

Utah's Integrated Data Resources for Furthering Science in the National Childhood Cancer Registry



Anne C. Kirchhoff, PhD^{1,2} Carol Sweeney, PhD,^{2,3} Ken R. Smith, PhD^{2,4}
 Andrew Post, PhD,² Heydon Kaddas, MPH,² Marjorie Carter, MPH,³
 Mark Fluchel, MD,¹ Jen Doherty, PhD^{2,3}



¹University of Utah, Department of Pediatrics, ²Huntsman Cancer Institute, ³Utah Cancer Registry,
⁴University of Utah, Department of Family and Consumer Studies

OVERVIEW

The **Huntsman Cancer Institute (HCI)** is home to unique data infrastructure that can provide new information to the National Childhood Cancer Registry. Our CCDI aims include:

- **Aim 1:** Integrate data on 4,504 pediatric cancer patients from HCI through incorporating robust treatment data, longitudinal data on residential history, healthcare utilization, outcomes, and survivorship, information on social determinants of health and family history of cancer, and genomic data.
- **Aim 2:** Perform data quality review and refine a data packaging and transfer process that will ensure the timely provision of high quality pediatric cancer data from HCI to the NCCR.

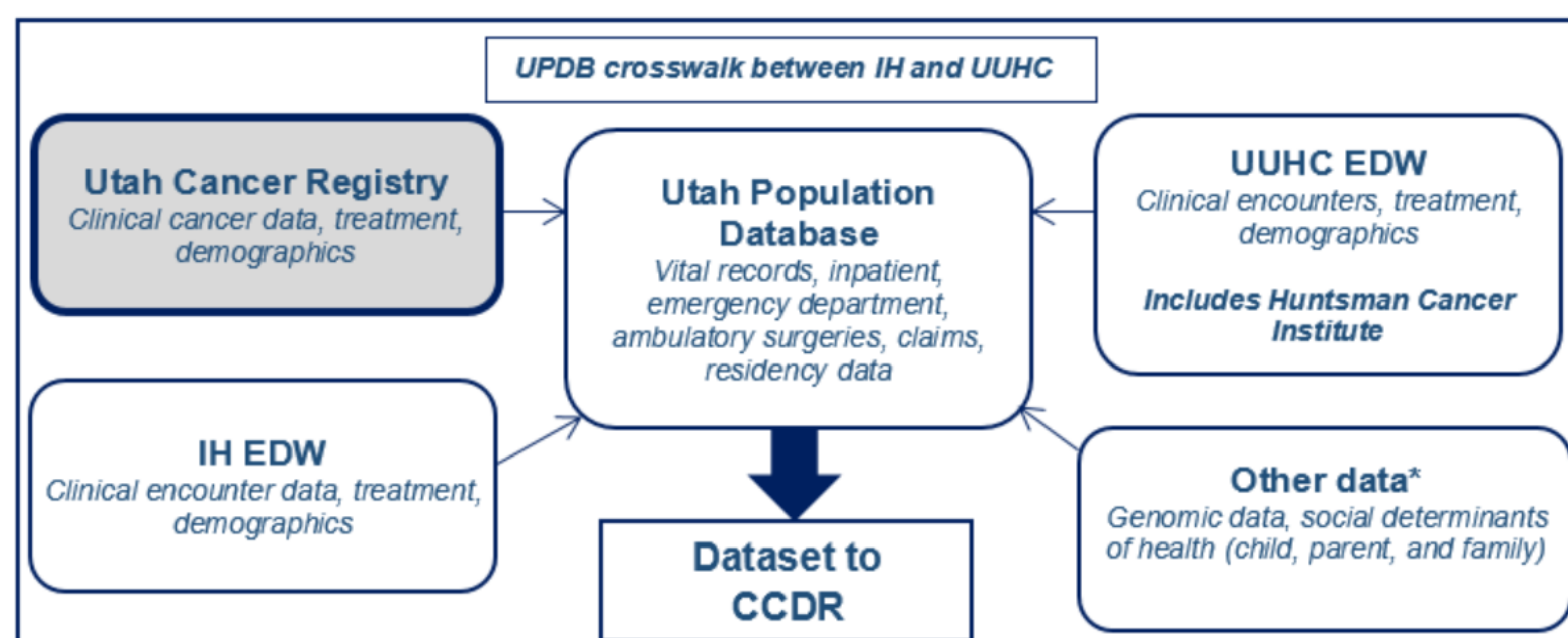
DATA SOURCES

HCI is uniquely poised to collect population-based, detailed data on pediatric cancer patients for several reasons:

- 1) Pediatric cancer patients in the region are diagnosed and treated centrally at **Primary Children's Hospital (PCH)**. PCH serves one of the largest geographic areas of any children's hospital in the United States.
- 2) The **Utah Cancer Registry (UCR)**, a Surveillance, Epidemiology and End Results (SEER) program registry since 1973, provides the backbone of curated data on pediatric cases.
- 3) HCI is home to the **Utah Population Database (UPDB)**, which links numerous data sources including Enterprise Data Warehouses from the two major healthcare systems accounting for ~90% of medical encounters in the state, allowing for rich longitudinal characterization of the Utah population.

The integration of these data resources is shown in **Figure 1**.

Figure 1: HCI and related cancer resources



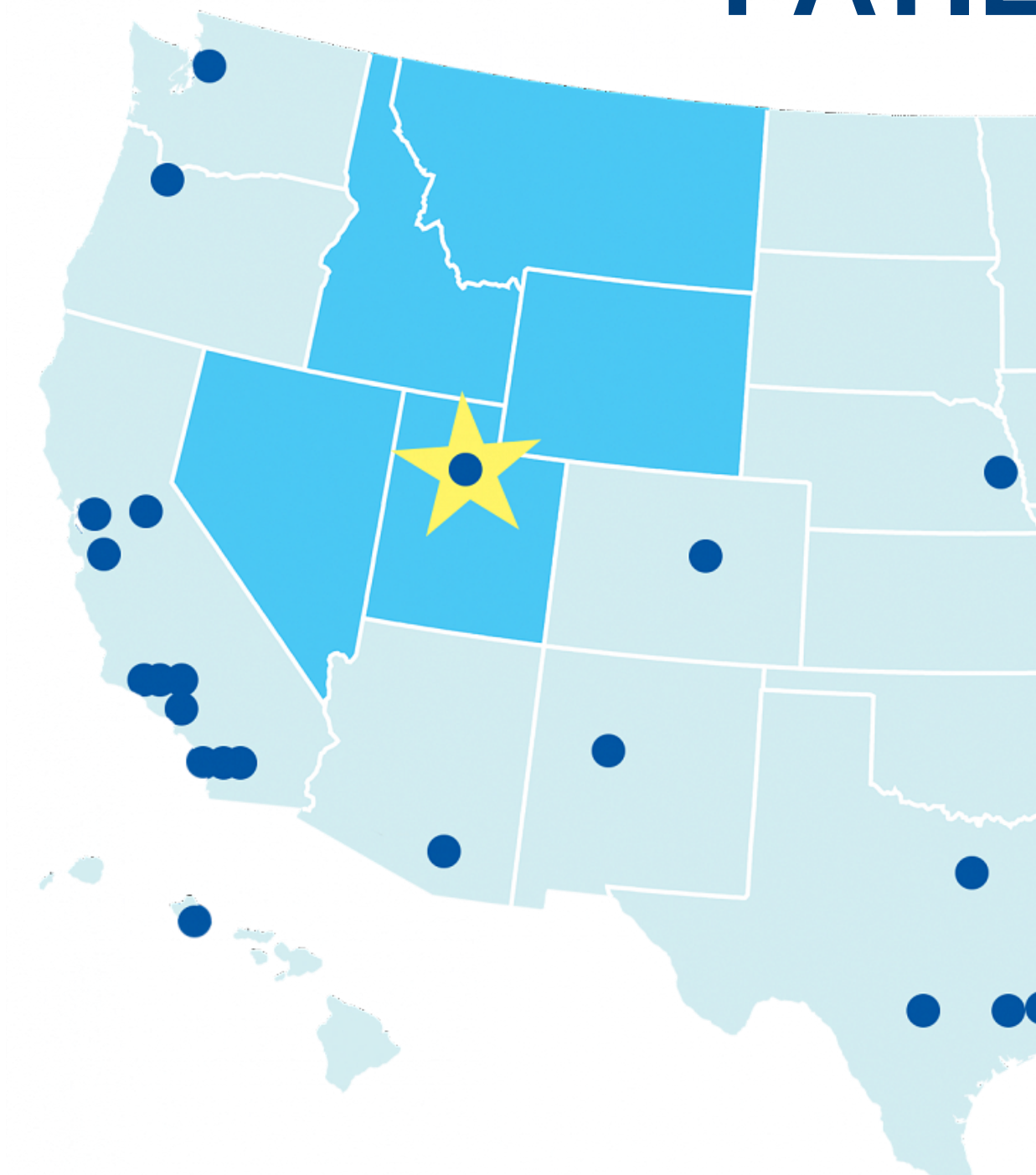
IH=Intermountain Healthcare; UUHC=University of Utah Health Care; EDW=Enterprise Data Warehouses

PROVOCATIVE QUESTIONS

CCDI records in aggregate would provide sufficient sample sizes to investigate questions among race, ethnicity, socioeconomic status, age, rural vs. urban, and distance to treatment subgroups. In particular, questions could be answered including:

- Can we identify patients at greatest risk for treatment toxicities and long-term health effects of cancer treatment? How to these vary by demographic and clinical factors?
- What environmental exposures could affect morbidity and mortality?
- What factors lead to delays in diagnosis?
- What are the determinants of clinical trial participation?

PATIENT POPULATION



PCH provides cancer care for over 85% of cases in Utah and many more from the five state region (**Figure 2**).

Figure 2.
HCI's Catchment Area

n=4,504, 1994-2017	Utah		Idaho		Other State	
	N	%	N	%	N	%
Total	3225		618		661	
Cancer Diagnosis						
Leukemia	856	27	154	25	172	26
<i>Acute lymphoblastic leukemia</i>	651		107		133	
<i>Other leukemia</i>	205		47		39	
Lymphomas	389	12	60	10	67	10
Central nervous system	792	25	159	26	151	23
Neuroblastoma	171	5	43	7	38	6
Retinoblastoma	51	2	16	3	16	2
Renal tumors	116	4	25	4	24	4
Hepatic tumors	56	2	<11	1	<11	1
Malignant bone cancers	175	5	38	6	49	7
Soft tissue sarcomas	210	7	48	8	60	9
Germ cell	161	5	23	4	35	5
Epithelial	150	5	26	4	21	3
Other/unspecified malignant	98	3	21	3	18	3
Sex						
Male	1709	53	329	53	361	55
Female	1516	47	289	47	300	45
Race/Ethnicity						
Non-Hispanic White	2577	80	524	85	531	80
Hispanic	498	15	73	12	86	13
Other Race/Ethnicity	150	5	21	3	44	7

DATA RESOURCES

- HCI expects to reuse data elements from the caDSR (cancer Data Standards Registry and Repository) and harmonize data to standards such as HL7 FHIR and V3 for provenance as well as family history and other pedigree-related data.
- The UPDB and HCI use industry standard databases and extract, transform, and load (ETL) tools to manage and prepare data.
- Best practices for data validation and preparation from other research data networks include All of Us Research Program and the NCATS Accrual to Clinical Trials Network.
- We will also employ locally developed software (Research Subject Registry) for capturing and managing germline test results.

GOVERNANCE

- Our local oversight includes involvement of the UPDB and the Resource for Genetic and Epidemiology Research (RGE).
- UPDB is used to assist with linkage of UCR data to medical record data at the University of Utah Health and Intermountain Healthcare.
- RGE is the review committee that approves the linkage and data access request. RGE also helps to ensure permission from the state's data contributors.
- For data transmission to NCCR, we are still evaluating UCR's ability to receive medical record data from health systems to then share with NCCR.
- Further assessment of state law surrounding UCR data receipt is ongoing. The ability to share data directly from the health systems to NCCR is a possible need.

CONTACT

- Anne Kirchhoff, PhD, MPH, Anne.Kirchhoff@hci.utah.edu (Project Director)
- Heydon Kaddas, MPH, Heydon.Kaddas@hci.utah.edu (Data Manager)
- Marjorie Carter, MPH, Marjorie.Carter@hsc.utah.edu (Project Manager)