Exploring the Feasibility of Establishing a Retrospective Cohort of Survivors of Adolescent and Young Adult Cancer to Study Long-Term Health Outcomes in an Integrated Managed Care Environment

Chun Chao, PhD,¹ Vicki Chiu, MS,¹ Lisa A. Mueller, MD,² and Robert Cooper, MD²

Purpose: To determine the feasibility of conducting survivorship research for long-term health outcomes with survivors of cancer diagnosed as an adolescent or young adult (AYA) and enrolled in Kaiser Permanente Southern California (KPSC), an integrated managed care organization.

Methods: Survivors diagnosed at ages 15–39 during 1990 and 2000 at KPSC were included. A 1:5 age-, gender-, and calendar-year-matched non-cancer KPSC comparison group was also identified. Date of cancer diagnosis was defined as the study baseline. KPSC insurance retention rate was calculated at 5 and 10 years post-baseline among survivors. Multivariable logistic regression was used to examine demographic and cancer characteristics associated with KPSC insurance retention at ≥ 5 years after baseline.

Results: A total of 6170 AYA cancer patients were identified: 4745 (77%) and 4471 (72%) survived at 5 and 10 years after diagnosis respectively. Of these survivors, 3654 (77%) and 2817 (63%) remained insured at KPSC at 5 and 10 years post-cancer diagnosis respectively. Those aged 20–29 years when diagnosed and those with stage 4 cancer were less likely to retain KPSC insurance than other survivors. For non-cancer comparison subjects, the KPSC insurance retention rate was lower: 66% at 5 years and 51% at 10 years post-baseline. Younger age, female gender, white race, and later calendar years of study baseline were associated with a lower likelihood of KPSC insurance retention.

Conclusion: These results demonstrate the feasibility and potential limitations of conducting survivorship research to characterize long-term health outcomes for survivors of AYA cancer in a large, integrated managed care organization.

Keywords: feasibility, survivorship cohort, managed care organization, late effects, long-term effects

THE INCIDENCE OF CANCER in adolescents and young adults (AYAs) has been steadily increasing over the past 25 years.¹ It is estimated that, in 2009, 68,400 new cancers were diagnosed among AYAs between the ages of 15 and 39 in the United States alone.² Common cancer types in AYAs include cancers of the breast, thyroid, testis, cervix, as well as melanoma and lymphoma.³ As with all cancer survivors, survivors of AYA cancer are known to experience increased risk of second malignancies,⁴ cardiotoxicity^{5,6} and infertility⁷ compared with the general population.

However, lack of research about the long-term health outcomes of survivors of AYA cancer has hindered the creation of long-term care guidelines focusing on this population, particularly for young adults diagnosed at the age of 21 and older. Addressing the knowledge gaps will not be

simple. Challenges to conducting survivorship research include the difficulties and costs associated with long-term follow-up and accruing sufficient sample sizes. In the United States, compliance with the Health Insurance Portability and Accountability Act complicates multi-institutional research.⁸ AYAs are also a relatively mobile and financially unstable population.⁹ As they often transition in and out of insurance, conducting survivorship research in an insurance-based setting can be difficult. Creating a prospective survivorship cohort can also be challenging, as AYAs generally do not participate in cancer clinical trials as often as older populations.^{10–12}

Kaiser Permanente is an integrated healthcare system that provides comprehensive health services for its members. Kaiser Permanente Southern California (KPSC) is the largest

¹Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, California.

²Department of Pediatrics, Los Angeles Medical Center, Kaiser Permanente Southern California, Los Angeles, California.

of the eight Kaiser Permanente regions, covering 3.5 million (approximately 1% of the U.S. population) racially and socioeconomically diverse enrollees who are broadly representative of residents of southern California.¹³ Within KPSC, individuals have relatively equal healthcare access, which allows health research to be conducted without the added variable of different health insurance levels. There is a centralized medical record system at KPSC that allows tracking of diagnoses, procedures, treatments, and other healthcare services. Recently, an integrated electronic medical record system replaced the paper-based system to serve the needs of daily clinical practice as well as health research better. With its comprehensive medical records and large and relatively stable population, KPSC offers a unique opportunity to study long-term health outcomes among survivors of AYA cancer.

In this study, we assessed the feasibility of conducting survivorship research for long-term health outcomes of AYA cancer in KPSC. We used the following criteria for assessing feasibility: (1) the ability to follow up for long-term health outcomes (e.g., ≥ 5 years post-cancer diagnosis), as indicated by insurance retention patterns and available follow-up time among survivors of AYA cancer; (2) the ability to evaluate the relative disease burden among survivors of AYA cancer compared to those without a history of cancer (including an evaluation of the insurance retention pattern among a matched sample of KPSC members without a history of cancer); (3) the power to study long-term health outcomes, as indicated by the size of the cancer survivor cohort; and (4) the opportunity to determine health outcomes based on routinely collected clinical data, as indicated by the frequency of provider visits made by study subjects. In addition, we also explored (5) the generalizability of our AYA cancer population compared to the Surveillance, Epidemiology and End Results (SEER) national statistics, and (6) the type of health outcomes that can be evaluated within the KPSC setting.

Methods

Study population

People who met the following criteria were included in the cohort of survivors of AYA cancer: (1) diagnosed with first cancer between the ages of 15 and 39 years old, (2) survived for at least 5 years post-cancer diagnosis, and (3) diagnosed at KPSC between 1990 and 2000. This time frame was chosen to allow a minimum of 5 years of follow-up after a person had become a 5-year cancer survivor at the time of analysis, and also because KPSC's cancer registry only started after 1988. The age range of 15–39 years for AYAs was based on the definition chosen by the U.S. National Cancer Institute (NCI)'s AYA Oncology Progress Review Group.¹⁴ For the comparison group without a history of cancer, a 1:5 yearly age-, gender-, and calendar-year-matched sample was selected from individuals at KPSC without a history of cancer. The non-cancer comparison sample for each survivor of AYA cancer was identified from those who were insured at KPSC in the year of each cancer survivor's corresponding cancer diagnosis (defined as the baseline in this study) and stayed alive for at least 5 years from that time. All data were collected from KPSC's electronic medical records, cancer registry, mortality files, and membership databases.

Statistical analysis

For survivors of AYA cancer, we calculated the distribution of demographic (age at diagnosis, gender, and race/ethnicity) and cancer characteristics (calendar year at diagnosis, stage at diagnosis, and cancer type). To assess the ability to follow up for long-term health outcomes, we calculated the KPSC insurance retention rate at 5, 10, 15 (for those diagnosed between 1990 and 1995), and 20 years (for those diagnosed in 1990 only) after cancer diagnosis. We conducted stratified analyses for age groups (15–19, 20–24, 25–29, 30–34, and 35–39 years) at diagnosis and by gender, race/ethnicity, cancer stage, cancer type, and calendar year of diagnosis. We also calculated the distribution of these cancer survivors' available follow-up (i.e., time before termination of KPSC insurance, or time in care at KPSC) through the end of 2010. To determine the ability to evaluate the relative risk of long-term health outcomes among survivors of AYA cancer compared to those without a history of cancer, we calculated KPSC insurance retention (at 5, 10, 15, and 20 years post-baseline) and distribution of available follow-up time through end of 2010 for the matched non-cancer comparison subjects.

We used multivariable logistic regression to assess factors associated with KPSC insurance retention at ≥ 5 years post-study baseline. Separate models were built for survivors of AYA cancer and for the matched non-cancer comparison group. Factors examined for survivors of AYA cancer included age at diagnosis, gender, race/ethnicity, calendar year at diagnosis, and cancer type; age, gender, race/ethnicity, and calendar year were examined for the non-cancer comparison subjects.

To evaluate the study power to assess disease burden in survivors of AYA cancer compared to that in persons without a history of cancer, we estimated the minimal detectable hazard ratio based on a two-sided log-rank test with 80% power and 5% type I error for health outcomes ranging from 0.999 to 0.900 cumulative survival at 10 years follow-up in non-cancer subjects, conservatively assuming a 50% loss to follow up in both cancer survivors and non-cancer subjects.

To assess the feasibility of detecting a health outcome of interest based on routinely collected clinical data, we calculated the rate of outpatient visits to any type of provider for the cohort of survivors of AYA cancer and the matched non-cancer subjects as the total number of outpatient visits divided by the cumulative person-time.

To evaluate potential generalizability of our AYA cancer population, we compared cancer type distribution and 5-year survival in our population to that reported by the national SEER statistics. The NCI's SEER program collects data on cancer cases from various geographic areas throughout the United States, representing 28% of the U.S. population.¹⁵ SEER-affiliated registries collect information on cancer incidence, prevalence, and survival. For this comparison, we used the most recently available published SEER statistics: years 2004–2008 for cancer incidence among persons aged 15–39 years, and 5-year survival statistics from the 1993–2000 period.¹⁶ KPSC members diagnosed with AYA cancer in 2004–2008 and 1993–2000 were included in these comparisons respectively to ensure a comparable calendar period.

All statistical analyses were conducted using SAS Enterprise Guide (v4.3, SAS Institute Inc., Cary, NC).

Results

Characteristics of study cohorts at KPSC

We identified 6170 AYA cancer patients diagnosed within KPSC between 1990 and 2000. Among them, 4745 (77%) survived at 5 years post-diagnosis (Table 1). Of the 4745 5-year cancer survivors, 41% were aged 35–39 at diagnosis, 28% were aged 30–34, 16% were aged 25–29, 9% were aged 20–24,

and 6% were aged 15–19 years old at diagnosis. More than half of the 5-year cancer survivors were female or white or had stage I cancer. Approximately 75% of survivors had one of the following cancer types (in order of highest frequency): breast cancer, melanoma, thyroid cancer, lymphoma, testicular cancer, or female genital cancer (cervical, ovarian, or uterine).

A total of 23,725 1:5 matched non-cancer comparison subjects were included in the study. They had the same age,

TABLE 1. CHARACTERISTICS OF PATIENTS AND SURVIVORS OF ADOLESCENT AND YOUNG ADULT CANCER AND THEIR INSURANCE RETENTION WITHIN KAISER PERMANENTE SOUTHERN CALIFORNIA

	<i>Incident adolescent and young adult cancer (N=6170)</i>	<i>5-year survivors (n=4745)</i>	<i>5-year^a insurance retention among 5-year survivors (n=3654 retained/4745 survivors: 77%)</i>	<i>10-year^a insurance retention among 10-year survivors (n=2817/4471: 63%)</i>	<i>15-year^a insurance retention among 15-year survivors (n=1161/2073: 56%)</i>	<i>20-year^a insurance retention among 20-year survivors (n=147/326: 45%)</i>
Age at cancer diagnosis, years	Number (%)					
15–19	394 (6.4%)	303 (6.4%)	245 (80.9%)	154 (52.2%)	52 (39.1%)	6 (24.0%)
20–24	511 (8.3%)	434 (9.2%)	295 (68.0%)	214 (51.4%)	88 (48.4%)	10 (34.5%)
25–29	937 (15.2%)	746 (15.7%)	532 (71.3%)	405 (56.8%)	173 (50.3%)	29 (39.2%)
30–34	1710 (27.7%)	1306 (27.5%)	999 (76.5%)	786 (64.1%)	329 (55.8%)	42 (50.0%)
35–39	2618 (42.4%)	1956 (41.2%)	1582 (80.9%)	1241 (68.2%)	511 (62.0%)	61 (53.5%)
Gender						
Female	3690 (59.8%)	3031 (63.9%)	2372 (78.3%)	1838 (64.6%)	780 (57.7%)	110 (51.6%)
Male	2480 (40.2%)	1714 (36.1%)	1281 (74.7%)	962 (59.2%)	373 (51.8%)	38 (33.6%)
Race/ethnicity						
White	3329 (54.0%)	2630 (55.4%)	1992 (75.7%)	1506 (60.3%)	653 (52.9%)	83 (39.7%)
African American	687 (11.1%)	461 (9.7%)	370 (80.3%)	282 (69.0%)	125 (62.5%)	9 (47.4%)
Hispanic	1679 (27.2%)	1298 (27.4%)	1020 (78.6%)	801 (65.6%)	278 (58.9%)	40 (55.6%)
Asian	452 (7.3%)	337 (7.1%)	261 (77.5%)	206 (63.6%)	96 (60.0%)	16 (69.6%)
Other/unknown	23 (0.4%)	19 (0.4%)	—	—	—	—
Calendar year of cancer diagnosis						
1990–1992	1634 (26.5%)	1171 (24.7%)	881 (75.2%)	698 (63.9%)	576 (54.6%)	148 (45.4%) ^b
1993–1995	1533 (24.9%)	1141 (24.1%)	918 (80.5%)	717 (67.4%)	577 (56.7%)	—
1996–1998	1702 (27.6%)	1358 (28.6%)	1060 (78.1%)	792 (61.8%)	—	—
1999–2000	1301 (21.1%)	1075 (22.7%)	794 (73.9%)	593 (57.4%)	—	—
Stage						
I	3077 (49.9%)	2822 (59.5%)	2142 (75.9%)	1690 (61.9%)	716 (55.6%)	99 (46.5%)
II	1278 (20.7%)	1051 (22.2%)	833 (79.3%)	626 (64.9%)	252 (55.8%)	33 (52.4%)
III	592 (9.6%)	409 (8.6%)	332 (81.2%)	248 (67.6%)	84 (54.6%)	8 (38.1%)
IV	1026 (16.6%)	354 (7.5%)	260 (73.5%)	176 (57.3%)	68 (55.7%)	4 (18.2%)
Unknown	197 (3.2%)	109 (2.3%)	—	—	—	—
Cancer type ^c						
Melanoma	762 (12.4%)	719 (15.2%)	509 (70.8%)	404 (57.4%)	181 (52.9%)	22 (35.5%)
Breast	1003 (16.3%)	788 (16.6%)	664 (84.3%)	501 (72.0%)	202 (61.4%)	28 (58.3%)
Thyroid	618 (10.0%)	614 (12.9%)	469 (76.4%)	381 (62.9%)	150 (57.5%)	22 (50.0%)
Lymphoma	661 (10.7%)	486 (10.2%)	374 (77.0%)	285 (61.7%)	118 (53.9%)	18 (50.0%)
Testicular	444 (7.2%)	431 (9.1%)	336 (78.0%)	266 (62.7%)	105 (52.5%)	10 (31.3%)
Cervical	314 (5.1%)	256 (5.4%)	197 (77.0%)	159 (63.9%)	76 (56.7%)	7 (35.0%)
Gastrointestinal	295 (4.8%)	182 (3.8%)	138 (75.8%)	107 (63.3%)	43 (61.4%)	3 (50.0%)
Ovarian	217 (3.5%)	182 (3.8%)	137 (75.3%)	103 (58.9%)	45 (52.3%)	7 (53.9%)
Brain	248 (4.0%)	160 (3.4%)	119 (74.4%)	74 (58.3%)	21 (42.9%)	3 (27.3%)
Leukemia	270 (4.4%)	128 (2.7%)	106 (82.8%)	70 (61.4%)	25 (65.8%)	0 (0%)
Uterine	130 (2.1%)	112 (2.4%)	93 (83.0%)	70 (65.4%)	36 (69.2%)	6 (66.7%)
Sarcoma	148 (2.4%)	111 (2.3%)	92 (82.9%)	67 (63.2%)	27 (58.7%)	4 (80.0%)
Other	1060 (17.2%)	576 (12.1%)	419 (72.7%)	313 (58.8%)	124 (50.2%)	18 (48.7%)

Note: Percentage may not add to exactly 100% due to rounding.

^aIn reference to time of cancer diagnosis. The number of 5-year cancer survivors and 10-year cancer survivors are based on those diagnosed between 1990 and 2000, the number of 15-year cancer survivors is based on those diagnosed between 1990 and 1995, and the number of 20-year cancer survivors is based on those diagnosed in 1990.

^bFor those diagnosed in 1990 only.

^cCancer types with at least 100 5-year survivors are shown. Cancer types with fewer than 100 5-year survivors are grouped into “Other.”

gender, and calendar-year distribution as the survivors of AYA cancer due to matching.

Feasibility for follow-up for long-term health outcomes

Table 1 shows the KPSC insurance retention for survivors of AYA cancer by age, gender, race/ethnicity, year of diagnosis, stage at diagnosis, and cancer type. Overall, 77%, 63%, 56%, and 45% of the 5-year, 10-year, 15-year, and 20-year survivors of AYA cancer remained insured at KPSC at 5 years, 10 years, 15 years, and 20 years post-cancer diagnosis respectively. Among those who remained insured at KPSC through the end of 2010, there was a median length of available follow-up (time in care at KPSC) of 6.9 years (interquartile range: 4.0 years, 10.3 years, maximum 22.6 years) after reaching the 5-year mark of survival.

Feasibility for studying the relative burden of long-term health problems in survivors of AYA cancer compared to non-cancer subjects

Compared to cancer survivors, the KPSC insurance retention rate was lower in the matched non-cancer subjects: 67% at 5 years, 53% at 10 years, 47% at 15 years, and 37% at 20 years post-baseline (data not shown). However, among those who remained insured at KPSC, the available length of follow-up (time in care at KPSC) through 2010 was comparable to that of the AYA cancer survivors, with a median of 6.9 years (interquartile range: 3.7 years, 10.4 years, maximum 18.7 years) after reaching the 5-year mark post-baseline.

In logistic regression, age 20–24 or 25–29 years at diagnosis, stage IV cancer diagnosis, and melanoma diagnosis were associated with lower likelihood of KPSC insurance retention among survivors of AYA cancer (Table 2). Among non-cancer comparison subjects, younger age, female gender, white race, and later years of study baseline were associated with lower likelihood of KPSC insurance retention at 5 years post-study baseline (Table 2).

Power calculation for detecting relative burden of health outcomes of interest

Based on 4745 survivors of AYA cancer and their 23,725 1:5 matched non-cancer comparison subjects, the minimal detectable hazard ratios for the effect of being a survivor of AYA cancer from a two-sided log-rank test with 80% power and 5% type I error were 3.5, 1.6, and 1.2 for health outcomes ranging from 10-year cumulative survival of 0.9990 (e.g., lung fibrosis, amputation), 0.9900 (e.g., cerebrovascular disease, liver disease), and 0.9000 (e.g., hypertension or dyslipidemia) in non-cancer comparison subjects respectively. These data suggest our study will be reasonably powered to study most health outcomes of interest.

Feasibility for detecting health outcomes using routinely collected clinical data

The median rate of outpatient provider visits was 3.5/person-year (interquartile range: 1.5/person-year, 6.9/person-year) among survivors of AYA cancer and 2.2/person-year (interquartile range: 0.7/person-year, 4.6/person-year) among the non-cancer comparison subjects during the follow-up period. These findings suggest that our study subjects are under regular medical surveillance at KPSC, and we can rely

on routinely collected clinical data for detecting health outcomes of interest.

Generalizability of KPSC AYA cancer population

The distribution of cancer types was similar for both the KPSC and SEER AYA cancer cohorts diagnosed between 2004 and 2008.¹⁶ For males, the three most common cancers were testicular cancer (22.2% vs. 20.5%), lymphoma (14.0% vs. 16.4%), and melanoma (9.2% vs. 10.8%) respectively. For females, the three most common cancers were breast cancer (22.2% vs. 24.4%), thyroid cancer (15.8% vs. 17.5%), and cervical/uterine cancer (11.8% vs. 10.9%) respectively (Table 3). Also shown in Table 3, the overall 5-year survival rate was similar between the KPSC cohort and national SEER statistics between 1993 and 2000.¹⁶

Type of health outcomes that can be evaluated in the KPSC setting

Since KPSC provides comprehensive healthcare services to its members and has a centralized medical record system, most health outcomes important to cancer survivors can be investigated. Health outcomes of interest for cancer survivorship study are those that may be late effects of cancer and its treatment, as well as common conditions in the AYA-age population in general. These include mortality, secondary cancers, cardiovascular or cerebrovascular diseases, metabolic or thyroid disorders, infertility, renal failure, pulmonary fibrosis, cirrhosis, hearing or vision loss, osteoporosis, amputation, and depression.⁷ All of these conditions can be captured by electronic diagnosis and procedure coding, such as the International Classification of Disease (ICD) and the Current Procedure Terminology (CPT) codes, laboratory test results, disease registries, or mortality files.

Discussion

We identified a cohort of survivors of AYA cancer within the KPSC integrated healthcare system that can be studied for survivorship and long-term health outcomes. The comparability of cancer type distribution and survival between our population and national SEER statistics offers some assurance of the generalizability of our cohort of survivors of AYA cancer. In addition, our ability to retain survivors of AYA cancer for long-term follow-up appears to be better than what was achieved in the AYA HOPE Study, a large NCI-led AYA cancer survivorship study with first-hand data collection. As is common in prospective studies, participant recruitment can be difficult despite the investment of considerable resources. Investigators for the AYA HOPE Study reported only a 43% response rate despite intensive recruitment effort.¹⁷ This points to the need to determine the feasibility of conducting survivorship studies in environments where passive clinical follow-up on long-term health outcomes is possible. Our results suggest such feasibility.

We found that several demographic and cancer characteristics—such as age, gender, and cancer type—were associated with KPSC insurance retention. Such non-random loss to follow-up may introduce bias when studying long-term health outcomes. Therefore, weighted or multivariable analyses taking into account factors associated with loss to

TABLE 2. ODDS RATIOS FROM MULTIVARIABLE LOGISTIC REGRESSION^a ON RETAINING KPSC INSURANCE STATUS AT 5 YEARS POST-CANCER DIAGNOSIS FOR SURVIVORS OF AYA CANCER OR 5 YEARS POST-BASELINE FOR MATCHED NON-CANCER COMPARISON SUBJECTS

Factors	5-year survivors of AYA cancer			Matched non-cancer comparison group		
	Odds ratio ^a	95% confidence interval	p-value ^b	Odds ratio	95% confidence interval	p-value ^b
Age at diagnosis or baseline, years						
15–19	Ref			Ref		
20–24	0.52	(0.36–0.74)	< 0.01	0.79	(0.63–0.98)	0.03
25–29	0.61	(0.43–0.86)	< 0.01	1.07	(0.87–1.31)	0.52
30–34	0.78	(0.56–1.09)	0.14	1.32	(1.09–1.61)	0.01
35–39	0.98	(0.71–1.36)	0.91	2.01	(1.66–2.44)	< 0.01
Gender						
Female	Ref			Ref		
Male	0.86	(0.72–1.03)	0.10	1.38	(1.25–1.52)	< 0.01
Race/ethnicity ^c						
White	Ref			Ref		
African American	1.18	(0.91–1.53)	0.21	1.46	(1.28–1.67)	< 0.01
Hispanic	1.18	(0.99–1.40)	0.07	1.71	(1.55–1.89)	< 0.01
Asian	1.01	(0.76–1.35)	0.94	1.62	(1.37–1.92)	< 0.01
Stage at diagnosis						
I	Ref					
II	1.01	(0.83–1.21)	0.94		—	
III	1.30	(0.99–1.71)	0.06		—	
IV	0.68	(0.50–0.93)	0.02		—	
Cancer type						
Melanoma	Ref					
Breast	1.73	(1.30–2.31)	< 0.01		—	
Thyroid	1.26	(0.97–1.64)	0.09		—	
Lymphoma	1.40	(1.04–1.87)	0.03		—	
Testicular	1.62	(1.19–2.20)	< 0.01		—	
Cervical	1.19	(0.84–1.70)	0.33			
Gastrointestinal	1.10	(0.73–1.64)	0.65	-		
Ovarian	1.09	(0.74–1.61)	0.66			
Brain	1.07	(0.71–1.62)	0.74	-		
Leukemia	2.68	(1.49–4.80)	< 0.01	-		
Uterine	1.50	(0.88–2.57)	0.14			
Sarcoma	1.89	(1.08–3.30)	0.03	-		
Other	1.06	(0.81–1.39)	0.67	-		
Calendar year of diagnosis or baseline						
1990–1992	Ref			Ref		
1993–1995	1.29	(1.05–1.58)	0.02	0.99	(0.87–1.12)	0.82
1996–1998	1.15	(0.95–1.39)	0.16	0.83	(0.73–0.94)	< 0.01
1999–2000	0.88	(0.73–1.08)	0.22	0.67	(0.59–0.76)	< 0.01

^aModel included all variables shown in the table. Higher odds ratio indicate higher likelihood of retention.

^b $p < 0.05$ is considered statistically significant. Significant p -value is shown in bold.

^cThe “Other/unknown” race/ethnicity group was excluded from the multivariable analysis.

AYA, adolescent and young adult; KPSC, Kaiser Permanente Southern California; Ref, reference.

follow-up should be considered to address this concern. Furthermore, in our analysis, loss to follow-up due to termination of KPSC membership was more prominent among non-cancer subjects. This may be in part due to the difference in overall health status among cancer survivors and those without a cancer history, as it has been shown that people with illnesses tend to have greater stability within a health plan.¹⁸ This phenomenon may introduce bias when comparing health outcomes between cancer survivors and the general population. Approaches such as the standardized prevalence ratio (which uses all active health plan enrollees), as opposed to using matched comparison subjects (who are affected by insurance retention) as the reference group, may be used as a sensitivity analysis to generate alternative estimates for such comparison.

In addition, in our cohort, we found that cancer survivors had more frequent healthcare system encounters than the non-cancer comparison subjects. This situation may lead to detection bias. For the health outcomes of interest that are severe and require immediate medical intervention, however, we do not suspect a strong detection bias despite the difference in the frequency of provider visits because the presence of these conditions will likely trigger a provider visit for its diagnosis. On the other hand, for milder health conditions that do not always immediately trigger a provider visit, the results should be interpreted in the context of potential detection bias, taking into consideration the pattern of office visits or laboratory testing.

With regard to identifying health outcomes of interest using electronic diagnosis coding, several efforts have been

TABLE 3. COMPARISON OF CANCER TYPE DISTRIBUTION AND 5-YEAR SURVIVAL AMONG PERSONS DIAGNOSED WITH CANCER AT AGES 15–39 YEARS IN KPSC AND IN THE SEER CANCER REGISTRIES

	KPSC	SEER
Calendar period ^a	2004–2008	2004–2008
<i>Cancer type by gender</i>		
Male		
Testicular	22.2%	20.5%
Lymphoma	14.0%	16.4%
Melanoma	9.2%	10.8%
Gastrointestinal tract	8.0%	11.0%
Central nervous system	7.9%	6.0%
Leukemia	6.7%	6.8%
Thyroid	5.6%	5.6%
Soft tissue sarcoma	4.6%	6.6%
Other	21.8%	16.3%
Female		
Breast	22.2%	24.4%
Thyroid	15.8%	17.5%
Cervical and uterine	11.8%	10.9%
Melanoma	9.2%	11.5%
Lymphoma	7.3%	8.3%
Gastrointestinal tract	4.8%	6.0%
Central nervous system	3.8%	2.8%
Leukemia	2.8%	3.1%
Soft tissue sarcoma	1.8%	2.9%
Other	14.6%	18.6%
Calendar period ^a	1993–2000	1993–2000
<i>5-year survival by age (in years) at diagnosis</i>		
15–19	79.2%	79.4%
20–24	86.1%	82.2%
25–29	81.5%	80.6%
30–34	79.2%	75.6%
35–39	76.1%	74.6%

^aCalendar period in which cancers were diagnosed. Calendar period chosen based on the published SEER statistics.

KPSC, Kaiser Permanente Southern California; SEER, Surveillance, Epidemiology and End Results program.

made at KPSC to validate the ICD diagnosis coding for various conditions. A high positive predictive value was reported for both hypertension diagnosis coding (88% for one code and 98% for having two codes)¹⁹ and heart disease diagnosis coding.²⁰ These data should be taken into consideration when designing an observational study on long-term health outcomes. Use of multiple diagnosis codes or the combination of diagnosis codes, laboratory test results, or pharmacy utilization when applicable may improve the accuracy of case identification. Additional validation effort is also needed to understand if diagnosis coding accuracy may differ among cancer survivors versus those without a history of cancer.

We also identified several limitations of this managed care population in regard to conducting cancer survivorship research. First, despite a large number of insured individuals at KPSC, the number of survivors of AYA cancer of most individual cancer types (except for breast cancer, melanoma, and thyroid cancer), as well as of cancers diagnosed in adolescents aged 15–18 years, remain small, thus limiting the ability to study survivorship within certain subgroups. Second, it should also be noted that a 1:5 cancer survivor versus non-

cancer comparison group ratio was chosen arbitrarily in this examination. A higher ratio can be used in a formal cohort study, depending on the expected prevalence of the health outcomes of interest. With regard to the measure of health outcomes, risk factors such as obesity, cigarette smoking, alcohol use, and exercise are important lifestyle factors for understanding the survivorship experience of cancer survivors. However, these factors can only be reasonably assessed at KPSC after the 2006 integrated electronic medical record system implementation and are not comprehensively documented in earlier years. Finally, socioeconomic status may be an important confounder to consider. At KPSC, information on socioeconomic status is available only at the census block level, thus limiting our ability to account for the influence of individual-level socioeconomic status.

Conclusion

The National Cancer Policy Board of the Institute of Medicine recommended the establishment of cancer survivorship as a distinct phase of cancer care and life-long follow-up as a key component to improving the quality of cancer care.⁸ A long-term care plan that includes prevention, early diagnosis, and intervention to limit the impact of late effects may significantly benefit the quality of life of survivors of AYA cancer. Our study demonstrates the feasibility and potential limitations of conducting observational AYA cancer survivorship research in a large, integrated managed care organization. These data should help inform the design of studies that are urgently needed to characterize long-term health outcomes for AYA cancer.

Author Disclosure Statement

No competing financial interests exist.

References

1. Bleyer A, O'Leary M, Barr R, Ries LAG. Cancer epidemiology in older adolescents and young adults 15 to 29 years of age, including SEER incidence and survival: 1975–2000 (NIH Pub. No. 06-5767). Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2006.
2. A snapshot of adolescent and young adult cancers. Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2012. Accessed February 5, 2013 from: www.cancer.gov/researchandfunding/snapshots/pdf/AYA-Snapshot.pdf
3. Bleyer A, Barr R. Cancer in young adults 20 to 39 years of age: overview. *Semin Oncol*. 2009;36(3):194–206.
4. Curtis RE, Freedman DM, Ron E, et al. (Eds). New malignancies among cancer survivors: SEER cancer registries, 1973–2000 (NIH Pub. No. 05-5302). Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2006. Accessed February 20, 2012 from: <http://seer.cancer.gov/publications/mpmono>
5. Adams MJ, Hardenbergh PH, Constine LS, Lipschultz SE. Radiation-associated cardiovascular disease. *Crit Rev Oncol Hematol*. 2003;45(1):55–75.
6. van Dalen EC, van der Pal HJ, Kok WE, et al. Clinical heart failure in a cohort of children treated with anthracyclines: a long-term follow-up study. *Eur J Cancer*. 2006;42(18):3191–8.

7. Soliman H, Agresta SV. Current issues in adolescent and young adult cancer survivorship. *Cancer Control*. 2008; 15(1):55–62.
8. Hewitt M, Greenfield S, Stovall E (Eds). *From cancer patient to cancer survivor: lost in transition*. Institute of Medicine and National Research Council, Committee on Cancer Survivorship: Improving Care and Quality of Life, Institute of Medicine, National Research Council. Washington, DC: National Academies Press; 2005.
9. They moved and went where: California's migration flow, 1995–2000. Sacramento, CA: California Department of Finance, Demographic Research Unit; June 2007.
10. Burke ME, Albritton K, Marina N. Challenges in the recruitment of adolescents and young adults to cancer clinical trials. *Cancer*. 2007;110(11):2385–93.
11. Bleyer A, Budd T, Montello M. Adolescents and young adults with cancer: the scope of the problem and criticality of clinical trials. *Cancer*. 2006;107(7 Suppl):1645–55.
12. Parsons HM, Harlan LC, Seibel NL, et al. Clinical trial participation and time to treatment among adolescents and young adults with cancer: does age at diagnosis or insurance make a difference? *J Clin Oncol*. 2011;29(30):4045–53.
13. Koebeck C, Langer-Gould A, Gould MK, et al. Socio-demographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *Perm J*. 2012;16(3):37–41.
14. Adolescent and Young Adult Oncology Progress Review Group. Closing the gap: research and cancer care imperatives for adolescents and young adults with cancer (NIH Pub. No. 06-6067). Bethesda, MD: Department of Health and Human Services, National Institutes of Health, National Cancer Institute, and the LIVESTRONG Young Adult Alliance; August 2006. Accessed May 16, 2012 from: http://planning.cancer.gov/library/AYAO_PRG_Report_2006_FINAL.pdf
15. National Cancer Institute. Surveillance Epidemiology and End Results. Accessed March 22, 2013 from: <http://seer.cancer.gov>
16. Howlader N, Noone AM, Krapcho M, et al. (Eds). Adolescent and young adult cancer by site, incidence, survival and mortality. In: *SEER cancer statistics review: 1975–2008*, based on November 2010 SEER data submission, posted to the SEER website 2011. Bethesda, MD: National Cancer Institute. Accessed February 7, 2013 from: http://seer.cancer.gov/csr/1975_2008/results_merged/sect_32_aya.pdf
17. Harlan LC, Lynch CF, Keegan TH, et al. Recruitment and follow-up of adolescent and young adult cancer survivors: the AYA HOPE Study. *J Cancer Surviv*. 2011;5(3):305–14.
18. Cutler D, Lincoln B, Zeckhauser R. Selection stories: understanding movement across health plans. *J Health Econ*. 2010;29(6):821–38.
19. Bhandari SK, Pashayan S, Liu IL, et al. 25-hydroxyvitamin D levels and hypertension rates. *J Clin Hypertens (Greenwich)*. 2011;13(3):170–7.
20. Shen AY, Yao JF, Brar SS, et al. Racial/ethnic differences in ischemic stroke rates and the efficacy of warfarin among patients with atrial fibrillation. *Stroke*. 2008;39(10):2736–43.

Address correspondence to:

Chun Chao, PhD

Department of Research and Evaluation

Kaiser Permanente Southern California

100 S Los Robles Avenue, 2nd floor

Pasadena, CA 91101

Email: chun.r.chao@kp.org