

Vergleich verschiedener Methoden zur Ersetzung von Fehlenden Werten

Regourd E¹, Tilke C^{1,2}, Schenkel I¹, Jacob G¹, Leverkus F¹

¹Pfizer Pharma GmbH, Karlsruhe

²Contract Statistician, Schöneck

elena.regourd@pfizer.com

Fehlende Werte sind ein bekanntes Problem bei klinischen Prüfungen und Anwendungsbeobachtungen, insbesondere wenn longitudinale Daten erhoben werden. Durch eine sorgfältige Planung kann die Anzahl der fehlenden Werte zwar vermindert, das Problem jedoch nicht eliminiert werden. Die Richtlinien ICH-E9[1] und CPMP-PtC[2] empfehlen, die fehlenden Werte zu ersetzen, um Selektionsbias und Powerverlust zu vermeiden.

Anhand der Daten aus zwei Projekten, einer nichtrandomisierten klinischen Studie und einer Anwendungsbeobachtung, wurden verschiedene Ersetzungsverfahren angewandt und miteinander verglichen. Neben der häufig benutzten LOCF (Last Observation Carry Forward)-Ersetzungstechnik wurden geeignete Multiple Imputation Methoden (SAS[®] Procedure PROC MI) - wie z. B. Markov Chain Monte Carlo (MCMC), Propensity Score, Regression, Predicted Mean Matching - angewandt. Da die meisten von der SAS[®] Prozedur PROC MI verwendbaren Methoden eine monotone Struktur der fehlenden Werte voraussetzen, wurden gängige Ersetzungsverfahren (z. B. LOCF, lineare Interpolation) vorgeschaltet, um die Daten in die notwendige Struktur zu bringen. In einem weiteren Modell wurde die Monotonie-Voraussetzung umgangen, indem die Propensity-Score-Methode sequentiell angewandt wurde [5]. Ferner wurden mehrere Strategien bei Subgruppenanalysen untersucht. Zum Vergleich der verschiedenen Methoden werden sowohl deskriptive Ergebnisse als auch Ergebnisse der geplanten statistischen Analysen dargestellt.

Literatur

- [1] International Conference on Harmonisation: "ICH Topic E9: Statistical Principles for Clinical Trials". September 1998 <http://www.emea.eu.int/pdfs/human/ich/036396en.pdf>
- [2] Committee for Proprietary Medicinal Products: "Points to Consider on Missing Data". November 2001 <http://www.emea.eu.int/pdfs/human/ewp/177699EN.pdf>
- [3] SAS[®] Version 9.1.3 <http://support.sas.com/onlinedoc/913/docMainpage.jsp>
- [4] Fairclough D.L., Design and Analysis of Quality of Life Studies in Clinical Trials, 2002 by Chapman & Hall/CRC
- [5] Li, X; D.V. Mehrotra; J. Barnard: Analysis of Incomplete Longitudinal Binary Data Using Multiple Imputation. Statistics in Medicine, in press