

Data and Statistical Center Update

Wendy Leisenring, Sc.D.

CCSS Investigator Meeting

June 19-20, 2019

Atlanta, GA

CCSS

Childhood Cancer
Survivor Study



FRED HUTCH
CURES START HERE™

Statistical Center

CCSS



Seattle

Wendy Leisenring, ScD
Lead Biostatistician



Memphis

Yutaka Yasui, PhD
Biostat/Epi Working Group Chair



Kumar Srivastava, PhD



Team Members: Seattle

CCSS

John Whitton

Wendy Leisenring

Kayla Stratton



Pam Goodman

Team Members: St Jude - Yutaka

ccss



Yutaka Yasui

Qi Liu



Weiyu Qiu



Yan Chen



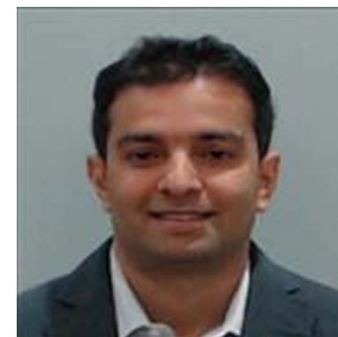
Cindy Im*



Nan Li



Yadav Sapkota*



*Focus on Genetic Analyses

Team Members: St Jude - Kumar

ccss



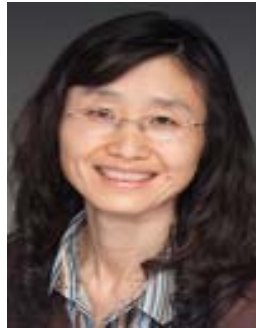
Kumar Srivastava



Sadigeh Mirzaei



Wei Liu



Chenghong Li



Mingjuan Wang

Key Activities

CCSS

- Database Management
- Proposal Development
- Statistical Analyses and Methodology
- Weekly Conference Call -
 - Statistical, Radiation Dosimetry and Coordinating Centers
- Regularly scheduled calls for ongoing projects

Database: Recent Updates

CCSS

- FU5 chronic health conditions graded, frozen (Complete Dec, 2019)
- Baseline survey and MRAF data for 4 new institutions cleaned, frozen (Complete Dec, 2019)
- Baseline chronic health conditions for 4 new institutions graded, frozen (Complete Dec, 2019)
- Subsequent Neoplasms cleaned through FU5 reports (Complete May, 2018)

Additional Data Related Activities

CCSS

- **GWAS Data:** Dissemination of de-identified phenotype data for merging with dbGaP available CCSS genotype data
- **External Linkages:** Society for Assisted Reproductive Technology (**SART**) linked with CCSS female cohort members (ages 15-50) with database.
- **Participation Weights:** Generated weights for use with inverse probability weighting (IPW) to ameliorate bias due to drop-out/non-participation. (Complex data – per survey, time-to-event)

Database: Planned updates

CCSS

- **Expansion Cohort Surgeries (treatment):** Currently cleaning data with ICD codes and dates for surgeries from medical records (coded by DFCI surgeons).
- **NDI** linkage underway - will include deaths through 2017
- **FU6** data cleaning and freezing in Fall/Winter, 2019.
- **Publicly available data set** a possibility in the coming funding period.

Concept Proposals, Ancillary studies

CCSS

- Assist with development of analysis proposals
 - Often helps to schedule a call with the working group chair and statistician.
- Collaborate on ancillary study grant proposals
- Please include us early in process

Manuscript and Analyses Volume

CCSS

January, 2017-June, 2019*:	
Published/In Press Manuscripts:	60
Submitted Manuscripts:	14
Manuscripts in Preparation:	19
Analyses Ongoing/Data sent:	55
Approved, waiting for specific data:	5
Approved Projects in Queue for analyst:	5
Proposals under Development:	5+

*Projects with a CCSS Statistician involved.

Priorities and Challenges

CCSS

Analyses:

- Maintain productivity and quality on analytic projects – distribution among statistical groups provides strengths and challenges

Methodological:

- Have generated participation weights and applied adjustments to several analyses to account for non-participation/drop-out.
 - Time consuming and complicates analyses available
 - Minimal effect on results (so far)
- Plan for infusion of eHealth data and analytic methods

Data Management:

- Continue to maintain high standard for data processes.
- Careful consideration and planning will need to be given to development of a publicly available data set.

Statistical Center SWOT

CCSS

Strengths

- **Personnel:** Strong faculty and staff; long historical memory; good range of technical expertise.
- **Growth:** Expanded analytic support via Yutaka and Kumar's groups (genetic, psychological foci).
- **Innovation:** Develop/apply innovative statistical methods for rigorous analytic results (e.g. IPW for dropout, cumulative mean count, multiple imputation missing data).

Opportunities

- **Ancillary Grants:** More formal involvement in ancillary grants - expand the base of funding and provide consistency in methodology.
- **Big Data:** With emphasis on GWAS, Whole Genome, mobile health data – we should expand our expertise and methodology for high impact research results.
- **Structure/Priorities:** Formal process for setting priorities between projects and more sophisticated system of project tracking, approvals, communication and program/data sharing.

Weaknesses

- **Keep up with Growth:** High demand for analyst time exceeds capacity and higher number of analysts can mean inconsistencies in data use / analytic methods.
- **Bottleneck:** Reliance on lead statistician approval of all papers and proposals creates bottleneck.

Threats

- **Funding:** We need a large group of statisticians to maintain productivity and currently rely on leveraging SJ funds – if we can't continue to do that, this is a threat.
- **Reputation:** Delays in progress or inconsistencies in projects can damage the reputation of CCSS.