

RESEARCH ARTICLE

Prospective evaluation of quality of life in children treated in UKALL 2003 for acute lymphoblastic leukaemia: A cohort study

C. Eiser¹ | C.B. Stride² | A. Vora³ | N. Goulden⁴ | C. Mitchell⁵ | G. Buck⁶ | M. Adams⁷ | M.E.M. Jenney⁷ | On behalf of the National Cancer Research Institute Childhood Leukaemia Sub-Group and UK Childhood Leukaemia Clinicians Network

¹Department of Psychology, University of Sheffield, Sheffield, United Kingdom

²Department of statistics, Management School, Conduit Road, University of Sheffield, Sheffield, United Kingdom

³Department of Paediatric Haematology, Sheffield Children's NHS Foundation Trust, Western Bank, Sheffield, United Kingdom

⁴Department of paediatric haematology, Great Ormond Street Hospital, London, United Kingdom

⁵Department of Paediatric Haematology/Oncology, John Radcliffe Hospital, Oxford, United Kingdom

⁶Clinical Trial Service Unit, University of Oxford, Oxford, United Kingdom

⁷Department of Paediatric Oncology, Children's Hospital for Wales, Heath Park, Cardiff, United Kingdom

Correspondence

M.E.M. Jenney, Children's Hospital for Wales, Heath Park, Cardiff CF14 4XW, United Kingdom
 Email: Meriel.Jenney@wales.nhs.uk

Grant sponsor: Medical Research Council (UK).

This article was corrected on 14 June 2017 after initial online publication because the copyright line needed to be changed.

Abstract

Background: Health-related quality of life (HRQoL) from diagnosis until end of treatment for children with acute lymphoblastic leukaemia was investigated, examining effects of age, gender, risk-stratified treatment regimen, and therapy intensity (one vs. two 'delayed intensifications' [DIs]).

Method: In a multi-centre prospective study, parents reported their child's generic and disease-specific HRQoL and their own care-giving burden at five time points. From 1,428 eligible patients, 874 parents completed questionnaires at least once during treatment.

Results: At each time point, generic HRQoL was significantly lower than equivalent norm scores for healthy children. HRQoL decreased significantly at the start of treatment, before recovering gradually (but remained below pre-treatment levels). Parents reported that older children worried more about side effects and their appearance, but showed less procedural anxiety than younger children. Concern for appearance was greater among girls than boys. Compared to Regimen B (i.e. additional doxorubicin during induction and additional cyclophosphamide and cytarabine during consolidation chemotherapy), patients receiving Regimen A had fewer problems with pain and nausea. There were no statistically significant differences in HRQoL by number of DI blocks received.

Interpretation: HRQoL is compromised at all stages of treatment, and is partly dependent on age. The findings increase understanding of the impact of therapy on children's HRQoL and parental care-giving burden, and will contribute to the design of future trials.

KEYWORDS

acute leukaemia, paediatric oncology, quality of life

1 | BACKGROUND

Acute lymphoblastic leukaemia (ALL) is the commonest childhood malignancy, affecting approximately 400 children in the United Kingdom annually. Over 80% now achieve long-term cure. Stratification by cytogenetics and minimal residual disease (MRD) identifies a 'low-risk' sub-group of patients,¹ with an excellent chance of cure (>90% 5-year

event-free survival). Those with persistent MRD at the end of induction (high risk) have a greater risk of early relapse.

UKALL 2003 was a randomised clinical trial testing whether the efficacy and toxicity of treatment for children and young adults with ALL could be optimised through MRD stratification. Following induction, sub-groups at low or high risk of relapse predicted by MRD were randomised to treatment reduction or intensification, respectively. Those

Abbreviations: ALL, acute lymphoblastic leukaemia; DI, delayed Intensification; HRQoL, health-related quality of life; MRD, minimal residual disease; QoL, quality of life; T1–T5, time points 1–5

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Pediatric Blood & Cancer* Published by Wiley Periodicals, Inc.

classified as 'MRD low-risk' (undetectable MRD at induction day 29 or detectable $<0.001\%$ leukaemic cells at day 29 becoming undetectable by week 11) were randomly assigned to one (experimental arm) or two (standard therapy) blocks of 'delayed intensification' (DI) chemotherapy prior to maintenance. This additional chemotherapy comprised doxorubicin, etoposide, cyclophosphamide over a 5-day period followed by a period of neutropenia, with blood count recovery by day 21. Patients with MRD $\geq 0.001\%$ at the end of induction were classified as 'MRD high-risk' and were randomised to continue standard therapy (Regimen A or B) or to intensify treatment further (Regimen C). There was an improvement in 5-year event-free survival to 87% in the trial overall, with no increase in relapse risk associated with de-escalation of treatment in the low-risk group² and a reduction in relapse risk in high-risk patients who received more intensive treatment.³

These excellent survival outcomes raise questions about how to balance treatment-related morbidity and health-related quality of life (HRQoL). Physical side effects of chemotherapy,⁴ repeated hospitalisations and associated limitations for social and physical opportunities⁵ compromise a child's HRQoL.

When assessing HRQoL, a distinction is made between generic and disease-specific measures.⁶ Generic measures enable comparison with the general population, while disease-specific measures focus on disease symptoms, and are considered more sensitive to evaluate different treatments. Although ratings of child HRQoL should be made by both child and parent,⁷ it is often necessary to rely on parents' proxy ratings, especially where children are too young or ill to respond themselves. Parents are under intense stress given the emotional and financial costs of caring for a sick child.⁸ Family care-giving burden therefore needs to be considered as an integral part of any comprehensive evaluation of HRQoL.

The aim of this study was to assess HRQoL of children treated in the UKALL 2003 trial from diagnosis until end of treatment. Parent proxy and patient reports (from children aged >8) were collected. Specifically, the following were determined: (i) generic HRQoL compared with population norms; (ii) changes in generic and disease-specific HRQoL over time, depending on child age, gender and treatment regimen; (iii) differences in HRQoL between low-risk patients randomised to treatment reduction and those receiving standard care; (iv) differences in HRQoL between high-risk patients randomised to treatment intensification and those receiving standard care.

2 | METHODS

From October 1, 2003 until June 30, 2011, children and young adults aged 1–25 with ALL (original age limit was 18 years, extended to 20 years in February 2006, and 25 years in August 2007) were recruited from 45 centres across the United Kingdom and Ireland (2), with 3,126 patients eligible for the main trial (ISRCTN 07355119). Inclusion criteria for the HRQoL study were as follows: (i) children were registered on UKALL 2003, and aged between 4 and 18; (ii) a parent was able to complete questionnaires in English and (iii) parents (or patients aged >16) gave signed informed consent. Recruitment to the HRQoL study

ceased ahead of the main trial on October 26, 2009. Ethical approval was granted by the Scottish Multicentre Research Ethics Committee. Clinic staff approached parents of eligible patients in clinic, gave written information and obtained signed consent for participation.

2.1 | UKALL 2003 treatment summary

Patients recruited to UKALL 2003 were initially stratified by clinical risk of relapse. Chemotherapy Regimens A (standard), B (intermediate) and C (high risk) were defined by National Cancer Institute (NCI) criteria, cytogenetics and morphological early response to treatment.² Standard- and intermediate-risk patients were assessed for MRD and randomised on the basis of MRD status at day 29, as described above. All patients received induction chemotherapy with vincristine, dexamethasone, asparaginase and intrathecal methotrexate. Patients on Regimen B received additional doxorubicin during induction (weeks 1–4), and additional cyclophosphamide and cytarabine during consolidation chemotherapy (weeks 7–15). Regimen C comprised two additional cycles of Capizzi maintenance (asparaginase and escalating doses of methotrexate) during weeks 15–22 and 31–38. Only patients receiving Regimens A and B were randomised between one versus two blocks of DI, as noted above. All patients on Regimen C received two DI blocks of chemotherapy. Following the second DI, all patients commenced maintenance chemotherapy comprising daily 6-mercaptopurine, weekly oral methotrexate and 4 weekly pulses of vincristine with 5 days of dexamethasone. Maintenance chemotherapy continued for 2 years in girls and 3 years in boys.²

2.2 | HRQoL study

Questionnaires were completed at five time points (time points 1–5, T1–T5) during scheduled clinic appointments.

T1: As soon as possible after diagnosis, with parents asked to provide a baseline assessment of their child's HRQoL before diagnosis (and care-giving burden immediately after diagnosis); T2: (week 4) end of induction chemotherapy; T3: immediately prior to maintenance therapy (week 23 for patients who received Regimen A and one DI to week 47 for those who received Regimen C); T4: completed at 18 months (during maintenance chemotherapy); T5: end of therapy.

2.3 | Measures and assessment strategy

2.3.1 | Child's HRQoL

The PedsQL 4.0 generic core⁹ is a 23-item scale that yields three scores for each time point (Total, Physical and Psychosocial HRQoL).¹⁰ Five-point response scales (0 = 'never a problem' to 4 = 'almost always a problem') were used for each item. Items are reverse-scored (and linearly transformed to a 0–100 scale) with higher scores indicating better HRQoL.

The PedsQL 3.0 Cancer Module¹¹ is a 27-item questionnaire comprising eight subscales to assess the impact of disease and treatment on pain and hurt, nausea, procedural anxiety, treatment anxiety, worry about side effects, cognitive problems, concern for appearance and communication. In order to ensure relevance to a U.K. sample, the Anglicised version of PedsQL which has been confirmed to be both

valid and reliable in the U.K. population¹² was used. Questions which were also present in the PedsQL generic version were removed from the cancer module in order to reduce the overall length of the questionnaire. This resulted in a 19-item scale assessing pain and hurt, nausea, procedural anxiety, worry about side effects, concern for appearance and communication. Five-point response scales (0 = 'never a problem' to 4 = 'almost always a problem') were used, with higher scores representing worse outcomes.

2.3.2 | Parental care-giving burden

A modified measure of parents' perceived care-giving burden in families with a child with asthma was used.¹³ The resulting scale comprised 11 items asking parents how often they were bothered about specific tasks associated with their child's illness. Responses were made on 7-point Likert scales (1 = 'all of the time' to 7 = 'none of the time'). Scores were reversed so that higher scores indicated greater care-giving burden.

Questionnaires were distributed in paper version to the treatment centres and given to patients by the leukaemia specialist nurses or data management team. Completed questionnaires were sent back to the UKALL 2003 Clinical Trials Centre (Sheffield, UK) and data entered by the trials data manager. The recall period as described in the questionnaire was 'over the last 4 weeks'. Parents answered the same questionnaires at all five time points, except that the Cancer Module was not administered at T1 since items were not relevant before treatment began (see for individual items).

2.3.3 | Statistical analysis

Responses were checked for temporal consistency (that response date fell close to scheduled time point (allowing for treatment schedules at T3, and gender at T5). Responses between 2 weeks before and 6 weeks after scheduled T1 or T2, and between 3 months before and 20 weeks after scheduled T3, T4 or T5 were considered acceptable. Responses outside these ranges were excluded as were those completed after relapse or stem cell transplant. These relatively large ranges were necessary given travel times to hospitals and differences in duration between treatment regimens.

HRQoL, demographics and treatment regimens grouped by response pattern over time were measured in order to compare non-responders with responders and detect bias related to non-response. Phi correlations and chi-square tests were used to determine relationships between non-response at one or more time points and gender, age, initial white blood count or treatment regimen. One-sample *t*-tests were used to compare mean Total, Physical and Psychosocial HRQoL at each time point against population norms.¹⁰

A multilevel growth model for longitudinal data¹⁴ was used to determine level, change over time and variation in change over time for each outcome variable (i.e. generic HRQoL: Total, Physical and Psychosocial subscales; six cancer-specific HRQoL subscales and care-giving burden), and the extent to which any variation was explained by age, gender and treatment. Multilevel modelling was considered preferable to repeated measures ANOVA, since it enabled the use of data from those who did not respond at every time point, and offered

flexibility in modelling intra-subject correlation (i.e. non-independence of observations) across time.

The modelling process comprised three stages for each outcome variable. First, the shape of change over time was determined by introducing response time point as a predictor (unconditional linear growth model). Response time point was treated as a factor due to varying temporal distance between time points and the possible non-linear pattern of change over the study period. The chosen dummy coding of time contrasted each time point against the preceding time point (i.e. repeated contrasts). Second, variables were added as predictors of HRQoL to determine main effects of age at registration (continuous variable), gender (male vs. female) and treatment (regimen received: A, B or C; number of DIs received: 1 vs. 2 and, where relevant, randomised allocation: low-risk 1 vs. 2 DIs, high-risk Regimen A/B vs. C). Lastly, the two-way interaction effects of each of these variables with time point were added to determine whether change in HRQoL over time varied by age, gender, treatment regimen or number of DIs. Analyses of treatment regimen received allowed for age and sex differences in regimens. Analyses of the effect of randomised allocation on HRQoL were restricted to T3–T5 as randomisation took place at T2. Similarly, analyses of the effect of DIs on HRQoL were restricted to T3–T5.

For each model an autoregressive (AR1) correlation structure was fitted to within-subject variance to account for non-independence of questionnaires over time. Correlations detected between adjacent time points were all positive and non-trivial across different outcome measures (range $0.12 < \rho < 0.28$), indicating efficacy of modelling within-subject 'nuisance' variation.

Analyses were conducted using SPSS v21.0 and SAS v9.3 software. Unstandardised regression coefficients (B) and significance levels of $P < 0.05$, $P < 0.005$ and $P < 0.0005$ are reported throughout. We used Bonferroni-corrected *P*-values and confidence intervals when assessing statistical significance.

3 | RESULTS

Internal consistency reliabilities for PedsQL were acceptable at each time point (Physical: $0.89 < \alpha$, T1–T5 < 0.93 ; Psychosocial: $0.79 < \alpha$, T1–T5 < 0.89) and for the cancer module (pain and hurt: $0.77 < \alpha$, T2–T5 < 0.86 ; nausea: $0.79 < \alpha$, T2–T5 < 0.85 ; procedural anxiety: $0.88 < \alpha$, T2–T5 < 0.90 ; worry about side effects: $0.87 < \alpha$, T2–T5 < 0.89 ; concern for appearance: $0.79 < \alpha$, T2–T5 < 0.84 ; communication $0.87 < \alpha$, T2–T5 < 0.90). For the Parental Care-Giving Burden scale, one item ('worried whether or not to ring the hospital') was removed after reliability analysis, leaving a 10-item scale with satisfactory internal consistency reliability at each time point ($0.87 < \alpha$, T1–T5 < 0.92).

Of 1,428 eligible patients, 904 (63%) were enrolled, that is, child or parent responded at least once (Fig. 1), but 11 cases included only child responses, giving 893 parents who responded at least once. After excluding questionnaires completed post-transplant, post-relapse or outside the acceptable response time frame (4–8% at each time point), there were 2,567 eligible questionnaires, from parents of 874 patients, across time points (T1, *N* = 681; T2, *N* = 601; T3, *N* = 526; T4,

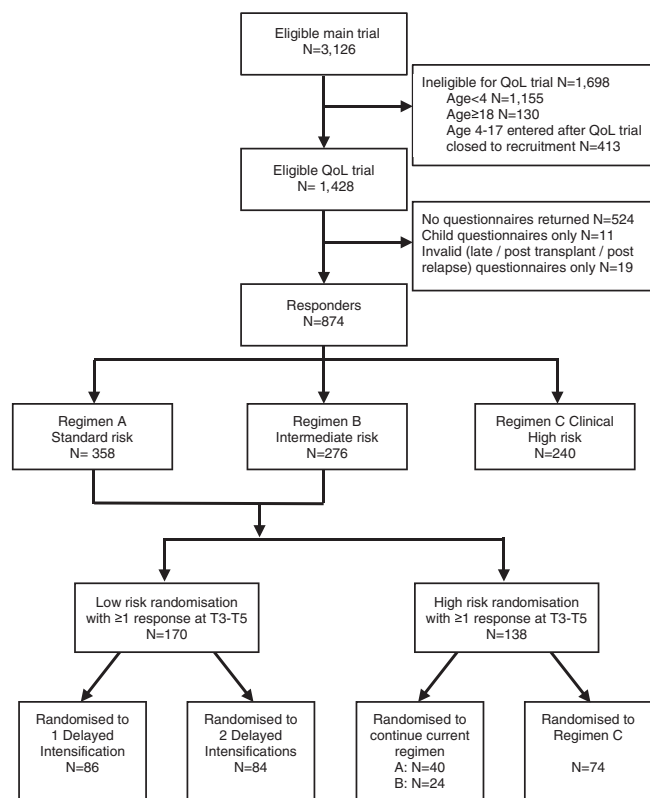


FIGURE 1 Flow chart of eligibility and response in the UKALL 2003 QoL trial

N = 448; T5, N = 311). Parents completed a median of three eligible questionnaires, 152 responded at all five time points. Table 1 outlines the number of responses received and the number of eligible participants at each time point.

Among eligible patients (N = 1,428), there were no demographic or clinical differences between the analysis sample (n = 874), and non-responders (including those who returned invalid questionnaires, n = 554), $P > 0.1$ in each case. The groups had similar demographic profiles (responders: 55% male, median age at registration = 8 years, median white blood cell count $10 \times 10^9/l$; non-responders: 59% male, median age at registration = 8 years, median white cell count $11 \times 10^9/l$) and treatment regimens (responders: 41% received Regimen A, 32% Regimen B, 27% Regimen C; non-responders: 43% Regimen A, 31% Regimen B, 27% Regimen C).

Responses were compared for each of the 10 outcome variables between those who responded close to the due date (within 2 weeks of T1 and T2, and 6 weeks of T3, T4 and T5) with those who responded later, but within the required time frame (numbers were too small to compare those that responded earlier than expected). No differences were found between these groups for any generic or cancer-specific HRQoL subscales except that at T1, later respondents reported higher care-giving burden (mean = 3.23) than early respondents (mean = 2.64, $P < 0.005$) and at T2, later respondents reported less problem with nausea (mean = 0.97) than earlier respondents (mean = 1.41, $P < 0.005$).

Mean scores for HRQoL subscales across time, and comparisons with norms for healthy children are shown in Table 2. At each time

point, Total, Physical and Psychosocial HRQoL scores were significantly lower than norms for healthy children.

3.1 | HRQoL change over time

Physical, Psychosocial and Total HRQoL scales varied across time points ($F = 397.35$, $P < 0.0005$; $F = 216.34$, $P < 0.0005$; $F = 391.37$, $P < 0.0005$, respectively). Comparing scores at T2–T5 against the preceding time point indicated a significant decrease in generic HRQoL on each scale between T1 and T2, followed by recovery between T2 and T3. Changes in generic HRQoL between T3, T4 and T5 were positive and significant, though the degree and rates of change were smaller (Table 2 and Fig. 2A).

Four cancer-specific HRQoL subscales (pain and hurt, procedural anxiety, communicating about illness and worries about side effects) showed a significant reduction between T2 and T3 before levelling between T3 and T5. Nausea increased significantly between T2 and T3 and then declined to T5. Concern for appearance showed no significant change over time (Fig. 2B). Care-giving burden was highest at T1 and T2, followed by a rapid then gradual decrease in scores reflecting a reduction in parental burden of care over time (Fig. 2C).

3.2 | Effects of age, gender and treatment on HRQoL and change over time

There were no significant differences in generic HRQoL or care-giving burden by treatment intensity (Regimen A, B or C). However, there were treatment differences in cancer-specific HRQoL. Parents of patients receiving Regimen A reported fewer problems with pain and nausea than Regimen B ($F = 7.98$, $P < 0.05$; $F = 9.32$, $P < 0.005$, respectively). There were no significant differences between Regimen B or C for any HRQoL variables. Interaction effects between treatment received and time point were significant for just one measure: communicating effects of illness (regimen by time-point interaction $F = 6.55$, $P < 0.05$). There was a reduction in overall problems reported between T2 and T3, but this only reached statistical significance for patients on Regimen A ($P < 0.005$). There were no significant differences in HRQoL or care-giving burden by the number of DIs received. Interaction effects between the number of DIs and time point for generic HRQoL, cancer-specific HRQoL and parental burden of care were likewise non-significant.

There were few gender effects, except that parents of girls consistently reported worse problems with 'concern for appearance' than parents of boys ($F = 29.08$, $P < 0.005$). Effects of gender on change over time (i.e. gender by time-point interaction) were only significant for nausea ($F = 7.33$, $P < 0.005$). There was a significant increase in nausea between T2 and T3 in boys ($P < 0.005$) but the increase was smaller and non-significant in girls. Conversely, nausea decreased significantly between T4 and T5 in girls ($P < 0.005$) but less in boys ($P = NS$).

Parents of older children were more likely to report that their child worried about side effects ($F = 113.14$, $P < 0.005$), and had concerns about appearance ($F = 26.12$, $P < 0.005$), but parents of younger children reported greater child procedural anxiety ($F = 33.29$, $P < 0.005$) especially at T2 (age by time-point interaction, $F = 8.41$, $P < 0.005$; age effect at T2, $P < 0.005$) and lessened with time.

TABLE 1 Number of eligible participants remaining at each time point and number of responses received

Time	Age <10 at QoL/ when QoL due	No questionnaires returned				Achieved remission and eligible at end of time period	No parent QoL but valid child QoL returned (or vice versa)	Questionnaire excluded as dates are out of range*	Questionnaire excluded as date is post-relapse or post-transplant	No. of eligible questionnaires
		Never achieved remission	Relapsed, bone marrow transplant (BMT) or died before end of time period	Last seen date is before end of time period						
Parent										
T1	N/A	5	8	0	687	14	33/714 (5%)	0		681/1,415 (48%)
T2	N/A	6	15	0	752	11	43/644 (7%)	0		601/1,407 (43%)
T3	N/A	6	95	1	761	14	21/547 (4%)	4		526/1,326 (40%)
T4	N/A	6	116	2	800	15	38/486 (8%)	3		448/1,304 (34%)
T5	N/A	6	160	7	900	27	16/327 (5%)	1		311/1,255 (25%)
Any										874
Child										
T1	889	4	5	0	245	32	12	0		241/530 (45%)
T2	886	5	9	0	273	39	18	0		198/528 (38%)
T3	869	5	58	0	271	33	13	1		178/496 (36%)
T4	790	5	78	0	343	36	13	1		162/555 (29%)
T5	659	5	108	2	451	36	27	1		139/656 (21%)
Any										363

BMT, bone marrow transplant; QoL, quality of life.

3.3 | Changes in HRQoL for patients randomised in the MRD low- and high-risk randomisations

Children with MRD low-risk ($n = 521$) were randomised between 1 and 2 DI and children with MRD high-risk ($n = 533$) were randomised between continuing on Regimen A/B and changing to Regimen C. Parents of 170 low-risk randomised children (84 with 2 DI, 86 with 1 DI) and parents of 138 high-risk randomised children (64 from Regimen A/B, 74 from Regimen C) completed HRQoL measures at least once between T3 and T5. There were no statistically significant effects of randomisation on parental care-giving burden or HRQoL, although both reduced subsample sizes and number of relevant time points should be considered when assessing these findings.

4 | DISCUSSION

This is a large prospective study of HRQoL in children and adolescents undergoing treatment for ALL, and it provides important outcome data charting the impact on child HRQoL. Children experience highly compromised HRQoL from diagnosis and up to 2 years later confirming previous cross-sectional findings.¹⁵ This study supports the previous study findings that children have a significant reduction in

HRQoL scores across all domains.¹⁶ One advantage of this study is the prospective design, which allows measurement over the whole treatment course demonstrating changes in HRQoL scores over time. The study also identifies important differences between younger and older children in their reactions.

The study extends previous findings that show that children treated for ALL experience very compromised HRQoL immediately after diagnosis,¹⁵ but HRQoL improves from 3 to 6 months and 1 year after diagnosis.¹⁷ Parent reports of child HRQoL were much lower at T2 (4 weeks after diagnosis) than pre-treatment, and parental care-giving burden levels were at their highest at T1 and T2, confirming this initial period as one of great stress for families.

Few gender differences were identified, except that parents of girls consistently reported more problems with 'concern for appearance'. These concerns among girls may be amenable to direct support and intervention throughout the treatment period. There was greater increase in nausea between T2 and T3 in boys, although there is no clear reason for this.

Parents reported that symptoms such as nausea affected all children regardless of age, but older children were more concerned about side effects and appearance than younger children. Parents reported more procedural anxiety among younger children.

TABLE 2 Parent-reported mean QoL scores (95% CIs) and population norm means (10)

Measure	Norm: healthy	Norm: acutely ill	Time 1, N = 1,088	Time 2, N = 977	Time 3, N = 850	Time 4, N = 779	Time 5, N = 532
Physical	89.32	81.81	74.47*** (73.39, 75.55)	26.78*** (25.08, 28.49)	51.41*** (50.01, 52.82)	56.45*** (55.22, 57.69)	62.69*** (61.15, 64.24)
Psychosocial	86.58	79.56	79.07 (77.53, 80.62)	47.56*** (45.99, 49.14)	62.58*** (60.95, 64.20)	64.14*** (62.58, 65.70)	67.88*** (65.92, 69.83)
Overall	87.61	80.42	78.10*** (77.09, 79.11)	42.23*** (40.48, 43.98)	60.08*** (58.72, 61.45)	62.43*** (61.21, 63.64)	66.77*** (65.27, 68.27)

Significance of difference versus acutely ill children norm: one-sample t-test: * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$. All means are significantly different from healthy children norm at $P < 0.0005$ level.

Parents of children receiving Regimen A were less likely to report procedural pain and nausea than parents of those receiving Regimen B. Related problems in Regimen C patients were more similar to Regimen B. Problems with communicating effects of illness in Regimen B did not change significantly over time, whereas A and C patients had non-significantly more problems at T2, but rates dropped between T2 and T3 and thereafter remained similar to rates seen in Regimen B.

No evidence was found to support the hypothesis that one DI was associated with better HRQoL than two DI, either for the whole group or MRD low-risk patients randomised to 1 versus 2 DI, nor were hypothesised differences between patients with MRD high-risk randomised to Regimen A/B versus C evident. It is possible that the HRQoL measures were not sufficiently sensitive to detect the impact of the different treatment regimens or that there was too much variability round assessment time points, especially in the small numbers randomised in the low- and high-risk randomisations. However, results are consistent with other work suggesting that current treatments have adverse implications for HRQoL¹⁸ and neurological functioning,¹⁹ and these may overwhelm more subtle differences related to treatment.

There were a number of limitations to this study. Firstly, recruitment rates differed between centres, from 94% of eligible patients responding at least once to 18%. In addition, particularly at the later time points, there were a significant number of questionnaires that were not completed and this could adversely affect the validity of the results obtained. These differences may partly reflect availability of research staff between centres. No funding was available to support recruitment. Despite this, no differences in demographic or treatment were found between patients in the study and those not, suggesting that differential recruitment rates do not challenge the integrity of the findings.

One limitation of this study is that the assessment of HRQoL at diagnosis is a retrospective judgement and may therefore be unreliable. However, the aim was to assess the impact of therapy on HRQoL and without this measure the baseline is entirely unknown. Many patients are unwell prior to diagnosis, and this may result in parents underestimating the child's pre-illness HRQoL.

The timing of recruitment around the five data collection time points was variable, especially T3 which was affected by allocated regimen (A, B or C), treatment delays and the number of DIs. T3 and T5 were chosen to represent particular points in the treatment protocol (pre-maintenance and end of treatment), rather than an absolute time point. Variation in response time was greater than anticipated and needs to be addressed in design of future similar trials. However, these data do not suggest that longer treatment for boys compromised their HRQoL more than girls in any measurable way, at least over the period of this study.

The age range of children recruited to the HRQoL study was 4–18 years, although the main trial included younger children and those up to 25 years of age. Difficulties were experienced obtaining responses from children, especially those in the younger age range, and no single HRQoL measure is sensitive across such a broad age range. As a result, younger children were not represented, even though they represent a significant proportion of the ALL population. This is an important

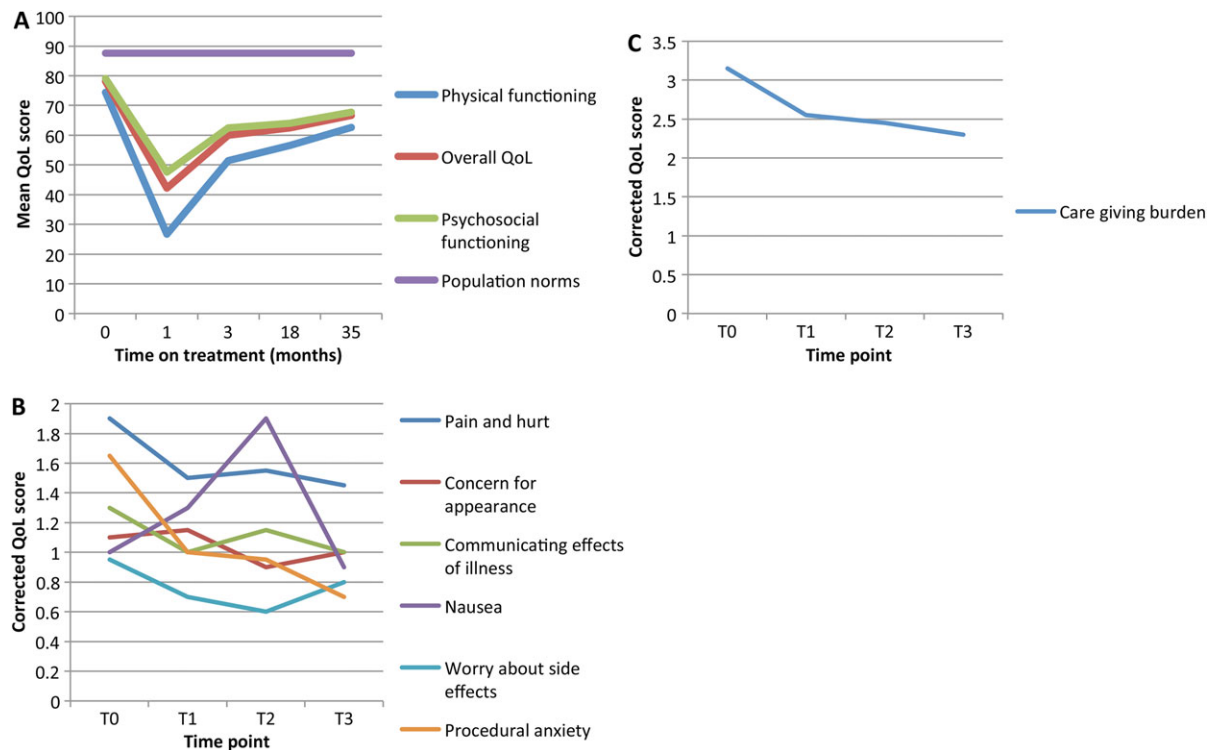


FIGURE 2 (A) Mean child generic (Total, Physical and Psychosocial) HRQoL by time (higher scores = higher functioning/QoL; (B) mean cancer-specific HRQoL subscale scores by time (higher scores = more problems); (C) mean parent's care-giving burden by time (higher score = higher burden)

limitation, as studies have shown that there are clear differences between child self-reports and parent proxy report, particularly in social and emotional domains. There is generally greater concordance between responses to questions regarding physical functioning.^{20,21} In addition, generic HRQoL was compared to U.S. population norms. These norms were chosen given the extensive validation work that has been reported,²² but may be sub-optimal for U.K. populations. The U.S. sample differs in age, sex and ethnic distribution compared to the UKALL 2003 sample, but given the way data were reported, it was impossible to allow for these differences in analysis. However, mean scores in the study population are so much lower than norms that these differences are unlikely to change the conclusions. The measure of care-giving burden was initially developed for work involving families of children with asthma. Although most items were relevant and the scale was acceptable to parents, a more specific care-giving measure might be more sensitive such as that developed by Wells et al.²³

5 | CONCLUSION

This prospective study demonstrates a significant impact of therapy on HRQoL for children receiving treatment for ALL. Excellent survival rates mean that it is possible to reduce treatment intensity for some patients in an attempt to improve quality of life (QoL). It is therefore vital that accurate and sensitive QoL measurement is undertaken.

Further trials are needed to confirm these findings from the perspective of the patient and to determine whether HRQoL can be enhanced by clinical or supportive interventions for patients and their

families. A fuller understanding of the impact of therapy on HRQoL in patients with ALL will make an important contribution to the development of patient-reported outcomes among young people.

ACKNOWLEDGEMENTS

This work was funded by the Medical Research Council (UK). We would like to thank Sue Richards who was responsible for the medical database until her retirement in 2012.

AV, NG, CM designed the clinical trial. CS and GB analysed the data and contributed to writing the report. CE and MJ were responsible for designing the QoL study and writing the report. MA contributed to the interpretation of the data and drafting the final manuscript.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

1. Vrooman LM, Silverman LB. Childhood acute lymphoblastic leukemia: update on prognostic factors. *Curr Opin Pediatr*. 2009;2:1–8.
2. Vora A, Goulden N, Wade R, et al. Treatment reduction for children and young adults with low-risk acute lymphoblastic leukaemia defined by minimal residual disease (UKALL 2003): a randomised controlled trial. *Lancet Oncol*. 2013;14:199–209.
3. Vora AJ, Goulden N, Mitchell CD, Hough R, Rowntree C, Richards SM. UKALL 2003, a randomised trial investigating treatment intensification for children and young adults with minimal residual disease defined high risk acute lymphoblastic leukaemia. *ASH Annu Meet Abstr*. 2012;120:136.

4. De Vries MAG, van Litsenburg RLR, Huisman J, et al. Effect of dexamethasone on quality of life in children with acute lymphoblastic leukaemia: a prospective observational study. *Health Qual Life Outcomes*. 2008;6:103–110.
5. Speyer E, Vuillemain A, Herbinet A, Chastagner P, Briancon S. Effect of adapted physical activity on health-related quality of life among hospitalized children and adolescents (the ACTIV'HOP randomized controlled trial): design and methods. *Contemp Clin Trials*. 2010;31:165–171.
6. Eiser C, Morse R. Quality of life measures in chronic diseases of childhood. *Health Technol Assess*. 2001;5(4):1–157.
7. Eiser C, Varni J. Health related quality of life and symptom reporting: similarities and differences between children and their parents. *Eur J Pediatr*. 2013;172:1299–1304.
8. Eiser C, Upton P. Costs of caring for a child with cancer: a questionnaire study. *Child Care Health Dev*. 2006;33:455–459.
9. Varni J, Seid M, Rode C. The PedsQL: measurement model for the pediatric quality of life inventory. *Med Care*. 1999;37(2):126–139.
10. Varni JW, Seid M, Kurtin PS. PedsQL™ 4.0: reliability and validity of the Pediatric Quality of Life Inventory™ Version 4.0 Generic Core Scales in healthy and patient populations. *Med Care*. 2001;39:800–812.
11. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. *Cancer*. 2002;94:2090–2106.
12. Upton P, Eiser C, Cheung I, et al. Measurement properties of the UK-English version of the Pediatric Quality of Life Inventory 4.0 (PedsQL) Generic Core scales. *Health Qual Life Outcomes*. 2005;3:22.
13. Juniper EF, Guyatt GH, Feeny DH, Ferrie PJ, Griffith LE, Townsend M. Paediatric asthma caregiver's quality of life questionnaire. *Qual Life Res*. 1996;5:27–34.
14. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence. New York: Oxford University Press; 2004.
15. Sung L, Yanofsky R, Klaassen RJ, et al. Quality of life during active treatment for pediatric acute lymphoblastic leukemia. *Int J Cancer*. 2011;128:1213–1220.
16. van Litsenburg RR, Kunst A, Huisman J, Ket JC, Kaspers GJ, Gemke RJ. Health status utilities in pediatrics: a systematic review of acute lymphoblastic leukemia. *Med Decis Making*. 2014;34(1):21–32.
17. Eiser C, Eiser JR, Stride CB. Quality of life in children newly diagnosed with cancer and their mothers. *Health Qual Life Outcomes*. 2005;3:29.
18. Landolt MA, Vollrath M, Niggli FK, Gnehm HE, Sennhauser FH. Health related quality of life in children with newly diagnosed cancer: a one year follow-up study. *Health Qual Life Outcomes*. 2006;4:63–71.
19. Van Litsenburg RR, Huisman J, Hoogerbrugge PM, Egeler RM, Kaspers GJ, Gemke RJ. Impaired sleep affects quality of life in children during maintenance treatment for acute lymphoblastic leukemia: an exploratory study. *Health Qual Life Outcomes*. 2011;18(9):25.
20. Eiser C, Morse R. Quality-of-life measures in chronic diseases of childhood. *Health Technol Assess*. 2001;5(4):1–157.
21. Waber DP, McCabe M, Sebree M, et al. Neuropsychological outcomes of a randomized trial of prednisone versus dexamethasone in acute lymphoblastic leukemia: findings from Dana-Farber Cancer Institute All Consortium Protocol 00-01. *Pediatr Blood Cancer*. 2013;60(11):1785–1791.
22. Varni JW, Burwinkle TM, Seid M, Skarr D. The PedsQL 4.0 as a pediatric population health measure: feasibility, reliability, and validity. *Ambul Pediatr*. 2001;3:329–341.
23. Wells DK, James K, Stewart JL, et al. The care of my child with cancer: a new instrument to measure caregiving demand in parents of children with cancer. *J Pediatr Nurs*. 2002;17:201–210.

How to cite this article: Eiser C, Stride C B, Vora A, et al. Prospective evaluation of quality of life in children treated in UKALL 2003 for acute lymphoblastic leukaemia: A cohort study. *Pediatr Blood Cancer*. 2017;00:e26615. <https://doi.org/10.1002/pbc.26615>