Description:

In Ayurveda, body constitution, tongue, and pulse examination all have a long history of use; however, there is still a lack of quantitative data on the reliability of these diagnostic methods. In this context, reliability refers to maintaining consistency of information, a critical factor because consistent diagnoses lead to consistent treatment and are important for clinical practice, education, and research. Thus, the aim of the present PhD study has been to assess the reliability of Ayurvedic diagnostic methods. Cohen's weighted kappa statistic was used as a measure of reliability. Permutation tests were used to test the hypothesis of homogeneous diagnosis (i.e., the doctor's diagnosis does not depend on the subject). The aim of study I was to assess the repeatability of pulse and body constitution assessment. A double-blind controlled study was conducted in Copenhagen. An Ayurvedic expert examined the pulses and body constitutions of 17 healthy participants twice in a random order without seeing them. A matrix of pulse and body constitution variables was developed. Moreover, it was discussed how the magnitude of the weighted kappa statistic may be interpreted using p-values calculated from random permutations of the data. The hypothesis of homogeneous classification was rejected on the 5% significance level (p-values of 0.02 and 0.001, respectively, for pulse and body constitution assessment). According to the LK scale, values of the weighted kappa for pulse examination (K = 0.42) and body constitution assessment (K= 0.65) correspond to "moderate" and "substantial" agreement, respectively. A reasonable level of consistency between the two pulse and body constitution assessments was observed.

The aim of study II was to assess the inter-rater and intra-rater reliability of pulse examination. In study II, 15 registered Ayurvedic doctors with 3-15 years of experience examined the pulses of 20 healthy subjects twice, making a total of 600 examinations. The examinations were performed blind and in a random order. The weighted kappa statistics were negative for two doctors and ranged from 0.03 and 0.56 for the other doctors. Overall, there was very little evidence against the hypothesis of homogeneous diagnosis. The kappa values were generally larger in the group of experienced doctors (p-value 0.04) and course takers. Thus, experience and proper training seem to play a role in pulse examination The aim of study III was to assess the inter-rater reliability of pulse, tongue and body constitution assessment. In this study, the same fifteen Ayurvedic doctors independently interviewed 20 healthy subjects and examined the body constitution, tongue, and pulse of each. Subjects completed self-assessment questionnaires and underwent software analysis for body constitution. Weighted kappa statistics for all 105 pairs of doctors were computed for the pulse, tongue, and body constitution data sets. According to the LK scale, the maximum pairwise kappas ranged from fair to moderate, slight to fair, and poor to slight for body constitution, tongue, and pulse assessment, respectively. For each data set and pair of doctors, the null hypothesis of random rating was rejected for just twelve pairs of doctors for body constitution, one pair of doctors for pulse examination and no pairs of doctors for tongue assessment. There was significant evidence against random software rating and the questionnaire that was used, and the diagnoses made by the majority of doctors. Body constitution assessment appears reliable when a questionnaire and software assessment is used, while other diagnostic methods have room for improvement.

The objective of study IV was to provide information about how the reliability studies can be designed and conducted for Ayurvedic diagnostic methods. In study IV, a review of literature was provided in order to illustrate relevant concepts of reliability studies of diagnostic methods and their implication in practice, education, and training. An introduction to reliability estimates, different study designs, and statistical analysis is given for future studies in Ayurveda.

In conclusion, this is the first study on the reliability of diagnostic methods in Ayurveda. The results showed that there is need for standardization and proper training to improve the reliability of diagnostic methods. The developed bio-statistical methodology might be beneficial for further reliability studies in Ayurveda.