

Titel:

Machine learning applications in the domain of male infertility: development of a prediction model for Klinefelter Syndrome in azoospermic men

Abstract:

Due to successful digitalization of documentation in hospitals over the past years, large sets of patient records labelled with diagnoses are now becoming available. These are facilitating the application of machine learning (ML) algorithm based models to predict conditions of new patients. With the in-house developed database Androbase [1], the Centre of Reproductive Medicine and Andrology (University Hospital Münster) has thousands of records on infertile men available that contain information on e.g. hormone levels, semen analysis, or ultrasound. The underlying reasons for infertility are manifold and range from genetic to endocrine causes. Especially the diagnosis and causes of azoospermia is clinically challenging due to the underlying heterogeneous pathology. ML based prediction models could present an ideal tool to improve diagnose precision and therefore patient treatment. In order to evaluate the potential for identifying the 47XXY karyotype (Klinefelter Syndrome, KS) in azoospermic men, multiple prediction models have been trained and analysed based on a set of different ML algorithms. For this task, a cohort of 1624 azoospermic patients (1239 controls with karyotype 46XY and 385 KS) has been extracted from androbase. Comparison of the models from different ML algorithms regarding sensitivity, specificity, classification error and AUC indicate that support vector machines, random forests, and linear classifiers all perform particularly well on this task. These results show that prediction models might be successfully implemented in the domain of male infertility given a large enough set of retrospective data and well defined diagnoses or conditions.