

## DOCTORAL THESIS

### Statistical inference for correlated binary data from bilateral studies

Pei, Yanbo

*Date of Award:*  
2009

[Link to publication](#)

#### General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# Statistical Inference for Correlated Binary Data from Bilateral Studies

PEI Yanbo

A thesis submitted in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

Principal Supervisor: Dr. Tang Man Lai

Hong Kong Baptist University

July 2009

# Abstract

In some medical comparative studies, outcomes are bilateral and dichotomous. Inference can be made by either treating subject as the fundamental unit or treating the paired outcomes as independent and ignoring the intra-class correlation information. The latter strategy is usually easier but may lead to inappropriate or wrong conclusions when correlation really exists. Taking the correlation into account can reduce variability and make comparison more precise. In my thesis, we base our statistical inference assuming that there is correlation in the bilateral binary data. In Chapter 2, we consider testing equality of the correlations of two paired binary responses from two treatment groups. Several statistics for testing the equality of correlations and sample size formulae for a two-arm randomized clinical trial are developed. Simulations are performed to evaluate the behaviors of various tests and the accuracy of different sample size formulae in terms of actual size and power. In Chapter 3, we construct and evaluate different confidence interval estimators for correlated proportion differences from bilateral binary data in two-arm randomized clinical trials under the dependence model with equal correlation coefficients between groups. We evaluate the performance of various confidence interval estimators with respect to exact coverage probabilities, exact confidence widths and non-coverage probabilities. In Chapter 4, we focus on model selection issues. We compare the performance of several goodness-of-fit tests for several dependence models used for bilateral binary data in the literature and demonstrate how one may choose an appropriate model in practice. In Chapter 5, we test the equality of the response rates between two treatments using combined unilateral and bilateral data. In particular, we propose several test statistics and investigate their performance. We also apply our proposed methods in each chapter to a real problem using real data sets. Chapter 6 contains conclusions and possible future work in this field.

**Keywords:** Correlated binary data, Paired data, Intra-class correlation, Score test, Wald test, Bootstrap method, Confidence interval estimation, Goodness of fit test, Model selection, Asymptotic tests, Systemic sclerosis, ophthalmology.

# Table of Contents

|   |            |
|---|------------|
| <b>Declaration</b>  | <b>i</b>   |
| <b>Abstract</b>   | <b>ii</b>  |
| <b>Acknowledgements</b>   | <b>iii</b> |
| <b>Table of Contents</b>  | <b>iv</b>  |
| <b>List of Tables</b>   | <b>vii</b> |
| <b>List of Figures</b>  | <b>x</b>   |
| <b>Chapter 1 Introduction</b>   | <b>1</b>   |
| <b>Chapter 2 Testing Equality of Correlations of Two Paired Binary Responses from Two Treated Groups in a Randomized Trial for Scleroderma Patients</b> | <b>6</b>   |
| 2.1 Introduction . . . . .  | 6          |
| 2.2 Testing Problem . . . . .   | 8          |
| 2.2.1 Test Statistics . . . . .   | 10         |
| 2.2.2 Simulation Study . . . . .  | 13         |
| 2.3 Sample Size Determination for Power Achievement . . . . .   | 17         |
| 2.3.1 Approximate Sample Size Formulae . . . . .  | 18         |
| 2.3.2 Simulation Study . . . . .  | 22         |
| 2.4 The Bovine Collagen Clinical Trial . . . . .  | 22         |

|   |   |           |
|---|---|-----------|
| 2.5   | Conclusion . . . . .  | 26        |
| <b>Chapter 3 Confidence Intervals for Correlated Proportion Differences<br/>from Paired Correlated Data</b> |   | <b>28</b> |
| 3.1   | Introduction . . . . .  | 28        |
| 3.2   | Statistical Model . . . . .                                     | 29        |
| 3.3   | Confidence Interval Estimators . . . . .                        | 31        |
| 3.4   | Simulation Study . . . . .                                      | 34        |
| 3.5   | Application to Reanalyze a Scleroderma Clinical Trial . . . . . | 36        |
| 3.6   | Conclusion . . . . .  | 38        |
| <b>Chapter 4 Goodness-of-fit Tests for Correlated Paired Binary Data</b>                                    |   | <b>49</b> |
| 4.1   | Introduction . . . . .  | 49        |
| 4.2   | Alternative Models for Paired Correlated Data . . . . .         | 51        |
| 4.2.1   | Saturated Model . . . . .                                       | 51        |
| 4.2.2   | Rosner's Model . . . . .  | 52        |
| 4.2.3   | Dallal's Model . . . . .  | 53        |
| 4.2.4   | Equal Correlation Coefficients Model . . . . .                  | 53        |
| 4.2.5   | Independence Model . . . . .                                    | 54        |
| 4.3   | Methods for Goodness-of-fit Test . . . . .                      | 55        |
| 4.4   | Simulation Study . . . . .                                      | 58        |
| 4.5   | Real Data Analysis and Model Selection . . . . .                | 59        |
| 4.6   | Conclusion . . . . .  | 61        |
| <b>Chapter 5 Testing the Equality of Two Proportions for Combined Uni-<br/>lateral and Bilateral Data</b>   |   | <b>67</b> |
| 5.1   | Introduction . . . . .  | 67        |
| 5.2   | Data Structure and Dependence Model . . . . .                   | 69        |
| 5.3   | Hypothesis and Asymptotic Tests . . . . .                       | 71        |
| 5.4   | Simulation Study . . . . .                                      | 74        |
| 5.5   | Real Example . . . . .  | 77        |

|   |           |
|---|-----------|
| 5.6 Conclusion . . . . .                    | 77        |
| <b>Chapter 6 Conclusion and Future Work</b> | <b>81</b> |
| <b>Bibliography</b>                         | <b>90</b> |
| <b>Curriculum Vitae</b>                     | <b>95</b> |