
Proceedings Paper:
Van Laar, M, Greenwood, D, Stark, D et al. (1 more author) (2013) Missing data and survival analysis of central nervous system tumours amongst children and young people in Yorkshire, 1990-2009. In: European journal of cancer. 17th ECCO / 38th ESMO / 32nd ESTRO European Cancer Congress on Reinforcing Multidisciplinarity , 27 September - 1 October 2013, Amsterdam, Netherlands. Elsevier , S305 - S305.

Missing data and survival analysis of central nervous system tumours amongst children and young people in Yorkshire, 1990–2009

M. Van Laar[1], D. Greenwood[2], D. Stark[3], R.G. Feltbower[1].

[1]University of Leeds, Division of Epidemiology, Leeds, United Kingdom;

[2]University of Leeds, Division of Biostatistics, Leeds, United Kingdom;

[3]Leeds Teaching Hospitals NHS Trust, St James Institute of Oncology, Leeds, United Kingdom

Background: Missing data is a common problem in many areas of research, including cancer registration. Severity of disease at diagnosis is often missing from medical records, and as a result this important prognostic variable is sometimes excluded from analysis. We investigated survival trends of central nervous system (CNS) tumours in children and young adults whilst using multiple imputation (MI) to impute missing values of disease severity (grade) and ethnicity.

Materials and Methods: Children and young people (<30 years) diagnosed with CNS tumours (1990–2009) were identified from a population based cancer register in Yorkshire. Logistic regression models were used to impute missing values of grade (WHO Classification, I−IV) and ethnicity (White, Asian and Other), this process was repeated 40 times. Age, sex, year of diagnosis, deprivation, diagnostic subgroup, relapse status and treatment data were all used as predictors of the missing data. Subsequently, a survival analysis was performed using Cox regression models for each set of imputed data, pooled hazard ratios (HR) were then obtained by averaging over all sets of results.

Results: There were 795 cases of CNS tumours diagnosed in Yorkshire between 1990 and 2009. Missing grade data occurred in 71 (8.9%) cases, and missing ethnicity occurred in 191 (24%) cases. Overall, missing data of one or both variables occurred in 242 (30%) cases. Kaplan–Meier survival curves were significantly lower for cases in which grade and ethnicity was missing, compared to those in which it was recorded (P<0.001). This implies that an analysis on only those with complete data (complete case analysis) would produce biased results towards improved survival over the whole cohort. The complete case survival analysis showed no difference in survival between ethnic groups or over time, however, it did show an increased risk of death for those with grade II, III and IV tumours compared to grade I (HR = 3.7, 7.02, 13.36 respectively). After MI, survival analysis showed an increased risk of death for ‘other’ compared to ‘white’ ethnicity (HR = 2.1; P = 0.034), and an increased risk of death for those with grade II, III and IV tumours compared to

grade I (HR = 3.5, 6.4 and 10.4 respectively). Additionally, survival improved significantly by 4% over the study period (P=0.001). Age, sex and deprivation did not significantly affect survival in the complete case or MI analysis. MI reduced standard errors of coefficients by an average of 18% when compared to a complete case analysis.

Conclusion: MI was used to minimise bias and enhance the precision of analyses. This technique offers considerable advantages over other approaches such as performing a complete case analysis. Survival rates varied by grade and ethnicity, and showed a significant improvement over time.

No conflict of interest.