

MISSING DATA

JH Loge^{1,2,3}, M Jordhøy², GI Ringdal², S Kaasa²

¹ University of Oslo,

² Norwegian University for Science and Technology

³ Ullevaal University Hospital

Data may be missing through attrition due to death, missing forms or missing items. In a palliative care trial¹, the number of missing forms and items (IES and EORTC QLQ-C30) were:

	At risk Completed forms IES mis. items
EORTC mis. items	
Base line	434 100% 5.5 % 1.2%
3 months	232 72% 10.7% 1.8%
6 months	149 72% 6% 2.4%

The same data were analysed from death and backwards:
From death

	At risk Completed forms IES mis. Items
EORTC mis. items	
0-30days	365 31% 13 % 5%
31-60days	247 54% 7.5% 1.4%
61-90days	200 68.5% 8.6% 1.5%

Missing forms and missing items were not by random. The same pattern was found for the HRQOL-data: the last scores of patients with missing forms were significantly poorer than the scores of patients with complete forms. Missing data can therefore bias results, potentially very seriously. Strategies to avoid missing data should be implemented in the planning and conduction of studies. Exclusion of cases with missing data assumes similarity between patients with and without missing data. Because they are dissimilar, imputation of data is preferable. Carrying last value forward or simple mean imputation are strategies for imputation of missing forms. Respondent's average score on completed items can be imputed for missing items within multi-item scales. For missing single items, no good method exists. Imputation is done to ease the statistical analysis, and the imputation's effects must be monitored.

¹ Jordhøy et al.. JCO 2001; 19: 3884-94.