Childhood Cancer Survivor Study
Analysis Concept Form

I. Title:

Exploring latent clusters of survivors using the BSI-18: A Childhood Cancer Survivor Study

II. Working Group:

The study will be within the Psychology Working Group. Proposed investigators will be:

Norma D'Agostino	(norma.dagostino@uhn.ca)
Kim Edelstein	(kim.edelstein@uhn.ca)
Christopher Recklitis	(Christopher_recklitis@dfci.harvard.edu)
Wendy Leisenring	(wleisenr@fhcrc.org)
Tara Brinkman	(tara.brinkman@stjude.org)
Greg Armstrong	(greg.armstrong@stjude.org)
Les Robison	(les.robison@stjude.org)
Kevin Krull	(kevin.krull@stjude.org)

III. Background and Rationale:

Survival rates in pediatric cancer have improved dramatically over the last three decades, leading to a growing population of adult survivors of childhood cancers. These survivors are at risk for physical, neurocognitive, and psychological late effects as a result of their disease and its treatment. Risk-based long-term monitoring of childhood cancer survivors for late effects has been recommended and is being instituted in multiple sites across North America and Europe [1].

Current guidelines highlight screening for psychological distress as an important part of comprehensive follow-up care of survivors [1]. While most survivors seem to adjust well, cross-sectional studies have identified subgroups of survivors experiencing significant levels of distress. For example, life-time prevalence rates of post-traumatic stress in childhood cancer survivors range from 20.5 to 35% [2, 3] and 13.9% of them report suicidal ideation [4].

The Childhood Cancer Survivor Study (CCSS) was designed to examine health outcomes over time in a large cohort of childhood cancer survivors in North America. The original cohort included 20,346 childhood cancer survivors diagnosed between 1970 and 1986 who were at least five years post-diagnosis, and approximately 4,000 siblings of survivors who serve as the comparison group. To account for changes in cancer treatment protocols, efforts are currently underway to follow an expanded cohort of roughly 15,000 survivors diagnosed between 1987 and 1999 and an additional 4,000 siblings.

The 18-item Brief Symptom Inventory [5] is used in the CCSS to screen for psychological distress symptoms experienced in the previous seven days. The BSI-18 is a self-report symptom checklist specifically designed to screen for distress and psychiatric disorders in medical and community populations. It consists of three 6-item subscales measuring anxiety, depression, and somatization, and also provides an overall summary score, the Global Severity Index (GSI). Recklitis et al [6] confirmed the 3-factor structure of the BSI-18 and its usefulness in assessing psychological distress in this population with data from the original CCSS cohort.

Survivors with a variety of cancer diagnoses are more likely to report elevated levels of overall distress as compared to siblings [7, 8] or population norms [9, 10]. Treatment factors such as radiation [11-14] and chemotherapy including alkylating agents and anthracyclines [12, 13, 15, 16] are related to the level and type of distress reported (for detailed review see Zeltzer 2009 [7]). In addition, reports of current poor physical health are associated with elevated global distress scores [7, 13]. Sociodemographic factors that contribute to overall levels of distress include female gender [8, 13], lower household income [10, 13, 17], lower educational attainment [13], unemployment [17], and not being married or in a significant relationship [7, 9, 13].

These associations have also been explored in terms of the specific types of distress reported. Increased somatization, depression, and anxiety have each been linked with diagnosis [7, 18, 19], treatment [11-16, 20], poor health status [8, 11, 13, 14, 21], unemployment [11, 13, 17, 21], low

income [11, 13, 17], and female gender [8-10, 13]. Somatization has also been linked with cancer pain [21] and older age [10, 13]. In terms of affective distress, both depression and anxiety have been associated with marital status [13, 14, 22] lower education [13, 14, 17], and Caucasian race [13]. Depression has also been linked to cancer pain [22], and fatigue and sleep disturbances [23].

A study examining distress in siblings in the CCSS cohort conducted by Buchbinder and colleagues [24] reported elevated GSI scores in 3.8% of the sample and even smaller proportions with elevated subscale scores (ranging from 1.1% for anxiety to 3.3% for depression). Risk factors identified for siblings include lower household income, being unmarried, lower education, unemployment, and self-reported poor health. In addition, survivor variables associated with elevated distress in siblings include the presence of a chronic health condition and psychological distress in the survivor. Younger sibling age at the time of second malignancy diagnosis or death of the survivor was also associated with increased distress in siblings. Being younger than the survivor and being the sibling of a male survivor was found to increase siblings' risk for distress.

Recent longitudinal findings from the CCSS indicate that while most survivors report consistently low levels of distress, some survivors report persistently elevated levels of distress [22]. Additionally, groups of survivors were identified with increasing levels of distress over time, while others displayed progressive reduction in distress over time. This four-group pattern of longitudinal distress was observed to be remarkably similar for symptoms of depression, anxiety and somatization independently. However, patterns of BSI-18 scores within individuals have not been examined. For example, whether the individuals who report high depression scores are the same as those who report high anxiety or somatization scores is not known. Examination of these patterns in adult survivors of childhood cancers from the CCSS may allow for identification of subgroups of individuals who report the highest levels of comorbid distress, and the variables that contribute to that distress. Identification of such comorbidity has important implications for treatment strategies.

Using latent profile analysis, we propose to identify subgroups based on patterns of single and comorbid distress in subscales from the BSI-18, and to identify which disease, treatment, and sociodemographic variables predict those patterns. By comparing clusters identified in the expanded cohort of survivors to those in the original cohort and to the siblings, we may be able to determine the extent to which changes to treatment protocols impact risk for and pattern of distress. Ultimately, we hope the results of this analysis will allow us to identify those survivors who are most in need of support, so that we can direct resources and develop specific interventions based on symptom profiles.

IV. Objective/Specific aims/Research Hypotheses:

Objective:

Proposed specific aims:

Aim 1. To identify latent clusters of survivors based on patterns on the three BSI-18 subscales (anxiety, depression, somatization) in a random sample of half the original and half of the expanded cohort of survivors. The remainder of the two cohorts will be used as the validation sample. We hypothesize that the following clusters of survivors would be identified:

1. Survivors with low distress on all three BSI-18 scores (well-adjusted)
2. Survivors with high distress on all three BSI-18 scores (highly distressed)
3. Survivors with predominantly high somatization relative to anxiety or depression (somatic distress)
4. Survivors with predominantly high anxiety or depression relative to somatization (affective distress)

Our predictions are based on a conceptual distinction between physical (somatic) versus affective (anxiety and depression) symptoms of distress. An evaluation of the BSI-18 using item response theory to determine which items are most strongly related to overall psychological distress in a variety of samples (students, clinical unipolar and bipolar mood disorder outpatients, prisoners) was conducted by Meijer, de Vries, and Bruggen [25]. Their findings suggest that in their clinical and prisoner samples the overall construct of psychological distress is particularly defined by the depression and anxiety items and less by the somatization items. They also found that the somatization scale did not discriminate between GSI scores. However, in survivors of childhood

cancer, Recklitis and colleagues [6] found that the somatization factor was most highly correlated with overall distress, followed by depression and then anxiety, consistent with our hypothesis that somatic distress may be a key differentiating factor between subgroups of cancer survivors. Regardless of our predictions, our analyses will allow for identification of our hypothesized clusters or any others, and the factors that distinguish between them.

Aim 2. To compare the patterns of the latent BSI-18 clusters observed in the expanded cohort to the original cohort and siblings.

We hypothesize that similar clusters would be identified in all three populations but that the relative proportion of each cluster would differ across the three groups.

1. Consistent with previous reports demonstrating that few survivors or siblings endorse elevated symptoms of distress, we expect that most survivors and siblings will be well-adjusted, with low scores on all three subscales.
2. Because of issues related to intensity of treatment, we expect that the original cohort would have the highest proportion of individuals with high distress scores on all three BSI-18 scales, followed by the expanded cohort. Very few siblings will exhibit this pattern [24].
3. Similarly, due to treatment, we expect a larger proportion of the original cohort of survivors with predominantly high somatization relative to anxiety or depression (somatic distress), followed by the expanded cohort, and finally siblings.
4. Because of the impact of sociodemographic factors on affective distress (described above), we do not expect group differences in the proportion of survivors who endorse this pattern of symptoms.

Aim 3. To identify disease, treatment and sociodemographic predictors of latent BSI-18 clusters.

We hypothesize that there will be 3 classes of risk factors influencing the distress patterns identified in the latent cluster analysis, and that these will be similar to those previously reported to predict high distress on BSI-18 subscales:

a) Disease Variables—specific diagnosis, age at diagnosis, time since diagnosis, current perceived health status, pain

b) Treatment Variables— chemotherapy [yes/no] for specific agents (antimetabolites, anthracyclines, alkylating agents, corticosteroids); radiation (cranial [yes/no], non-cranial [yes/no])

c) Sociodemographic Variables—sex, race, current age, educational attainment, employment, relationship status, household income

Aim 4. To explore cluster frequency by sex.

V. Analysis Framework:

Population:
We propose to conduct our analysis on the original CCSS survivor cohort and their siblings in addition to the survivors in the new expanded cohort. The original and expanded cohorts of survivors will be combined and randomly divided into two samples. One sample will be used as the derivation cohort for the latent clusters, and the other will be used as the validation sample.

Subject population.

Survivors from the original and expanded cohort and siblings from the original cohort who self-completed the BSI-18 at Baseline are eligible for this analysis.

Outcome measures.

- BSI-18 subscales
  - Baseline Expansion Survivor survey items K1-K18
  - Baseline Original Survivor survey items J16-J35 (minus J25 and J28)
  - Baseline Original Sibling survey items J16-J35 (minus J25 and J28)
- Anxiety, depression and somatization T-scores

Covariates:

- Disease, treatment, and sociodemographic variables listed above.

Disease Variables:

D1. Diagnosis
D2. Age at diagnosis
D3. Time since diagnosis
D4.D4. Perceived health status (O21 expanded survivor baseline, N15 original survivor and sibling baseline). We will use health status and chronic health conditions (i.e., CTCAE grade) for the expanded cohort, if available
D5. Pain (J3, J4, J9 expanded survivor baseline; J6, J7, J13 original survivor and sibling baseline)

Treatment Variables:

T1. Chemotherapy (yes/no) for specific agents (antimetabolites, anthracyclines, alkylating agents, corticosteroids)
T2. Radiation (yes/no) for cranial, non-cranial other

Sociodemographic Variables:

S1. Sex (A2 baseline expanded, original, sibling)
S2. Race (A5, A5a for expanded baseline, A4, A4a for original and sibling baseline)
S3. Age at survey (A1 date of birth expanded, original, and sibling baseline and date of survey front cover)
S4. Educational attainment (R1-R2 expanded baseline, O1-O2 original and sibling baseline)
S5. Employment status (S1-S3a expanded baseline, O5-O7 original and sibling baseline) Only overlapping information is S1 and O5 which is a yes/no question "Have you ever had a job".
S6. Marital/relationship status (M2-M3 for expanded baseline, L1-L2 for original and sibling baseline). We will look at: Ever married (yes/no) and current marital status (married, living as married, widowed, etc.).
S7. Household income (T1& T3 expanded baseline, Q8-Q9 original and sibling baseline). Total household income less than 19999, 20000-39999, 40000-59999, over 60000) and personal income over the past year (none, less than 19999, 20000-39999, 40000-59999, over 600000)

Statistical Analyses:

Aim 1: Identify latent clusters

We will use latent profile analysis to identify clusters of survivors based on patterns on the three BSI-18 subscales (anxiety, depression, somatization) in a random sample of 50% of the original and 50% of the expanded cohort of survivors (Aim 1). We will use the average silhouette width to evaluate the quality of clustering by taking the compactness and separation of the clusters into account. The larger the silhouette width, the better the separation of cluster compactness. We will also examine the Gap statistic to assess separation between clusters.

The remaining 50% of the original and 50% of the expanded cohort of survivors will be used for a stability evaluation of the clusters. We will classify this cross-validation sample to the identified clusters by the nearest centroid method. We will then use the Adjusted Rand Index (ARI) to measure the reliability between cluster solutions (original cluster and derived cluster). ARI values range from -1 to 1, with the larger ARI demonstrating better agreement between partitions.

We expect that several classes will be identified (e.g. low distress on all three scores; high distress on all three scores; high somatization relative to anxiety or depression [somatic distress]; high anxiety or depression relative to somatization [affective distress]), but will not pre-specify a set number of classes. We will require a minimum cluster size of 5% of the cohort.

Aim 2: Compare cluster patterns

To compare the patterns of the latent BSI-18 clusters observed in the expanded cohort to the original cohort and siblings, the clusters identified in Aim 1 will be applied to the expanded survivor, original survivor and sibling cohorts independently. The proportion of each cohort within each cluster will then be tested for equivalency. To assure adequate power for subsequent analyses we will require that each latent class include no less than 5% of the respective cohort.

Aim 3: Examine predictors of cluster membership

Multinomial logistic regression will then be used to determine whether disease and treatment variables predict cluster membership in the expanded cohort. The disease, treatment and sociodemographic variables identified above will be used as predictors, with outcomes determined by the identified clusters identified in Aim1. We will examine collinearity among all predictors and select variables to be included based on reasonable groupings (e.g. adequate cell size). The possibility of conducting several separate logistic regression models (i.e., evaluate each cluster separately) will be explored based on interpretability of the multinomial regression results. Because this approach can underestimate the associations between covariates and classes we will use a multiple step maximum likelihood based correction model. We will also provide a figure showing the variability in patterns of individuals within each class to illustrate how well the class structure fits. This will provide evidence that the class grouping does not simply represent an average, but is actually representative of the majority of subjects in the class.

Aim 4. To explore cluster frequency by sex.

Because self-reported symptoms of anxiety and depression tend to be higher in females in most samples, we will also look at cluster frequency by gender. In order to do this we will need to explore the choice of cut-off scores for the BSI, in particular whether the choice of $T \geq 63$ is too high for males or females. In general, women report higher levels of depression on rating scales, and are more likely to be diagnosed with depression than men. The BSI cut off t-score is set at greater or equal to 63, meaning that the top 9% of men and women will be classified as depressed, regardless of their absolute depression scores. Consequently, this results in some men who are classified as depressed who actually report fewer depressed symptoms than some women who are classified as not depressed, and vice versa. An absolute cut-off score, by contrast, would specify a number of symptoms reported (or the appropriate matching non-gender adjusted t-score) and this would ensure that men and women would be classified as depressed/not depressed by the same level of depressed symptoms. In order to determine the best way to classify performance on each of the three BSI scales, we will compare unadjusted raw scores, sex-adjusted t-scores, and unadjusted t-scores to understand if and how the distribution changes with adjustment, and then come to a consensus of which method to use in the cluster analysis.

Table 1. Characteristics of participants who completed the BSI-18 at baseline in original cohort and expanded cohort

	Model Building		Validation Cohort		Siblings	
	N	%	N	%	N	%
Gender						
Male						
Female						
Race						
Caucasian						
Black						
Other						
Age at time of Survey						
18-29 yrs.						
30-39 yrs.						
40-54 yrs.						
Educational Attainment						
< High School						
High School						
College grad						
Employment Status						
Employed						
Student						
Unemployed						
Annual Household Income						
≤20000						
20001-59999						
≥60000						
Relationship Status						
Single						
Married/living as married						
Divorced/Separated/Widowed						
Health						
Good/very good/excellent						
Fair/Poor						
Pain						
Yes						
No						
Diagnosis						
Leukemia						
CNS tumor						
Hodgkin lymphoma						
Non-Hodgkin’s lymphoma						
Neuroblastoma						
Wilms Tumor						
Soft tissue sarcoma						
Bone tumor						
Age at cancer diagnosis						
0-3 yrs.						
4-9 yrs.						
10-14 yrs.						
15-20 yrs.						
Chemotherapy						
Antimetabolites						
Anthracyclines						
Alkylating agents						
Corticosteroids						
Radiation						
Cranial						
Other						
None						
Survival Time Years						
<20						

20-29					
30+					

Table 2: BSI-18 subscale scores for survivors and siblings. Data presented are group means ± standard error of the mean, range of scores and proportion T≥63 are in parentheses

Group		GSI	Depression	Anxiety	Somatization
Original Cohort	Survivors				
	Siblings				
Expanded Cohort	Survivors				

Table 3a: Latent clusters for distress in original cohort

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	M (SD)	M (SD)	M (SD)	M (SD)
Anxiety				
Depression				
Somatization				

Table 3b. Latent clusters for distress in expansion cohort

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	M (SD)	M (SD)	M (SD)	M (SD)
Anxiety				
Depression				
Somatization				

Table 3c. Latent clusters for distress in sibling cohort

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	M (SD)	M (SD)	M (SD)	M (SD)
Anxiety				
Depression				
Somatization				

Table 4: Proportion of cohort by latent cluster

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	N (%)	N (%)	N (%)	N (%)
Expansion Survivor				
Original Survivor				
Original Sibling				

Table 5. Multinomial logistic regression model of cluster membership prediction

	Cluster 1		Cluster 2		Cluster X	
	OR	95% CI	OR	95% CI	OR	95% CI
Gender						
Male						
Female						
Race						
Caucasian						
Black						
Other						
Age at time of Survey						
18-29 yrs.						
30-39 yrs.						
40-54 yrs.						
Educational Attainment						
< High School						
High School						
College grad						
Employment Status						
Employed						
Student						
Unemployed						
Annual Household Income						
≤20000						
20001-59999						
≥60000						
Relationship Status						
Single						
Married/living as married						
Divorced/Separated/Widowed						
Pain						
Yes						
No						
Health						
Good/very good/excellent						
Fair/Poor						
Diagnosis						
Leukemia						
CNS tumor						
Hodgkin lymphoma						
Non-Hodgkin’s lymphoma						
Neuroblastoma						
Wilms Tumor						
Soft tissue sarcoma						
Bone tumor						
Age at cancer diagnosis						
0-3 yrs.						
4-9 yrs.						
10-14 yrs.						
15-20 yrs.						
Chemotherapy						
Antimetabolites						
Anthracyclines						
Alkylating agents						
Corticosteroids						
Radiation						
Cranial						
Other						
None						
Survival Time Years						
<20						
20-29						
30+						

References

1. Landier, W., et al., *Development of risk-based guidelines for pediatric cancer survivors: the Children's Oncology Group Long-Term Follow-Up Guidelines from the Children's Oncology Group Late Effects Committee and Nursing Discipline.* J Clin Oncol, 2004. 22(24): p. 4979-90.
2. Bruce, M., *A systematic and conceptual review of posttraumatic stress in childhood cancer survivors and their parents.* Clin Psychol Rev, 2006. 26(3): p. 233-56.
3. Taieb, O., et al., *Posttraumatic stress symptoms after childhood cancer.* Eur Child Adolesc Psychiatry, 2003. 12(6): p. 255-64.
4. Recklitis, C., T. O'Leary, and L. Diller, *Utility of routine psychological screening in the childhood cancer survivor clinic.* J Clin Oncol, 2003. 21(5): p. 787-92.
5. Derogatis, L.R., ed. *BSI 18 Brief Symptom Inventory 18, Administration, Scoring, and Procedurals Manual.* 2000, NCS Pearson Minneapolis.
6. Recklitis, C.J., et al., *Factor structure of the brief symptom inventory--18 in adult survivors of childhood cancer: results from the childhood cancer survivor study.* Psychol Assess, 2006. 18(1): p. 22-32.
7. Zeltzer, L.K., et al., *Psychological status in childhood cancer survivors: a report from the Childhood Cancer Survivor Study.* J Clin Oncol, 2009. 27(14): p. 2396-404.
8. Gianinazzi, M.E., et al., *Adolescent survivors of childhood cancer: are they vulnerable for psychological distress?* Psychooncology, 2013. 22(9): p. 2051-8.
9. Wenninger, K., et al., *Coping in long-term survivors of childhood cancer: relations to psychological distress.* Psychooncology, 2013. 22(4): p. 854-61.
10. Kim, M.A. and J. Yi, *Psychological distress in adolescent and young adult survivors of childhood cancer in Korea.* J Pediatr Oncol Nurs, 2013. 30(2): p. 99-108.
11. Zeltzer, L.K., et al., *Psychosocial outcomes and health-related quality of life in adult childhood cancer survivors: a report from the childhood cancer survivor study.* Cancer Epidemiol Biomarkers Prev, 2008. 17(2): p. 435-46.
12. Schultz, K.A., et al., *Behavioral and social outcomes in adolescent survivors of childhood cancer: a report from the childhood cancer survivor study.* J Clin Oncol, 2007. 25(24): p. 3649-56.
13. Zebrack, B.J., et al., *Psychological distress in long-term survivors of solid tumors diagnosed in childhood: a report from the childhood cancer survivor study.* Pediatr Blood Cancer, 2007. 49(1): p. 47-51.
14. Zebrack, B.J., et al., *Psychological outcomes in long-term survivors of childhood brain cancer: a report from the childhood cancer survivor study.* J Clin Oncol, 2004. 22(6): p. 999-1006.
15. Hudson, M.M., et al., *Health status of adult long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study.* Jama, 2003. 290(12): p. 1583-92.
16. Zebrack, B.J., et al., *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, 2002. 110(1 Pt 1): p. 42-52.
17. Gurney, J.G., et al., *Social outcomes in the Childhood Cancer Survivor Study cohort.* J Clin Oncol, 2009. 27(14): p. 2390-5.
18. Ness, K.K., et al., *The impact of limitations in physical, executive, and emotional function on health-related quality of life among adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study.* Arch Phys Med Rehabil, 2008. 89(1): p. 128-36.
19. Nathan, P.C., et al., *Health-related quality of life in adult survivors of childhood Wilms tumor or neuroblastoma: A report from the childhood cancer survivor study.* Pediatr Blood Cancer, 2007. 49(5): p. 704-15.
20. Krull, K.R., et al., *Screening for neurocognitive impairment in pediatric cancer long-term survivors.* J Clin Oncol, 2008. 26(25): p. 4138-43.
21. Brinkman, T.M., et al., *Longitudinal patterns of psychological distress in adult survivors of childhood cancer.* Br J Cancer, 2013. 109(5): p. 1373-81.
22. Brinkman, T.M., et al., *Prevalence and predictors of prescription psychoactive medication use in adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study.* J Cancer Surviv, 2013. 7(1): p. 104-14.
23. Mulrooney, D.A., et al., *Fatigue and sleep disturbance in adult survivors of childhood cancer: a report from the childhood cancer survivor study (CCSS).* Sleep, 2008. 31(2): p. 271-81.
24. Buchbinder, D., et al., *Psychological outcomes of siblings of cancer survivors: a report from the Childhood Cancer Survivor Study.* Psychooncology, 2011. 20(12): p. 1259-68.
25. Meijer, R.R., R.M. de Vries, and V. van Bruggen, *An evaluation of the Brief Symptom Inventory-18 using item response theory: which items are most strongly related to psychological distress?* Psychol Assess, 2011. 23(1): p. 193-202.