Citation:

Franklin, J. (2003). A generalized linear model incorporating measurement error and heterogeneity applied to meta-analysis of published results in Hodgkin's disease. In R. Schulze, H. Holling & D. Böhning (Eds.), *Meta-analysis: New developments and applications in medical and social sciences* (pp. 197-206). Hogrefe & Huber.

13

A Generalized Linear Model Incorporating Measurement Error and Heterogeneity Applied to Meta-Analysis of Published Results in Hodgkin's Disease

Jeremy Franklin

Biometrics Group
First Department of Internal Medicine
University of Cologne

Summary

For a meta-analysis based on comparisons between studies rather than controlled comparisons within studies, it is especially important to explain, estimate, and to allow for the heterogeneity in results between studies. The danger of bias in the use of "historical controls" is well known. The aim of the present investigation was to develop a simple method to analyze differences in results between paediatric and adult clinical trials in Hodgkin's disease, through meta-analysis of a full and extensive collection of published results. Patient and treatment characteristics were included as possible explanatory factors in a generalized linear model. Sampling errors in the Kaplan-Meier estimates derived in the studies as well as heterogeneity between studies were estimated iteratively in order correctly to weight the observations and assess significance while fitting the model and testing effects. A significant superiority of treatment results in children, compared with adults, was demonstrated, allowing for

patient and treatment characteristics. A generalized linear model incorporating heterogeneity and explanatory factors was found to be a practicable and flexible method for a "between-studies" meta-analysis, suitable for investigations where controlled comparisons are not possible or not available.

13.1 INTRODUCTION

The "standard" type of meta-analysis combines the results of several randomized studies, each of which makes the same (randomized) comparison as the meta-analysis. The technique has been extended to non-randomized investigations, for example, of diagnostic methods or prognostic factors (see Chapters 11 and 6 in this volume). Again, such meta-analyses combine comparative information (relative frequencies, correlations, etc.) from each study. The meta-analysis addresses the same question as each component study, its advantage lying in the greater (combined) sample size and therefore power, and in its greater representativity. In practice, many questions and hypotheses in clinical research have not been – or cannot be – investigated within studies. Tentative inferences are then made using comparisons *between* studies, for instance, a historical comparison between a former treatment and a new treatment. Such comparisons are liable to suffer from hidden or unquantifiable biases due to systematic differences between the patients, or their treatment, in the compared studies. Two techniques may help to improve the reliability of between-study comparisons: firstly, avoiding selection bias and "averaging-out" of chance differences by systematic inclusion of a large number of relevant studies; secondly, modeling the influence of known study characteristics to make allowance for biases. This report applies these techniques to the comparison of treatment results in Hodgkin's disease between paediatric and adult institutions.

13.2 OBJECTIVE

In the development and optimization of therapy for malignant lymphoma it has been widely observed during the last decades that paediatric treatment results, as a whole, are superior to those achieved in adults. Since there are systematic differences in treatment strategy between paediatric and non-paediatric institutions (emphasis on combined chemo-radiotherapy even for early stages, lower radiation doses, new and more intensive chemotherapies for children), the reason for this superiority and the role of patient age were unclear. Should the type of therapy rather than age per se be responsible for the good paediatric results, then a rethinking of adult therapy and a borrowing of ideas from children's institutions might be fruitful (Magrath, 1997). The aim of the present analysis was to model the dependence of cure rates in paediatric and non-paediatric Hodgkin's disease patient cohorts on the factors age-range of patients, distribution of disease characteristics in the cohort and type of ther-

apy. The size of differences, if any, in cure rates between children's and adults' cohorts with similar disease characteristics and therapy were to be evaluated.

13.3 METHODS

The publications of studies from which data were to be extracted were selected using a systematic search in the medical literature database Medline (1980–1997) followed by the application of several predefined criteria (first-line treatment, sample size at least 30, chemotherapy as main therapy component, adequate information and follow-up). Pure radiotherapy trials were omitted since children are rarely treated with radiation alone. The patients reported in each paper were, where appropriate and as far as the sample size and the reported details allowed, divided into homogeneous cohorts according to disease characteristics and therapy, avoiding subgroups of less than 30 patients.

Data concerning the type of institution or trial group, sample size, distribution of disease characteristics, type of therapy and Kaplan-Meier estimates of cure rates (disease free survival (DFS) and overall survival (SV) rates) were extracted. Cure rates were adjusted to the time point 5 years after first diagnosis, this adjustment being based on results of a linear regression on pooled data at multiple time points from all those publications where a Kaplan-Meier plot covering an adequate time span was given.

The form of this meta-analysis was a generalized linear model (McCullagh & Nelder, 1989; for some further developments see Nelder, 1998) with cure rate (S) as response variable:

$$\log\left(-\log(E(S))\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k.$$

In order to restrict predicted rates to between 0 and 100%, the complementary log-log link function was chosen to relate response variable S to the linear combination of explanatory variables X and unknown parameters β. The observed response is assumed to vary normally about the expected value with variance comprising two components, namely the sampling error of the Kaplan-Meier estimate σ_P^2 (P = variation between patients) and the heterogeneity σ_C^2 (C = variation between cohorts):

$$S \sim \mathcal{N}\left(E(S);\, \sigma_P^2 + \sigma_C^2\right).$$

Nine potentially explanatory variables were considered, namely:

- type of institution (single-centre, oligocentric, multicentric)
- recruitment period
- sample size
- proportion with advanced disease (stages III and IV)
- proportion with systemic symptoms
- treatment modality (chemotherapy or combined chemo-radiotherapy)

- type of chemotherapy
- number of different drugs
- number of cycles of chemotherapy.

The observations were weighted according to their estimated variance: the standard error of the (iteratively estimated) cure rate according to the formula of Greenwood (1926) plus a second component allowing for the heterogeneity between cohorts. The amount of heterogeneity was estimated iteratively from the sum of squared model residuals, allowing for the contribution due to σ_P^2 (iterative reweighting).

Model fitting was performed in SAS by iterative use of the procedure GENMOD. After each model fitting, the fitted values of cure rate and the residuals were used to estimate the Greenwood standard errors and the heterogeneity respectively. These estimates were combined to estimate the variance of each observation and hence its weight for the next fit, until (after typically 4–5 iterations) the results were stable (Figure 13.1). Using stepwise regression techniques, the effect of including or excluding each explanatory factor was assessed by the change in log-likelihood. Thus, an "optimal" model including only the significant explanatory factors was selected.

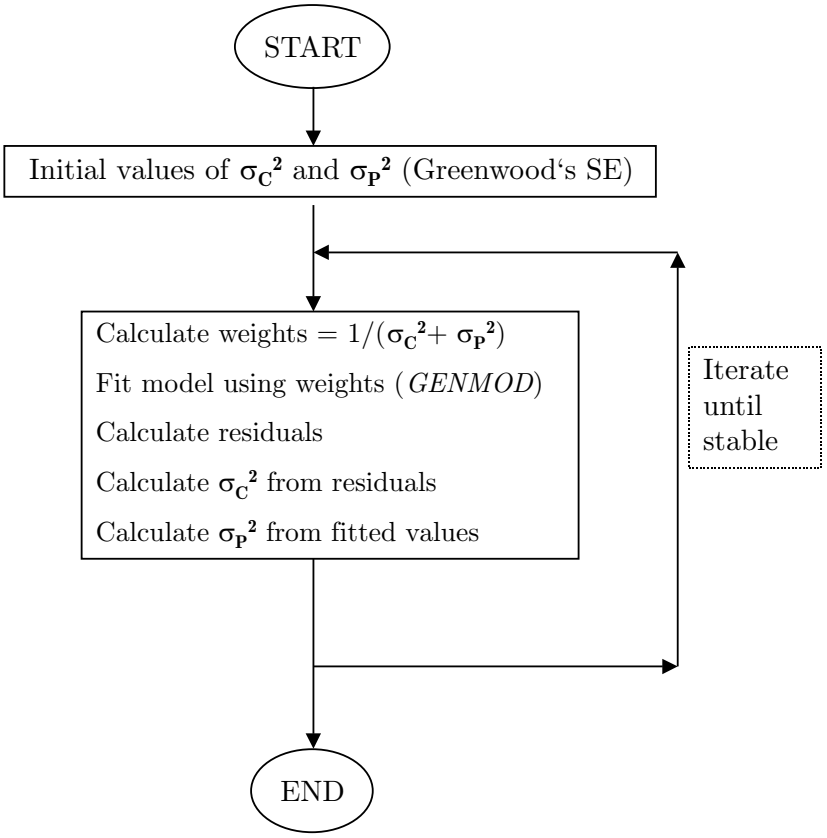

Figure 13.1 Iteratively reweighted fitting of the generalized linear model using the SAS procedure GENMOD.

The sensitivity of the results to small changes in the model (choice of explanatory factors, logistic link function, uniform weighting, weighting proportional to sample size) was investigated. The model was also fitted to several subgroups of the available cohorts (combined modality treatments, pure chemotherapy treatments, particular chemotherapy schemes, early stages, advanced stages, larger cohorts) to assess the generalizability of the results.

In order to correct for the poorer prognosis of the elderly patients who were represented in almost all adult cohorts, the age-specific cure rates of patients in the multicentre German Hodgkin's Lymphoma Study group were analyzed. The effect of older patients in lowering the cure rate of the whole cohort was estimated. Allowing for this effect enabled us to use the meta-analysis results to compare children with younger adults alone.

13.4 RESULTS

Thirty-eight paediatric and 85 adult cohorts were selected for inclusion. The distribution of disease characteristics was similar, on average, in paediatric and non-paediatric cohorts. However, consistent differences in type of cohort and type of therapy were seen (Tables 13.1 and 13.2). Due to the lower incidence of Hodgkin's disease in children, the paediatric cohorts were on average smaller, more often multicentric, and less often randomized. Children more often received combined modality therapy and a lower radiation dose, and therapy more often included a more modern, ABVD[1]-type regimen.

The heterogeneity of both DFS and SV rates between cohorts was significant, according to the Q-statistic of DerSimonian and Laird (1986). For DFS, heterogeneity represented about two-thirds of the residual variation and for SV, about one-third.

Cure rates were consistently better, on average, for paediatric cohorts than for adults. Figures 13.2 – 13.4 show examples of the distribution of paediatric and adult DFS plotted against three of the potential explanatory factors, allowing paediatric and adult results (with respect to the chosen factor) to be compared. This graphical method of comparison is limited to single explanatory factors.

Using the generalized linear modeling technique with all the explanatory factors listed above, highly significant differences of circa 13% in DFS (95% confidence interval: 6–19%) and 12% in survival (95% confidence interval: 8–15%) were found. These differences in cure rates were calculated from the estimated regression coefficients for the factor paediatric/adult via the link function (see above). Figure 13.5 shows estimated differences, which (due to the curvature of the link function) vary according to the level at which the cure rates lie. Alongside the factor paediatric/adult, the following factors were selected as significant for DFS by the stepwise fitting procedure: type of institu-

[1] Adriamycin, Bleomycin, Vinblastine, Dacarbazine

Table 13.1 Characteristics of Paediatric and Adult Cohorts

		Pediatric		Adult	
Type of Cohort	Single Centre	14	37%	50	59%
	2-4 Centres	–	–	7	8%
	Multicentre	24	63%	28	33%
Randomized	No	29	76%	40	47%
Study?	Yes	9	24%	45	53%
Number of	< 40	9	24%	15	18%
Patients	41-60	11	29%	19	22%
	61-100	15	40%	27	32%
	101-200	3	8%	17	20%
	> 200	–	–	7	8%
Total		38		85	

Note. Entries represent number and percentage of cohorts.

Table 13.2 Treatment in Paediatric and Adult Cohorts

		Pediatric		Adult	
Chemotherapy	MOPP or similar	8	21%	42	49%
	ABVD or similar	7	18%	5	6%
	MOPP/ABVD or similar	9	24%	13	15%
	OPPA	9	24%	–	–
	Other	5	13%	25	29%
Number of	2–3	7	20%	17	21%
Chemotherapy	4	6	17%	5	6%
Cycles	5–7	16	44%	40	49%
	8	5	14%	14	17%
	> 8	2	6%	6	7%
Radiotherapy	None	5	13%	26	31%
Fields	Localized	27	71%	18	21%
	Extended	6	16%	33	39%
	Various	–	–	7	8%
Total		38		85	

Note. Entries represent number and percentage of cohorts. MOPP = Mustargen, Vincristine, Procarbacine, Prednisone, ABVD = Adriamycin, Bleomycin, Vinblastine, Dacarbazine, OPPA = Vincristine, Procarbacine, Prednisone, Adriamycin.

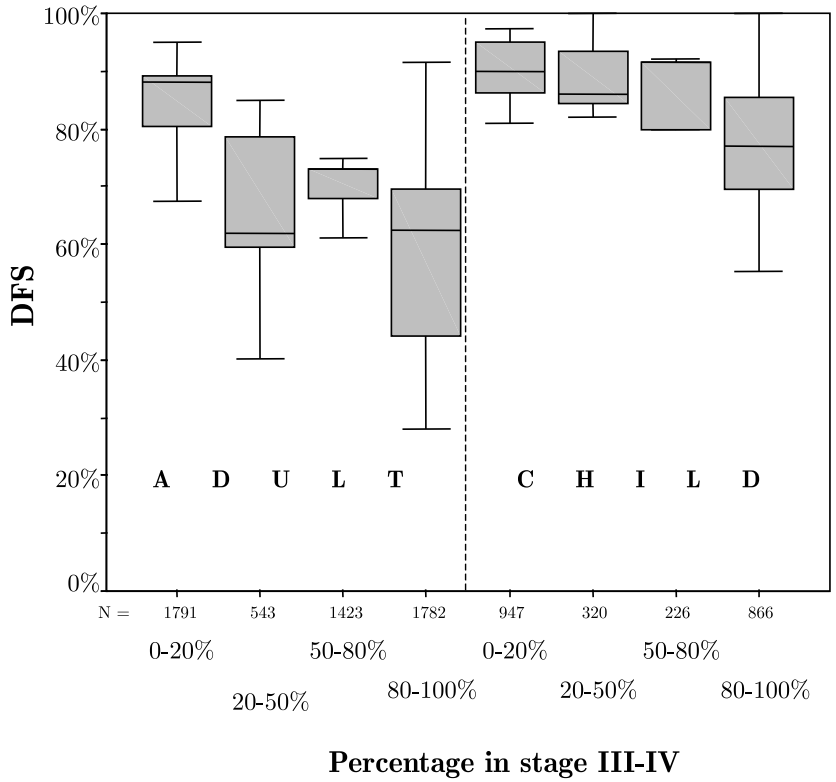

Figure 13.2 Boxplots of (5-year-adjusted) disease free survival rates according to the proportion of advanced stage patients in the cohort, for adult (left) and paediatric (right) cohorts.

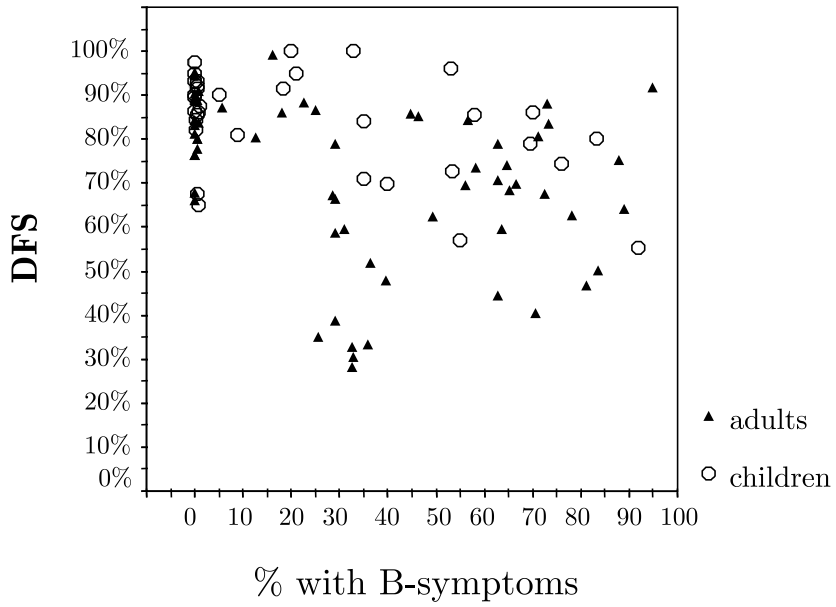

Figure 13.3 Scatterplot of disease free survival (5-year adjusted) against proportion of patients with systemic (B) symptoms in the cohort, for paediatric and adult cohorts.

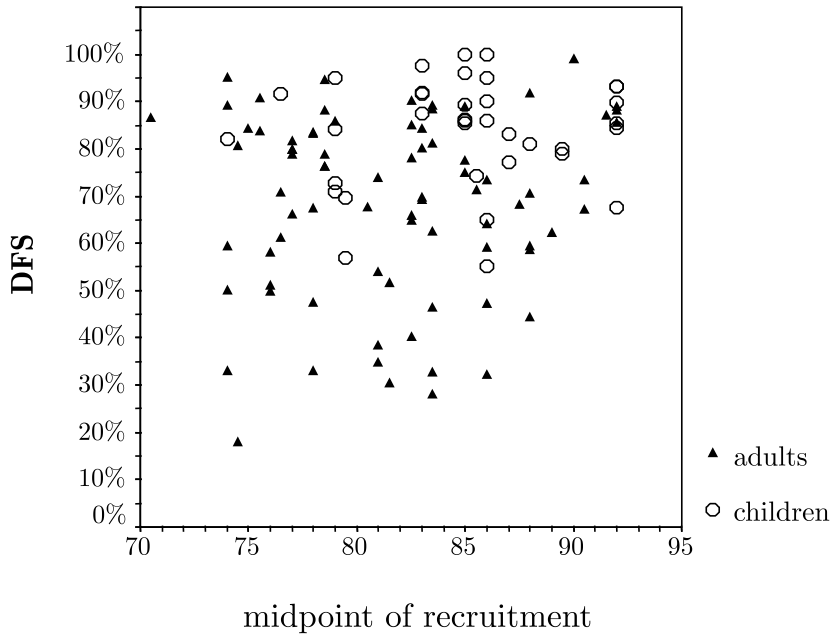

Figure 13.4 Scatterplot of disease free survival (5-year adjusted) against year at midpoint of recruitment period, for paediatric and adult cohorts.

tion, proportion with disease stage III-IV, proportion with systemic symptoms, treatment modality).

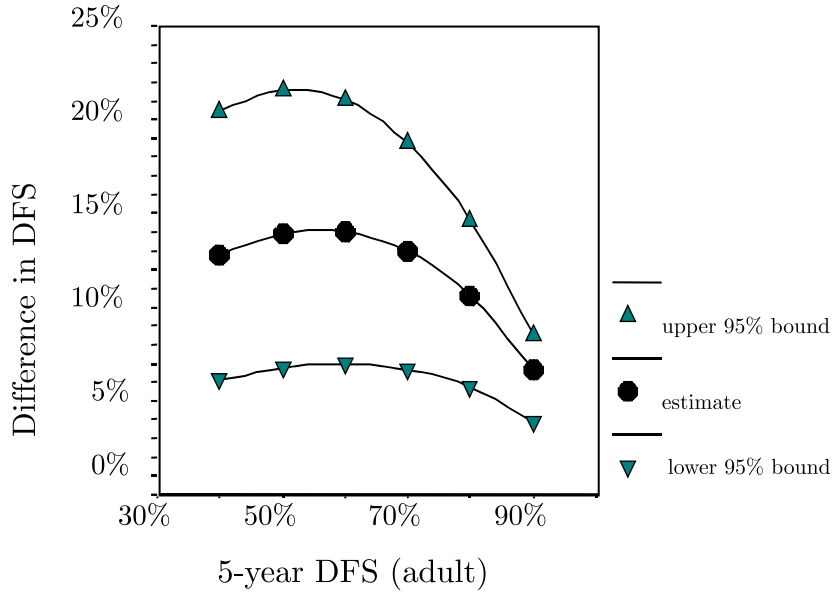

Figure 13.5 Estimated differences in disease free survival rates between paediatric and adult cohorts, calculated using the estimated parameter $\beta = 0.442$ from the generalized linear model.

The paediatric/adult difference was not restricted to certain types of cohorts but reappeared as significant in all main subgroups of cohorts, including the

cohorts of predominantly early stage patients, the cohorts of predominantly advanced stage patients, the cohorts receiving combined chemo-radiotherapy, the cohorts of size over 80, the multicentre cohorts, and so forth. Small variations in modeling methods (weighting scheme, form of link function) did not qualitatively change the results.

The reduction in DFS and SV due to the presence of patients over 45 years old in the adult cohorts was estimated as 3% and 4% respectively. The remainder of 9% in each case therefore represents a difference between children and young adults. This difference could not be accounted for by therapy-related or other factors. It could be due to an intrinsic biological difference or to hidden confounding factors such as quality of care in paediatric institutions.

13.5 CONCLUSIONS

A statistical model for the dependence of treatment results on explanatory factors relating to treating institution, patient cohort and type of therapy was constructed and fitted, with the aim of estimating the difference in treatment results attributable to the age range of the patients (paediatric or adult, respectively).

The generalized linear model allows an appropriate form of dependence and an appropriate specification of error to be incorporated. Heterogeneity between cohorts was an important part of the random variation in both endpoints. The iterative estimation of heterogeneity together with an approximate calculation of the standard error of each cohort-based Kaplan-Meier estimate leads to an error structure which allows for both types of error and thus to a plausible weighting scheme. In the application to Hodgkin's disease, the size of the effect of interest (superiority of paediatric cure rates) could be estimated, although the precision was not high. The results were not sensitive to small changes in modeling methods or inclusion criteria.

A more sophisticated, integrated approach would make use of a maximum likelihood technique to fit a generalized linear model with variance components, the heterogeneity between cohorts being represented by a cohort random effect. Aitkin (1999) lists alternative techniques and proposes a nonparametric approach.

The presence of hidden or non-quantifiable differences which occur systematically between paediatric and adult *cohorts* could influence the results of such an analysis despite the attempt to explain and allow for heterogeneity. The possible influence of such effects should be carefully assessed. In the present analysis, intrinsic differences in curability between children and adults may be confounded with different treatment strategies adopted by paediatric compared with adult institutions. The inclusion of treatment type as a factor in the model may only partially allow for such confounding. Furthermore, it has been suggested that children systematically receive a more thorough staging, treatment administration and care, factors which are not available for inclusion in the model. Thus, the conclusion that paediatric results are superior, for com-

parable patient characteristics and treatment schemes, applies only under the current standards of care and management in paediatric and adult institutions respectively.

The credibility of the results of any meta-analysis, but especially one based on comparisons *between* studies, depends on the unbiased and representative selection of studies for inclusion. In the present analysis, credibility was sought through systematic inclusion of a large number of studies.

REFERENCES

Aitkin, M. (1999). A general maximum likelihood analysis of variance components in generalized linear models. *Biometrics*, *55*, 117–128.

DerSimonian, R., & Laird, N. M. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188.

Greenwood, M. (1926). The errors of sampling in the survivorship tables (Appendix 1). In *Reports on public health and statistical subjects.* London: HMSO.

Magrath, I. T. (1997). The treatment of pediatric lymphomas: Paradigms to plagiarize? (Supplement 1). *Annals of Oncology*, *8*, 7–14.

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). London: Chapman & Hall.

Nelder, J. A. (1998). A large class of models derived from generalized linear models. *Statistics in Medicine*, *17*, 2747–2753.

Acknowledgments

The author is grateful to the chairmen of the German paediatric and adult Hodgkin's disease study groups, respectively Professors G. Schellong and V. Diehl, for motivating the work, and to several colleagues in Cologne for fruitful discussions.