*Martin Bernstorff | #201307263Supervisor: Anders Foldspang*

*18/10-16*

A comparison of the Edinburgh Postnatal Depression Scale (EPDS) and the Postpartum Depression Screening Scale (PDSS) for peripartum depression screening

# Bachelorprojekt

# Abstract

# Problem Statement

· Does EPDS or PDSS offer the largest area under the curve in a receiver-operating-characteristics-curve?

· What are the trade-offs in deciding on an appropriate cut-off value for each questionnaire in this setting?

# Introduction

The basic background to the question you will work with, ending with a  brief and clear statement of the aim of your work, one aim being better than more aims (!). In this section you may cite individual articles, reviews and other (hopefully) reliable sources (e.g. textbooks). Brevity and clarity are basic virtues.

Major depressive disorder (MDD)

# Methods

Balancing with the aim(s), a description of how the source literature was found and considered for inclusion in the project. PubMed is recommended as the primary database for sourcing original articles.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Criterion** | **Beck et al. (2001)** | | **Zhao et al. (2015)** | |
| **Sample size** | 150 | | 842 | |
| **Inclusion criteria** | · Age ≥ 18  · Able to speak and read English  · 2-12 weeks post-partum  · Delivered a live, healthy infant | | · Obstetric complication | |
| **Gold-standard test** | DSM-IV diagnostic interview | | M.I.N.I. | |
| **Country** | United States | | China | |
| **Language** | English | | Chinese | |
|  | **EPDS** | **PDSS** | **EPDS** | **PDSS** |
| **Cut-off (MDD)** | 12/13 | 79/80 | 12/13 | 79/80 |
| **Cronbach’s α(entire test)** | 0.89 | Not reported | 0.78 | 0.95 |
| **AUC (MDD)** | 0.96\* | 0.98\* | 0.983\*\* | 0.898\*\* |
| **Interviewer** | Nurse psychotherapist | | Not reported | |
| **Blinding** | Yes (interviewer blind to scores) | | No (only high-risk women interviewed) | |

# Findings

*\* EPDS vs. PDSS not statistically significant (p = 0.4)*

*\*\* EPDS vs. PDSS statistically significant (p < 0.001)*

*Focus areas from the aim statement are investigated in depth based upon the findings of original research articles. Keep the strict connection to the aim(s)!*

Blablabl

# Discussion

For a comparison of articles to make sense, the methodology of the articles must be adequately similar.

The articles employ different confirmatory tests. The confirmatory tests have been reported as comparable with a Cohen’s kappa of 0.84 for MDD{Sheehan:1998ua}. This indicates that the confirmatory tests are comparable.

Both the language of the confirmatory test and the screening test differ between the studies.

For the confirmatory tests, the Chinese version of the M.I.N.I. has been shown to correlate with the DSM-IV interview in Chinese{TianMei:2009ww}. However, the comparability of the Chinese and English versions of the DSM-IV interview is not sufficiently accounted for. This weakens a comparison of the studies.

Fortsæt her

For an analysis of a screening tool to be meaningful, a suitable confirmatory test must be used. In the case of depression, the gold standard is a DSM-structured or semi-structured diagnostic interview. Whether this choice is psychometrically valid is outside the scope of this thesis. The interview appears to have sufficient inter-rater reliability with Cohen’s kappas between .7 and 1 for each dimension1,2.

However, every interviewer can affect the interview differently. It is therefore vital that the interviewer is blind to the scores of the screening test and, in the case of multiple interviewers, their concordance is assessed. If the interviewer preferentially diagnoses patients as depressed if they scored highly on one test, the test comparison can be skewed to favour this test.

In Beck et al. the interviewer is a nurse psychotherapist. There is no explicit information on whether multiple therapists are used. The interviewer is blind to the screening results.

In Zhao et al. confirmatory testing is done with the M.I.N.I. by the trained research assistant. The scale is designed for use with minimal training and the selection of interviewer is therefore appropriate. There is no explicit statement of blinding of the interviewer. Therefore, there’s a risk of misclassification. If this misclassification by the interviewer is non-random, i.e. the interviewer exhibits preference towards the results of either questionnaire, it can increase the specificity and sensitivity of this questionnaire, while decreasing the corresponding values for the other questionnaire.

# Conclusion

*Based on the aims, methods, findings and discussion, a very brief summary of the research evaluated in the project as well as an opportunity to suggest future directions for the research area you have analysed – just a few lines with statements – no discussion (!).*

# References

1. Maffei, C. *et al.* Interrater reliability and internal consistency of the structured clinical interview for DSM-IV axis II personality disorders (SCID-II), version 2.0. *J. Pers. Disord.* **11,** 279–284 (1997).

2. What is the relaibility of the SCID-II? *scid.org* Available at: http://www.scid4.org/psychometric/scidII\_reliability.html. (Accessed: 26 October 2016)