Big Data in Early Childhood: Using Integrated Data to Guide Impact

Rob Fischer, Ph.D. & Beth Anthony, Ph.D.
Center on Urban Poverty & Community Development

Rebekah Dorman, Ph.D.
Cuyahoga County Office of Early Childhood/Invest in Children
Overview

• Early childhood period as crucial opportunity
• Understanding childhood influences in communities
• Creating an integrated child data system as a community resources
• Using integrated data to inform policy and practice
Big data presents tremendous opportunities and challenges

“Analogous to what it must have been like when they first handed out microscopes to microbiologists, social scientists are getting to the point in many areas at which enough information exists to understand and address major previously intractable problems that affect human society.” Gary King, MIT

*Science*, 331, 11 February, 2014, 719
Big data defined

**Definition:**
Digital information that challenges existing technology due to size, complexity, analytic demands

**Common Features**
- **Origin:** Generated as by product of other activity
- **Proportionality:** Encompass all data from a particular source
- **Dynamic:** Capture events over time
- **Variability:** Many types and varying forms
- **Velocity:** Real time data capture
Digital processes generate social data: E-Records, E-transactions, digital media, E-government

Susan LaMotte
@SusanLaMotte

Everyone's talking about #bigdata in #HR. Everyone! read.bi/ZE8SZR

↩️ Reply  Retweet  Favorite  More
In early childhood...many systems at work
But Big Data can be a heavy lift

- Ownership and access
- Transmission
- Data quality
- Data management and storage
- Analytic tools and models
- Visualization
- Applications

A growing field

- Actionable Intelligence for Social Policy (AISP) Network (11 sites) – U. Pennsylvania
  - States: MI, SC, WA, WI, FL*
  - Counties: Los Angeles, CA; Allegheny, PA; Cuyahoga, OH
  - Cities: Chicago, NYC, Philly
- Site hosts vary
  - Government, University, Nonprofits
- Children’s Data Network
  - http://www.datanetwork.org/about-us/
Building a data infrastructure in Cuyahoga County

Since 1999, Cuyahoga County has developed and launched a continuum of supports for young children and their families

• Proven strategies
  – High quality and targeted home visiting
  – High quality early childhood education + educator supports
  – Medical insurance outreach and medical home strategies

• Promising strategies
  – Early childhood mental health services
  – Early literacy services
  – Low cost lead abatement & parent education
Origination of CHILD System - ChildHood Integrated Longitudinal Data (CHILD) System

- Began with child registry of all children born in Cuyahoga County since 1992
- Draws on routinely available administrative data to monitor program delivery and outcomes
- Brings together data at the child level for longitudinal analyses
- Now expanding data to reflect experiences throughout childhood
CHILD System Data Sources
- 400,000 children and counting

- Birth certificates
  - Maternal health
  - Low birth weight births
- Death certificates
  - Infant mortality
- Child welfare
  - Abuse/neglect investigations
  - Placements and services
- Public school data
  - Attendance
  - Kindergarten readiness
  - Proficiency tests
  - Graduation
- Homeless Services
- Public Assistance Receipt
  - Medicaid receipt
  - Food Stamp receipt
  - TANF receipt
  - Child care voucher receipt
- IIC program participation
  - Home visiting services
  - Special needs child care
  - Early mental health
  - Universal pre-k
- Juvenile Justice
  - Delinquency Filings
- Geo-data
  - Neighborhood and environmental conditions
IDS provides longitudinal perspective

Birth - Early Childhood - School Age - Youth - Young Adult

Data Integration

Moms First Home Visiting (Prenatal)

Out-of-Home Placement

Child Care Voucher

Access to Early Learning

Out-of-school time

Juvenile Court Filing

Birth Certificates

Child Abuse & Neglect

Medicaid

Universal Pre-K

CMHA Tenant

Help Me Grow Home Visiting

Early Childhood Mental Health

Public Assistance

Kindergarten Readiness (KRA-L)

Ohio Graduation Test

Newborn Home Visiting

Special Needs Child Care

SNAP

Public School

Postsecondary

Longitudinal Pipeline
IDS approach: CHILD system

ChildHood Integrated Longitudinal Data (CHILD) System

- Teen births
- Low weight birth
- Infant mortality
- Elevated Blood Lead
- Medicaid
- Food Stamp
- TANF
- Child care voucher
- Attendance
- Kindergarten readiness
- Proficiency test
- Graduation test
- Disability
- Abuse/neglect reports
- Foster care
- Home visiting
- Special needs child care
- Early childhood mental health
- Universal pre-k

Public Assist

Birth Certificate

Child Medical Data

Public School

Services

Child Welfare
Child Well-Being: Poverty

Child Poverty in Cuyahoga County
Percentage of Children under 6 years old living below the poverty level, 2005-2009

- < 10%
- 10%- 20%
- 21% - 30%
- 31% - 40%
- over 40%
- No children under 6
- Cleveland Boundary

Source: American Community Survey data, 2005-2009
Prepared by: Center on Urban Poverty and Community Development, MSASS, Case Western Reserve University
Child Well-Being: Public Assistance

Percent of children under the age of 6 receiving public assistance (unduplicated annual totals), Cuyahoga County

- Medicaid
- Food Stamps
- Childcare Vouchers
- Cash Welfare

Years: 2000 to 2010
Child Well-Being: Medicaid

Percent of Projected Eligible Children Receiving Medicaid in 2011 by Cleveland Neighborhood and Suburban Municipality

- Well Below Saturation (<75%)
- Below Saturation (75% to 99%)
- Near Saturation (100+%)
- Suppressed

Note: Values are suppressed when there are 10 or fewer children projected to be under 200% of poverty.

Source: CRIS-E Individual Extract File Cuyahoga Employment and Family Services
Prepared by: Center on Urban Poverty and Community Development, MSASS, Case Western Reserve University
Child Well-Being: Low Birth Weight Births

Low Birth Weight in Cuyahoga County, 2009
Percent of births weighing less than 2500 grams

- Cleveland Boundary
- 3.00% - 8.00%
- 8.01% - 12.00%
- 12.01% - 20.00%
- Suppressed (< 20 Births)

Note: 190 births had missing birth weight or address, or the address could not be geocoded.

Prepared by: Center on Urban Poverty and Community Development, MSASS, Case Western Reserve University
Source: NEO CANDO system, Center on Urban Poverty and Community Development, MSASS, Case Western Reserve University (http://neocando.case.edu)
Child Well-Being: Child Maltreatment

Percent of children under the age of 6 with a report, a substantiated/indicated report, and a report transferred to ongoing, 2000-2010. Cuyahoga County

- **2000**: 6.5% (all reports), 3.0% (substantiated/indicated), 1.5% (transferred to ongoing)
- **2001**: 6.7% (all reports), 3.2% (substantiated/indicated), 1.6% (transferred to ongoing)
- **2002**: 6.8% (all reports), 3.3% (substantiated/indicated), 1.7% (transferred to ongoing)
- **2003**: 7.0% (all reports), 3.5% (substantiated/indicated), 1.8% (transferred to ongoing)
- **2004**: 7.1% (all reports), 3.6% (substantiated/indicated), 1.9% (transferred to ongoing)
- **2005**: 7.3% (all reports), 3.7% (substantiated/indicated), 2.0% (transferred to ongoing)
- **2006**: 7.5% (all reports), 3.9% (substantiated/indicated), 2.1% (transferred to ongoing)
- **2007**: 7.7% (all reports), 4.1% (substantiated/indicated), 2.2% (transferred to ongoing)
- **2008**: 7.9% (all reports), 4.3% (substantiated/indicated), 2.3% (transferred to ongoing)
- **2009**: 8.0% (all reports), 4.4% (substantiated/indicated), 2.4% (transferred to ongoing)
- **2010**: 8.2% (all reports), 4.6% (substantiated/indicated), 2.5% (transferred to ongoing)

**Legend**
- Blue: All reports
- Red: Substantiated/indicated
- Green: Transferred to ongoing

**Source**
Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University
The Power of Integrated Data

• Data helps inform our understanding of the early childhood system
• Individuals and families interact with multiple systems and services, so integrated data offers a more complete view of reality
• Understanding of how systems work and how to better meet existing needs can be informed by integrated data
• Service models emphasize long term and collective impact, so data needed across services and over time
Using Integrated Data

Examples

• Tracking well-child visits
• Understanding school readiness
• Informing ‘Pay for Success’ planning
Child Health Example

• Despite dramatic increase in health insurance coverage for children ages 0-6, only 43% of poor children got all the recommended well-child visits in the first year of life.
• Integrated data showed that 49% of these families were involved with other services close to birth.
• A preventive approach developed by having dedicated staff at clinics reach out to families.
• Result - Medical Home Pilot with one Patient Advocate each at NEON & Neighborhood Family Practice
  – 86% of families completed scheduled well-child visits, double the rate for children born on to Medicaid in Cuyahoga County.
School Readiness example

- State of Ohio passed Third Grade “Reading Guarantee” in 2012
- Used integrated data to understand early child influences on kindergarten readiness
  - Cohort of all children entering CMSD 2007-2010
    - Early childhood data examined retrospectively to birth
    - Early school performance data examined forward to 3rd grade

Figure 1: Cohort-based Study Design

- Showed negative influence of maternal education, poverty, as well as home visiting and high-quality preschool experiences
Children from UPK Sites Entering CMSD Are More Prepared

The Kindergarten Readiness Assessment-Literacy (KRA-L) has a value of 0-29 and has three score bands: Band 1 0-13 (Assess broadly for intense instruction), Band 2 14-23 (Assess for targeted instruction), and Band 3 24-29 (Assess for enriched instruction). Children scoring Band 1 may be at serious risk of being unprepared for kindergarten.

The Kindergarten Readiness Assessment-Literacy (KRA-L) has a value of 0-29 and has three score bands: Band 1 0-13 (Assess broadly for intense instruction), Band 2 14-23 (Assess for targeted instruction), and Band 3 24-29 (Assess for enriched instruction). Children scoring Band 1 may be at serious risk of being unprepared for kindergarten.
Pay for Success Example

• CHILD system data were used to explore the experiences of homeless families with child welfare services, particularly the use of foster care services.
• The population was sufficiently large and costly to merit new intervention.
• Cuyahoga County launched the first county-level Pay For Success initiative with a focus on reducing foster care stays for children whose families experience homelessness.
## Disseminating Results

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Births</td>
<td>1,443</td>
<td>16,246</td>
</tr>
<tr>
<td># Teen Births, mother's age 10 - 14 (rate per 1,000)</td>
<td>2 (1)</td>
<td>42 (1)</td>
</tr>
<tr>
<td># Teen Births, mother's age 15 - 19 (rate per 1,000)</td>
<td>124 (39)</td>
<td>2,031 (41)</td>
</tr>
<tr>
<td>% of Mothers without High School diploma</td>
<td>14%</td>
<td>19%</td>
</tr>
<tr>
<td>% Low Birth Weight</td>
<td>9%</td>
<td>10%</td>
</tr>
<tr>
<td>% Premature Low Weight Births</td>
<td>6%</td>
<td>7%</td>
</tr>
<tr>
<td>% Mothers w/adequate prenatal care</td>
<td>52%</td>
<td>53%</td>
</tr>
<tr>
<td>% Mothers w/out prenatal care</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>% Healthy Births</td>
<td>53%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Infant Deaths (rate per 1,000 births)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Births</td>
<td>1,877</td>
<td>16,246</td>
</tr>
<tr>
<td># Teen Births, mother's age 10 - 14 (rate per 1,000)</td>
<td>12 (2)</td>
<td>42 (1)</td>
</tr>
<tr>
<td># Teen Births, mother's age 15 - 19 (rate per 1,000)</td>
<td>358 (79)</td>
<td>2,031 (41)</td>
</tr>
<tr>
<td>% of Mothers without High School diploma</td>
<td>32%</td>
<td>19%</td>
</tr>
<tr>
<td>% Low Birth Weight</td>
<td>14%</td>
<td>10%</td>
</tr>
<tr>
<td>% Premature Low Weight Births</td>
<td>9%</td>
<td>7%</td>
</tr>
<tr>
<td>% Mothers w/adequate prenatal care</td>
<td>42%</td>
<td>53%</td>
</tr>
<tr>
<td>% Mothers w/out prenatal care</td>
<td>2%</td>
<td>1%</td>
</tr>
<tr>
<td>% Healthy Births</td>
<td>36%</td>
<td>49%</td>
</tr>
<tr>
<td>Infant Deaths (rate per 1,000 births)*</td>
<td>29 (15)</td>
<td>164 (10)</td>
</tr>
<tr>
<td>#(%') Children with Elevated Blood Lead Levels (≥5µg/dl)</td>
<td>637 (16%)</td>
<td>3,951 (16%)</td>
</tr>
<tr>
<td>#(%') Children with Elevated Blood Lead Levels (≥10µg/dl)</td>
<td>168 (4%)</td>
<td>1,174 (5%)</td>
</tr>
</tbody>
</table>
Informing system responses

• What does the CMSD kindergarten class look like?

<table>
<thead>
<tr>
<th>Family Well Being Indicators</th>
<th>CMSD Kindergarten Class 2008-09(^1)</th>
<th>Cleveland(^2)</th>
<th>Cuyahoga County(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Teen Births, mother’s age 10 – 14</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>% Teen Births, mother’s age 15 – 19</td>
<td>22.4</td>
<td>16.7</td>
<td>9.8</td>
</tr>
<tr>
<td>% of Mothers without H.S. diploma</td>
<td>41.7</td>
<td>30.2</td>
<td>15.9</td>
</tr>
<tr>
<td>% Low Birth Weight</td>
<td>12.6</td>
<td>11.6</td>
<td>9.4</td>
</tr>
<tr>
<td>% Premature Low Weight Births</td>
<td>8.7</td>
<td>8.2</td>
<td>6.7</td>
</tr>
<tr>
<td>% Mothers w/ adequate prenatal care (Kessner Index)</td>
<td>63.1</td>
<td>69.4</td>
<td>81.3</td>
</tr>
<tr>
<td>% Mothers w/out prenatal care</td>
<td>1.9</td>
<td>1.7</td>
<td>0.9</td>
</tr>
<tr>
<td>% Healthy Births (based on information about Apgar, prenatal care, birthweight, gestational age)</td>
<td>56.4</td>
<td>61.5</td>
<td>70.9</td>
</tr>
<tr>
<td>% Children with a substantiated or indicated report of abuse/neglect by age 4</td>
<td>12.1</td>
<td>9.6</td>
<td>5.1</td>
</tr>
<tr>
<td>% Children referred to ongoing services with Child &amp; Family Services by age 4</td>
<td>19.8</td>
<td>14.7</td>
<td>7.6</td>
</tr>
<tr>
<td>% Children with any report of abuse/neglect by age 4, including substantiated and unsubstantiated</td>
<td>35.2</td>
<td>26.7</td>
<td>14.7</td>
</tr>
<tr>
<td>% Children in households receiving Food Stamps in 2008</td>
<td>76.9</td>
<td>51.1</td>
<td>28.8</td>
</tr>
<tr>
<td>% Children in households receiving Cash Assistance in 2008</td>
<td>19.0</td>
<td>11.3</td>
<td>6.1</td>
</tr>
</tbody>
</table>
Weaknesses of early childhood re: Big Data

• Organizational culture and politics not sufficiently data driven and data savvy
• Big data requires long term investments, but funding is tough to raise, and short term results come quicker from “one off” data collection
• Proportion of early childhood research effort devoted to big data small relative to survey and observational data
• Too few early childhood folks with Big Data skills
Challenges to pursuing Big Data

• Funding!
• Data access – data sharing, privacy concerns
• Data depth – matching variables, service measures, outcome measures
• Data frequency/recency
  o how frequent is enough?
  o Capacity issues for data partners
Navigating Big Data challenges

• Defining the “system”
• Securing and sustaining funding
• Promoting collaboration
• Using data to inform decisionmaking
• Managing partner expectations
What's Next?

- Undertake more study of factors that predict longitudinal outcomes for children
- Explore mechanisms to make aggregate data from more available to the policy and practice communities
- Extend CHILD System to include more early childhood data (e.g., Head Start, developmental disabilities) and data from later adolescence (e.g., juvenile court involvement, school systems)