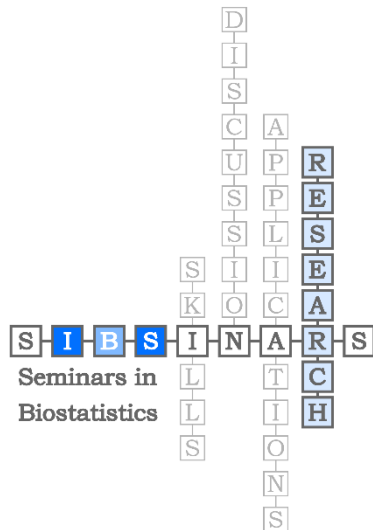




24. April 2026, 13 Uhr

Equivalence tests for assessing consistency of treatment effects in the Cox proportional hazards model

Frank Bretz, Bjoern Bornkamp, Holger Dette, Lukas Koletzko, Zoe Kristin Lange

In confirmatory Phase III clinical trials, researchers and regulators routinely assess whether treatment effects are consistent across important patient subgroups. Such consistency analyses typically rely on interaction tests, although they can only reject the null hypothesis of equality and cannot confirm consistency. Furthermore, interaction tests often have low power but can also be overly sensitive with large samples, leading to statistically significant results that may lack clinical relevance.

We propose a formal framework for equivalence testing to assess the consistency of treatment effects in time-to-event data using the Cox proportional hazards model. Consistency between subgroup and overall treatment effects is assessed using equivalence hypotheses formulated in terms of subgroup-by-treatment interactions. Two testing procedures are developed: a classical two-one-sided test (TOST) and a restricted parametric bootstrap test that controls the type I error while achieving superior power for finite samples. We establish the asymptotic validity of both procedures and compare their performance characteristics through extensive simulations. The methods are illustrated using a real case study.