

Proposal No: 98-21
Topic: Loss to Follow-Up

Lead CCSS Investigator: Roberta Sposato-Wells

Collaborators: Yasui, Mertens, Robison

Submitted to Publications Committee: 11/20/98

Approved by Publications Committee: 1/20/99

Priority Rating: 1.5

CHILDHOOD CANCER SURVIVOR STUDY Analysis Concept Form

Submitted: October 1998

1. **Title:** Predicting Loss to Follow-up and Successful Tracing in a Cohort of Childhood Cancer Survivors
2. **Working Group and Investigators:** This proposed publication will be within the Epidemiology/ Biostatistics Working Group. Proposed investigators (name/ e-mail/ fax) will include:

Roberta Sposato Walls Rsposato@uhc.com 612-936-7270

3. **Background and Rationale:** In order for the results of research studies to be generalizable to the population of childhood cancer survivors, follow-up on a sample of this sample of this population needs to be as complete as possible. Follow-up of cohort members is a major issue in research studies. To keep bias to a minimum and ensure that the sample is representative of the population, researchers must undertake major efforts to find all eligible members. Depending upon the methods of the study, contact information can be obtained from the study subject at the time of the survey or from another source such as a medical record or hospital database.

Follow-up has been studied in relation to the methods used to locate individuals and the differences between those who were lost to follow-up and those who were not. The factors that have been found to be related to loss to follow-up include age, mobility, health status, socioeconomic factors, and gender.

4. **Specific Aims and Hypotheses:** The purpose of this paper is to explore the issues of age, mobility, SES, race, and gender in relation to tracing a cohort of childhood cancer survivors.
 - I. Identify the factors that predict the need to trace certain individuals.
Hypothesis: Individuals that need tracing are most likely going to be diagnosed early (prior to 1975), be older now, and be women.
 - II. Identify the factors that predict successful tracing.
Hypothesis: Individuals that are successfully traced are going to have a social security number on record at registration, reside in the same state, be alive, have a middle name or initial on record at registration, and have an uncommon last name.
 - III. Compare individuals successfully traced to individuals that did not need tracing.
Hypothesis: Individuals that need tracing will have fewer adverse health related outcomes, be in a minority group, and have lower socioeconomic status.

5. Analysis Framework:

- a. Outcome of interest: tracing status, follow-up status, and completion status.
- b. Subject population: all CCSS cases
- c. Exploratory variables: gender, age, age at diagnosis, year of diagnosis, diagnosis type, years since last contact, alive/dead status, commonness of last name, presence of middle initial or name at registration, presence of social security number at registration, mobility (moved to different state), race/ ethnicity, educational level, personal income, specific medical conditions (by system), and health related conditions reported in the last year.
- d. Specific tables: *Analysis was run with 20848 individuals and will be updated with complete data.* Please refer to the following example tables.

Table 1: Demographics of entire cohort.

	All eligible cases	Males	Females
sample size	20846	11533	9313
mean current age (in years), (range)	26.85 (11.45 - 48.89)	26.92 (11.46 - 48.86)	26.76 (11.45 - 48.89)
mean age at diagnosis (in years)	8.29	8.39	8.17
YEAR OF DIAGNOSIS			
1970-1975, number (%)	5036 (24.16)	2754 (23.88)	2282 (24.50)
1976-1981, number (%)	7786 (37.35)	4335 (37.59)	3451 (37.06)
1982-1986, number (%)	8023 (38.49)	4443 (38.53)	3580 (38.44)
DIAGNOSIS**			
Leukemia, # (%)	6817 (32.70)	3680 (31.91)	3137 (33.69)
CNS tumor, # (%)	3031 (14.54)	1687 (14.63)	1344 (14.43)
Hodgkin's disease, # (%)	2749 (13.19)	1579 (13.69)	1170 (12.56)
Non-Hodgkin's lymphoma, # (%)	1089 (7.36)	1089 (9.44)	446 (4.79)
Kidney tumors, # (%)	1761 (8.45)	836 (7.25)	925 (9.93)
Neuroblastomas, # (%)	1352 (6.49)	696 (6.04)	656 (7.04)
Soft tissue sarcoma, # (%)	1855 (8.90)	1018 (8.83)	837 (8.99)
Bone cancer, # (%)	1744 (8.37)	947 (8.21)	797 (8.56)

Table 2: Demographics between those study subjects who did or did not need tracing.

	All cases	Needed Tracing	Did not need tracing
sample size	20848	7378	13130
current mean age# (in years), (range)	26.85 (11.45-48.89)	27.43 (11.45-48.81)	26.51 (11.46-48.86)
mean age at diagnosis (in years)	8.29	8.36	8.25
GENDER**			
Male, Number (%)	11346 (55.33)	4212 (57.09)	7134 (54.34)
Female, Number (%)	9313 (44.67)	3166 (42.91)	5995 (45.66)
YEAR OF DIAGNOSIS**			
1970-1975, number (%)	4950 (24.16)	2056 (27.87)	2894 (22.04)
1976-1981, number (%)	7654 (37.35)	2826 (38.30)	4828 (36.77)
1982-1986, number (%)	7903 (38.49)	2496 (33.83)	5407 (41.18)
DIAGNOSIS**			
Leukemia, # (%)	6727 (32.70)	2274 (30.82)	4453 (33.92)
CNS tumors, # (%)	2978 (14.54)	1133 (15.36)	1845 (14.05)
Hodgkin's disease, # (%)	2679 (13.19)	991 (13.43)	1688 (12.86)
Non-Hodgkin's lymphoma, # (%)	1504 (7.36)	538 (7.29)	966 (7.36)
Kidney tumors, # (%)	1736 (8.45)	621 (8.42)	1115 (8.49)
Neuroblastomas, # (%)	1333 (6.49)	512 (6.94)	821 (6.25)
Soft tissue sarcomas, # (%)	1830 (8.90)	689 (9.34)	1141 (8.69)
Bone Cancer, # (%)	1719 (8.37)	620 (8.40)	1099 (8.37)

t-test p<.05

* Pearson chi-square p<.05

** Pearson chi-square p<.001

339 study subjects are pending at this time in regards to tracing.

Table 3: Univariate Analysis of Factors related to the need for tracing.

potential factors	OR	CI	p-value
gender M-F	1.118	1.056-1.184	0.001
year of diagnosis			
diagnosed 1970-1975	1.539	1.430-1.657	0.0001
diagnosed 1976-1981	1.269	1.187-1.356	0.0001
diagnosed 1982-1986	1	-	-
age			
ages 10-14	1	-	-
ages 15-19	0.904	0.768-1.063	0.2218
ages 20-24	1.051	0.896-1.234	0.5387
ages 25-29	1.189	1.015-1.392	0.0323
ages 30-34	1.365	1.162-1.604	0.0002
ages 35-39	1.315	1.109-1.559	0.0017
over 40	1.168	0.963-1.416	0.1150
years since last contact			
less than 5	1	-	-
6 - 10	2.016	1.869-2.176	0.000
more than 11	4.743	4.337-5.188	0.000

Note: p-values are from Wald chi-square.

Table 4: Demographics of those study subjects traced comparing those found and not found.

	All traced	Found	Not found
sample size	7378	2638	4740
mean current age (in years), (range)	27.44 (11.45-48.81)	27.92 (11.72-48.68)	27.17 (11.45-48.81)
mean age at diagnosis (in years)	8.36	8.81	8.10
GENDER			
Male, Number (%)	4212 (57.09)	1492 (56.56)	2720 (57.38)
Female, Number (%)	3166 (42.91)	1146 (43.44)	2020 (42.62)
YEAR OF DIAGNOSIS			
1970-1975, number (%)	2056 (27.87)	753 (28.54)	1303 (27.49)
1976-1981, number (%)	2826 (38.30)	984 (37.30)	1842 (38.86)
1982-1986, number (%)	2496 (33.83)	901 (34.15)	1595 (33.65)
DIAGNOSIS			
Leukemia, # (%)	2274 (30.82)	794 (30.10)	1480 (31.22)
CNS tumor, # (%)	1133 (15.36)	413 (15.66)	720 (15.19)
Hodgkin's disease, # (%)	991 (13.43)	387 (14.67)	604 (12.74)
Non-hodgkin's lymphoma, # (%)	538 (7.29)	192 (7.28)	346 (7.30)
Kidney tumor, # (%)	621 (8.42)	217 (8.23)	404 (8.52)
Neuroblastoma, # (%)	512 (6.94)	175 (6.63)	337 (7.11)
Soft tissue sarcoma, # (%)	689 (9.34)	239 (9.06)	450 (9.49)
Bone cancer, # (%)	620 (8.40)	221 (8.38)	399 (8.42)

Table 5: Univariate Analysis of Factors related to successful tracing:

potential factors	OR	CI	p-value
gender M-F	0.968	0.879-1.065	0.501
year of diagnosis			
1970-1975	1.023	0.906-1.155	0.713
1976-1981	0.946	0.845-1.058	0.331
1982-1986	1	-	-
age			
ages 10-14	1	-	-
ages 15-19	1.290	0.958-1.737	0.093
ages 20-24	1.177	0.879-1.575	0.274
ages 25-29	1.343	1.007-1.791	0.045
ages 30-34	1.428	1.068-1.909	0.016
ages 35-39	1.597	1.179-2.164	0.003
ages 40+	1.650	1.178-2.311	0.004
years since last contact			
less than 5	1	-	-
6 - 10	0.799	0.707-0.902	0.000
more than 11	0.782	0.693-0.882	0.000
Middle name or initial	1.016	0.923-1.118	0.750
Alive	0.799	0.686-0.884	0.001
Social security number	1.079	0.971-1.199	0.157
Common last name	0.670	0.588-0.765	0.001

Note: p-values are from Wald chi-square

Table 6: Demographics of study subjects lost to follow-up and with known addresses

	All eligible cases	Lost to follow-up	Address known
Sample size	20507	4740	15767
Mean current age (in yrs), (range)	26.85 (11.45-48.89)	27.23 (11.51-48.87)	26.80 (11.52-48.92)
Mean age at diagnosis (in yrs)	8.29	8.10	8.34
GENDER**			
Male, Number (%)	11345 (55.33)	2720 (57.38)	8625 (54.71)
Female, Number (%)	9161 (44.67)	2020 (42.62)	7141 (45.29)
YEAR OF DIAGNOSIS**			
1970 - 1975, number (%)	4950 (24.2)	1303 (27.49)	3647 (23.13)
1976- 1981, number (%)	7653 (37.4)	1842 (38.86)	5811 (36.86)
1982 -1986, number (%)	7904 (38.5)	1595 (33.65)	6309 (40.01)
DIAGNOSIS			
Leukemia, # (%)	6727 (32.70)	1480 (31.22)	5247 (33.28)
CNS tumor, # (%)	2979 (14.54)	720 (15.19)	2259 (14.33)
Hodgkin's disease, # (%)	2679 (13.19)	604 (12.74)	2075 (13.16)
Non-Hodgkin's lymphoma, # (%)	1504 (7.36)	346 (7.30)	1158 (7.34)
Kidney tumor, # (%)	1736 (8.45)	404 (8.52)	1332 (8.45)
Neuroblastoma, # (%)	1333 (6.49)	337 (7.11)	996 (6.32)
Soft tissue sarcoma, # (%)	1830 (8.90)	450 (9.49)	1380 (8.75)
Bone cancer, # (%)	1718 (8.37)	399 (8.42)	1319 (8.37)

Note: 340 study subjects with tracing status pending have been excluded from analysis.

** Pearson chi-square $p < .001$

Table 7: Univariate Analysis of Factors related to loss to follow-up.

potential factors	OR	CI	p-value
Gender M-F	1.115	1.044-1.190	0.001
Year of diagnosis			
1970 - 1975	1.413	1.300-1.537	0.000
1976 - 1981	1.254	1.162-1.353	0.000
1982 - 1986	1	-	-
Age			
ages 10-14	1	-	-
ages 15-19	0.804	0.667-0.968	0.022
ages 20-24	0.970	0.809-1.164	0.743
ages 25-29	1.001	0.836-1.199	0.992
ages 30-34	1.090	0.907-1.309	0.359
ages 35-39	0.996	0.820-1.210	0.969
over 40	0.898	0.720-1.120	0.342
Years since last contact			
less than 5	1	-	-
6 - 10	1.927	1.770-2.098	0.000
more than 11	3.627	3.311-3.972	0.000

Note: p-values are from Wald chi-square.

Table 8: Mobility, race and SES on study subjects who completed the questionnaire.

	All returned questionnaires	Returned after tracing	Returned without tracing
Sample size	12910	2127	11118
Moved to different state**	15.26	22.62	13.85
RACE/ ETHNICITY*			
White	89.61	87.96	89.89
Hispanic	4.59	5.21	4.49
African American	4.58	5.91	4.36
Asian	1.21	0.93	1.26
EDUCATION^*			
Did not complete HS	21.04	18.31	21.53
Completed high school	19.83	21.13	19.60
Post HS training	33.36	34.74	33.12
Completed College	25.76	25.82	25.75
PERSONAL INCOME^**			
None	8.24	7.66	8.34
Less than \$9,999	36.65	28.92	37.99
\$10,000 - \$19,999	20.84	20.05	20.97
\$20,000 - \$39,999	24.01	27.11	23.48
\$40,000 - \$59,999	6.49	9.21	6.02
Over \$60,000	3.77	7.06	3.20

^ Calculated using study subjects over 19 years of age.

* Pearson chi-square $p < .05$

** Pearson chi-square $p < .001$

Table 9: Health-related data by system on study subjects who completed the questionnaire.

	All returned questionnaires	Returned after tracing	Returned without tracing
Sample size	12910	2127	11118
Accessed health care**	90.13	88.30	90.43
SPECIFIC MEDICAL CONDITIONS by system			
Hearing	17.39	16.98	17.46
Vision	31.06	32.50	30.82
Urinary	16.73	18.13	16.50
Hormonal	24.46	25.10	24.36
Heart and Circulatory	19.90	21.61	19.62
Respiratory	59.83	61.40	59.57
Digestive**	34.86	38.05	34.34
Brain/ Nervous System**	47.34	52.15	46.55
Recurrence or second malignancy**	18.71	21.74	18.21
Subsequent recurrence or malignancy**	3.93	5.88	3.61

Values are expressed as the percentages of individuals responding yes to the question unless indicated.

* Pearson chi-square $p < .05$

** Pearson chi-square $p < .001$

Table 10: Health-related data of conditions reported in the last year on study subjects who completed the questionnaire.

	All returned questionnaires	Returned after tracing	Returned without tracing
Sample size	12910	2127	11118
Diabetes	0.99	0.94	1.00
Epilepsy	1.52	1.57	1.51
Repeat seizures/blackouts	2.98	3.07	2.97
Repeat kidney infections	1.35	1.69	1.30
Migraine*	10.30	12.16	10.01
Frequent headaches	19.52	19.34	19.55
Gallstones	0.77	0.88	0.75
Cirrhosis of the liver	0.20	0.25	0.19
Hepatitis	0.84	1.19	0.79
Jaundice	0.34	0.56	0.31
Ulcer	1.92	2.07	1.89
Frequent indigestion*	10.80	12.34	10.56
Diverticulitis	0.18	0.31	0.16
Colitis	0.65	0.44	0.68
Frequent constipation	5.74	6.77	5.58
Bronchitis	9.26	8.78	9.33
Hay fever	15.49	14.96	15.58
Tonsillitis	5.39	5.33	5.40
Emphysema	0.11	0.13	0.11
Pleurisy	0.74	0.81	0.73

Note: Values expressed as percentages of individuals who answered yes to the question unless indicated.

* Pearson chi-square $p < .05$

** Pearson chi-square $p < .001$