Using [linked] Data to Shape Outcomes...

Implications for practice and policy in early childhood settings

Emily Putnam-Hornstein, PhD
Associate Professor & Director
USC Children’s Data Network

Regan Foust, PhD
Research Scientist & Director of Strategic Partnerships
USC Children’s Data Network
Harnessing the scientific potential of linked, administrative data to inform children's programs and policies.
overview

1. CDN 101  *(and why record linkage rocks)*
2. Birth records  *(and how they can be used...)*
3. Family support projects  *(research)*
4. Real time applications  *(PRM)*
CDN 101
(and why you should love record linkage!)
Overview
What is linked administrative data? And why is it important?

http://www.datanetwork.org/about-us/
Probabilistic de-duplication and cross-program linkage of records using an algorithm trained / customized for California data (machine learning)
Bloomberg -- The U.S. government canceled one of its most ambitious health research projects, an effort to follow 100,000 children from before birth through adolescence, after spending about $1.3 billion since 2007 without it ever really getting off the ground.
Birth Records

(so much more than what parents bring home from the hospital...)

section 2
"Each person in the world creates a Book of Life. This Book starts with birth and ends with death. Its pages are made up of the records of the principal events in life. Record linkage is the name given to the process of assembling the pages of this Book..."

Halbert L. Dunn, 1946
1. Universally collected *(good for research & real-time applications)*
2. Nationally standardized and well-documented fields
3. Information for three individuals *(child, mother, father)*
4. Provides a population-base *(spine)* for developing prospective studies
5. Health, demographic, financial, and service information
Birth Cohort

Vital birth records

Population-based information

Substance Use Disorder

CPS Report

Death

Child A

Child B

Child C

Child D

Home Visiting

Subsidized Childcare
Calculating Cumulative Rates of Child Welfare Involvement
Cumulative Percentage of Children Reported for Alleged Abuse / Neglect
Children born in LA County in 1999, by race/ethnicity

- Black
- Hispanic
- White
- Asian / PI

Age in Years

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Documenting intergenerational maltreatment dynamics
Family Support Projects

(in various stages of development...)

section 3
structure

- Medi-Cal
- Child Welfare
- CalWORKs
- Developmental
- WIC
- CalFresh
- FPACT
- IHSS

(today) Research & Development

(in development) Transactional

Child Welfare
- Medi-Cal
- CalWORKs
- CalFresh
- WIC
- FPACT
- IHSS
CHHS Clients [2016]

<table>
<thead>
<tr>
<th>Service</th>
<th>Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medi-Cal</td>
<td>18,176,928</td>
</tr>
<tr>
<td>CalWORKs</td>
<td>1,840,427</td>
</tr>
<tr>
<td>CalFresh</td>
<td>5,808,535</td>
</tr>
<tr>
<td>Child Welfare</td>
<td>331,986</td>
</tr>
<tr>
<td>Developmental</td>
<td>365,114</td>
</tr>
<tr>
<td>WIC</td>
<td>2,000,480</td>
</tr>
<tr>
<td>FPACT</td>
<td>1,928,708</td>
</tr>
<tr>
<td>IHSS</td>
<td>603,626</td>
</tr>
</tbody>
</table>
## Linkage Matrix

<table>
<thead>
<tr>
<th></th>
<th>Medi-Cal</th>
<th>CalWORKs</th>
<th>CalFresh</th>
<th>Child Welfare</th>
<th>Developmental</th>
<th>WIC</th>
<th>FPACT</th>
<th>IHSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medi-Cal</td>
<td></td>
<td>10.1%</td>
<td>30.9%</td>
<td>1.5%</td>
<td>1.7%</td>
<td>10.0%</td>
<td>4.5%</td>
<td>3.3%</td>
</tr>
<tr>
<td>CalWORKs</td>
<td>99.9%</td>
<td></td>
<td>91.2%</td>
<td>5.2%</td>
<td>1.2%</td>
<td>24.8%</td>
<td>3.3%</td>
<td>0.2%</td>
</tr>
<tr>
<td>CalFresh</td>
<td>95.5%</td>
<td>28.5%</td>
<td></td>
<td>2.5%</td>
<td>0.8%</td>
<td>16.5%</td>
<td>3.6%</td>
<td>1.1%</td>
</tr>
<tr>
<td>Child Welfare</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>child</td>
<td>93.0%</td>
<td>48.4%</td>
<td>48.6%</td>
<td></td>
<td>8.2%</td>
<td>30.1%</td>
<td>1.5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>parent</td>
<td>64.2%</td>
<td>14.3%</td>
<td>39.6%</td>
<td></td>
<td>1.2%</td>
<td>7.3%</td>
<td>7.4%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Developmental</td>
<td>82.6%</td>
<td>5.9%</td>
<td>13.4%</td>
<td>3.6%</td>
<td></td>
<td>16.4%</td>
<td>0.4%</td>
<td>21.9%</td>
</tr>
<tr>
<td>WIC</td>
<td>89.2%</td>
<td>23.1%</td>
<td>49.1%</td>
<td>2.9%</td>
<td>3.1%</td>
<td></td>
<td>6.2%</td>
<td>0.2%</td>
</tr>
<tr>
<td>FPACT</td>
<td>45.5%</td>
<td>3.4%</td>
<td>12.2%</td>
<td>1.0%</td>
<td>0.1%</td>
<td>7.1%</td>
<td></td>
<td>0.1%</td>
</tr>
<tr>
<td>IHSS</td>
<td>99.9%</td>
<td>0.5%</td>
<td>10.3%</td>
<td>0.6%</td>
<td>13.3%</td>
<td>0.5%</td>
<td>0.3%</td>
<td></td>
</tr>
</tbody>
</table>
Cross-Program Data
More than half of Medi-Cal enrollees did not receive services from any other CHHS programs.

Two-thirds of CalWORKs clients received benefits from 3 CHHS programs in 2016; another quarter were involved with 4 programs.

1 in 10 individuals involved with the child welfare system touched at least 5 CHHS programs in 2016.
Information universally registered at birth can be used to document assets available to each California newborn at birth. Although assets and conditions at birth are not destiny, thoughtful supports and services may be required to ensure that children with fewer assets find themselves on equal footing with their peers in California. Monitoring the distribution of assets among newborns in different communities can help ensure our investments are intentional and equitable.
The CASSi score is generated by summing the total number of assets (1-12) as coded from each child’s (de-identified) birth record.

Child-specific scores are then aggregated by residential census tract and a mean for census tract asset score is calculated.

Means for all census tracts in California are placed along a continuum, and then grouped into quintiles (e.g., 20% groupings), each representing where that census tract’s mean CASSi score falls relative to all other census tracts in California.
Risk stratification for service delivery / maltreatment prevention
<table>
<thead>
<tr>
<th>At-Risk Birth Score</th>
<th>All Birth Records for Los Angeles County</th>
<th>All Birth Records for XXX XXX Hospital (N=5,129)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>col%</td>
</tr>
<tr>
<td>1</td>
<td>15,429</td>
<td>12%</td>
</tr>
<tr>
<td>2</td>
<td>14,420</td>
<td>11%</td>
</tr>
<tr>
<td>3</td>
<td>13,563</td>
<td>10%</td>
</tr>
<tr>
<td>4</td>
<td>12,153</td>
<td>9%</td>
</tr>
<tr>
<td>5</td>
<td>12,639</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>11,805</td>
<td>9%</td>
</tr>
<tr>
<td>7</td>
<td>12,063</td>
<td>9%</td>
</tr>
<tr>
<td>8</td>
<td>12,867</td>
<td>10%</td>
</tr>
<tr>
<td>9</td>
<td>12,675</td>
<td>10%</td>
</tr>
<tr>
<td>10</td>
<td>12,613</td>
<td>10%</td>
</tr>
</tbody>
</table>
Projected Percentage (and count) of Children Referred to CPS by Ventile for 2013-2015 birth cohorts

- 30.6% of All Children Projected to be Referred by Age 5
- 5,618 Children Projected to be Referred
- 9,308 Children Projected to be Referred
- 102,400 Tier 1
- 12,800 Tier 3

- 30.6% of All Children Projected to be Referred by Age 5
- 5,618 Children Projected to be Referred
- 9,308 Children Projected to be Referred
- 102,400 Tier 1
- 12,800 Tier 3

- 30.6% of All Children Projected to be Referred by Age 5
- 5,618 Children Projected to be Referred
- 9,308 Children Projected to be Referred
- 102,400 Tier 1
- 12,800 Tier 3
Real Time Applications

(not just research...)

section 4

[ www.datanetwork.org ]

Summary

The increased availability and quality of administrative data during the last several decades have led to growing interest in tools and statistical models that can be deployed in real time to predict future events. Predictive risk modeling (PRM) is one such class of tools. PRM is used to automatically generate a risk score for each individual in a given data system, providing an efficient means of screening populations without requiring any additional data entry.

The goal of the project is to establish whether the statistical modeling of historical child protection records can be used to improve the initial screening and triaging of child abuse and neglect referrals. Although this project will not result in a tool without future technological investments, it will lead to the development of data that can inform (in an open and transparent fashion) the opportunities

FUNDERS

California Department of Social Services (CDSS)
Office of Child Abuse Prevention (OCAP)
Laura and John Arnold Foundation (LJAF)
[system] Outcomes

- Follow-up Referral
- ★ Follow-up Referrals (3+ total)
- ★ Follow-up Substantiation
- ★ Follow-up Inconclusive / Substantiation
- Case Opening
- ★ Foster Care Placement

Victim Demographics
- Victim History
- Victim Current
- Referral
- Siblings
- Maternal
- Paternal
- Other

CWS / CMS

0 6 12 18 24 months

long-arc risk...
We use the algorithm to assign each child/referral into 10 evenly sized groups (or deciles) based on the predicted probability of system involvement.
We then examined how well the algorithm risk-stratified children (in unique family and referral events) by looking at how many children were placed in foster care within 24 months (test set, statewide average: 8.0%).
Questions?

ehornste@usc.edu
rfoust@usc.edu