

Cancer Research Network: Data Availability Pre- and Post- Cancer Diagnosis

J Chubak¹, R Ziebell¹, G Hart¹, R Greenlee², S Honda³, M Hornbrook⁴, L Lamerato⁵, K Mazor⁶, L Nekhlyudov⁷, P Pawloski⁸, D Ritzwoller⁹, V Quinn¹⁰, VP Doria-Rose¹¹, L Kushi¹²

¹Group Health Research Institute, ²Marshall Clinic Research Foundation, ³Kaiser Permanente (KP) Center for Health Research – Hawaii, ⁴KP Center for Health Research – Northwest, ⁵Henry Ford Health System, ⁶Meyers Primary Care Institute, ⁷Harvard Pilgrim Health Care Institute, ⁸HealthPartners Institute for Education and Research, ⁹KP Colorado Institute for Health Research, ¹⁰KP Department of Research & Evaluation, ¹¹National Cancer Institute, ¹²KP Division of Research



Introduction

- The Cancer Research Network (CRN) is a collaboration among institutions within the HMO Research Network and the National Cancer Institute.
- Successful research within the CRN benefits from understanding the underlying populations and availability of relevant data.
- We focus here on describing the data required for three important research study designs possible within the CRN: retrospective studies of risk factors; retrospective studies of outcomes; and prospective studies.

Methods

- All nine funded and two of six affiliate CRN health plans (“sites”) participated in this analysis.
- A centrally developed SAS® program was run against each site’s Virtual Data Warehouse (VDW) during February 2014. The program accessed data from the enrollment, demographics, tumor, and death tables.
- Each site’s results were returned to the lead programmer and combined for analysis.



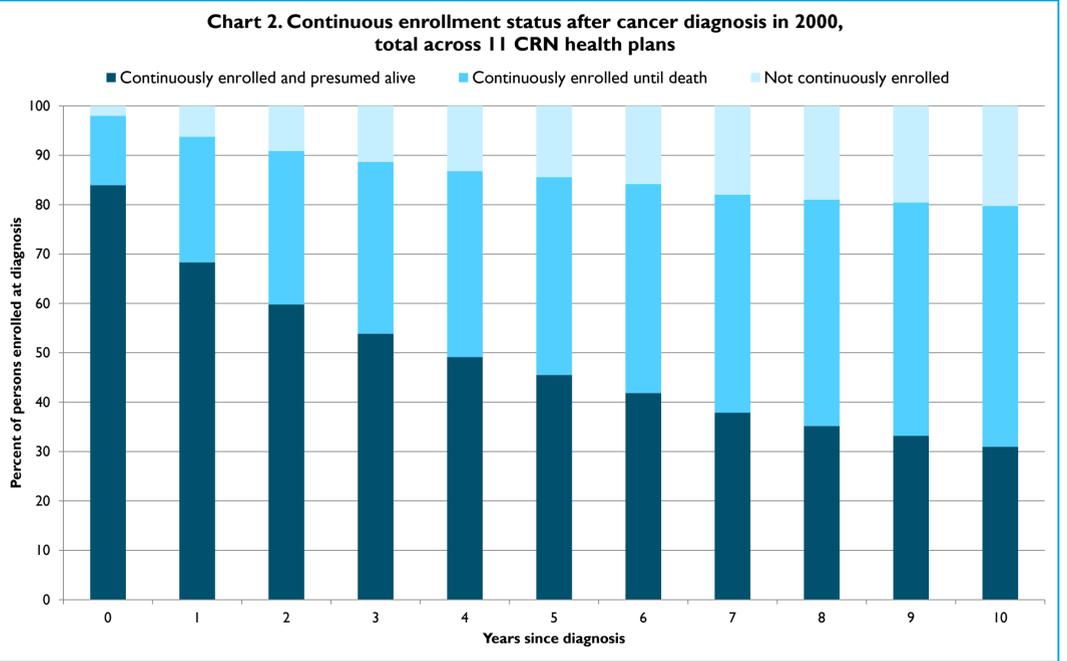
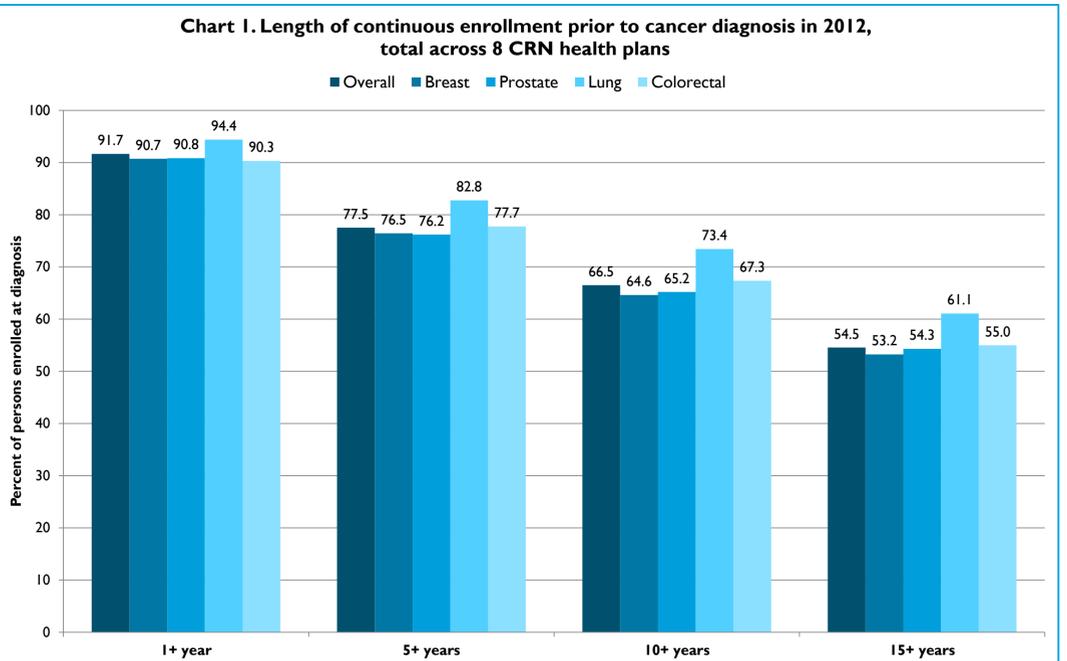
Results

Table 1. Current Health Plan Enrollment

	N	%
Total Enrollment	11,125,724	100.0
Age (Yrs)		
<18	2,428,478	21.8
18–44	3,945,362	35.5
45–64	3,175,539	28.5
≥65	1,576,345	14.2
Race		
American Indian/Alaska Native	36,698	0.3
Asian	1,044,909	9.4
Black/African American	712,108	6.4
Native Hawaiian/Other Pacific Islander	61,366	0.6
White	4,857,406	43.7
Multiple	174,152	1.6
Unknown	4,239,085	38.1
Hispanic		
Yes	2,083,641	18.7
No	3,262,162	29.3
Unknown	5,779,921	52.0

Table 2. Years of VDW Completeness (Approx.)

CRN Health Plan	Enrollment	Tumor Diagnoses	Vital Status (Tumor Registry)	Vital Status (Other Sources)
GHC	1988–2013	1974–2012	1974–2012	1977–2011
KPNW	1960–2013	1960–2012	1960–2012	1995–2012
KPMC	1997–2012	1972–2012	1973–2012	1966–2012
KPHI	1958–2013	1960–2012	1960–2012	1960–2011
KPCO	1994–2013	1987–2012	1987–2012	1970–2011
SHP	1992–2013	1960–2012	1960–2012	2004–2011
HAP	1997–2012	1985–2011	1985–2012	1990–2011
FCHP	1999–2012	1996–2009	1997–2009	2006–2012
HP	1994–2013	1960–2012	1960–2012	1980–2011
HPHC	2000–2012	1997–2009	2000–2009	2000–2010



Summary

- Population (Table 1)**
- The CRN represents >11 million persons.
 - Age distribution is similar to that of U.S. population.
 - Racial/ethnic distribution is broad, but data are incomplete.
- Data availability (Table 2)**
- All sites have complete enrollment data from the year 2000 onward, tumor diagnoses from 1997, and vital status data from 2000.
- Pre-diagnosis enrollment (Chart 1)**
- Cancer patients tend to be enrolled for many years prior to diagnosis.
 - Length of pre-diagnosis enrollment varies by CRN site (not shown) but not much by anatomic location.
- Post-diagnosis enrollment (Chart 2)**
- Cancer patients tend to remain enrolled after diagnosis.
 - Length of post-diagnosis enrollment varies by CRN site but not by anatomic location (not shown).

Conclusions

- Retrospective studies of cancer risk factors** are possible due to long-term pre-diagnosis enrollment.
- Retrospective studies of cancer outcomes** are possible due to long post-diagnosis enrollment.
- Prospective studies of cancer outcomes** are also enhanced by long-term post-diagnosis enrollment.
- The extent and diversity of CRN data** provide unique opportunities for cancer research; however, understanding data availability is key to study planning.

Chart 1. Persons diagnosed with malignant primary tumors while enrolled in 8 CRN health plans* in 2012, by duration of continuous enrollment prior to diagnosis and anatomic site

Anatomic Site	Total	Pre-Dx Enrollment							
		1+ Year		5+ Years		10+ Years		15+ Years	
		N	%	N	%	N	%	N	%
Overall	41,468	38,013	91.7	32,141	77.5	27,573	66.5	22,620	54.5
Breast	7,328	6,648	90.7	5,603	76.5	4,736	64.6	3,902	53.2
Prostate	6,018	5,467	90.8	4,586	76.2	3,924	65.2	3,269	54.3
Lung	3,841	3,626	94.4	3,179	82.8	2,821	73.4	2,346	61.1
Colorectal	3,604	3,255	90.3	2,802	77.7	2,427	67.3	1,982	55.0

*Limited to 8 sites with 15+ years of available enrollment data prior to most recent tumor diagnosis year (i.e., excluding HFHS, HPHC, and MPCI)

Chart 2. Percent of persons enrolled at diagnosis after being diagnosed with incident malignant primary tumors in 2000, by follow-up year, 11 CRN health plans

Cumulative Percent Disenrolled (Not Due to Death) by Follow-up Year										
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
2.0	6.2	9.1	11.3	13.2	14.4	15.8	17.9	19.0	19.5	20.3

Cumulative Percent Continuously Enrolled Until Death by Follow-up Year										
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
14.1	25.5	31.1	34.8	37.6	40.0	42.3	44.2	45.8	47.3	48.8

Cumulative Percent Continuously Enrolled and Presumed Alive by Follow-up Year										
2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
83.9	68.3	59.8	53.9	49.1	45.5	41.9	37.9	35.2	33.2	31.0