)



Nina Vøllestad

University of Oslo

131 PUBLICATIONS 6,396 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



[Qigong View project](#)



[Akupunktur View project](#)

[Journals](#)[Jobs](#)[Education](#)[Decision support](#)[Quality improvement](#)[Community](#)[BMJ](#)

This site uses cookies. By continuing to browse the site you are agreeing to our use of cookies. [Find out more here](#)

Acupuncture in Medicine

Acupuncture in Medicine

aim.bmj.com

Acupunct Med doi:10.1136/acupmed-2013-010473

Original paper

Poor multi-rater reliability in TCM pattern diagnoses and variation in the use of symptoms to obtain a diagnosis

Oddveig Birkeflet¹, Petter Laake², Nina K Vøllestad¹

[+](#) Author Affiliations

Correspondence to

Oddveig Birkeflet, Institute of Health and Society, University of Oslo, N-0318 Oslo, Norway; oddveig.birkeflet@medisin.uio.no

Received 12 October 2013

Accepted 10 April 2014

Published Online First 8 May 2014

Abstract

Background Pattern differentiation and diagnosis are fundamental principles of Traditional Chinese Medicine (TCM). Studies have shown low inter-rater reliability in TCM pattern diagnoses. This variability may originate from both the identification and the interpretation of symptoms and signs.

Objective To examine the inter-rater reliability in TCM pattern diagnoses made in the style of Maciocia for 25 case histories by eight acupuncturists and to explore the impact of demographic factors on the diagnostic conclusion. Further, the association between the diagnosis and the presence of symptoms was examined for a single TCM diagnosis.

Methods Eight acupuncturists independently diagnosed 25 women (15 fertile, 10 infertile) based on written case histories. Descriptive statistics, logistic regression and inter-rater reliability (κ) were used.

Results Poor inter-rater reliability on TCM patterns ($\kappa < 0.20$) and large variation in the number of TCM pattern diagnoses were found. Sex, duration of practice and education had a highly significant effect ($p < 0.001$) on the use of TCM patterns and working hours had a significant effect ($p = 0.029$). There was considerable intra- and inter-rater variation in the use of symptoms to make a diagnosis. Symptoms occurring frequently as well as infrequently were inconsistently used to diagnose Liver Qi Stagnation. The study was limited by a small sample size.

Conclusions The results showed extensive variation and poor inter-rater reliability in TCM diagnoses. Demographic variables influenced the frequency of diagnoses and symptoms were used inconsistently to set a diagnosis. The variability shown could impede individually tailored treatment.

