HEALTH RELATED QUALITY OF LIFE IN CHILDHOOD ACUTE MYELOID LEUKEMIA SURVIVORS: A COMPARISON OF MOOD DISTURBANCE OUTCOME IN PATIENTS TREATED WITH CHEMOTHERAPY AND BONE MARROW TRANSPLANTATION

by

Saunders Hsu, MD

A THESIS

Presented to the Department of Public Health and Preventive Medicine
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of
Master of Public Health
May 2007

HEALTH RELATED QUALITY OF LIFE IN CHILDHOOD ACUTE MYELOID LEUKEMIA SURVIVORS: A COMPARISON OF MOOD DISTURBANCE OUTCOME IN PATIENTS TREATED WITH CHEMOTHERAPY AND BONE MARROW TRANSPLANTATION

by

Saunders Hsu, MD

A THESIS

Presented to the Department of Public Health and Preventive Medicine
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of
Master of Public Health
May 2007

Oregon Health & Science University
School of Medicine

CERTIFICATE OF APPROVAL

This is to certify that the Master's thesis of

Saunders C. Hsu, MD

has been approved.

[redacted]

Mentor/Advisor: Donald Austin, MD MPH

[redacted]

Member: H Stacy Nicholson, MD MPH

[redacted]

Member: Tomi Mori, PhD

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Abbreviations

ALL	Acute lymphoid leukemia
Allo-BMT	Allogeneic bone marrow transplant
AML	Acute myeloid leukemia
Auto-BMT	Autologous bone marrow transplant
BMT	Bone marrow transplantation
CCG	Children's Cancer Group
CCS	Childhood cancer survivor
CCSS	Childhood Cancer Survivor Study
CCT	Consolidation chemotherapy
cGVHD	chronic graft versus host disease
CHQ	Child Health Questionnaire
COG	Children's Oncology Group
DFS	Disease free survival
EFS	Event free survival
GVHD	Graft versus host disease
HRQL	Health related quality of life
QOL	Quality of life
OS	Overall survival
POG	Pediatric Oncology Group
POMS	Profile of Mood States
SF-36	Short Form 36
TBI	Total body irradiation
TMD	Total Mood Disturbance
WHO	World Health Organization

ACKNOWLEDGEMENTS

I would like to thank Dr. H. Stacy Nicholson for getting me involved in this study, and allowing me to take it on as my MPH thesis project. Dr. Donald Austin has provided useful insight, and has been most supportive in this endeavor. Thanks to Dr. Tomi Mori for offering her statistical expertise. Tianni Zhou and Nelson Chen at the COG Operations Office were very responsive to my requests and were a great help. Finally, I would like to thank the Kiwanis Clubs of Oregon for supporting me during my studies.

ABSTRACT

Since 1979, children and adolescents enrolled on Children Cancer Group (CCG) clinical trials with a complete HLA-matched sibling were assigned to undergo allogeneic bone marrow transplantation (allo-BMT) for consolidation therapy. Patients who did not have a matched sibling went on to receive consolidation chemotherapy (CCT). Some studies, but not all, have shown that patients treated with allo-BMT have a significantly higher rate of survival compared to consolidation with CCT. However, allo-BMT is also associated with significant long-term side effects unique to transplantation, such as chronic graft versus host disease (cGVHD).

This study will contribute to our understanding of the long-term impact of treatment in survivors of pediatric AML. It is hypothesized that post-induction allo-BMT is associated with a diminished QOL in survivors compared to treatment with CCT. Because the two therapies are marginally different in terms of overall survival, knowledge about quality of life (QOL) outcome may be useful to physicians and families when deciding between the two options.

The research reported herein was part of a multi-dimensional study of HRQL sponsored by the Children's Cancer Group (CCG) titled "Quality of Life Following Successful Therapy for Acute Myelogenous Leukemia: A Comparison of Bone Marrow Transplant and Chemotherapy." One of the instruments was the Profile of Mood States (POMS), a 65-item, adjective rating scale used to measure six mood states (tension/anxiety, depression, anger, confusion, vigor, and fatigue) and calculate the total mood disturbance (TMD) score. The dependent variables were the POMS scores, and the principle independent variable was the treatment variable (allo-BMT or non-allo-BMT). Additional independent variables include the demographic and clinical data.

There were 124 Profile of Mood States (POMS) questionnaires scored; 51 underwent allo-BMT, and 73 received autologous BMT or CCT. No statistically significant differences were found in the POMS scores between the two treatment groups. The subjects' gender was associated with the vigor score, but none of the other POMS scores. Race, age at diagnosis, time elapsed from diagnosis, and history of relapse were not associated with POMS scores. Among those that received allo-BMT, relapse was significantly associated with the outcomes depression, anger, and TMD; cGVDH was significantly associated with the outcomes depression, confusion, and TMD; TBI was not associated with any POMS outcomes. POMS scores in this sample of AML survivors were similar to those in the general population.

Based on this study, it is reassuring to know that childhood AML survivors are not having excess problems with mood disturbance compared to their peers, and there is no statistically significant difference in mood disturbance score whether one received an allo-BMT or not. Therefore, this aspect of HRQL need not factor into the decision process for the type of post-remission treatment a patient receives.

INTRODUCTION

The treatment of childhood cancer is often cited as one of the triumphs of modern medicine. Once a fatal condition in the vast majority of patients, childhood cancer can now be cured in almost 80% of cases. So for many, pediatric cancer is rightfully considered a life-threatening condition with potential long-term consequences rather than a terminal disease. The criteria used for evaluating therapy have most commonly been tumor response, survival, and measurement of side-effects. Increasingly, endpoints that measure quality of life (QOL) are becoming important in assessing the value of medical advances.

Children and adolescents diagnosed with acute myeloid leukemia (AML) are initially treated with intensive chemotherapy ("induction") in order to achieve remission, followed by further treatment to prevent relapse ("consolidation"). Since 1979, children and adolescents enrolled on Children Cancer Group (CCG) clinical trials who had a complete HLA-matched sibling were assigned to undergo allogeneic bone marrow transplantation (allo-BMT) for consolidation therapy. Patients who did not have a matched sibling went on to receive consolidation chemotherapy (CCT). Some studies, but not all, have shown that patients treated with allo-BMT have a significantly higher rate of survival compared to consolidation with CCT. However, allo-BMT is also associated with significant long-term side effects unique to transplantation, such as chronic graft versus host disease (cGVHD).

The chance of cure for a child with acute myeloid leukemia (AML) is still only around 45%. More effective and less toxic treatment is needed. This study will contribute to our understanding of the long-term impact of treatment in survivors of pediatric AML. It is hypothesized that post-induction allo-BMT is associated with a diminished QOL in survivors compared to treatment with CCT. Because the two therapies are marginally different in terms of overall survival, knowledge about QOL outcome may be useful to physicians and families when deciding between the two options.

BACKGOUND

In this section I hope to give the reader sufficient background information to understand the context of this thesis topic. An outline of the background section is provided here:

I. Acute Myeloid Leukemia
    1. Background
    2. Treatment
    3. The BMT vs. chemotherapy debate
II. Cancer treatment late-effects overview
III. Quality of life research
    1. Definitions
    2. Purpose of HRQL
    3. Measurement of HRQL outcomes
        a. Short Form 36
        b. Profile of Mood States
    4. HRQL literature review
        a. Introductory comments
        b. Overview of HRQL research in pediatric oncology
        c. Childhood ALL
        d. BMT recipients
        e. Chemotherapy vs. BMT
        f. Childhood AML

Acute myeloid leukemia

Cancer during infancy and childhood is not common. There are approximately 8700 newly diagnosed cases in the US per year in children 0 to 14 years old, corresponding to approximately 1 in every 7000 children in this age range (Pizzo and Poplack 2002). Acute leukemia is the most common malignancy in childhood, with acute lymphoblastic leukemia (ALL) being the most common type. Acute myeloid leukemia (AML) accounts for approximately 20% of childhood leukemia, and in the US there are approximately 500 new cases diagnosed per year.

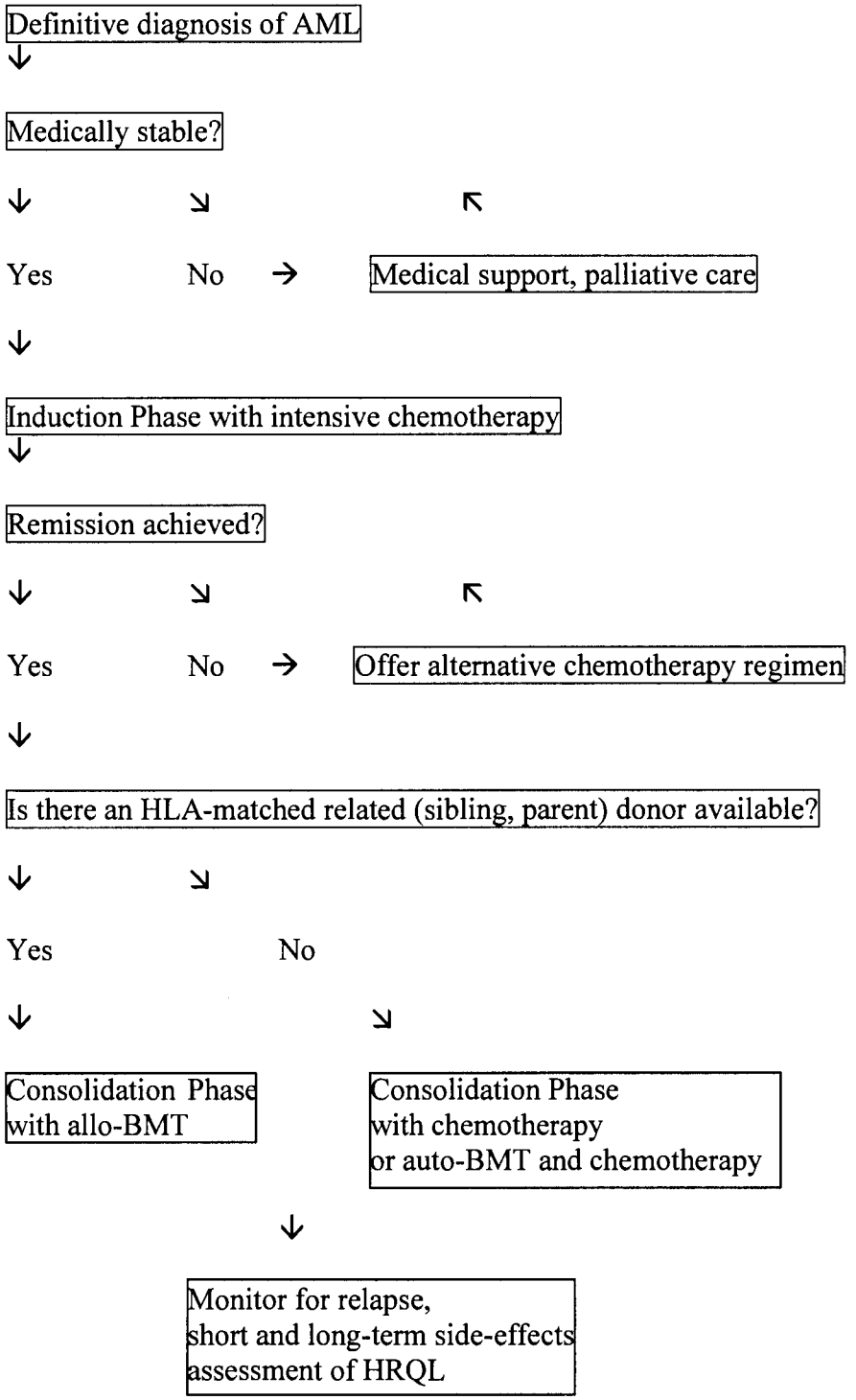
AML refers to the group of hematologic malignancies that arise from the precursors of myeloid, monocyte, erythroid, and megakaryocytic cell lineages. Environmental factors (including ionizing radiation and certain chemical exposures), inherited conditions (e.g. Down syndrome), and bone marrow failure syndromes (e.g. aplastic anemia) predispose one to develop AML. Exposure to certain chemotherapy agents can also lead to secondary AML. However, in the majority of cases, no predisposing factor is identified. The molecular pathogenesis of AML is not completely understood. In many cases, AML is associated with chromosomal translocations, which can result in the generation of novel fusion genes. Transcription factor genes are the most commonly recognized class of genes that are rearranged. Proteins encoded by these genes are important in regulating hematopoietic cell development, and its aberrant expression can lead to leukemic transformation.

The clinical presentation of children with AML can vary widely, from minimal symptoms to life-threatening complications. Fever occurs in about a third of patients. Replacement of the normal bone marrow with leukemia cells can cause bone pain, as well as lead to signs and symptoms related to the decrease in marrow function. Pallor, fatigue, headache, dypsnea, and congestive heart failure can be caused by a decrease in the red cell mass. Bleeding can be the result of decreased platelets or disseminated intravascular coagulation. Other presenting signs and symptoms include hepatosplenomegaly, lymphadenopathy, skin nodules (leukemia cutis), and involvement of the central nervous system.

Treatment

Without treatment, AML is uniformly fatal. Systemic therapy is started as soon as possible after definitive diagnosis is made and the patient is medically stabilized (see Figure 1). The goal for the first part of therapy is to induce a remission, which is defined as a morphologically normal appearing bone marrow, with <5% presence of leukemia cells. Induction therapy consists of the use of multiple chemotherapy agents. Treatment is very intensive and has been associated with significant toxic mortality. Over the years, various induction regimens have been studied to determine which has the best induction and the lowest mortality rate. In addition, the "quality" of the induction may impact the long term survival as well. Following the induction of remission, patients go on to receive further treatment (termed "consolidation," "post-induction," or "post-remission" therapy), consisting of conventional chemotherapy or high-dose chemotherapy ± radiation therapy followed by BMT.

Figure 1. Schematic diagram of AML treatment



Abbreviations: AML, acute myeloid leukemia; allo-BMT, allogeneic bone marrow transplantation; auto-BMT, autologous bone marrow transplantation; HRQL, health-related quality of life

There were two different types of BMT procedures used in the treatment of AML for patients included in this study. One is called allogeneic BMT, the other autologous BMT. In an allo-BMT, the donor is an HLA-matched sibling or relative (matched, related donor). The chance that a full sibling is a match is 1:4. Marrow from unrelated donors was not used in this study. In an auto-BMT, the patients own bone marrow is collected, stored, and re-infused at a later time.

When chemotherapy and/or radiation are given in high doses, the patient's native bone marrow is ablated and will not spontaneously recover in an acceptable period of time. The transplantation of donor (allogeneic) marrow cells, or the re-infusion of autologous marrow cells, allows the patient to recover marrow function after receiving the high dose therapy. Transplant-related complications are a significant cause of death, higher in allo-BMT than auto-BMT. The theoretical advantage of BMT is that it allows the use of a preparative regimen consisting of high doses of chemotherapy ± radiation (to kill leukemia cells) that would otherwise be lethal. In the case of allo-BMT, the patient's native immune system is replaced by that of the donor. Another potential benefit of allo-BMT is that the transplantation of a donor marrow can result in a "graft versus leukemia" (GVL) effect, which helps to eliminate remaining leukemia cells and decrease the chance of recurrence. However, the new immune system can also react to other cells in the recipient's body and cause "graft versus host disease" (GVHD). This condition can lead to long-term side-effects, and can be fatal. GVL and GVHD are not seen in auto-BMT, because marrow is the patient's own, and the immune system in not replaced.

Overall survival for childhood AML in the 1960s was less than 10%, and today, approximately 45-50% can expect to be cured of their disease. As in other areas of pediatric oncology, this improvement has been made possible by clinical trials conducted by cooperative study groups. The subjects in this study were enrolled in one of four Children's Cancer Group (CCG) trials conducted between 1979 and 1995 (Table 1) (Woods, Kobrinsky et al. 1993; Nesbit, Buckley et al. 1994; Wells, Woods et al. 1994; Woods, Neudorf et al. 2001). Research has focused on the role of intensive induction therapy to improve remission rates as well as overall survival, and the role of aggressive post-remission therapy, including BMT (Woods 2006). In these four clinical trials, subjects were "biologically randomized" to what consolidation treatment they would receive. In other words, if the subject had a fully matched sibling or relative, they would be assigned to undergo allo-BMT. Otherwise, they would be assigned to chemotherapy in studies CCG-251, 213, 2861, and 2891, or potentially assigned to chemotherapy followed by auto-BMT in CCG-2861 or 2891 (see Figure 2). These four trials have shown an overall survival (OS) advantage in patients who underwent allo-BMT (Woods 2006). Furthermore, it was also demonstrated that post-remission auto-BMT has no advantage over conventional chemotherapy (Woods, Neudorf et al. 2001).

BMT vs. chemotherapy debate

Despite the evidence of improved OS with allo-BMT in the CCG studies, not everyone in the pediatric oncology community agrees that allo-BMT is superior. Other investigators have reported impressive results utilizing conventional chemotherapy regimens (Creutzig, Ritter et al. 2001), or no survival advantage with allo-BMT compared to

Table 1. CCG AML treatment studies for CCG-L9704

Study Number	Year of enrollment	N	Description of therapy	Survival outcome
251	1979-83	341	**Induction:** Cytarabine/doxorubicin **BMT Conditioning:** TBI/cyclophosphamide **Chemo:** 18 Gy cranial RT; two maintenance regimens were used for up to 3 yrs	5 year OS (1) BMT: 50% Chemo: 35%
213	1986-89	591	**Induction:** Randomized to cytarabine/daunomycin *versus* DCTER (dexamethasone, cytarabine, thioguanine, etoposide, daunomycin) **BMT Conditioning:** TBI/cyclophosphamide **Chemo:** High-dose cytararbine/asparaginase and two courses of PATCO (thioguanine, VCR, cytarabine, cyclophosphamide, azacytidine). Half randomized to continue maintenance with PATCO for 18 additional months.	5 year OS (2) BMT: 54% Chemo: 37%
2861	1988-89	142	**Induction:** DCTER **BMT Conditioning:** Busulfan/cyclophosphamide **Chemotherapy:** Bu/Cy followed by auto-BMT	3 year OS (3) BMT: 55% Auto: 51%
2891	1989-95	537	**Induction:** Randomized to standard *versus* intensive timing DCTER **BMT Conditioning:** Busulfan/cyclophosphamide **Chemotherapy:** Randomized to Bu/Cy followed by auto-BMT *versus* high-dose cytararbine/asparaginase	8 year OS (4) BMT: 60% Auto: 48% Chemo: 53% P<0.05

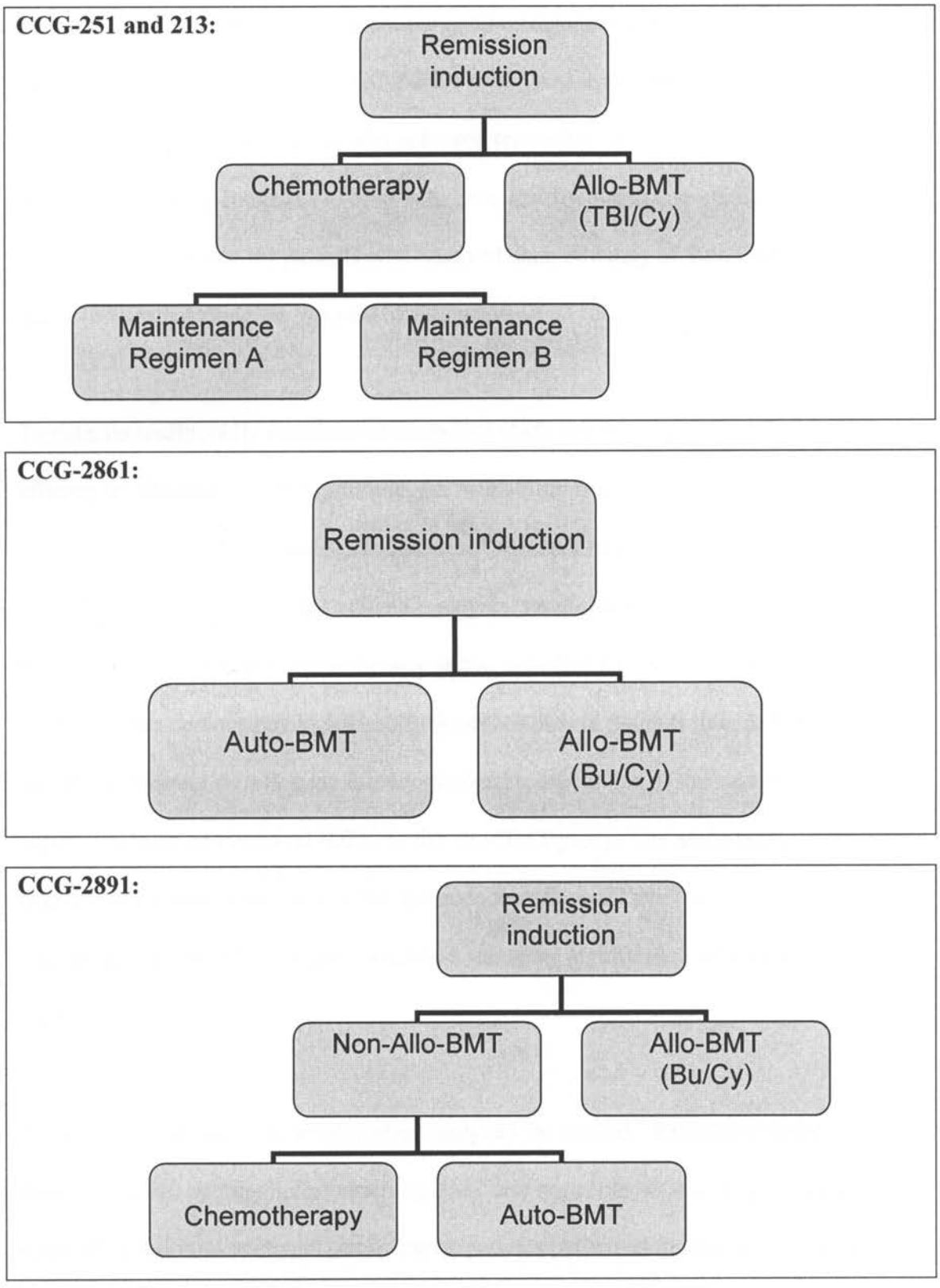
(1) Nesbitt, 1994; (2) Wells, 1994; (3) Woods, 1993; (4) Woods, updated 2006

Table 1. CCG AML treatment studies for CCG-L9704

Study Number	Year of enrollment	N	Description of therapy	Survival outcome
251	1979-83	341	**Induction:** Cytarabine/doxorubicin **BMT Conditioning:** TBI/cyclophosphamide **Chemo:** 18 Gy cranial RT; two maintenance regimens were used for up to 3 yrs	5 year OS (1) BMT: 50% Chemo: 35%
213	1986-89	591	**Induction:** Randomized to cytarabine/daunomycin *versus* DCTER (dexamethasone, cytarabine, thioguanine, etoposide, daunomycin) **BMT Conditioning:** TBI/cyclophosphamide **Chemo:** High-dose cytararbine/asparaginase and two courses of PATCO (thioguanine, VCR, cytarabine, cyclophosphamide, azacytidine). Half randomized to continue maintenance with PATCO for 18 additional months.	5 year OS (2) BMT: 54% Chemo: 37%
2861	1988-89	142	**Induction:** DCTER **BMT Conditioning:** Busulfan/cyclophosphamide **Chemotherapy:** Bu/Cy followed by auto-BMT	3 year OS (3) BMT: 55% Auto: 51%
2891	1989-95	537	**Induction:** Randomized to standard *versus* intensive timing DCTER **BMT Conditioning:** Busulfan/cyclophosphamide **Chemotherapy:** Randomized to Bu/Cy followed by auto-BMT *versus* high-dose cytararbine/asparaginase	8 year OS (4) BMT: 60% Auto: 48% Chemo: 53% P<0.05

(1) Nesbitt, 1994; (2) Wells, 1994; (3) Woods, 1993; (4) Woods, updated 2006

Figure 2. Schematic diagrams of CCG treatment protocols.



chemotherapy (Stevens, Hann et al. 1998). Furthermore, allo-BMT is associated with long-term morbidity and mortality unique to allogeneic transplant recipients. With more than 10 years of follow-up in the CCG-2891 trial, there is still an OS advantage in the allo-BMT group among patients who achieved remission (Woods 2006). However, the survival probability continues to drop with time, mostly due to late-effects from GVHD, while the survival rate for patients who received chemotherapy or auto-BMT has been stable (no deaths) since the 4th year of follow-up.

To date, no traditionally randomized controlled study has been performed to test the efficacy of allo-BMT in AML. Instead, the availability of a HLA-matched sibling donor has been used as a "biological randomization." A recent review discussed the 9 prospective pediatric trials that utilized biological randomization that compared allo-BMT with post-remission chemotherapy and/or auto-BMT (Chen, Alonzo et al. 2002). All the studies demonstrate significantly fewer relapses in patients intended to receive allo-BMT. Despite significantly more toxic deaths, all but one of the studies showed a superior disease free survival (DFS) in the allo-BMT group. The advantage of allo-BMT over chemotherapy is less clear when the endpoint is overall survival (OS) from end of induction, as several studies did not achieve statistical significance when analyzed by intent-to-treat.

Selection bias can alter the results when analyzed "as treated." Patients with the worst disease may fail therapy before reaching BMT and not counted (selecting for patients responding to initial treatment); conversely, patients perceived to have a better prognosis

may not be referred for BMT, or may refuse BMT. To protect against selection bias, results should be analyzed by intent-to-treat, recognizing that this approach reduces a study's power (i.e. true differences between groups are diluted by patients who cross over).

The UK study group compared allo-BMT and chemotherapy, and found a lower rate of relapse with allo-BMT, but no survival advantage (Stevens, Hann et al. 1998). The chemotherapy used was different in this study, potentially negating the benefit of BMT. The UK trial was a smaller study with less power to detect a difference if one truly exists. Furthermore, any comparison between studies must be approached with caution, as there may be different inclusion criteria, definition of endpoints, and/or differences in patient population. Nevertheless, this study has been cited by investigators who caution the use of allo-BMT and/or are in favor of CCT.

In summary, there isn't an international consensus for the "standard of care" in the treatment of AML. This is still an area of active clinical research. Currently, the COG still recommends allo-BMT in first remission if the patient has a HLA-matched sibling.

Late effects of treatment in childhood cancer survivors

Childhood cancer survival rates have improved over the past four decades. Overall, the cure rate now stands at 78% (Institute of Medicine 2003). As a result, there has been an increase in the number of childhood cancer survivors (CCS) reaching adulthood. An estimated 270,000 Americans are CCS, translating to 1 in 640 adults between 20 and 39

(Oeffinger and Hudson 2004). Cured of their cancer, survivors may suffer from late effects from the disease or the treatment given. Survivors have an excess risk of second malignancies, endocrine abnormalities, infertility, neurological complications, and loss of bone density, any of which can adversely affect QOL (Pizzo and Poplack 2002). Approximately two-thirds of CCS report having a late-effect, with one forth having serious conditions resulting from late-effects (Oeffinger and Hudson 2004). Another study reported that over half of survivors had at least one chronic medical problem (Stevens, Mahler et al. 1998).

The Childhood Cancer Survivor Study (CCSS) is an ongoing, longitudinal cohort study that is tracking the outcome of over 20,000 long-term survivors of childhood cancer (Robison, Mertens et al. 2002). Study participants are patients diagnosed with cancer at one of 26 participating institutions between 1970 and 1986, age less than 21 years at the time of diagnosis, and living 5 or more years from the time of diagnosis. This study has made significant contributions to the literature on CSS late effects, and continues to enhance our understanding of the survivor experience.

Late affects can be systematically approached from the vantage point of the type of cancer, the therapeutic exposure, or the organ system affected. The potential late effects of chemotherapy used in AML therapy are provided in Appendix I.

Stem cell transplantation

Preparative regimens for BMT utilize high dose (alkylator) chemotherapy and total body irradiation (TBI). Radiation can exacerbate the toxicity of certain agents. For example, it increases the gonadal and pulmonary toxicity of alkylators, and the cardiac toxicity of anthracyclines. Radiation exposure is mutagenic, and recipients are at increased risk to developing a whole host of other malignancies. Skeletal growth can also be affected.

Unique to allogeneic BMT recipients is graft versus host disease (GVHD). Donor cells recognize the host cells as "foreign" and mount an immunologic attack. GVHD can affect any organ system, and the range of severity spans the spectrum from no symptoms to life-threatening. The most common cause of non-relapse mortality following BMT is due to complications of GVHD.

Psychosocial aspects of survivorship

There are a host of psychological or psychosocial outcomes that have been assessed in CCS, including cognitive functioning, specific neuropsychological functioning (e.g. attentiveness), mood/anxiety, self-esteem, coping, post-traumatic stress, social functioning, educational achievement, marriage and employment rates. While some studies paint a rather bleak picture for CSS in terms of risk of having adjustment issues, post-traumatic stress, and emotional difficulties, others do not show these problems to be more common in CSS compared to the general population. In most of these studies, different instruments were used, and often there were small sample sizes and a heterogeneous mix of cancer diagnoses, time since diagnosis, and age of subjects at

diagnosis when tested, making comparison between studies problematic. (Eiser, Hill et al. 2000).

Among 9535 young adult survivors in the CCSS, 17% had depressive, somatic, or anxious symptoms (Hudson, Mertens et al. 2003). Thirteen percent expressed frequent fears related to their cancer experience. Although almost half stated they had physical impairments and activity limitations, only 10% thought that their health was "fair" or "poor." Zebrack *et al.* compared psychological outcomes in a large sample of childhood leukemia (ALL and AML not distinguished) and lymphoma survivors with sibling controls (Zebrack, Zeltzer et al. 2002). Psychological health status was assessed using a 20 item questionnaire selected from the Brief Symptom Inventory. They found that although survivors were 1.6 times more likely than controls to report symptomatic levels of depression and somatic distress, the actual rate of symptomatic levels of depression (5.4%) and distress (12.7%) was not elevated compared to rates in the general population, allowing the authors to conclude that the majority of childhood cancer survivors are psychologically healthy. In multivariate analysis, intensive chemotherapy (as defined in their study) was significantly associated with an increased risk for symptomatic depression and distress. Subjects with AML in their cohort were considered to have received intensive treatment; BMT recipients were not separately analyzed.

Quality of life research in childhood cancer survivors

Quality of life (QOL) research is subset of late effect research. By its nature, QOL outcomes are a function of the various late effect outcomes. However, documenting late

effects alone does not fully describe QOL, as QOL specifically refers to the impact of these late effects *as perceived by the survivor*. Therefore, in order to understand QOL, it must be studied as a distinct entity. Traditionally, health research garnering the most attention has been disease pathology, morbidity/mortality, and functional impairment. It is increasingly accepted that this offers a limited view of overall life satisfaction. Today, there is general agreement that in addition to prolonging life, QOL as an important and valid outcome of medical therapy to study. However, challenges and obstacles in studying QOL are significant, in part due to its inherently subjective nature, and also due to the mix of approaches and definitions that have been used in the past. Progress in this field will be achieved by adopting standard methods as well as continued innovation.

Quality of life: definitions and terminology

The phrase "quality of life" is in common use, and each person defines QOL for him or herself for the various aspects of ones life. For these reasons, it isn't surprising that QOL is defined many different ways in the literature. The World Health Organization (WHO) defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO 1948) and QOL as "the individual's perception of their position in life in the context of the cultural and value systems in which they live and in relation to their goals, expectations, standards, and concerns" (WHO 1993). QOL can be categorized as health- and non-health related. Health related QOL refers to aspects of life that can be influenced by the health care system. It has been used interchangeably (although not always correctly) with health status, functional status, physical functioning, perceived health status, subjective health, health perceptions,

symptoms, or physical disability (Sobo and Kurtin 2003). Examples of non-health related QOL include environmental factors such as climate, social and economic factors, and personal attributes such as coping skills. There are many areas of overlap. For example, unemployment may restrict one's ability to comply with medical advice or obtain a nutritious diet, or lack of clean drinking water increases the risk of diarrhea illness. For the purposes of this thesis, only HRQL will be considered, and the terms HRQL and QOL will be used interchangeably.

In the context of the WHO definition of health, it follows that the HRQL is composed of different parts, or domains. The most commonly cited domains are physical and occupational functioning, psychological state, social interaction, and somatic sensation. Physical and occupational functioning refers to questions about strength and energy, and the ability to carry out various tasks. An assessment of psychological state includes the presence or absence of anxiety, depression, or fear. Social interaction refers to one's ability to conduct this function among different groups, such as family, close friends, colleagues, and the general community. Somatic sensation encompasses pain, nausea, or other unpleasant physical feelings that detract from quality of life. These four domains do not cover the total spectrum of HRQL, but are the ones most commonly used in research.

Purpose of measuring HRQL

The three primary purposes for measuring HRQL are discrimination, evaluation and prediction (Feeny, Furlong et al. 1999). When one uses an HRQL measure for discrimination, the purpose is to detect differences in morbidity among groups or

individuals at a point in time (e.g. a cross-sectional study comparing two groups). Evaluation involves the assessment of change in HRQL within individuals over time (e.g. longitudinal study). Prediction involves the use of an HRQL measure to predict another outcome at some point in time.

HRQL measures can be used to identify those who should be targeted for interventions, aid in decision making for individual patients, stratify groups of patients for treatment based on the prognostic significance of an HRQL score, and potentially help choose among alternative therapies.

Measuring HRQL

While there is broad agreement that HRQL is an important and useful clinical outcome for patients, clinicians, and health administrators alike, it is meaningful only if it can be accurately and reliably measured. Unlike an outcome such as "disease-free survival" or "infertility," HRQL is a hypothetical construct, and has no widely agreed upon definition or measure (Koot and Wallander 1991). A discussion of its definition was presented above. The *operational* definition of HRQL is one that is inferred from its measures—i.e. those instruments designed to be meaningful and tangible indices of the underlying construct. The most important qualities of a measure are its reliability and validity. The former refers to how consistently an instrument measures something within the same individual, and among similar individuals (recognizing that by its very nature, HRQL is subjective and varies over time). The latter refers to how well the instrument measures what it claims to be measuring, and is the core of any measurement. The process of

establishing a valid instrument first involves the articulation of the theoretical concepts and their interpretations, followed by the development of instruments to measure this construct, and lastly by empirical testing and evaluation of the relationship between theory and what was observed.

Medical Outcomes Study Short Form 36 Health Survey (SF-36)

The SF-36 is one of the most widely used instruments in studying HRQL (Ware and Gandek 1998). It was in fact one of the first a measurement tools developed for HRQL research. It consists of 36 questions, used to score eight health concepts. These health concepts were selected because they represented concepts previously shown in other health surveys to be most frequently affected by disease and treatment. These are physical functioning, role-physical, bodily pain, general health, vitality, social functioning, role-emotional, and mental health. The first four concepts are used to form a summary measure of physical health, and the latter four to form the summary measure of mental health.

Despite the widespread use of the SF-36, there is only one published study of its use in adult survivors of childhood cancer (Veenstra, Sprangers et al. 2000). This study from the Netherlands compared the HRQL outcome in bone tumor survivors compared to published population norms. They found that survivors' physical functioning was worse than healthy peers, but better than those with chronic illness.

The Child Health Questionnaire measures the same domains as the SF-36, and consists of a child-completed version for older children and a parent completed version for younger children.

HRQL outcome: Psychological functioning

Profile of Mood States (POMS) is a 65-item, adjective rating scale used to measure six mood states: tension/anxiety, depression, anger, confusion, vigor, and fatigue. A copy of the instrument is provided in Appendix II. It has been used extensively, and has a high reliability and validity (McNair and Heuchert 2005). Briefly, six analytic studies were conducted in the development and validation of the POMS. The test groups were male psychiatric patients, male college students, and male and female outpatient clinic patients.

Subjects are asked to describe, on a scale of 0 ("not at all") to 4 ("extremely"), the extent to which the adjectives describe the way they have been feeling over a specified period (e.g. past week, right now). Except for vigor, higher scores indicate greater mood disturbance. A total mood disturbance score is obtained by summing the scores across the six factors, weighing vigor negatively.

It has been used in two studies of adult survivors of childhood cancer (Gray, Doan et al. 1992; Zeltzer, Chen et al. 1997). Gray *et al.* reported that 62 survivors with a range of diagnoses were similar to their 51 healthy age-matched peers. Zeltzer *et al.* found that in 580 ALL patients that survivors had a greater negative mood, more tension, depression,

anger and confusion than 396 matched sibling controls. The female survivors reported the highest mood disturbance. Scores were not as high as were found in a psychiatric sample.

HRQL and pediatric oncology

For a number of reasons, relatively little is known about HRQL in pediatric oncology patients who are undergoing treatment and among survivors, and there has been less interest in the development of HRQL measures for children with cancer (Pollock 1999). Listed below are other potential obstacles gleaned from the literature and personal experience. Some deal with perceptions about HRQL research, while others point to the nature of HRQL research and the state of HRQL research as it stands today.

1. Pediatric trials are generally designed with curative intent, contributing to a reduced emphasis on the development of strategies to improve HRQL.
2. Pediatric oncologists deal with life-threatening conditions, and patient survival is often viewed with paramount importance.
3. Long-term cancer survivorship in a large population is a relatively recent phenomenon.
4. Parents are guiding decisions for young pediatric cancer patients and may feel responsible to maximize survival with less consideration of HRQL.
5. The relative value and balance between quality and quantity of life is different for children and adults.
6. HRQL outcomes are often subjective and difficult to grasp compared to other outcome measures such as survival or medical late effects.

7. Measuring QOL can be more problematic in children, and they must take developmental stage into consideration; furthermore, the validity of proxy assessments is debatable.
8. In large part, the HRQL measurement instruments for children with cancer have not been universally accepted or longitudinally validated.
9. The effects of treatment that impact HRQL may not be apparent for many years, when contact between the patient and treating physician is minimal or non-existent.
10. Interventions to prevent the risk of decreased HRQL are not well defined, not validated, and/or perceived as ineffective.

The decision making process for treating cancer involves the consideration of many parameters. The treatment plan will of course depend upon of the patient's cancer diagnosis, risk factors, and overall medical condition. However, factors such as the patient's age, cultural background, expectations, and understanding may be important as well. Patients, families, and medical providers are encouraged to ask questions about how far treatment can be justified in order to increase the chance of survival, especially when HRQL is compromised.

The impact of cancer on the HRQL of a child can be profound, depending on the specific diagnosis, type of treatment, response to treatment, and individual variation. During active therapy, most children experience some degree of physical limitation. This can range for decrease in energy to loss of a limb. Some forms of treatment are well known to

cause changes in mood and behavior. Anxiety and sleep disturbance is not uncommon. These are generally temporary, but can impact social functioning and educational achievement. Radiation to the brain can have significant long-term effects on cognitive functioning. Many children will experience unpleasant side-effects from treatment, such as nausea and pain. Although with good intent, some parents may be "over-protective" of their child with cancer. This can restrict the child's activities and have negative effects on HRQL.

After completing therapy, many childhood cancer survivors (CCS) are free of medical late effects, re-establish their developmental trajectory, and go on become "normal" members of society. However, the cure for childhood cancer is a relatively recent development, and we are only beginning to study this cohort of long-term childhood cancer survivors in the population. CSS and their families should be counseled about the possibility of late effects and impact on HRQL. In addition, survivors may face other issues, including follow-up schedules for health monitoring, emotional aspects of surviving cancer, challenges in areas such as educational achievement, employment, personal relationships and fertility concerns, transition from pediatric to adult care, and discrimination in obtaining health insurance.

AML, BMT, and HRQL

In the remainder of the background, I will present a review of the literature relevant to the study of HRQL in survivors of pediatric AML. First, I will briefly review HRQL research in CCS. I will then discuss HRQL research in survivors of pediatric acute lymphoblastic

lymphoma (ALL) in more detail. ALL is the most common pediatric cancer, and HRQL research similar to this study has already been conducted. Next, I will describe late effects research in pediatric BMT recipients. Then, I will review studies that examined HRQL in survivors of AML given chemotherapy or stem cell transplantation for post-induction treatment. Finally, I will discuss HRQL research in survivors of pediatric AML.

HRQL in childhood cancer survivors

Compared to the number of studies in adult cancer patients, there are relatively few HRQL studies in childhood cancer survivors. In general, most studies are small in size, combine different cancer types, use proxy respondents, and/or study HRQL during or shortly after treatment (Langeveld, Stam et al. 2002). As a result, there isn't a clear or consistent picture about HRQL outcomes in CSS. Overall, in terms of physical functioning, the majority of survivors reported they were in good health, with the exception of bone tumor survivors who were more likely to perceive their health as fair to poor (Langeveld, Stam et al. 2002). In most studies, long-term survivors did not have greater emotional or psychological problems compared to controls. Zeltzer *et al.* did find more depression among ALL survivors than sibling controls, and Lansky *et al.* found that the prevalence of treated depression was higher in survivors compared to the general population (Zeltzer, Chen et al. 1997; Lansky, List et al. 1986). Female gender, racial minority, older age at follow-up, relapse, severe functional impairment, and CNS irradiation were associated with an increased risk for emotional problems in some studies. Three studies assessed self-esteem, and no differences between survivors and

controls were found. In terms of social functioning, brain tumor and ALL survivors were found to be at increased risk of lower educational attainment, which was associated with CNS irradiation and younger age at diagnosis. Most research showed that survivors did not differ from controls in employment status, although some survivors reported job discrimination or difficulties in obtaining work. Survivors leave home at an older age than controls. More survivors never marry; male CNS tumor survivors are at greatest risk. Fewer survivors have children compared to controls.

HRQL in survivors of pediatric ALL

The Children's Cancer Group (CCG) has studied late effects in survivors of the most common pediatric malignancy, acute lymphoblastic leukemia (ALL). It has been shown, for example, that central nervous system radiation therapy (CRT) causes a significant decrease in IQ in survivors of childhood ALL (Meadows, Gordon et al. 1981). Survivors were more likely to utilize special education services, and less likely to attend college compared to sibling controls (Haupt, Fears et al. 1994). There are conflicting reports on the psychosocial outcome in CSS. Some studies have shown that this population has more psychological impairment, psychiatric symptoms and/or behavioral problems compared with control groups, while others have shown no adjustment problems or mood disorders (Seitzman, Glover et al. 2004). As discussed above, Zebrack *et al.* found that survivors were more likely to have symptomatic levels of depression and somatic distress compared to sibling controls, but not more than would be expected in the general population (Zebrack, Zeltzer et al. 2002).

Zeltzer *et al.* published a study of the psychological outcome using the POMS questionnaire in adult survivors of childhood ALL (Zeltzer, Chen et al. 1997). The study group consisted of subjects who were age 0-20 years when diagnosed with ALL, in remission, no longer on treatment, and alive at least 2 years after diagnosis (95% of the sample was alive at least 5 years after diagnosis). Subjects were at least 18 years old by October 1990. There were 580 subjects and 396 sibling controls. As mentioned previously, higher POMS scores indicates greater mood disturbance. By multivariable regression, controlling for age, sex, and survivor status, the POMS score was significantly higher in survivors compared to controls for total score, and subscales for tension, depression, anger, and confusion. The differences in vigor and fatigue were not significantly different. The symptoms of negative mood in survivors were higher than a normative control group, but less than a control group of outpatient psychiatric patients. Interactions between demographic factors and survivor status were assessed. Female, minority, and unemployed survivors reported the highest total mood disturbance. Additionally, significantly more survivors who were not students were unemployed or working less than half-time.

Taken together, these studies show that CSS seem to be at an increased risk of problems in the interdependent cognitive, neuropsychological, and psychosocial domains. Decreases in IQ, educational achievement, self-esteem, and mood state can all negatively impact HRQL in this population.

Zeltzer *et al.* published a study of the psychological outcome using the POMS questionnaire in adult survivors of childhood ALL (Zeltzer, Chen et al. 1997). The study group consisted of subjects who were age 0-20 years when diagnosed with ALL, in remission, no longer on treatment, and alive at least 2 years after diagnosis (95% of the sample was alive at least 5 years after diagnosis). Subjects were at least 18 years old by October 1990. There were 580 subjects and 396 sibling controls. As mentioned previously, higher POMS scores indicates greater mood disturbance. By multivariable regression, controlling for age, sex, and survivor status, the POMS score was significantly higher in survivors compared to controls for total score, and subscales for tension, depression, anger, and confusion. The differences in vigor and fatigue were not significantly different. The symptoms of negative mood in survivors were higher than a normative control group, but less than a control group of outpatient psychiatric patients. Interactions between demographic factors and survivor status were assessed. Female, minority, and unemployed survivors reported the highest total mood disturbance. Additionally, significantly more survivors who were not students were unemployed or working less than half-time.

Taken together, these studies show that CSS seem to be at an increased risk of problems in the interdependent cognitive, neuropsychological, and psychosocial domains. Decreases in IQ, educational achievement, self-esteem, and mood state can all negatively impact HRQL in this population.

HRQL after BMT

As mentioned already, the BMT recipient is at risk for numerous late effects of treatment. Survivors of BMT are at increased risk of second malignancies, growth disturbances, hypothyroidism, gonadal dysfunction, infertility, leukencephalopathy, and loss of bone density (Oeffinger and Hudson 2004). Any of these can have an impact on HRQL. For the most part, however, studies examining the effect of BMT on organ function have not addressed the impact on HRQL directly.

A QOL instrument specific for patients treated with BMT has been developed for use with adults, but not children. One study of late effects and "qualitative aspects of daily life" among allo-BMT recipients included 50 patients who were transplanted at age 17 years or younger (median age at interview 9.5 years, range 7-27) (Schmidt, Niland et al. 1993). The qualitative aspects of daily life assessment involved a subjective rating of overall quality of life on a scale of 1 (low) to 10 (high), and additional levels to rate appetite, sleeping habits, and appearance. For children younger than 14 years at interview, parents served as proxy respondents. Only 16% of the children were rated as a 7 or less on the QOL outcome, and the vast majority had either excellent or normal ratings for the other 3 categories.

Younger age at the time of transplant is associated with an improved QOL (Baker, Wingard et al. 1994; Andrykowski, Greiner et al. 1995). In the Schmidt study above, children had a higher QOL score than adults, although the difference was not statistically significant.

Matthes-Martin *et al.*, reported on 155 patients who underwent allo-BMT during childhood at a single institution between 1980 and 1996 (Matthes-Martin, Lamche et al. 1999). The overall survival rate was 52% (81/155), with 73 patients observed at least one year. They conducted a cross-sectional survey a median of 4.6 years after transplant. Various elements of HRQL were examined. There was a strong association with QOL and the presence of cGVHD (present in 33% of this sample). All but one patient had a Karnofsy or Lansky score of >80%. Eight percent responded that they had severe restriction of their social contacts (secondary to immunosuppression), 4% had severe restriction of mobility and normal life activities, and 3% classified themselves as severely physically handicapped. Seventy-five percent reported no physical or psychological impairment. Common late effects in this group were growth retardation (38% at three years) and delayed puberty (33% of females requiring estrogen replacement). Patients who underwent total body irradiation (TBI) were more likely to report growth retardation three years after transplantation (68% vs. 8%). Neither the length of time since treatment, stem cell source (matched sibling donor vs. matched unrelated donor), or underlying disease (malignant vs. non-malignant) was significantly associated with HRQL.

Increased time from transplant is associated with improved QOL (Andrykowski, Greiner et al. 1995; Bush, Haberman et al. 1995), although this isn't a uniform finding (Matthes-Martin, Lamche et al. 1999). A substantial portion do well in tested domains, and in one study, 88% stated that the benefits of BMT outweighed the side-effects (Bush, Haberman et al. 1995). In a study of 135 BMT recipients, most survivors had above average satisfaction with their major life domains (Baker, Wingard et al. 1994). They were least

satisfied with their bodies, level of physical strength, and ability to attain sexual satisfaction. They also found that GVHD was associated with diminished QOL.

In summary, BMT recipients are at risk for numerous late effects, which can have a negative impact on HRQL. Research to date has not found uniform factors that predict HRQL outcomes, although absence of GVHD, younger age at transplant, and the passage of time seem to be associated with better HRQL measures. The majority of BMT recipients have a high level of function and a favorable HRQL outcome.

HRQL: chemotherapy versus BMT

The intensive treatment used in AML therapy can have a substantial impact on HRQL. This burden is most apparent at the time of diagnosis and during the course of therapy. Most research in adults has shown, however, that the majority long-term survivors of AML recover completely and enjoy HRQL comparable to the general population (Redaelli, Stephens et al. 2004).

In 50 long term survivors of BMT (32% who had AML), Malassiotis *et al.* found physical symptoms to be the strongest predictor of HRQL (Molassiotis, Boughton et al. 1995). Substantial physical dysfunction was reported in 14.4% of auto-BMT, and 11.5% of allo-BMT recipients who were at least 6 months post-transplant. Symptoms commonly reported were dry mouth and fatigue. Lesko *et al.* studied a group of acute leukemia survivors (53% with AML) an average of 5 years after completing treatment and found that 94% of BMT patients and 93% of chemotherapy patients had near-normal

levels of physical functioning (Lesko, Ostroff et al. 1992). In contrast, Zittoun *et al.* observed in a group of 98 AML survivors (35 allo-BMT, 29 auto-BMT, 34 chemotherapy) an average of 4 years in remission that 49% reported difficulties taking a long walk, or felt limited while doing work or household chores (Zittoun, Suciu et al. 1997). Overall physical condition and HRQL was significantly worse for allo-BMT patients. In their cohort, 34% reported cGVHD, with more than half being moderate or severe. Somatic symptoms that were reported significantly more often in allo-BMT patients included mouth sores, cough, hair loss, and pain during sexual intercourse (see below).

Two studies have shown reduced sexual function in AML patients who underwent BMT compared to consolidation chemotherapy (CCT). The United Kingdom Medical Research Council AML10 trial compared QOL in 479 patients one year from the end of treatment in patients who received post-induction allo-BMT, auto-BMT, or CCT (Watson, Buck et al. 2004). Patients who underwent BMT had worse sexual and social relationships, and diminished professional and leisure activities. Patients who had an auto-BMT were less impacted than allo-BMT. Significantly more BMT patients reported a decrease in sexual interest (48% vs. 24%), activity (53% vs. 35%), pleasure (36% vs. 18%) and ability (53% vs. 35%). However, patients undergoing BMT had higher risk disease, and this may have biased the result. Zittoun *et al.* also reported that sexual function was significantly more impaired after allo-BMT than auto-BMT or CCT (sexual relationships worsened in 50%, 29%, and 13%, respectively).

Infertility is a recognized complication after treatment for AML, and can negatively impact HRQL. According to Wellisch *et al.*, 50% of CCT patients, compared to 66% of BMT patients reported that infertility was their most significant problem (Wellisch, Centeno et al. 1996). Zittoun *et al.* reported infertility rates of 63% in allo-BMT, 48% in auto-BMT, and 3% for CCT. Watson *et al.* reported similar rates of 64% in allo-BMT, 51% in auto-BMT, and 10% for CCT (Watson, Wheatley et al. 1999).

Long-term psychological outcome was studied in 206 leukemia survivors (77% with AML) using the Psychosocial Adjustment to Illness scale (Greenberg, Kornblith et al. 1997). This study showed that younger age at time of testing (20-30 years) and lower educational achievement (high school or less) was associated with greater psychological distress when measured on average 5 years after treatment. In contrast, the study by Lesko *et al.* showed higher levels of psychiatric distress in subjects who were male and highly educated. Poor social adjustment was associated with female gender, treatment with BMT, and high level of education.

In summary, multiple HRQL domains are affected by AML and its treatment, which may include BMT. Because different HRQL instruments have been used, and the populations studied were not uniform, it can be difficult to compare the results among studies, with some findings apparently contradicting each other. In general, however, the long-term HRQL outcome is generally good among AML survivors, although BMT recipients may fare worse. Medical late effects associated with BMT, including infertility and symptoms of GVHD, can have a negative impact on HRQL.

Quality adjusted survival after treatment of childhood AML

The Pediatric Oncology Group (POG) conducted a trial comparing CCT and auto-BMT for post-remission therapy in AML (Ravindranath, Yeager et al. 1996). Allo-BMT was offered to patients in complete remission who had an available matched sibling donor. The intent-to-treat analysis found no significant difference in disease free survival (DFS) in CCT and auto-BMT. OS and DFS were marginally better in allo-BMT compared to CCT. Outcome and toxicity data from this study were retrospectively analyzed using the quality-adjusted time without symptoms (Q-TWiST) method in order to characterize and compare the QOL experience in these patients (Parsons, Gelber et al. 1999). Using this method, three clinical health states are defined: TOX, the period with treatment-related side-effects with a severity of grade 3 or more, and deemed by the clinician to have functional impact on the patient; TWiST, the period with no toxicities from treatment or symptoms of disease; REL, the period after relapse. For patients receiving allo-BMT, there is a fourth health state, GVHD, the period with symptoms of severe GVHD. Pair-wise comparisons were made between CCT and auto-BMT groups, and CCT and allo-BMT groups. Patients who received CCT had less time in TOX and more in TWiST, relapse-free, and alive than the auto-BMT group. Although patients who received allo-BMT had more time in TOX compared to the CCT group, they had more time in TWiST, relapse free, and alive. In order to make a summary comparison between pairs, Q-TWiST was calculated by multiplying each term by a "utility coefficient" and summing these values to obtain a final score. A variety of utility coefficients were systematically assessed ("utility threshold analysis"). Accounting for time spent with severe GVHD in

the allo-BMT recipients, they found that this group had more quality-adjusted time than the other two modalities.

HRQL in childhood AML survivors

The research reported herein was part of a multi-dimensional study of HRQL in childhood AML survivors. Data from the SF-36 and CHQ were recently analyzed (Nicholson, Zhou et al. 2005). QOL summary scores for the 82 survivors treated with BMT (autologous and allogeneic) did not differ from that of the 124 treated with CCT. None of the physical or mental SF-36 or CHQ subscales significantly differed by treatment, although physical subscale domains trended toward being better in the CCT group. They concluded that QOL in long-term survivors is not diminished by BMT and that the primary concern in assigning treatment remains survival.

RESEARCH DESIGN AND METHODS

Overview

The data used in this study are from Children's Cancer Group (CCG) study L9704, titled "Quality of Life Following Successful Therapy for Acute Myelogenous Leukemia: A Comparison of Bone Marrow Transplant and Chemotherapy." The principal investigator was Dr. H. Stacy Nicholson. The study was open for accrual from May 17, 1999 to April 25, 2005.

Research question

Among long-term survivors of pediatric acute myeloid leukemia (AML), is treatment with allogeneic bone marrow transplantation (allo-BMT) associated with a diminished quality of life (QOL) compared to treatment with chemotherapy or autologous BMT (auto-BMT)?

Specific aims and hypotheses

1. The first aim was to determine the impact of initial therapy (allo-BMT versus non-allo-BMT) on QOL outcome, as defined by a questionnaire used to measure mood disorders in long-term survivors of childhood and adolescent AML. We hypothesized that QOL, in terms of mood disorder outcome, was more adversely affected in patients treated with allo-BMT compared to patients treated with chemotherapy or auto-BMT.
2. The next aim was to determine whether initial patient characteristics (e.g. age at diagnosis, sex, race), disease and treatment factors (e.g. use of total body

irradiation), and survivor characteristics (e.g. age, disease status, late effects) are associated with QOL outcome. We hypothesized that female gender, age at diagnosis, disease relapse, and adverse treatment effects was associated with diminished QOL in survivors.

3. The final aim was to define risk factors for diminished QOL, so that interventions may be undertaken to improve QOL in survivors of AML.

Eligibility Criteria

In order to participate, subjects had to meet the following criteria:

1) Enrolled on CCG AML clinical trial 251, 213, 2861, or 2891. These therapeutic clinical trials enrolled previously untreated patients diagnosed with AML age 0-21 years old from 1979-1995.
2) Alive and without active disease at least 5 years after completing treatment.
3) Currently living in the U.S. or Canada
4) English or Spanish speaking

Enrollment procedure

A list of eligible subjects was generated by the CCG operations office using the criteria listed above. Permission to contact the survivor was obtained from the responsible investigator at the patient's treating institution. Contact was then made by the principal investigator of CCG-L9704 or a physician at the treating institution. For potential participants with invalid addresses, tracing was performed using the following sources: postal service, state motor vehicle registry, and on-line telephone directories.

Size and characteristics of the study population

The number of participants meeting the above criteria as of January 1999 was estimated as follows: 200 allo-BMT participants, 90 auto-BMT participants, and 320 chemotherapy participants, for a total of 610 potential participants. Based on prior experience in conducting similar research, it was estimated that there would be a 9% tracing failure rate, 7% refusals rate, and 4% rate of those not able to participate for other reasons. The estimated number of participants was therefore 488, classified as follows: 160 allo-BMT participants, 72 auto-BMT participants, and 256 chemotherapy participants. Based on the number of participants meeting the above criteria, the proportion of males is expected to be 55%. The expected ethnic group proportions are as follows: white, non-Hispanic, 64%; Hispanic, 18%; Black, 8%; Asian, 6%; Native American, 1%; other, 3%.

Statistical power estimation

The primary comparison was the difference in QOL test score between allo-BMT and non-allo-BMT survivors. A difference in mean score of one half of one standard deviation was considered clinically significant. Using the estimates above, there would be approximately twice the number of subjects in the chemotherapy group (2N) than in the allo-BMT group (N). The table below shows the sample size needed in the allo-BMT group (N) to detect a difference of one-third of one standard deviation.

Power	N
0.80	106
0.85	121
0.90	142
0.95	175

Using the estimate above (N=160), there is a 90-95% power to detect a very small difference in QOL score between these two groups.

Data collected and instruments used

Clinical Data

Clinical data was obtained from the CCG Statistical and Data Center, including the subtype of AML, sex, race, and age at diagnosis. Treatment data include the assigned protocol and treatment regimen, occurrence of relapse, occurrence of GVHD, and history of relapse or second cancer. Questions about the clinical data were resolved by contacting the treating physician.

The Childhood Cancer Survivors Study (CCSS) Late Effects Questionnaires

The CCSS (described previously) questionnaire is based on questions contained in the National Health Interview Survey (NHIS). They have been extensively validated and can generate expected numbers of events. The questionnaire contains items related to the following general areas: demographics, access to and use of medical care, medical conditions (i.e., hearing/vision/speech, hormonal, brain and nervous system, heart and circulatory system, respiratory system, anxiety, genetic conditions, and congenital defects), family history including first degree relatives, reproductive history, offspring, health habits, physical activity, and sociodemographic factors (i.e., education, marital status, employment and insurance).

SF-36 Health Survey and Child Health Questionnaire

The SF-36 has been validated for use in adult populations and accurately measures QOL in adolescents who are at least 14 years of age. The Child Health Questionnaire has recently been validated for use in children over four years of age. This instrument measures the same domains as the SF-36, and consists of a child-completed version for older children and a parent-completed version for younger children. The Child Health Questionnaire was utilized for younger participants.

Measurement of mood disorders instrument (Profile of Mood States)

Mood disturbance was measured using the Profile of Mood States (POMS), described previously (see also Appendix II). The measure has been shown to be both reliable and valid.

The test instrument consists of 65 questions, used to calculate 6 factor scores shown here:

Factor	**Number of questions**
Tension	9
Depression	15
Anger	12
Vigor	8
Fatigue	7
Confusion	7

Each question consisted of a word (e.g. tense, bitter, forgetful), and the subject was asked to describe how they have been feeling with respect to each word during the past week using the following scale:

0	Not at all
1	A little
2	Moderately
3	Quite a bit
4	Extremely

For each factor score, except vigor, the higher the score, the greater the degree of mood disturbance. The factor scores were added up, with vigor weighted negatively, to obtain a "Total Mood Disturbance" (TMD) score. The possible score range for each factor and TMD are as follows:

Factor	**Score Range**
Tension	0 to 36
Depression	0 to 60
Anger	0 to 48
Vigor	0 to 32
Fatigue	0 to 28
Confusion	0 to 28
TMD	-32 to 200

If an item was missed or skipped, an adjusted factor score was calculated using the completed items. If more than two items were missed within a factor, the factor score was not calculated and marked as incomplete. If more than 4 factors needed to be adjusted, or if a factor score could not be calculated, the TMD score was not calculated.

Data collection

The CCSS questionnaires were mailed to participants who are not already enrolled in the CCSS. Telephone reminders were utilized for participants late in returning questionnaires following standard procedures of the CCSS. The SF-36 and Child Health Questionnaire were administered by trained personnel via telephone interview in either English or Spanish. The POMS and Harter Self-Perception Profile were mailed to participants.

Statistical analysis

For the analysis reported here, the dependent variables are the POMS scores. The principal independent variable will be the treatment variable (allo-BMT, auto-BMT, or chemotherapy). Additional independent variables include the demographic and clinical data.

The primary comparison was between allo-BMT and non-allo-BMT (specific aim 1). Additional comparisons were between allo-BMT and auto-BMT, and between auto-BMT and chemotherapy. Some patients treated with allo-BMT received total body irradiation, and others did not. Comparison between these groups was performed. Associations between the test scores (POMS and SF-36) and demographic and clinical variables, as well as the presence or absence of late effects of therapy were examined (specific aim 2). Chi-square or exact tests were used to compare proportions when the dependent and independent variables are categorical. T-test was used when the dependent variable is continuous and independent variable is categorical. Linear regression was used when both are continuous. For all analyses, two-sided tests of significance were used. Multivariable regression models were utilized to examine potential predictive variables of outcome (specific aim 3).

RESULTS

Study population

Four hundred and eighty-eight subjects were targeted for enrollment, based on the expected number of eligible survivors and expected participation rate. Due to unanticipated recruitment difficulties (personal communication), the actual number of patients enrolled was 206 subjects. There were 124 patients who received post-remission chemotherapy, 54 who received allogeneic bone marrow transplantation (allo-BMT), and 28 who received autologous bone marrow transplantation (auto-BMT). Their treatment allocation for this study is based on treatment actually received, and may differ from how they were randomized on the study (e.g. if a subject was randomized to allo-BMT but refused this procedure, they would be categorized as "chemotherapy" in this study). Subject characteristics are shown in Table 2 below. Median age at testing was 19 years (range 8-39 years). Median age of the subjects at the time of acute myeloid leukemia (AML) diagnosis was 3 years (range 0-20 years), and the average time elapsed between diagnosis and enrollment was 13.4 years. Forty-eight percent of the subjects were male and 87% were white. Sixty-three were enrolled on CCG-213 (31%), 48 on CCG-251 (23%), 8 on CCG-2861 (4%), and 87 on CCG-2891 (42%).

POMS

There were 130 Profile of Mood States (POMS) questionnaires submitted. One form was incomplete and unscorable, and five forms could not be linked with treatment data, leaving 124 subjects for all subsequent analysis. There were substantially fewer completed POMS questionnaires compared to study subjects, presumably because the

Table 2. Characteristics of subjects enrolled on L9704 overall and by post-remission therapy received.

	All (N=206)	Auto-BMT (N=28)	Allo-BMT (N=54)	Chemotherapy (N=124)	*P*
Age at enrollment for L9704, median (range)	19 (8 – 39)	13.5 (8 - 25)	21.5 (8 - 38)	19.5 (8 - 39)	0.001
Age at diagnosis of AML, median (range)	3 (0 – 20)	3 (0 – 16)	8 (0 – 20)	3 (0 – 19)	<0.001
Time since diagnosis of AML, mean (SE)	13.42 (0.31)	9.79 (0.78)	12.83 (0.56)	14.49 (0.37)	<0.001
Gender					
Male	98 (48%)	13 (46%)	28 (52%)	57 (46%)	0.76
Female	108 (52%)	15 (54%)	26 (48%)	67 (54%)	
Race					0.24
White (incl. Hispanic)	181 (87%)	25 (89%)	48 (89%)	108 (87%)	
African American	9 (4%)	1 (4%)	2 (4%)	6 (5%)	
American Indian, Aleutian, Eskimo	1 (1%)	0 (0%)	0 (0%)	1 (1%)	
Others or Unknown	15 (8%)	2 (7%)	4 (7%)	9 (7%)	
Treatment Study					
CCG-213	63 (31%)	0 (0%)	11 (21%)	52 (42%)	
CCG-251	48 (23%)	0 (0%)	14 (26%)	34 (27%)	
CCG-2861	8 (4%)	5 (18%)	3 (6%)	0 (0%)	
CCG-2891	87 (42%)	23 (82%)	26 (48%)	38 (31%)	

(Table adapted from Nicholson 2005)
Abbreviations: Allo-BMT, allogeneic bone marrow transplantation; auto-BMT, autologous bone marrow transplant

POMS required the subject to return the form by mail (versus telephone interview for the Short Form-36 instrument). Characteristics for subjects completing the POMS are shown in Table 3 below (note: subjects treated with chemotherapy and auto-BMT were combined and labeled "non-allo-BMT"). Median age at testing was 21 years (range 12-39

years). Median age of the subjects at the time of AML diagnosis was 8.4 years (range 0.1-20.1 years), and the average time elapsed between diagnosis and enrollment was 14.8 years. Forty-one percent of the subjects were male, and 89% were white. Compared to the overall study group, POMS respondents were approximately the same age, although there were more females and older at time of diagnosis. Nearly all of the allo-BMT enrollees returned the POMS (51 out of 54), whereas the return rate for the non-allo-BMT subjects was substantially lower (73 out of 152).

Table 3. Characteristics of subjects completing the POMS questionnaire overall and by post-remission therapy.

	Overall (N= 124)	**Allo-BMT (N=51)**	**Non-allo-BMT (N=73)**	***P***
Age at enrollment for L9704				
Mean (yrs)	22.5	23.0	22.2	0.486
Median	21.0			
Range	12 – 39			
Std. Dev.	6.2	6.1	6.3	
Age at diagnosis of AML				
Mean (yrs)	8.4	8.9	8.0	0.354
Median	8.4			
Range	0.1 – 20.1			
Std. Dev.	5.6	5.9	5.3	
Time since diagnosis of AML				
Mean (yrs)	14.1	14.0	14.2	0.846
Median	14.8			
Range	5.6 – 21.7			
Std. Dev.	4.1	4.3	4.1	
Sex				
Male (%)	51 (41.1%)	22 (43.1%)	29 (39.7%)	0.704
Race				
White	110 (88.7%)	44 (86.3%)	66 (90.4%)	0.668
Black	7 (5.6%)	3 (5.9%)	4 (5.5%)	
Other	7 (5.6%)	4 (7.8%)	3 (4.1%)	
Relapse				
Yes (%)	18 (14.5%)	7 (13.7%)	11 (15.1%)	0.835

Abbreviations: Allo-BMT, allogeneic bone marrow transplantation

Univariate analysis

The data shown in Table 4 are the POMS scores for the group overall, and then grouped according to post-remission treatment received. The *P*-values shown are for the mean scores for the allo-BMT and non-allo-BMT groups using a two-sided t-test. There are no significant differences in the mean scores for any of the POMS domains or Total Mood Disturbance (TMD) score. When the means are compared for the three post-remission treatment group separately (data not shown), the only statistically significant score is for Anger (P=0.031), which is higher in the Auto-BMT group (mean 13.21, SE 2.6) compared to the other two groups (allo-BMT: mean 8.21, SE 0.8; chemotherapy: mean 7.36, SE 0.031).

Table 4. POMS scores for overall group and post-remission therapy (univariate analysis).

	Overall		**Allo-BMT**		**Auto-BMT**		**Chemo-therapy**		**Non-allo-BMT**		***P****
	N=124		**N=51**		**N=14**		**N=59**		**N=73**		
	Mean, SE		**Mean, SE**		**Mean, SE**		**Mean, SE**		**Mean, SE**		
Tension	8.84	0.4	**8.51**	0.5	10.20	1.6	8.80	0.7	**9.07**	0.6	**0.526**
Depression	8.43	0.8	**8.35**	0.9	13.07	3.7	7.41	0.8	**8.49**	1.2	**0.932**
Anger	8.56	0.6	**8.21**	0.8	13.21	2.6	7.76	1.0	**8.81**	1.0	**0.648**
Vigor	17.47	0.6	**18.69**	0.9	16.36	1.5	16.70	0.9	**16.63**	0.8	**0.084**
Fatigue	7.46	0.5	**6.70**	0.6	8.57	1.7	7.85	0.8	**7.99**	0.7	**0.211**
Confusion	7.38	0.4	**7.42**	0.5	8.43	1.4	7.10	0.5	**7.36**	0.5	**0.933**
TMD	23.21	2.6	**20.48**	3.0	37.13	10.1	22.22	4.0	**25.08**	3.8	**0.382**

Abbreviations: Allo-BMT, allogeneic bone marrow transplantation; auto-BMT, autologous bone marrow transplant; TMD, Total Mood Disturbance
* t-test, comparing means in Allo-BMT and non-allo-BMT groups.

The histograms for POMS scores, paneled by treatment group (allo-BMT and non-allo-BMT), are shown in Figures 3. By inspection, the distribution and means appear similar for all POMS outcomes.

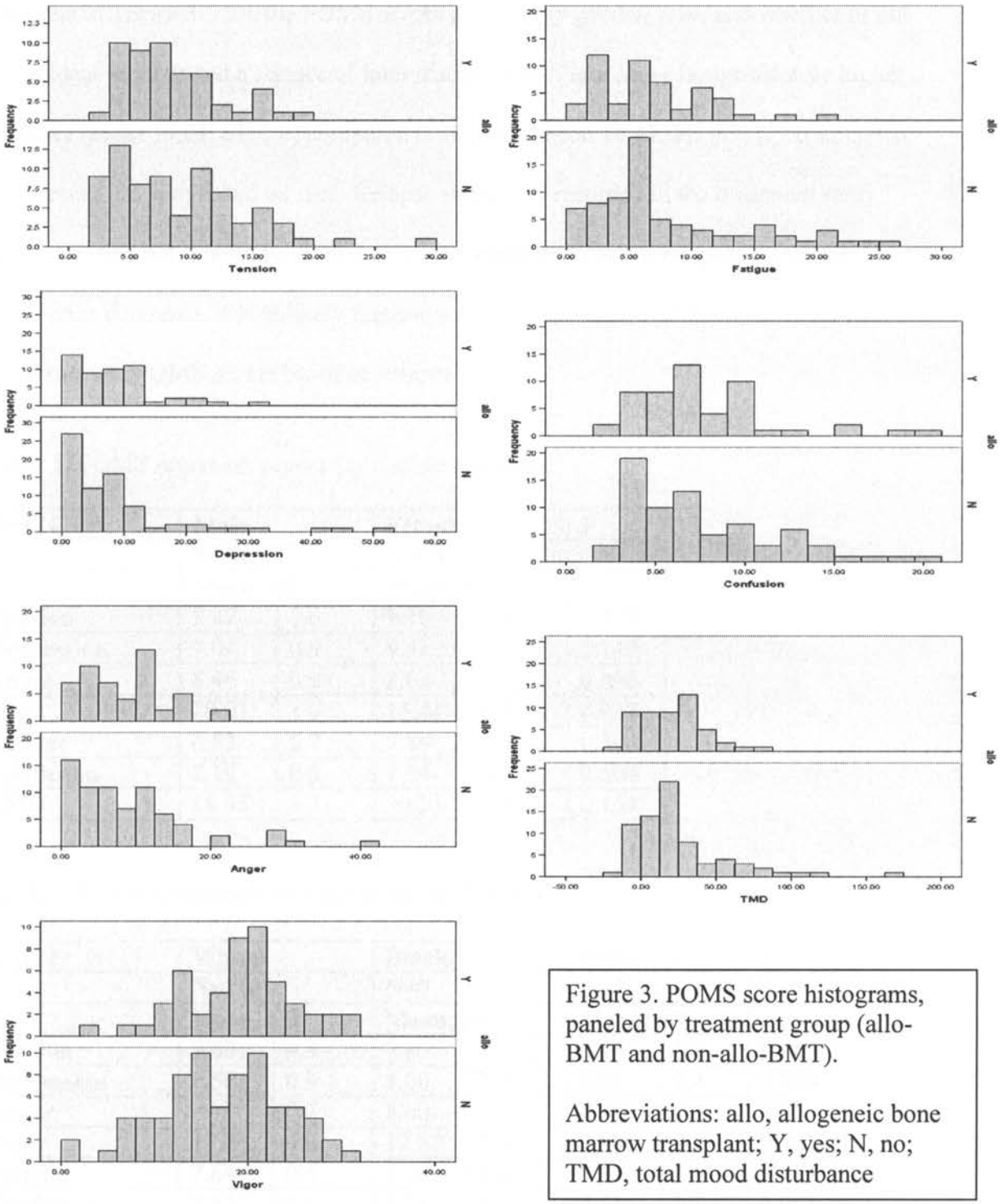


Figure 3. POMS score histograms, paneled by treatment group (allo-BMT and non-allo-BMT).

Abbreviations: allo, allogeneic bone marrow transplant; Y, yes; N, no; TMD, total mood disturbance

The data in Tables 5-7 are the POMS scores grouped by gender, race, and whether or not the subject experienced a relapse of their disease. The Vigor score is significantly higher in males (mean 19.01, SE 1.0) compared to females (mean 16.41, SE 0.7) (*P*=0.027). No differences are seen based on race. Relapse status was recorded in the *treatment* study data form, and not this study. Because the vast majority of AML relapses occur within 5 years from diagnosis, it is unlikely that cases of relapse were missed. There are no differences in POMS scores based on relapse status.

Table 5. POMS scores by gender (univariate analysis).

N=124	**Male**		**Female**		***P***
	N=51		**N=73**		
	Mean, SE		**Mean, SE**		
Tension	8.47	0.6	9.10	0.6	0.475
Depression	7.04	0.9	9.41	1.2	0.147
Anger	8.46	0.9	8.63	0.9	0.896
Vigor	19.01	1.0	16.41	0.7	**0.027**
Fatigue	6.83	0.7	7.89	0.7	0.306
Confusion	7.16	0.5	7.54	0.5	0.604
TMD	18.95	3.2	26.23	3.8	0.164

Table 6. POMS scores by race (univariate analysis).

N=124	**White**		**Black**		**Other**		***P***
	N=110		**N=7**		**N=7**		
	Mean, SE		**Mean, SE**		**Mean, SE**		
Tension	9.02	0.5	7.86	1.7	6.95	0.9	0.465
Depression	8.56	0.9	8.00	2.3	6.92	1.4	0.889
Anger	8.73	0.7	8.00	1.9	6.57	1.5	0.727
Vigor	17.26	0.6	17.57	1.9	20.86	2.8	0.367
Fatigue	7.64	0.5	4.50	1.8	7.57	2.1	0.364
Confusion	7.42	0.4	6.71	1.7	7.43	0.8	0.904
TMD	24.13	2.8	17.50	8.6	14.58	5.9	0.599

Table 7. POMS scores by history of relapse (univariate analysis).

N=124	Yes		No		*P*
	N=18		N=106		
	Mean, SE		Mean, SE		
Tension	9.22	1.2	8.77	0.5	0.714
Depression	7.56	2.3	8.58	0.9	0.654
Anger	8.31	2.1	8.61	0.7	0.871
Vigor	16.71	1.9	17.61	0.6	0.588
Fatigue	7.72	1.6	7.41	0.5	0.830
Confusion	7.78	0.8	7.31	0.4	0.653
TMD	23.88	7.7	23.09	2.7	0.914

Among the 51 allo-BMT recipients, 35 (68.6%) received total body irradiation (TBI), and 8 (15.7%) had chronic GVHD (cGVHD). Univariate analysis based on TBI and cGVHD are shown in Tables 8 and 9. There is no difference in POMS scores based on TBI. The diagnosis of cGVHD was recorded in the *treatment* study data form, and does not necessarily mean that the subject had cGVHD at the time of L9704 enrollment. It is very unlikely that a patient would develop cGVHD *de novo* beyond the treatment study data capture period of 5 years; it is possible, but unlikely, that cases of cGVHD were missed. The Depression (P=0.002), Fatigue (P=0.026), Confusion (P=0.036), and TMD (P=0.003) scores are significantly higher in subjects who had cGVHD compared to those who did not.

Table 8. POMS scores by total body irradiation (univariate analysis).

N=51	Yes		No		*P*
	N=35		N=16		
	Mean, SE		Mean, SE		
Tension	8.37	0.5	8.81	1.3	0.712
Depression	8.78	1.0	7.44	2.0	0.516
Anger	8.45	0.9	7.69	1.5	0.645
Vigor	17.94	1.0	20.31	1.7	0.205
Fatigue	6.13	0.7	7.94	1.2	0.180
Confusion	7.27	0.6	7.75	1.0	0.679
TMD	21.03	3.42	19.31	6.0	0.791

Table 9. POMS scores by chronic graft versus host disease (univariate analysis).

N=51	Yes		No		*P*
	N=8		N=43		
	Mean, SE		Mean, SE		
Tension	10.63	1.4	8.12	0.6	0.096
Depression	14.88	3.5	7.14	0.8	**0.002**
Anger	10.63	2.3	7.74	0.8	0.167
Vigor	15.75	2.4	19.23	0.9	0.143
Fatigue	9.88	1.9	6.10	0.6	**0.026**
Confusion	10.00	1.8	6.94	0.5	**0.036**
TMD	40.25	8.8	16.72	2.8	**0.003**

For the continuous independent variables age at diagnosis, age at testing, and time elapsed, a Pearson correlation coefficient was calculated and significance value obtained. Results are presented in Table 10 below. The only statistically significant correlation found is between Fatigue and age at diagnosis (r=0.183, *P*=0.042).

Table 10. Correlations between POMS scores and age at diagnosis, age at testing, and years elapsed (univariate analysis).

N=124	Age at diagnosis		Age at testing		Years elapsed	
	Pearson	Sig.	Pearson	Sig.	Pearson	Sig.
Tension	-0.076	0.403	-0.056	0.539	0.018	0.845
Depression	-0.102	0.259	-0.065	0.470	0.039	0.671
Anger	-0.153	0.092	-0.137	0.130	-0.003	0.978
Vigor	-0.005	0.956	-0.078	0.387	-0.111	0.218
Fatigue	0.183	**0.042**	0.139	0.124	-0.036	0.689
Confusion	-0.079	0.385	-0.088	0.330	-0.027	0.764
TMD	-0.057	0.530	-0.031	0.731	0.029	0.748

Correlations between the POMS and SF-36 scores are shown in Table 11. Correlation between the POMS and SF-36 Mental score are all highly significant (*P*<0.001). The sign of the correlation coefficient corresponds to the characteristics of the instruments (i.e.

higher POMS scores correlate with lower SF-36 scores, except for Vigor). The Tension, Vigor, Fatigue, and TMD scores are significantly correlated with the SF-36 Physical score, again in the expected direction (i.e. high Vigor correlated with a high SF-36 Physical score, and a high Fatigue score correlated with a low SF-36 Physical score).

Table 11. Correlations between POMS and SF-36 scores.

N=124	**SF-36 Physical**		**SF-36 Mental**	
	Pearson	**Sig.**	**Pearson**	**Sig.**
Tension	-0.194	**0.031**	-0.463	**<0.001**
Depression	-0.173	0.055	-0.533	**<0.001**
Anger	-0.089	0.325	-0.418	**<0.001**
Vigor	0.272	**0.002**	0.472	**<0.001**
Fatigue	-0.239	**0.007**	-0.478	**<0.001**
Confusion	-0.054	0.548	-0.378	**<0.001**
TMD	-0.229	**0.011**	-0.608	**<0.001**

Abbreviation: SF-36, Short Form 36

Multivariable analysis

Multivariable linear regression was performed for each of the six POMS domains and TMD using the variables shown in Table 12. *TBI* and *cGVHD* were not used in the model when using all of the subjects, as these predictor variables are not distributed between the two main effect groups (i.e. these risk factors only apply to subjects who received allo-BMT). Analysis with these two variables was performed only among subjects who underwent allo-BMT.

The linear regression model with all variables entered into the equation yielded significance values as shown in Table 13. The only independent variables found to be significant are for the associations between *sex* and Vigor and between *age_dx* and

Table 12. Variables examined in multivariable linear regression model.

	Variable name	Variable description	Property / grouping
Dependent variable	POMS score	Test score for each mood domain and TMD	Continuous
Fixed effects	*allo*	Post-remission treatment	Allo-BMT, non-allo-BMT
	sex	Subject gender	Male, Female
	race	Subject race	White, Black, Other
	relapse	If a subject relapsed	Yes, No
	study	Treatment study number	213, 251, 2861, 2891
	FAB	AML subtype	M0 - M7, other/unknown
	TBI	If a subject received TBI	Yes, No
	cGVHD	If a subject had cGVHD	Yes, No
Covariates	*age_dx*	Age at diagnosis	Continuous
	age_test	Age at POMS testing	Continuous
	yrs_elapsed	Years since diagnosis	Continuous

Abbreviations: TMD, Total Mood Disturbance; FAB, French-American-British; TBI, total body irradiation; cGVHD, chronic graft versus host disease; Allo-BMT, allogeneic bone marrow transplantation

Fatigue. Males have significantly higher Vigor scores than females (P=0.042), and older age at diagnosis is associated with a greater Fatigue score (P=0.035). These findings parallel those in the univariate analysis. More importantly, the variable *allo* is not associated with any of the POMS scores. This again reflects the finding in univariate analysis.

Table 13. Significance values for variables in multivariable linear regression model for POMS scores.

N=124	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
sex	.450	.133	.762	**.042**	.385	.473	.162
race	.341	.843	.498	.343	.820	.931	.510
study	.766	.483	.708	.414	.069	.469	.643
age_dx	.625	.358	.266	.521	**.035**	.470	.844
yrs_elapse	.886	.661	.514	.771	.217	.401	.767
relapse	.708	.747	.696	.472	.806	.794	.912
FAB	.301	.306	.096	.283	.717	.593	.195
allo	.665	.934	.820	.138	.141	.885	.484

Based on the results above, a simplified model retaining the variable *allo* (because this is the independent variable of interest) and *sex* was performed (Table 14). Again, the only significant finding is between *sex* and Vigor. Furthermore, when treatment was considered in three categories (*consolid*) no significant associations were found except between *consolid* and Anger, reflecting the significantly higher Anger score in the auto-BMT group. Similar analyses were performed comparing allo-BMT and chemotherapy recipients (i.e. without the auto-BMT subjects) with findings consistent to those described above.

Table 14. Significance values for variables in simplified multivariable linear regression models for POMS scores.

N=124	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
sex	.489	.150	.912	**.030**	.325	.603	.175
allo	.542	.972	.653	.093	.224	.919	.413

N=124	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
sex	.441	.102	.750	**.029**	.309	.543	.131
consolid	.487	.077	**.031**	.230	.420	.515	.122

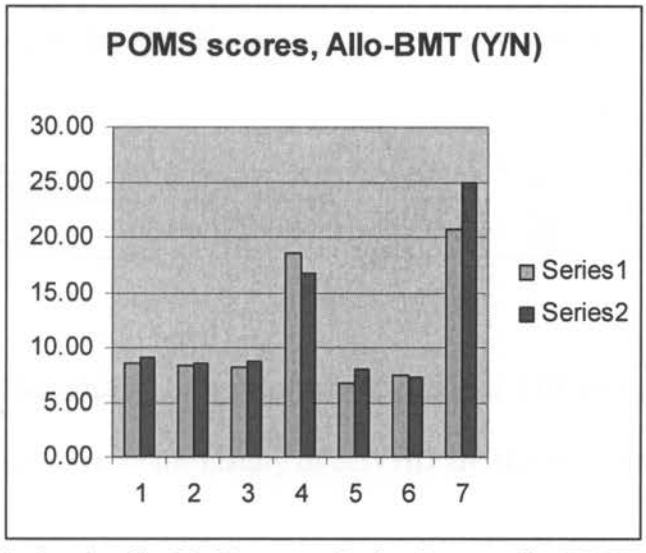
The adjusted mean scores (with *allo* and *sex* in the model) for subjects in the allo-BMT and non-allo-BMT groups are shown in Table 15 and Figure 4. The effect size was calculated by dividing the difference of the adjusted means by the unadjusted standard deviation of the non-allo-BMT group (Cohen 1988).

In order to examine the association between the variables *TBI* and *cGVHD* and POMS scores, it was necessary to perform multivariable linear regression only among subjects who received allo-BMT. With all variables in the equation, the significance values are as

Table 15. Adjusted POMS scores for the allo-BMT and non-allo-BMT groups.

Adjusted mean	Allo- BMT	Non-allo-BMT	*P*	Effect size
Tension	8.52	9.06	0.542	-0.10
Depression	8.40	8.46	0.972	-0.01
Anger	8.21	8.80	0.653	-0.07
Vigor	18.64	16.67	0.093	0.30
Fatigue	6.72	7.97	0.224	-0.20
Confusion	7.43	7.35	0.919	0.02
TMD	20.66	24.95	0.413	-0.13

Figure 4. Adjusted POMS scores for the allo-BMT and non-allo-BMT groups.



Series 1: allo-BMT group Series 2: non-allo-BMT group.
1=Tension, 2=Depression, 3=Anger, 4=Vigor, 5=Fatigue, 6=Confusion, 7=TMD.
All differences NS

Table 16. Significance values for variables in multivariable linear regression model for POMS scores among subjects who received allo-BMT.

N=51	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
age_dx	.982	.806	.289	.797	.902	.919	.766
age_test	.982	.806	.290	.797	.902	.919	.766
yrs_elapse	.983	.804	.290	.799	.901	.918	.768
sex	.849	.520	.289	.434	.883	.767	.758
relapse	.442	**.048**	**.007**	.405	.086	.884	**.036**
TBI	.484	.331	.717	.363	.990	.496	.439
cGVHD	.098	**.000**	.067	.058	.058	**.043**	**.001**

shown in Table 16 above. *Relapse* is significantly associated with the outcomes Depression, Anger, and TMD; *cGVDH* is significantly associated with the outcomes Depression, Confusion, and TMD. *TBI* is not associated with any POMS outcomes. Based on these results, a simplified model retaining the variables *TBI* and *cGVHD* (because these were independent variables of interest) and *relapse* was performed (Table 17). The same associations were seen, with the addition of the association between *cGVHD* and Fatigue.

Table 17. Significance values for variables in a simplified multivariable linear regression model for POMS scores among subjects who received allo-BMT.

N=51	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
relapse	.449	**.037**	**.023**	.242	.081	.882	**.032**
TBI	.944	.166	.421	.110	.319	.949	.361
cGVHD	.111	**.001**	.132	.085	**.040**	**.044**	**.002**

The adjusted mean scores (with *relapse, TBI,* and *cGVHD* in the model) for subjects with and without the history of cGVHD are shown in Table 18.

Table 18. Adjusted POMS scores for subjects with and without the history of cGVHD.

N=51	**Yes** (N=8)		**No** (N=43)		**Difference of means**	***P***
	Mean	SE	Mean	SE	Difference [95% CI]	
Tension	10.60	1.4	8.12	0.6	-2.47 [-5.53, 0.59]	0.111
Depression	15.29	2.1	7.06	0.9	-8.22 [-12.86, -3.59]	**0.001**
Anger	10.78	1.8	7.72	0.8	-3.07 [-7.09, 0.96]	0.132
Vigor	15.23	2.1	19.33	0.9	+4.10 [-0.59, 8.79]	0.085
Fatigue	9.60	1.5	6.16	0.6	-3.44 [-6.72, -0.16]	**0.040**
Confusion	9.99	1.3	6.94	0.6	-3.05 [-6.01, -0.09]	**0.044**
TMD	40.95	6.7	16.58	2.9	-24.37 [-39.12, -9.63]	**0.002**

An analysis of potential interactions between *sex* and the other independent variables was performed. No significant interactions were found, including *allo*sex*. Among the subjects who received allo-BMT, the interactions *sex*relapse* and *sex*cGVHD* were not significant. However, interactions between *sex* and *TBI* are significant or borderline significant (see Table 19 and Figure 5) for multiple POMS scores. The interaction plots show that POMS scores are higher with TBI recipients in females, with the opposite effect in males for Tension, Depression, Confusion, and TMD. For Vigor, female scores went down with TBI, and male scores slightly increased.

Table 19. Significance values for multivariable linear regression with *sex* and *TBI* interaction in the model for POMS scores among subjects who received allo-BMT.

N=51	**POMS outcome**						
	Tension	Depression	Anger	Vigor	Fatigue	Confusion	TMD
cGVHD	.069	**.000**	.123	**.040**	**.029**	**.024**	**.001**
sex	.231	.075	.502	.340	.374	.188	.079
TBI	.321	.655	.359	.633	.159	.273	.425
relapse	.768	.155	**.015**	.747	.197	.547	.143
*sex * TBI*	.198	*.060*	.726	**.048**	.305	.152	*.056*

Power

The original study proposal estimated a sample size of 488 participants. The anticipated ratio of non-allo-BMT to allo-BMT participants was 2:1 (see Table 20). Based on these initial assumptions, the power to detect a difference of one-third of one standard deviation (assuming equal variances) is 93% (Table 21; also see Methods section). With an enrollment of 51 allo-BMT participants, the estimated power is 52%. However, as stated in the study proposal and supported in the literature (Norman, Sloan et al. 2003), a difference of one-half of one standard deviation is considered clinically significant. In this situation, with 51 participants the power to detect a difference this size is 83%.

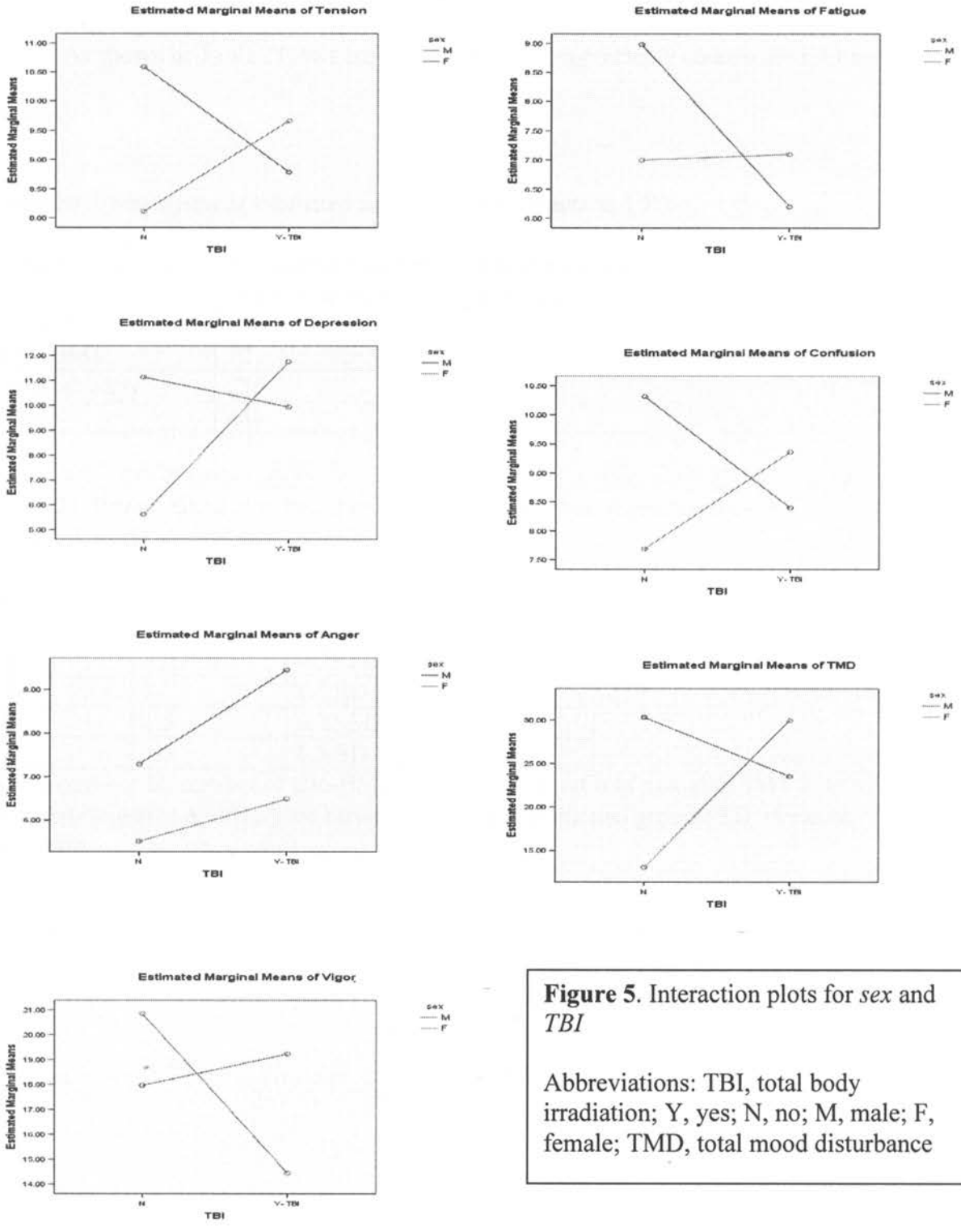


Figure 5. Interaction plots for *sex* and *TBI*

Abbreviations: TBI, total body irradiation; Y, yes; N, no; M, male; F, female; TMD, total mood disturbance

The actual ratio of non-allo-BMT to allo-BMT participants enrolled in this study is closer to 3:2. As shown in Table 21, this adjustment does not appreciably change the estimated power.

Table 20. Comparison of estimated and actual participants in L9704.

Treatment group	**Estimated number of participants**	**Actual number of participants**
Allo-BMT	160	51
Auto-BMT	72	14
Chemotherapy	256	59
Total	488	124

Table 21. Power calculation based on different study group characteristics and assumptions on clinical significance.

N	**k**	**Δ**	**Power**
160	2	0.33 SD	0.93
51	2	0.33 SD	0.52
51	2	0.5 SD	0.83
51	1.5	0.33 SD	0.47
51	1.5	0.5 SD	0.79

Abbreviations: N, number of allo-BMT participants; k, ratio of non-allo-BMT to allo-BMT participants; Δ, difference between the means in the two groups; SD, standard deviation

DISCUSSION

It is well established that allogeneic bone marrow transplant (allo-BMT) recipients are at significant risk for late effects (Oeffinger 2004). Although not as well studied, especially in pediatric populations, allo-BMT recipients may be at risk for a negative health related quality of life (HRQL) outcome in multiple domains (Schmidt 1993; Baker 1994; Andrykowski 1995; Matthes-Martin 1999). Furthermore, chronic graft versus host disease (cGVHD) has been shown to be associated with diminished HRQL (Baker 1994; Matthes-Martin 1999). Five studies among survivors of acute myeloid leukemia (AML) diagnosed in adulthood compared HRQL outcome among allo-BMT, autologous BMT (auto-BMT), and consolidation chemotherapy (CCT) treatment recipients (Molassiotis 1995; Lesko 1992; Zittoun 1997; Wellisch 1996; Stalfelt 1994). Three did not show a difference between allo-BMT and chemotherapy recipients, while two did, with a worse outcome in allo-BMT compared to auto-BMT and chemotherapy. Little is known about the HRQL outcome in pediatric AML survivors, and if there is a difference in HRQL outcome based on the type of post-remission treatment received.

The Children's Cancer Group (CCG) study L9704 was undertaken to study HRQL outcome in childhood AML survivors, and to compare between patients who received allo-BMT and those who did not. It was a cross-sectional study of childhood AML survivors alive $\geq$5 years from diagnosis who were enrolled in one of 4 CCG AML treatment protocols. The study utilized multiple HRQL instruments, including the Medical Outcomes Study Short-form 36 (SF-36), Profile of Mood States (POMS), and Harter Self-perception Profile. Based on the analysis presented here, there is no

difference in self-reported mood disturbance between the allo-BMT and chemotherapy/auto-BMT (non-allo-BMT) groups. Additionally, no differences were found in self-reported physical or mental health (Nicholson 2005), or with issues of self-esteem (unpublished data).

Parsons *et al.* (1999) looked at HRQL in childhood AML survivors and quantified the quality-adjusted survival for the first 60 months following diagnosis. When comparing subjects randomized to allo-BMT and CCT, they found that patients who received allo-BMT spent more time with severe treatment toxicity, but they experienced more time toxicity-free, relapse free, and alive. Based on utility threshold analysis, they found that the allo-BMT group had more quality-adjusted time than the other two modalities. The relative value a person assigns for the different health states of course differs, and this analysis may not apply for a given individual (e.g. if someone wants to diminish time spent with severe toxicity at all costs, thereby weighing this factor heavily, the utility-weighted quality-adjusted survival may favor CCT).

The literature does not give a consistent picture on what the childhood cancer survivor (CCS) can expect in terms of their likely psychological and psychosocial outcome. Some have shown that survivors are at significant risk for having adjustment issues, post-traumatic stress, and emotional difficulties, while others have found that these problems are not more common in CSS compared to the general population (Eiser 2000). However, many of these previous studies had a heterogeneous mix of cancer diagnoses, time elapsed since diagnosis, age of subjects at diagnosis and age when tested. Sample sizes

were often small, and different instruments were used. Although the sample size in this study was not large (N=124), it has the advantage of focusing on a single cancer diagnosis, with the comparison groups coming from the same population. Although there was a heterogeneous mix of ages and time points from diagnosis in this study, these factors were examined and controlled for in statistical analysis, and not found to be significant. Time since allo-BMT has been shown to be significantly associated with HRQL in some, but not all previous studies (Andrykowski 1995; Bush 1995; Matthes-Martin 1999).

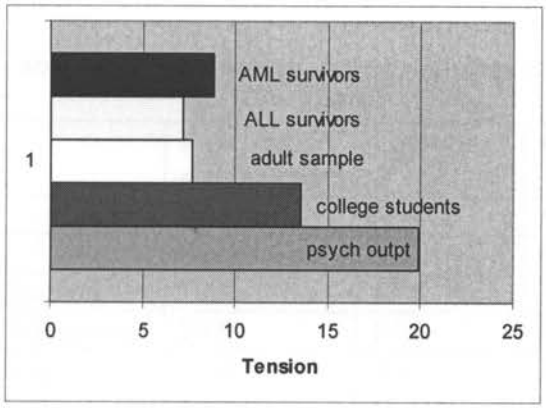
Among the allo-BMT survivors, a history of the diagnosis of cGVHD was associated with increased mood disturbance in multiple domains (depression, fatigue, confusion, and total mood disturbance). This association is consistent with multiple studies in the past, including one of the largest in pediatric BMT recipients (Matthes-Martin 1999). The symptoms of cGVHD vary widely, can affect any organ system, and result in impairment in any functional domain. For example, past cGVHD may result in physical impairment, easy fatigue, issues with self-image, and problems with sexuality. Active treatment of cGVHD with immunosuppressant drugs may interfere with social functioning.

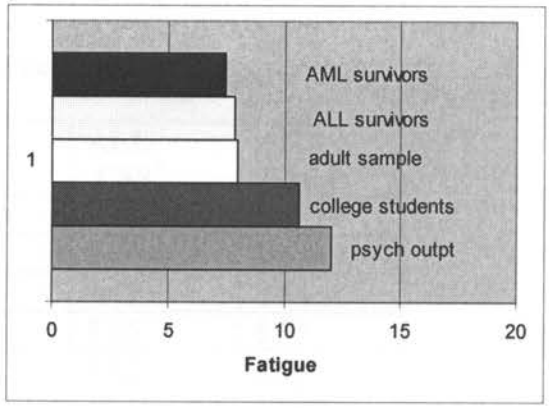
An interesting interaction between sex and exposure to total body irradiation (TBI) was found among the allo-BMT recipients. TBI itself was not associated with POMS scores, but there was a consistent trend that females who received TBI had higher scores than females who did not. In males, the opposite trend was observed. Radiation effects on reproductive organs are sex, age, and dosage dependent. Females are at greater risk of

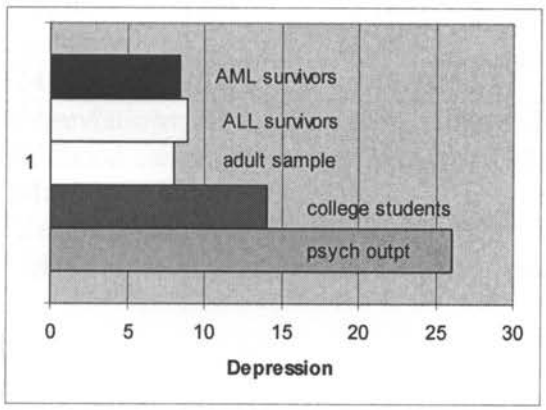
gonadal failure requiring hormone replacement compared to males (Pizzo and Poplack 2002), and this may explain this finding. Although preparative regimens utilizing TBI are no longer favored in AML therapy, this finding may have implications in other diseases.

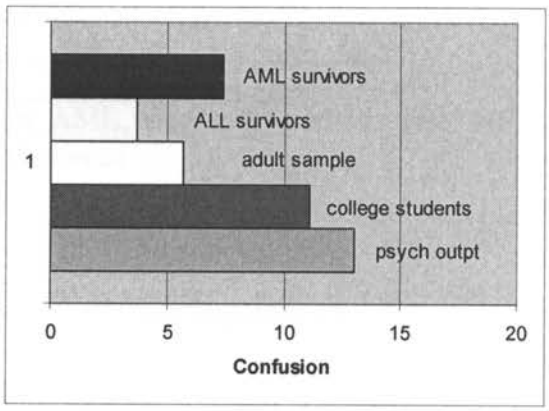
When this sample was analyzed as three groups (allo-BMT, auto-BMT, and chemotherapy), the anger score was associated with the type of post-remission therapy received. For reasons that are not easily explained, the anger score was significantly higher in the auto-BMT group than the chemotherapy group (P=0.029), and marginally significantly higher than the allo-BMT group (P=0.057). The unadjusted mean score among auto-BMT recipients (N=14) was 13.2, which is higher than the adult normative sample, but still below that for psychiatric outpatients. Whether or not auto-BMT recipients truly have more problems with anger or other mood disturbances is an interesting question. However, in the case of treatment for childhood AML, auto-BMT is no longer recommended for use in first remission.

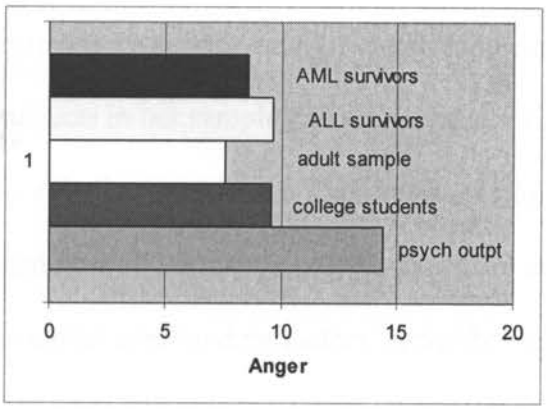
In addition to the finding that POMS outcome is not associated with the type of post-remission treatment received, this study also provides the opportunity to compare childhood AML survivors score with other sample populations (Figure 6, Table 22). The subjects in our sample compared similarly to those in the adult normative sample, and compared favorably to college and psychiatric outpatient samples. The psychiatric outpatient sample was obtained from a series of consecutive initial visits to a medical

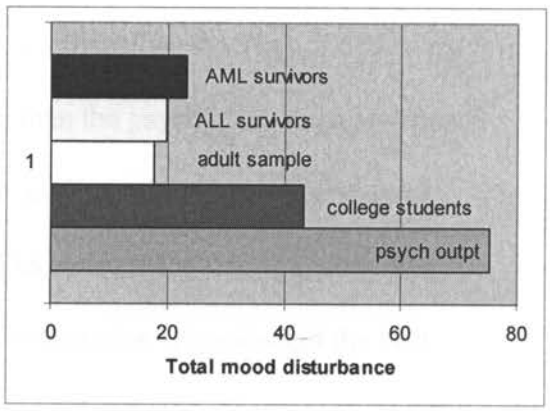












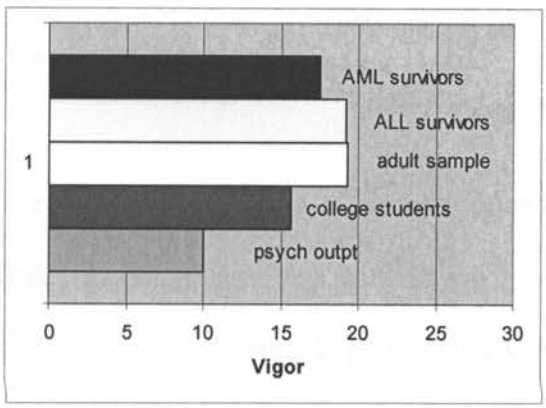


Figure 6. POMS scores for normative groups and leukemia survivors.

Table 22. POMS scores for normative groups and pediatric leukemia survivors.

	Psychiatric outpatients[1]	**College**[1]	**Adult**[1]	**ALL CSS**[2]	**AML CSS**
	N=1000	N=856	N=400	N=580	N=124
	Mean	Mean	Mean	Mean	Mean*
Tension	19.9	13.5	7.7	7.17	8.84
Depression	26.0	14.1	8.0	8.95	8.43
Anger	14.4	9.6	7.6	9.62	8.56
Vigor	10.0	15.6	19.3	19.19	17.48
Fatigue	12.0	10.6	8.0	7.87	7.46
Confusion	13.0	11.1	5.7	3.67	7.38
TMD	75.3	43.3	17.7	19.9	23.21

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; CSS, childhood cancer survivor; TMD, Total Mood Disturbance

[1] McNair and Heuchert 2005

[2] Zeltzer, Chen et al 1997

* Mean score of study group, adjusted for sex

center psychiatry clinic in the Eastern U.S in the late 1960s that excluded alcoholics, extreme psychotics, and those requiring emergency admissions (McNair 2005). Very few subjects in our sample had scores equal or greater than the psychiatric outpatient mean (4.8% of subjects had a TMD score $\geq$75.3). There are limitations in comparing our sample with normative groups, as we are not able to account for sex, age, and other potential confounding factors. Nevertheless, this information suggests that the vast majority of long term childhood AML survivors are not suffering from an excess of mood disturbance symptoms.

The POMS instrument was used in a prior CCG study that compared acute lymphoblastic leukemia (ALL) survivors with sibling controls (Zeltzer 1997). In that study the POMS scores were significantly higher in survivors than controls for tension, depression, anger, confusion, and total mood disturbance. But again, scores in the ALL survivors and

controls were similar to the adult normative sample, and well below the scores for psychiatric outpatients (Table 22). Additionally, childhood ALL and AML survivors fare similarly with respect to mood disturbance.

As just described, the mean POMS scores for the allo-BMT and non-allo-BMT groups were not statistically different, and well within the range of the normal adult population. Thus, although there isn't a compelling need to consider this issue, another way of looking at the "clinical significance" of an observed difference between two treatment groups is the effect size (Cohen 1988). By convention, effect size of 0.2 to 0.49 is considered "small," 0.5 to 0.79 as "medium," and 0.8 or higher as "large." Further, it has been suggested that an effect size of 0.5 corresponds to the minimum difference considered clinically significant (Norman 2003). By this rule, none of the differences between the two groups are clinically significant, reinforcing the conclusions above.

Strengths and limitations

One of the strengths of this study is that it is composed of subjects enrolled and treated on successive Children's Cancer Groups studies. Open from 1979 to 1995, these studies had similar eligibility criteria and drew from the same population. All utilized a "biological randomization" scheme to assign patients for post-remission chemotherapy. In other words, if the subject had a fully matched sibling or relative, they would be assigned to undergo allo-BMT. Otherwise, they would be assigned to the non-allo-BMT arm that utilized chemotherapy or auto-BMT. Randomization limits selection bias based on (perceived) disease severity at the time of diagnosis or based on response to treatment.

Treatment given for induction, post-remission chemotherapy, and allo-BMT conditioning differed among the protocols. However, our hypothesis is that allo-BMT is the most important factor influencing HRQL outcome, subordinating these differences. Alternatively, we could have combined the allo-BMT and auto-BMT recipients, and compared them to the chemotherapy group. Although auto-BMT treatment is myeloablative, potentially resulting in longer hospitalization during treatment and incurring greater acute toxicity, recipients are not at risk for GVHD. Thus, once beyond the transplant procedure, auto-BMT recipients are more like the chemotherapy patients than allo-BMT recipients. However, acute treatment complications can have long term physical and psychological consequences. This factor was not considered in this study.

There was very good correlation between SF-36 and POMS scores. The SF-36 mental summary score combines questions that measure the subjects' perceptions on vitality, social functioning, emotional adaptation, and mental health (e.g. anxiety, depression). The statistically significant and strong negative correlation (except for vigor, which was positively correlated) for this pair of instrument scores demonstrates internal consistency in this study. Furthermore, both instruments have been extensively used and shown to have good reliability and validity (McNair 2005; Ware 1998). Taken together, this suggests that the findings reported here, too, are reliable and valid.

When originally conceived, this study was estimated to have a 90-95% power to detect a very small difference between the treatment groups. There were unforeseen issues with recruitment, and the statistical power of this study with the actual number of subjects

enrolled was around 50%. However, if a difference in the means of at least one-half of a standard deviation is considered "clinically significant" then this study actually has a power of 83% to detect such a difference if it truly exists.

This was a cross-sectional study at a single point in time. For the POMS questionnaire, the respondent is asked to quantify how they have been feeling with respect to each descriptive term during the past week. Because mood can be labile, one-time testing may inaccurately estimate the actual burden of mood disturbance in the survivor population. This error should occur equally between the two groups. Misclassification bias of the outcome could lead to a null result when a difference between the groups actually exists. Although the POMS instrument has been used extensively in many different populations (McNair 2005), as commented upon by Eiser, perhaps complex issues such as psychological state cannot be "reduced to standardized questionnaires" (Eiser 2002). Rather, qualitative methods and in-depth interviews may be preferable for this kind of research.

In the four CCG treatment protocols, subjects with HLA-matched siblings who were assigned to the allo-BMT regimen may have refused the transplant procedure, in which case they would be treated per the chemotherapy arm of the protocol. In CCG-2891, the refusal rate was 18%, which was lower than prior CCG AML studies (Woods 2001). In the study reported here, patients designated as "allo-BMT" did in fact receive a BMT. Thus, this analysis is not "intent to treat," but "as treated" in this regard. Because of legitimate concerns for selection bias when analyzed "as treated" (particularly in bone

marrow transplant studies), the CCG *treatment* studies were appropriately analyzed by "intent to treat" (Chen 2002). In this study, however, there is no reason to be concerned about selection bias with regard to the decision to undergo allo-BMT or not, as the hypothesis is that the treatment received (allo-BMT or not), and not disease factors or family/physician preference during treatment, is associated with HRQL outcome. Therefore, correct assignment of the exposure is paramount, as the statistical advantage is clearly an increased power to detect a difference if one exists.

It needs to be noted, however, that if a patient relapsed and subsequently received either an autologous or allo-BMT, this information was not captured for this study, and misclassification of exposure may have occurred. Because the number of subjects who relapsed was low (N=18), it is unlikely that this would alter the results, as illustrated in Table 23. Scheme A shows the results based on the original assignments. Scheme B shows the results when the 18 subjects who relapsed were not included in the analysis. In Scheme C, it was assumed that all patients who relapsed received an allo-BMT (either initially, or after relapse). Out of the 18 subjects who relapsed, 11 were in the non-allo-BMT treatment group; thus, 11 subjects "crossed-over" from the non-allo-BMT group to the allo-BMT group. The scores did not change appreciably among schemes, and mean scores between allo-BMT and non-allo-BMT subjects were not significantly different within a given scheme.

Clearly there is a potential selection bias concern, as participation in this study was voluntary. This was a self-selected sample, and we do not know the characteristics of the

Table 23. POMS scores by post-remission therapy (univariate analysis) for different assumptions for subjects who relapsed

		Scheme A	Scheme B	Scheme C
Tension	Allo-BMT	8.51	8.68	8.64
	Non-allo-BMT	9.07	8.82	8.64
Depression	Allo-BMT	8.35	9.07	8.63
	Non-allo-BMT	8.49	8.24	8.24
Anger	Allo-BMT	8.21	8.89	8.72
	Non-allo-BMT	8.81	8.41	8.41
Vigor	Allo-BMT	18.69	18.30	17.83
	Non-allo-BMT	16.63	17.12	17.12
Fatigue	Allo-BMT	6.70	7.13	7.30
	Non-allo-BMT	7.99	7.62	7.62
Confusion	Allo-BMT	7.42	7.39	7.51
	Non-allo-BMT	7.36	7.26	7.26
TMD	Allo-BMT	20.48	22.89	23.18
	Non-allo-BMT	25.08	23.24	23.24

Scheme A	Treatment groups as assigned (original analysis)
Scheme B	Analysis without subjects who relapsed
Scheme C	Analysis assuming that all relapse subjects underwent an allo-BMT procedure (primary treatment, and/or at relapse)

Abbreviations: allo-BMT, allogeneic bone marrow transplant; TMD, Total Mood Disturbance

group that didn't participate. One can speculate that eligible survivors who did not participate were faring worse than those that did. If, as we hypothesized, allo-BMT recipients have a worse HRQL than non-allo-BMT recipients, this study may underestimate the burden of mood disturbance more so in the allo-BMT group. Thus, self-selection bias may be preventing us from detecting a true difference between the groups. Although it isn't possible to test the population that did not participate, there is some suggestion that this form of selection bias did not occur. If the risk of mood disturbance is actually higher in allo-BMT recipients, and mood disturbance decreased the likelihood of participation in this kind of study, then a lower proportion of allo-BMT recipients would be expected. However, the proportion of allo-BMT recipients who

participated in this study is greater than the proportion predicted (51/124 vs. 160/488). Further, there did not appear to be this form of selection bias among L9704 enrollees, as the POMS questionnaire completion rate was higher in the allo-BMT group (51/54 vs. 73/152).

This study had a limited ability to control for potential confounding factors. Socioeconomic factors such as employment status, household income, and education may be associated with mood outcome in this study population. The latter two were associated with depression, and all three were associated with somatic distress in multivariate analysis in childhood leukemia and lymphoma survivors (Zebrack, 2002). This information on our subjects was captured with the Childhood Cancer Survivor Study questionnaire and can be incorporated into these results when these data become available.

The incidence of cGVHD was relatively low in this group of long-term survivors (15.7%). This may explain why a difference in HRQL outcome was not seen between the groups (i.e. if the rate of cGVHD were higher, HRQL may have been significantly worse in the allo-BMT group compared to the non-allo-BMT group). It should be noted that the variable used in this analysis for cGVHD was taken from the treatment study data collection form. Therefore, it does not indicate if a subject was having active symptoms of cGVHD at the time of POMS testing, however, it is very unlikely that a patient would develop cGVHD beyond the follow-up data collection period (approximately 5 years from the time of enrollment), or that the presence of cGVHD would go unrecorded.

SUMMARY AND CONCLUSIONS

Compared to "hard' outcomes such as event-free survival and specific late effects, HRQL outcomes may seem ambiguous or indeterminate. I believe that a direct comparison such as this misses the point about the utility of HRQL research. Information on "hard" and "soft" outcomes complement each other and give the patient, family, and clinician valuable information about the true benefit of a particular therapy. As important as it is to know the relapse rate and chance that a survivor will have a particular late effect, ultimately it is the patients' own perception of their life that matters the most. Based on this study, it is reassuring to know that childhood AML survivors are not having excess problems with mood disturbance compared to their peers, and there is no difference in mood disturbance whether one received an allo-BMT or not. Therefore, this aspect of HRQL need not factor into the decision process for the type of post-remission treatment a patient receives.

REFERENCES

Andrykowski, M. A., C. B. Greiner, et al. (1995). "Quality of life following bone marrow transplantation: findings from a multicentre study." Br J Cancer **71**(6): 1322-9.

Baker, F., J. R. Wingard, et al. (1994). "Quality of life of bone marrow transplant long-term survivors." Bone Marrow Transplant **13**(5): 589-96.

Bush, N. E., M. Haberman, et al. (1995). "Quality of life of 125 adults surviving 6-18 years after bone marrow transplantation." Soc Sci Med **40**(4): 479-90.

Chen, A. R., T. A. Alonzo, et al. (2002). "Current controversies: which patients with acute myeloid leukaemia should receive a bone marrow transplantation?--an American view." Br J Haematol **118**(2): 378-84.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Creutzig, U., J. Ritter, et al. (2001). "Improved treatment results in high-risk pediatric acute myeloid leukemia patients after intensification with high-dose cytarabine and mitoxantrone: results of Study Acute Myeloid Leukemia-Berlin-Frankfurt-Munster 93." J Clin Oncol **19**(10): 2705-13.

Eiser, C., J. J. Hill, et al. (2000). "Examining the psychological consequences of surviving childhood cancer: systematic review as a research method in pediatric psychology." J Pediatr Psychol **25**(6): 449-60.

Feeny, D., W. Furlong, et al. (1999). "A framework for assessing health-related quality of life among children with cancer." Int J Cancer Suppl **12**: 2-9.

Gray, R. E., B. D. Doan, et al. (1992). "Psychologic adaptation of survivors of childhood cancer." Cancer **70**(11): 2713-21.

Greenberg, D. B., A. B. Kornblith, et al. (1997). "Quality of life for adult leukemia survivors treated on clinical trials of Cancer and Leukemia Group B during the period 1971-1988: predictors for later psychologic distress." Cancer **80**(10): 1936-44.

Haupt, R., T. R. Fears, et al. (1994). "Educational attainment in long-term survivors of childhood acute lymphoblastic leukemia." JAMA **272**(18): 1427-32.

Hudson, M. M., A. C. Mertens, et al. (2003). "Health status of adult long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study." JAMA **290**(12): 1583-92.

Institute of Medicine (2003). Childhood Cancer Survivorship. Washington, DC, The National Academy Press.

Koot, H. M. and J. L. Wallander (1991). Quality of life in child and adolescent illness: concepts, methods and findings. New York, Taylor & Francis.

Langeveld, N. E., H. Stam, et al. (2002). "Quality of life in young adult survivors of childhood cancer." Support Care Cancer **10**(8): 579-600.

Lansky, S. B., M. A. List, et al. (1986). "Psychosocial consequences of cure." Cancer **58**(2 Suppl): 529-33.

Lesko, L. M., J. S. Ostroff, et al. (1992). "Long-term psychological adjustment of acute leukemia survivors: impact of bone marrow transplantation versus conventional chemotherapy." Psychosom Med **54**(1): 30-47.

Matthes-Martin, S., M. Lamche, et al. (1999). "Organ toxicity and quality of life after allogeneic bone marrow transplantation in pediatric patients: a single centre retrospective analysis." Bone Marrow Transplant **23**(10): 1049-53.

McNair, D. and J. Heuchert (2005). Profile of Mood States Technical Update, MHS.

Meadows, A. T., J. Gordon, et al. (1981). "Declines in IQ scores and cognitive dysfunctions in children with acute lymphocytic leukaemia treated with cranial irradiation." Lancet **2**(8254): 1015-8.

Molassiotis, A., B. J. Boughton, et al. (1995). "Comparison of the overall quality of life in 50 long-term survivors of autologous and allogeneic bone marrow transplantation." J Adv Nurs **22**(3): 509-16.

Nesbit, M. E., Jr., J. D. Buckley, et al. (1994). "Chemotherapy for induction of remission of childhood acute myeloid leukemia followed by marrow transplantation or multiagent chemotherapy: a report from the Childrens Cancer Group." J Clin Oncol **12**(1): 127-35.

Nicholson, H. S., T. Zhou, et al. (2005). "Quality of Life in Survivors of Childhood and Adolescent Acute Myelogenous Leukemia (AML) Does Not Differ by Treatment (Bone Marrow Transplant (BMT) vs. Chemotherapy). A Report from the Children's Oncology Group (COG)." Blood (ASH Annual Meeting Abstracts) 2005 106: Abstract 701.

Norman, G. R., J. A. Sloan, et al. (2003). "Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation." Med Care **41**(5): 582-92.

Oeffinger, K. C. and M. M. Hudson (2004). "Long-term complications following childhood and adolescent cancer: foundations for providing risk-based health care for survivors." CA Cancer J Clin **54**(4): 208-36.

Parsons, S. K., S. Gelber, et al. (1999). "Quality-adjusted survival after treatment for acute myeloid leukemia in childhood: A Q-TWiST analysis of the Pediatric Oncology Group Study 8821." J Clin Oncol **17**(7): 2144-52.

Pizzo, P. A. and D. G. Poplack (2002). "Principles and Practice of Pediatric Oncology."

Pollock, B. H. (1999). "Obstacles and opportunities for the use of health-related quality-of-life assessment in pediatric cancer clinical trials (discussion)." Int J Cancer Suppl **12**: 151-3.

Ravindranath, Y., A. M. Yeager, et al. (1996). "Autologous bone marrow transplantation versus intensive consolidation chemotherapy for acute myeloid leukemia in childhood. Pediatric Oncology Group." N Engl J Med **334**(22): 1428-34.

Redaelli, A., J. M. Stephens, et al. (2004). "Short- and long-term effects of acute myeloid leukemia on patient health-related quality of life." Cancer Treat Rev **30**(1): 103-17.

Robison, L. L., A. C. Mertens, et al. (2002). "Study design and cohort characteristics of the Childhood Cancer Survivor Study: a multi-institutional collaborative project." Med Pediatr Oncol **38**(4): 229-39.

Schmidt, G. M., J. C. Niland, et al. (1993). "Extended follow-up in 212 long-term allogeneic bone marrow transplant survivors. Issues of quality of life." Transplantation **55**(3): 551-7.

Seitzman, R. L., D. A. Glover, et al. (2004). "Self-concept in adult survivors of childhood acute lymphoblastic leukemia: a cooperative Children's Cancer Group and National Institutes of Health study." Pediatr Blood Cancer **42**(3): 230-40.

Sobo, E. J. and P. S. Kurtin (2003). Child health services research: applications, innovations, and insights. San Francisco, Jossey-Bass.

Stevens, M. C., H. Mahler, et al. (1998). "The health status of adult survivors of cancer in childhood." Eur J Cancer **34**(5): 694-8.

Stevens, R. F., I. M. Hann, et al. (1998). "Marked improvements in outcome with chemotherapy alone in paediatric acute myeloid leukemia: results of the United Kingdom Medical Research Council's 10th AML trial. MRC Childhood Leukaemia Working Party." Br J Haematol **101**(1): 130-40.

Veenstra, K. M., M. A. Sprangers, et al. (2000). "Quality of life in survivors with a Van Ness-Borggreve rotationplasty after bone tumour resection." J Surg Oncol **73**(4): 192-7.

Ware, J. E., Jr. and B. Gandek (1998). "Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) Project." J Clin Epidemiol **51**(11): 903-12.

Watson, M., G. Buck, et al. (2004). "Adverse impact of bone marrow transplantation on quality of life in acute myeloid leukaemia patients; analysis of the UK Medical Research Council AML 10 Trial." Eur J Cancer **40**(7): 971-8.

Watson, M., K. Wheatley, et al. (1999). "Severe adverse impact on sexual functioning and fertility of bone marrow transplantation, either allogeneic or autologous, compared with consolidation chemotherapy alone: analysis of the MRC AML 10 trial." Cancer **86**(7): 1231-9.

Wellisch, D. K., J. Centeno, et al. (1996). "Bone marrow transplantation vs. high-dose cytorabine-based consolidation chemotherapy for acute myelogenous leukemia. A long-term follow-up study of quality-of-life measures of survivors." Psychosomatics **37**(2): 144-54.

Wells, R. J., W. G. Woods, et al. (1994). "Treatment of newly diagnosed children and adolescents with acute myeloid leukemia: a Childrens Cancer Group study." J Clin Oncol **12**(11): 2367-77.

Woods, W. G. (2006). "Curing childhood acute myeloid leukemia (AML) at the half-way point: promises to keep and miles to go before we sleep." Pediatr Blood Cancer **46**(5): 565-9.

Woods, W. G., N. Kobrinsky, et al. (1993). "Intensively timed induction therapy followed by autologous or allogeneic bone marrow transplantation for children with acute myeloid leukemia or myelodysplastic syndrome: a Childrens Cancer Group pilot study." J Clin Oncol **11**(8): 1448-57.

Woods, W. G., S. Neudorf, et al. (2001). "A comparison of allogeneic bone marrow transplantation, autologous bone marrow transplantation, and aggressive chemotherapy in children with acute myeloid leukemia in remission." Blood **97**(1): 56-62.

Zebrack, B. J., L. Zeltzer, et al. (2002). "Psychological outcomes in long-term survivors of childhood leukemia, Hodgkin's disease, and non-Hodgkin's lymphoma: A report from the Childhood Cancer Survivor Study." Pediatrics **110**(1): 42-52.

Zeltzer, L. K., E. Chen, et al. (1997). "Comparison of psychologic outcome in adult survivors of childhood acute lymphoblastic leukemia versus sibling controls: a cooperative Children's Cancer Group and National Institutes of Health study." J Clin Oncol **15**(2): 547-56.

Zittoun, R., S. Suciu, et al. (1997). "Quality of life in patients with acute myelogenous leukemia in prolonged first complete remission after bone marrow transplantation (allogeneic or autologous) or chemotherapy: a cross-sectional study of the EORTC-GIMEMA AML 8A trial." Bone Marrow Transplant **20**(4): 307-15.

Appendix I. Late-effects associated with chemotherapy

Alkylating agents
There is a dose-related risk of gonadal dysfunction. Ovaries in prepubertal girls are more resistant to injury compared with adults. Most girls treated with conventional doses of alkylators will retain ovarian function, although higher doses, including those used in transplant preparative regimens, often result in ovarian failure. In boys, infertility is common following alkylator therapy. Leydig cells are more resistant to damage, and androgen replacement is uncommon.

Alkylator-associated pulmonary disease has been most commonly linked with busulfan, which has been used in transplant preparative regimens. Females are more susceptible than males. Lung damage is characterized by interstitial fibrosis and bronchopulmonary dysplasia.

Genitourinary tract complications are most commonly reported after administration of cyclophosphamide or ifosfamide. Exposure can cause hemorrhagic cystitis and renal damage. Long term monitoring of renal function, hypertension, and bone mineralization is suggested.

Secondary AML following alkylating agents usually occurs 5 to 7 years after exposure following a prodromal myelodysplastic phase. Certain agents are more potent leukemogens (e.g. nitrogen mustard), and the use of more contemporary agents such as cyclophosphamide has reduced the incidence of secondary AML.

Anthracyclines
Anthracyclines are well-known to cause late-onset cardiomyopathy. The risk is dose-related, increased when treated at a young age, in females, and when combined with chest radiation. An additional high-risk category is women in the latter stages of pregnancy. Although most survivors with echocardiographic evidence of dysfunction are asymptomatic, a significant proportion will progress to congestive heart failure (CHF). In one study of 607 children treated with anthracycline, the cumulative incidence of CHF was 4.8% after 15 years (Kremer, van Dalen et al. 2001).

Corticosteroids
Corticosteroids are associated with a wide array of acute toxicities, as well as potential late-effects. The etiology of decreased bone mineral density in survivors is multifactorial. Steroids interfere with bone metabolism, thus reducing the peak bone mass attained. It is associated with avascular necrosis of the femoral head, which can lead to debilitating hip pain and loss of function.

Epipodophyllotoxins
Etoposide can cause a distinctive secondary AML characterized by brief time of onset and a chromosomal translocation involving the MLL gene.

Appendix II.

Profile of Mood States

Subject's Initials _____
Birth date _____
Date _____
Subject Code No. _____

Directions: Describe HOW YOU FEEL RIGHT NOW by checking one space after each of the words listed below:

FEELING	Not at all	A little	Mod.	Quite a bit	Extremely
Friendly	1	2	3	4	5
Tense	1	2	3	4	5
Angry	1	2	3	4	5
Worn Out	1	2	3	4	5
Unhappy	1	2	3	4	5
Clear-headed	1	2	3	4	5
Lively	1	2	3	4	5
Confused	1	2	3	4	5
Sorry for things done	1	2	3	4	5
Shaky	1	2	3	4	5
Listless	1	2	3	4	5
Peeved	1	2	3	4	5
Considerate	1	2	3	4	5
Sad	1	2	3	4	5
Active	1	2	3	4	5
On edge	1	2	3	4	5
Grouchy	1	2	3	4	5
Blue	1	2	3	4	5
Energetic	1	2	3	4	5
Panicky	1	2	3	4	5
Hopeless	1	2	3	4	5
Relaxed	1	2	3	4	5
Unworthy	1	2	3	4	5
Spiteful	1	2	3	4	5
Sympathetic	1	2	3	4	5
Uneasy	1	2	3	4	5
Restless	1	2	3	4	5
Unable to concentrate	1	2	3	4	5
Fatigued	1	2	3	4	5

Helpful	1	2	3	4	5
Annoyed	1	2	3	4	5
Discouraged	1	2	3	4	5
Resentful	1	2	3	4	5
Nervous	1	2	3	4	5
Lonely	1	2	3	4	5
Miserable	1	2	3	4	5
Muddled	1	2	3	4	5
Cheerful	1	2	3	4	5
Bitter	1	2	3	4	5
Exhausted	1	2	3	4	5
Anxious	1	2	3	4	5
Ready to fight	1	2	3	4	5
Good-natured	1	2	3	4	5
Gloomy	1	2	3	4	5
Desperate	1	2	3	4	5
Sluggish	1	2	3	4	5
Rebellious	1	2	3	4	5
Helpless	1	2	3	4	5
Weary	1	2	3	4	5
Bewildered	1	2	3	4	5
Alert	1	2	3	4	5
Deceived	1	2	3	4	5
Furious	1	2	3	4	5
Effacious	1	2	3	4	5
Trusting	1	2	3	4	5
Full of pep	1	2	3	4	5
Bad-tempered	1	2	3	4	5
Worthless	1	2	3	4	5
Forgetful	1	2	3	4	5
Carefree	1	2	3	4	5
Terrified	1	2	3	4	5
Guilty	1	2	3	4	5
Vigorous	1	2	3	4	5
Uncertain about things	1	2	3	4	5
Bushed	1	2	3	4	5