Statistica Sinica **35** (2025), 1-18
doi:https://doi.org/10.5705/ss.202023.0107

ZIKQ: AN INNOVATIVE CENTILE CHART METHOD FOR UTILIZING NATURAL HISTORY DATA IN RARE DISEASE CLINICAL DEVELOPMENT

Tianying Wang*, Wenfei Zhang and Ying Wei

Colorado State University, Sarepta Therapeutics and Columbia University

Abstract: Utilizing natural history data as external control plays an important role in the clinical development of rare diseases, since placebo groups in double-blind randomization trials may not be available due to ethical reasons and low disease prevalence. This article proposed an innovative approach for utilizing natural history data to support rare disease clinical development by constructing reference centile charts. Due to the deterioration nature of certain rare diseases, the distributions of clinical endpoints can be age-dependent and have an absorbing state of zero, which can result in censored natural history data. Existing methods of reference centile charts can not be directly used in the censored natural history data. Therefore, we propose a new calibrated zero-inflated kernel quantile (ZIKQ) estimation to construct reference centile charts from censored natural history data. Using the application to Duchenne Muscular Dystrophy drug development, we demonstrate that the reference centile charts using the ZIKQ method can be implemented to evaluate treatment efficacy and facilitate a more targeted patient enrollment in rare disease clinical development.

Key words and phrases: Kernel estimation, natural history data, quantile regression, zero-inflated data.

1. Introduction

A rare disease is defined as a disease or condition that affects less than 200,000 persons in the United States, according to Section 526(a)(2)(A) of the Federal Food, Drug, and Cosmetic Act (FD&C Act) [21 USC 360bb] (U.S. Congress, 1934). There are approximately 7,000 recognized rare diseases, cumulatively affecting about 1 in 10 people in the United States. However, most rare diseases do not have approved therapies owing to their complexity and the challenges in clinical development. Most prominently, the gold standard randomization, commonly used in clinical trials (Ingram et al., 1997; Rubenstein et al., 1984), is often unethical and impractical for rare diseases. To determine the treatment efficacy, one has to rely on external controls, which are commonly determined from natural history studies, i.e., preplanned observational studies that "collect

*Corresponding author. E-mail: Tianying.Wang@colostate.edu.