The Science of Good Policy

Integrated Data Systems

Pathways to Informed Decision-Making

Milbank Memorial Fund
Reforming States Group
Regional Meetings - Fall 2016

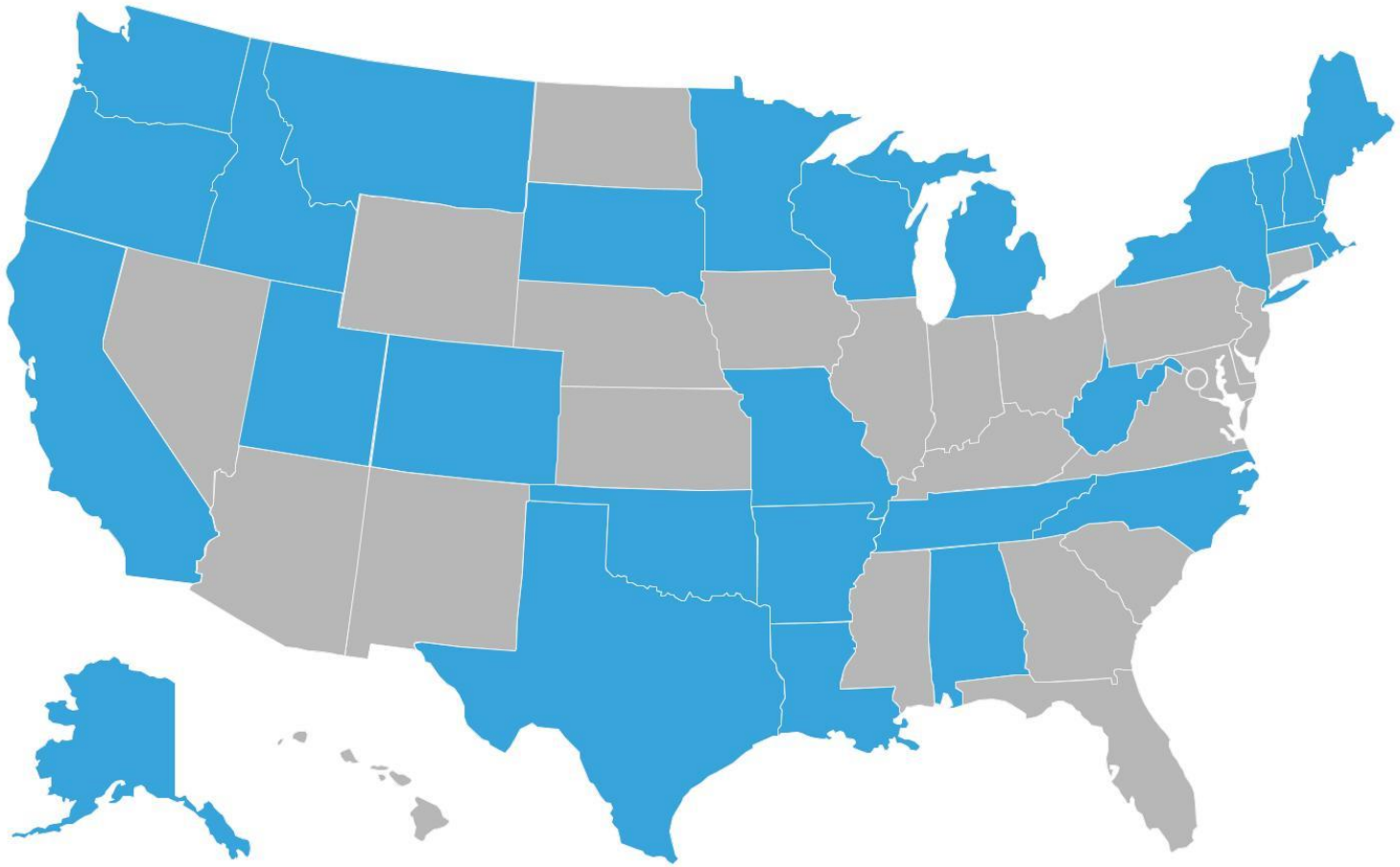
Christopher Kelleher | Center for Evidence-based Policy

Who are we?

Center for Evidence-based Policy

- Based at Oregon Health & Science University
- Established in 2003
- Address policy challenges by applying evidence and improving collaboration
- Convene research collaboratives to drive effective policy and program reforms with proven returns on investment

Who do we serve?



EVIDENCE-BASED POLICY CENTER FOR

Integrated Datasets

Oregon Data Sources

Integrated from 2001 forward

- Birth Record
- Child Welfare
- Human Services Programs
- Self Sufficiency Payments
- Addiction Treatment
- Education
- Juvenile Justice and Corrections
- Home Visiting Programs

Integrated Datasets

Denominator

- All children born since 2001, matched to their parents
- Enables predictive modeling

Geographic Specificity

- Data coded at the census-block level

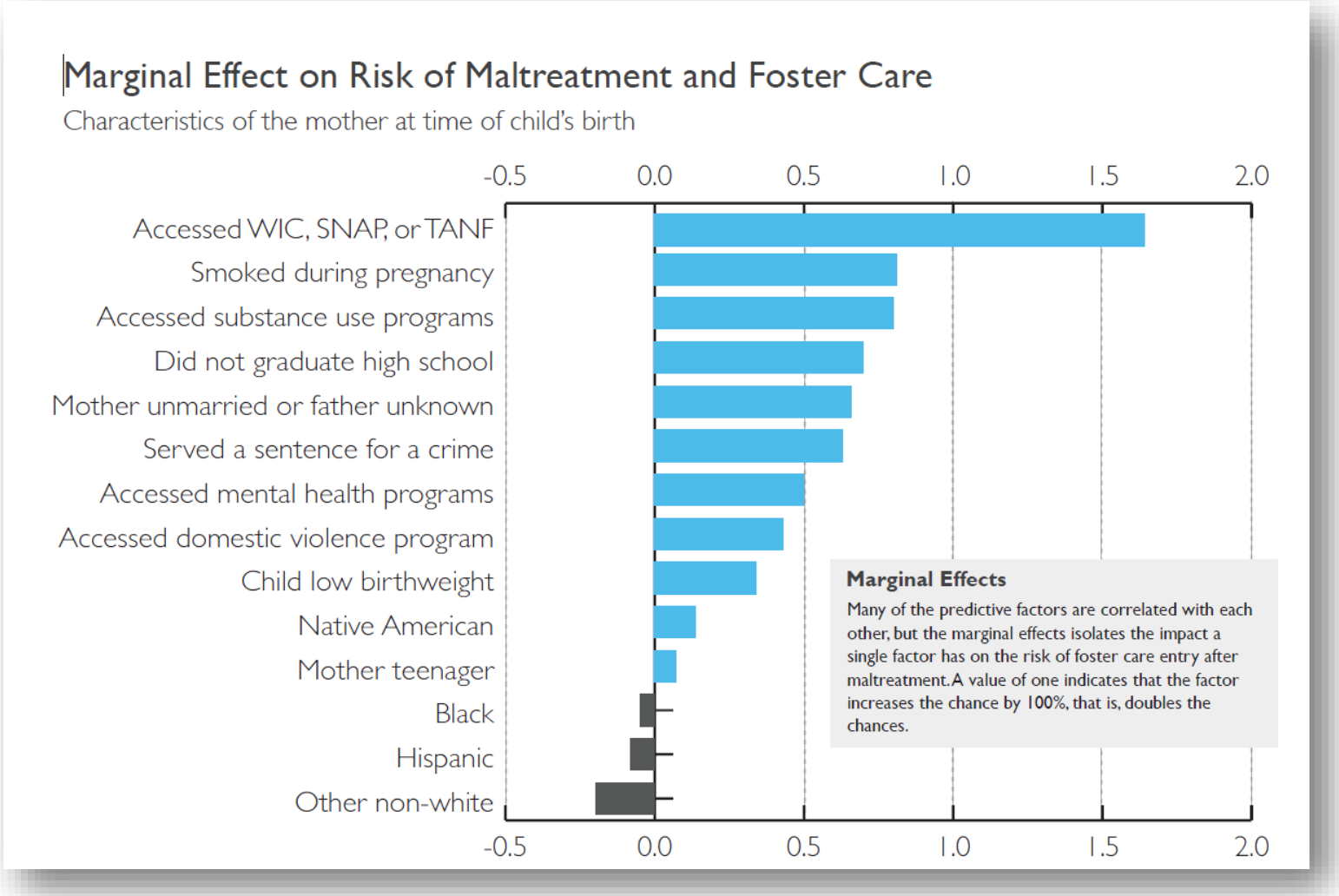
Longitudinal

- Tracing paths through the system over time

Agency Cost Structures

- Supports modeling of cost avoidance scenarios

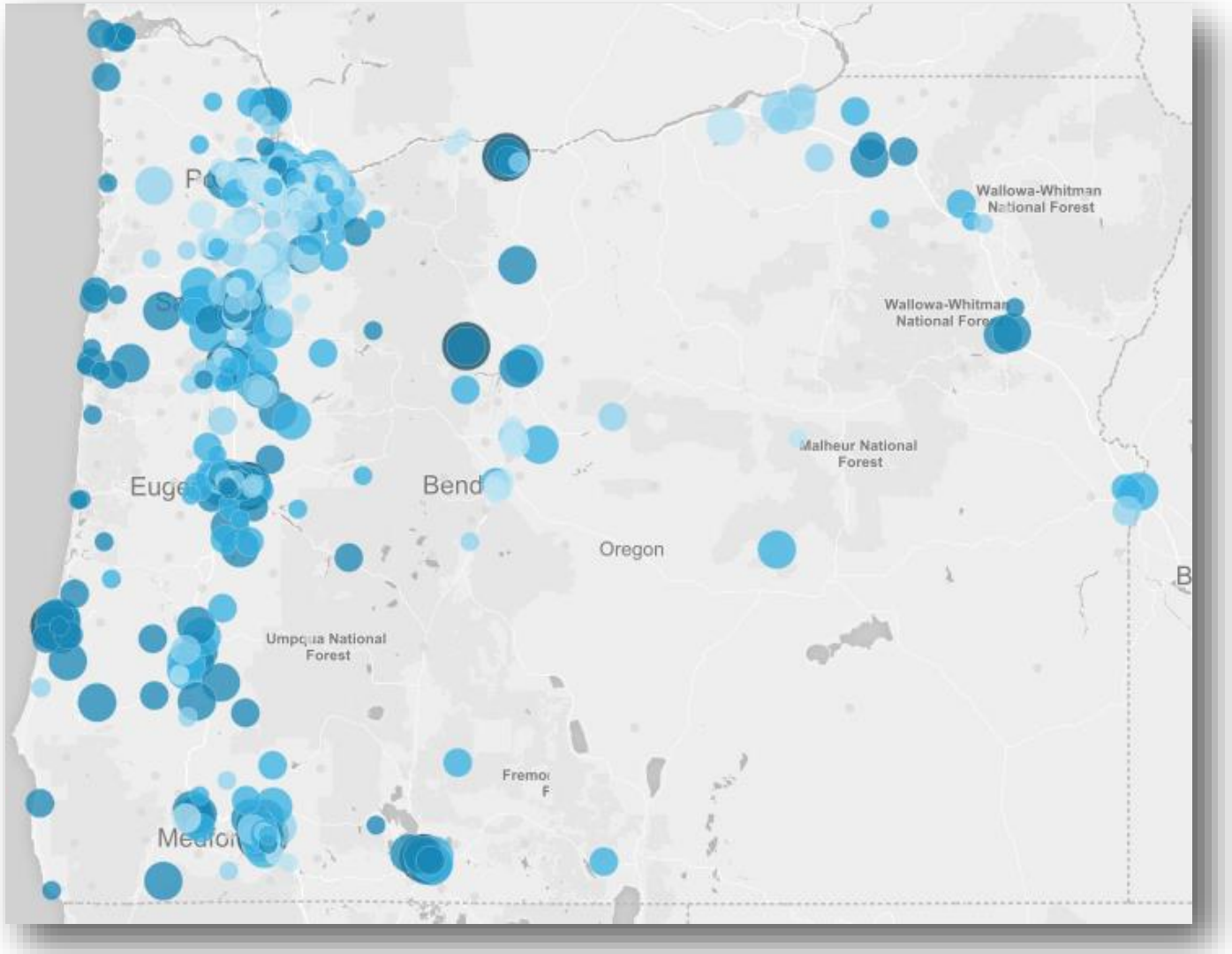
Why Integrated Data?



Identify risk factors for foster care placement

CENTER FOR EVIDENCE-BASED POLICY

Why Integrated Data?



Visualize hotspots

EVIDENCE-BASED POLICY CENTER FOR

Why Integrated Data?

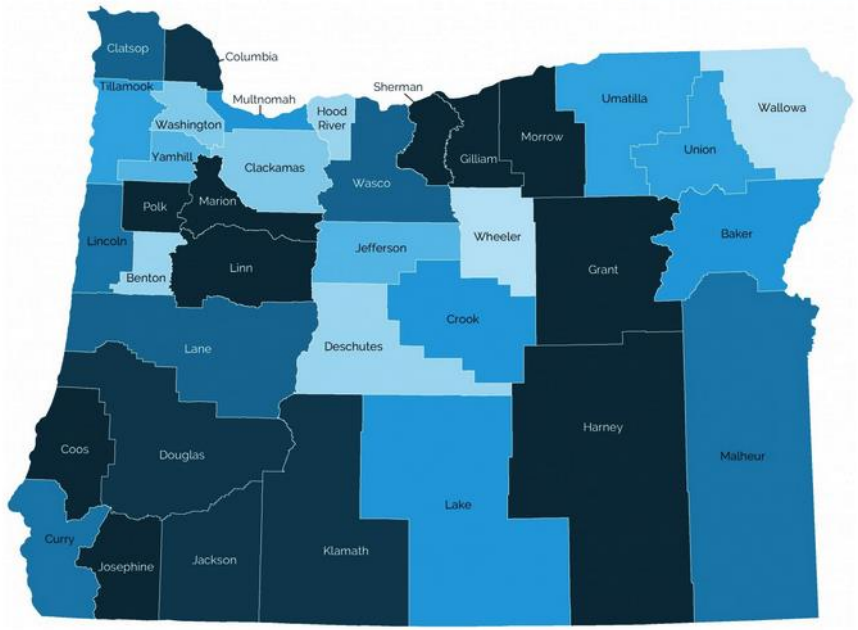
ALL AGES | 0 - 4 YEARS | 5 - 9 YEARS | 10 - 14 YEARS | 15 - 19 YEARS

DOWNLOAD PDF

FOSTER CARE ENTRY RATES

(Per 1,000 age-equivalent children)

- 0.0 - 1.5
- 1.5 - 3.0
- 3.0 - 4.5
- 4.5 - 6.0
- 6.0 - 7.5
- 7.5 - 9.0
- 9.0 - 10.5
- 10.5 - 12.0
- 12.0 - 13.5
- 13.5 +



START

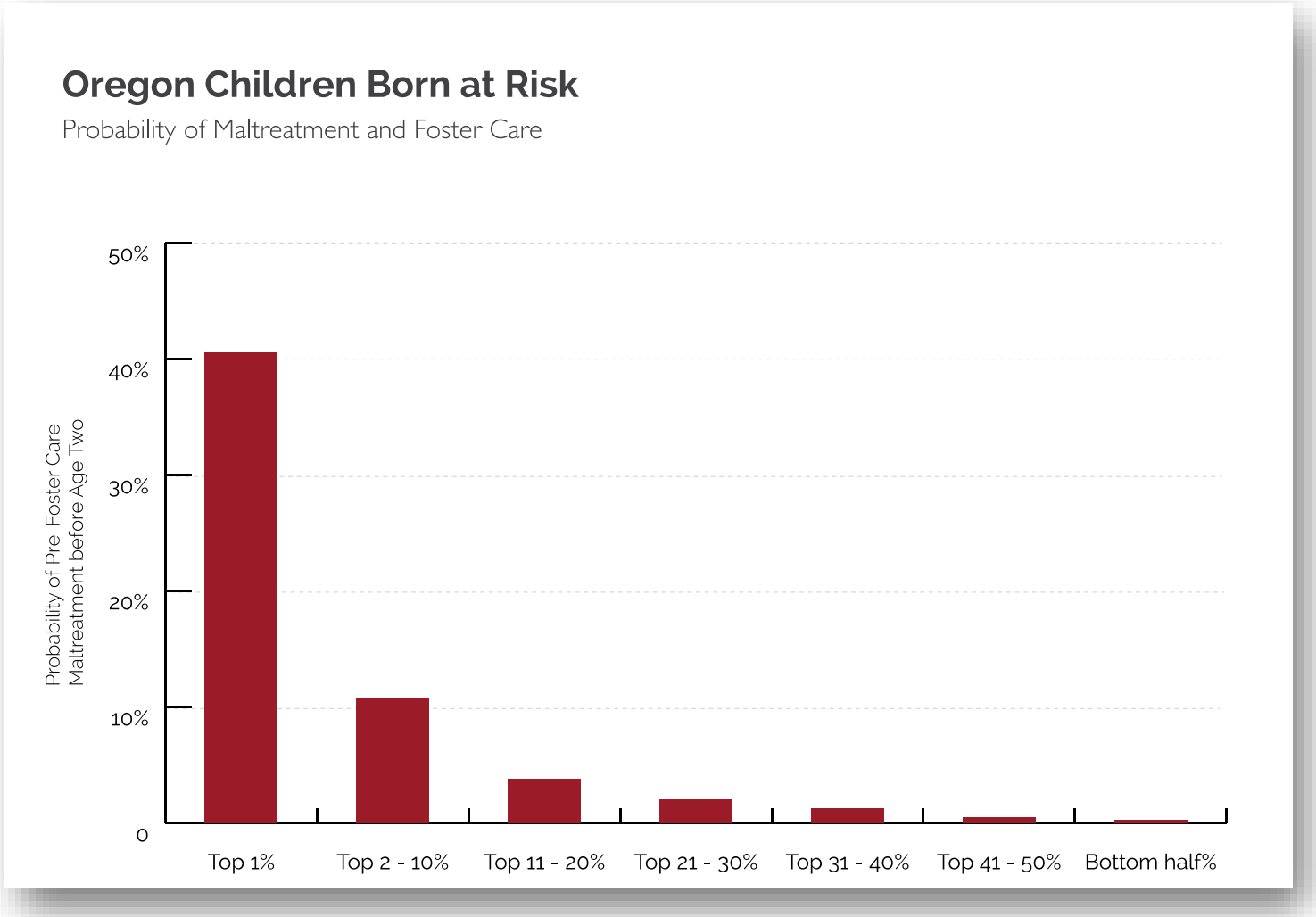
2007

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 All Years

Map changes over time, place, and subgroups

CENTER FOR EVIDENCE-BASED POLICY

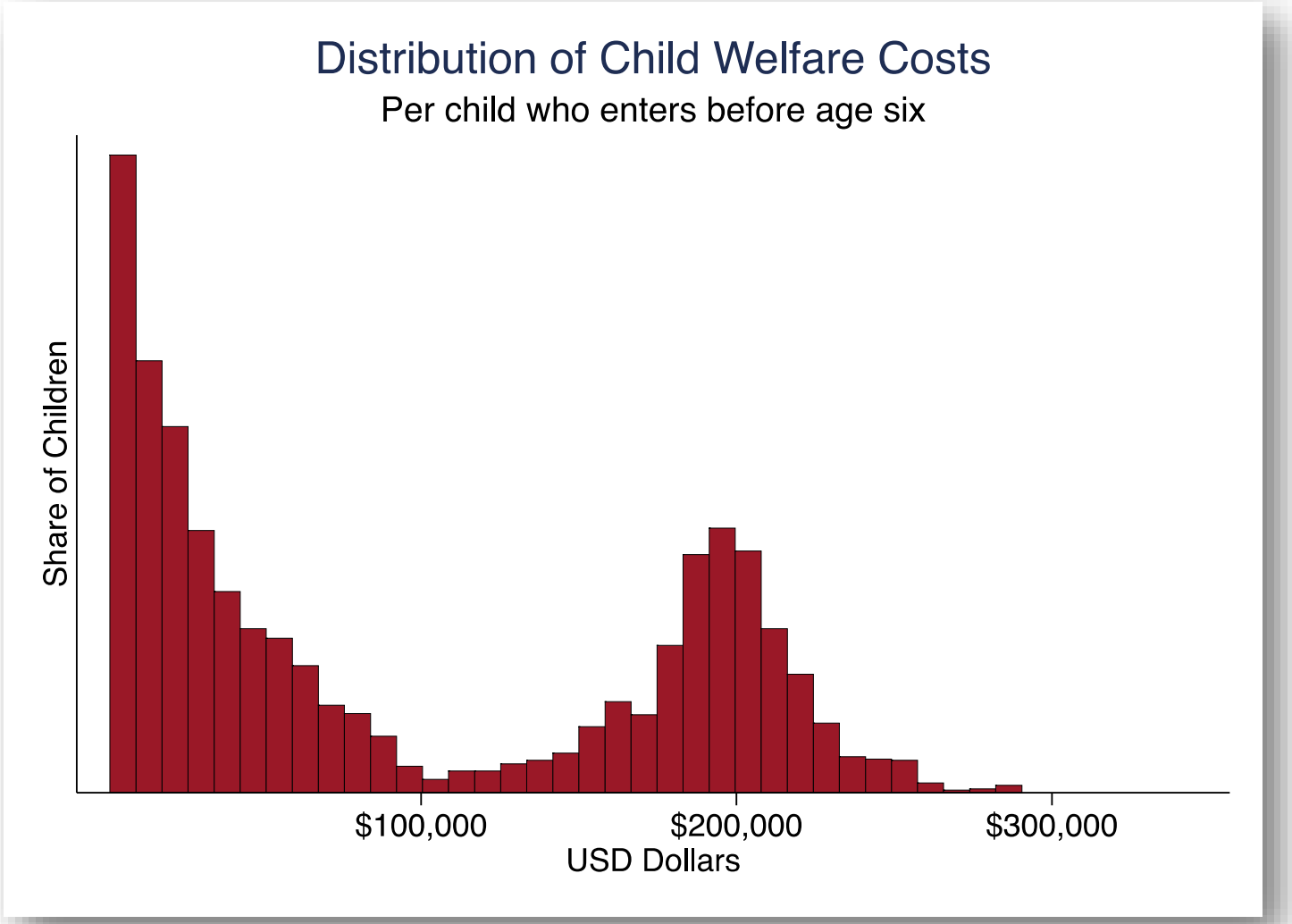
Why Integrated Data?



Target interventions for the most at-risk children

EVIDENCE-BASED POLICY CENTER FOR

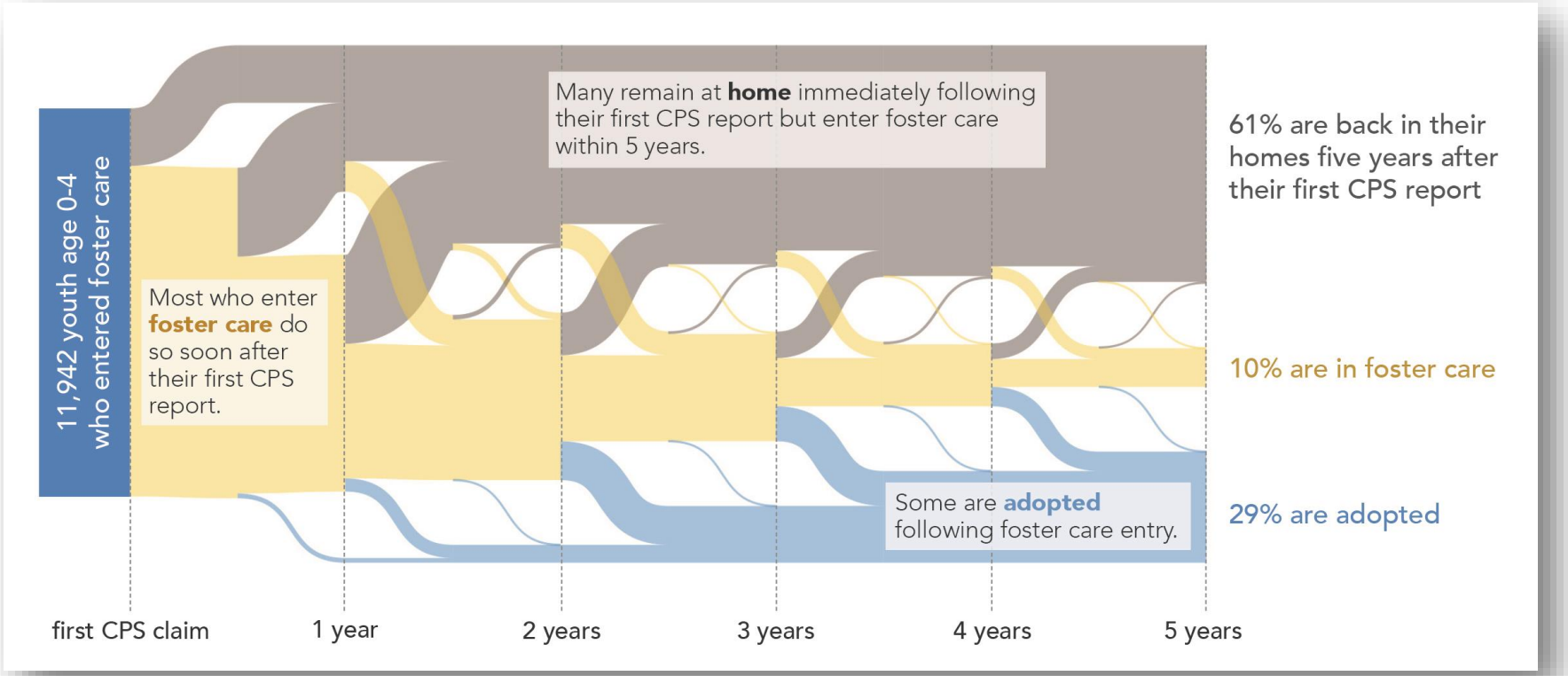
Why Integrated Data?



Understand service-demand patterns and program costs

EVIDENCE-BASED POLICY CENTER FOR

Why Integrated Data?



Trace paths through the system

Why Integrated Data?

OREGON CHILDREN AT RISK

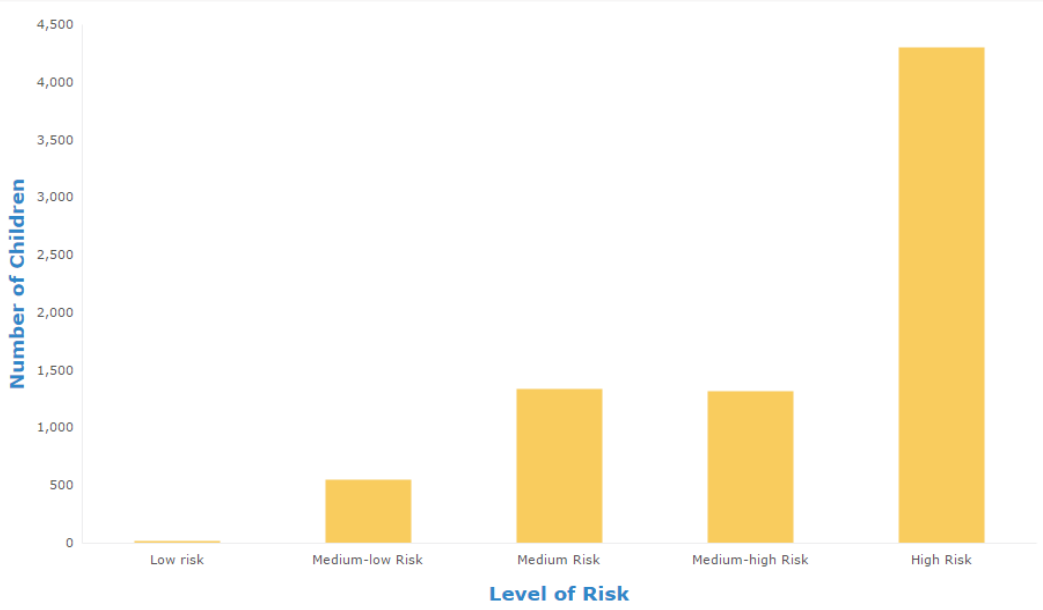
Between 2001 and 2010, almost half a million children were born in Oregon. Over 2% of these children entered the foster care system before age four. Many of these children were at foreseeable risk. P4P research and modeling has identified key characteristics of children and their families that predict the likelihood of childhood abuse and neglect. Children without these characteristics are at much lower risk.

These data show all children born in Oregon between 2001 and 2010. Children with predictive factors are at much higher risk of maltreatment and entering foster care than children with none of these factors.

Oregon Children 7,527

Select and unselect any combination of the following risk factors:

- [x] Poverty
- [ ] Parental education
- [ ] Parental substance abuse
- [x] Parental criminal activity
- [ ] Parental mental health
- [ ] Family instability



Low Risk = x2 the chances of being maltreated and entering foster care relative to someone with no predictive factors. Medium-low = x10. Medium = x20. Medium-high = x40. High = x80.

Expand access to actionable information for public, private and community decision-makers

CENTER FOR EVIDENCE-BASED POLICY

The "Use Case" Pathway

Step 1
Plan and Organize the Use Case

Step 2
Collect and Integrate the Data

Step 3
Analyze and Apply Results

1.1. Identify a real-world research or policy question

1.2. Engage partners to collect, integrate, and analyze data

2.1. Identify data required to answer the research or policy question

2.2. Match identifiers across datasets

3.1. Analyze integrated data to answer research questions

3.2. Capture learning to build a business case for a full IDS

CENTER FOR EVIDENCE-BASED POLICY

The “IDS” Pathway

4. Secure executive authorization for a project charter and governance structure

5.1. Identify datasets and acquire data use agreements from data partners

5.2. Build data collection, processing, and integration infrastructure

5.3. Incorporate analysis, modeling, and decision-support capabilities

6. Institutionalize IDS as a tool for planning and policy-making

CENTER FOR EVIDENCE-BASED POLICY

The Science of Good Policy

Chris Kelleher
Center for Evidence-based Policy
kellehch@ohsu.edu

Dan Vizzini
Center for Evidence-based Policy
vizzinid@ohsu.edu

Extra Slides

The “Use Case” Project

Step 1. Plan and Organize								
Task 1.1 Research Questions								
Task 1.2 Engage Partners								
Step 2. Collect and Integrate Data								
Task 2.1 Obtain Required Data								
Task 2.2 Match Data Identifiers								
Step 3. Analyze and Apply Results								
Task 3.1 Analyze Integrated Data								
Task 3.2 Capture Lessons Learned								
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8

Timeline

18-24 months

Cost

$250,000 - $400,000

Costs scale with extent of integration, analysis, reporting

Assumptions

- No executive commitment
- Available project staff
- Capacity to match data with a unique, reliable identifier
- One-time data pull
- Limited scope

CENTER FOR EVIDENCE-BASED POLICY

Step 1.1: Define The Questions

- ☑ What are the risk factors for entry into foster care?
- ☑ What short-term and long-term outcomes are strongly associated with entry into foster care?
- ☑ What interventions and programs are most effective at reducing maltreatment and/or foster-care entry?
- ☑ What are the fiscal costs and benefits of effective interventions?

Step 1.2: Engage Partners

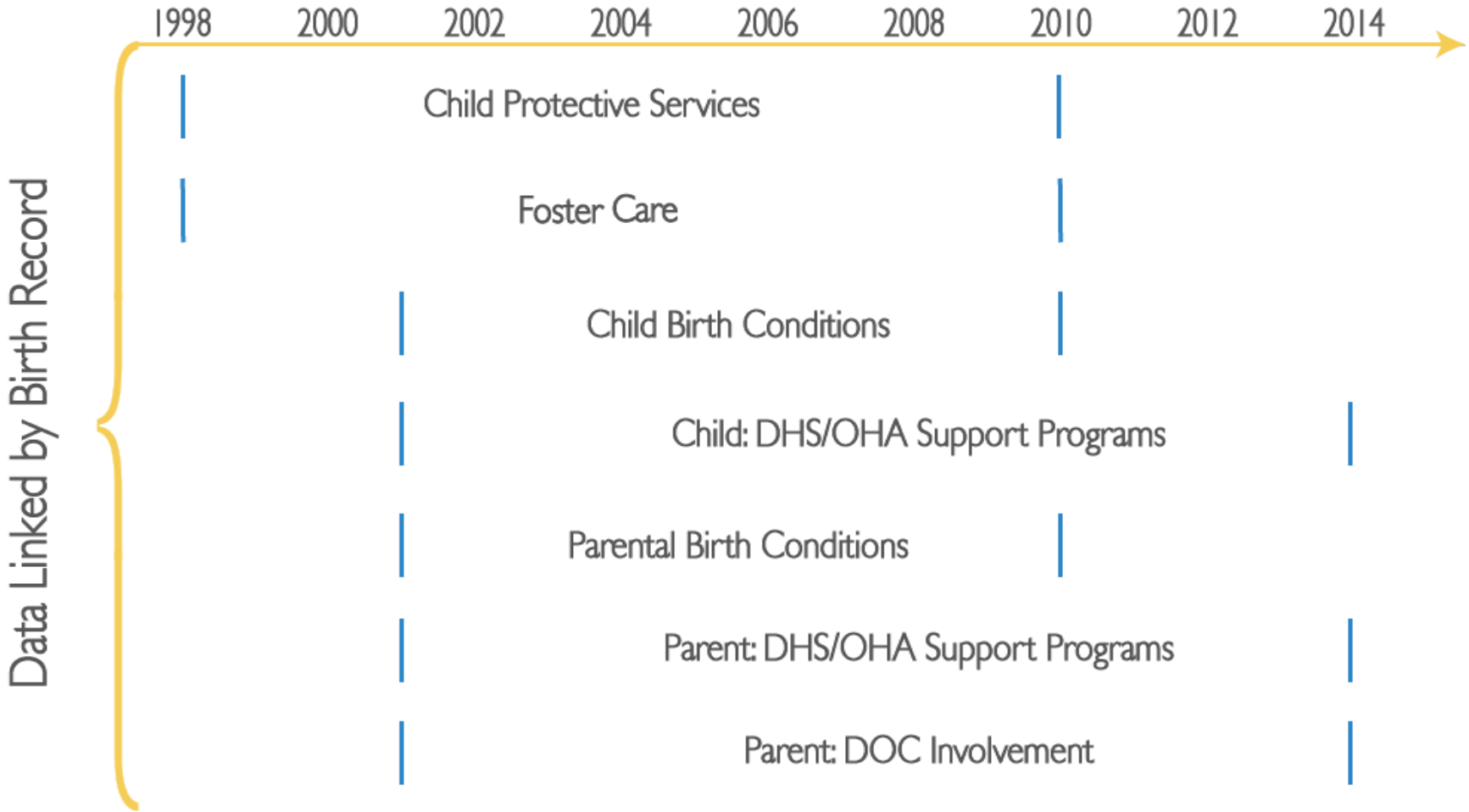
☑ Data Owners

- Vital Records
- DHS Child Welfare: Child Protective Services and Foster Care
- DHS/OHA Child Support Services

☑ Analysis Partners

- ECONorthwest: Statistical and Economic Analysis, Modeling and Visualization

Step 2.1: Identify Data



Data Use Agreements

Data Use Agreements (DUAs) & Institutional Review Board (IRB)

- Agreements with each data partner
- Sequence
  - Negotiate DUA language
  - Submit study protocol to the IRB
    - Revise DUAs if requested by the IRB
  - Send IRB approval to agencies. Then sign DUAs

Data Privacy

Privacy Protections

- De-identified dataset
- 10/50 Rule: No release of information when . . .
  - Cell size is 10 or fewer
  - Where the measured population totals 50 or fewer
  - When the aggregated data may lead to the identification of one or more specific individuals

Data Hygiene

Cleaning and Interpreting the Data

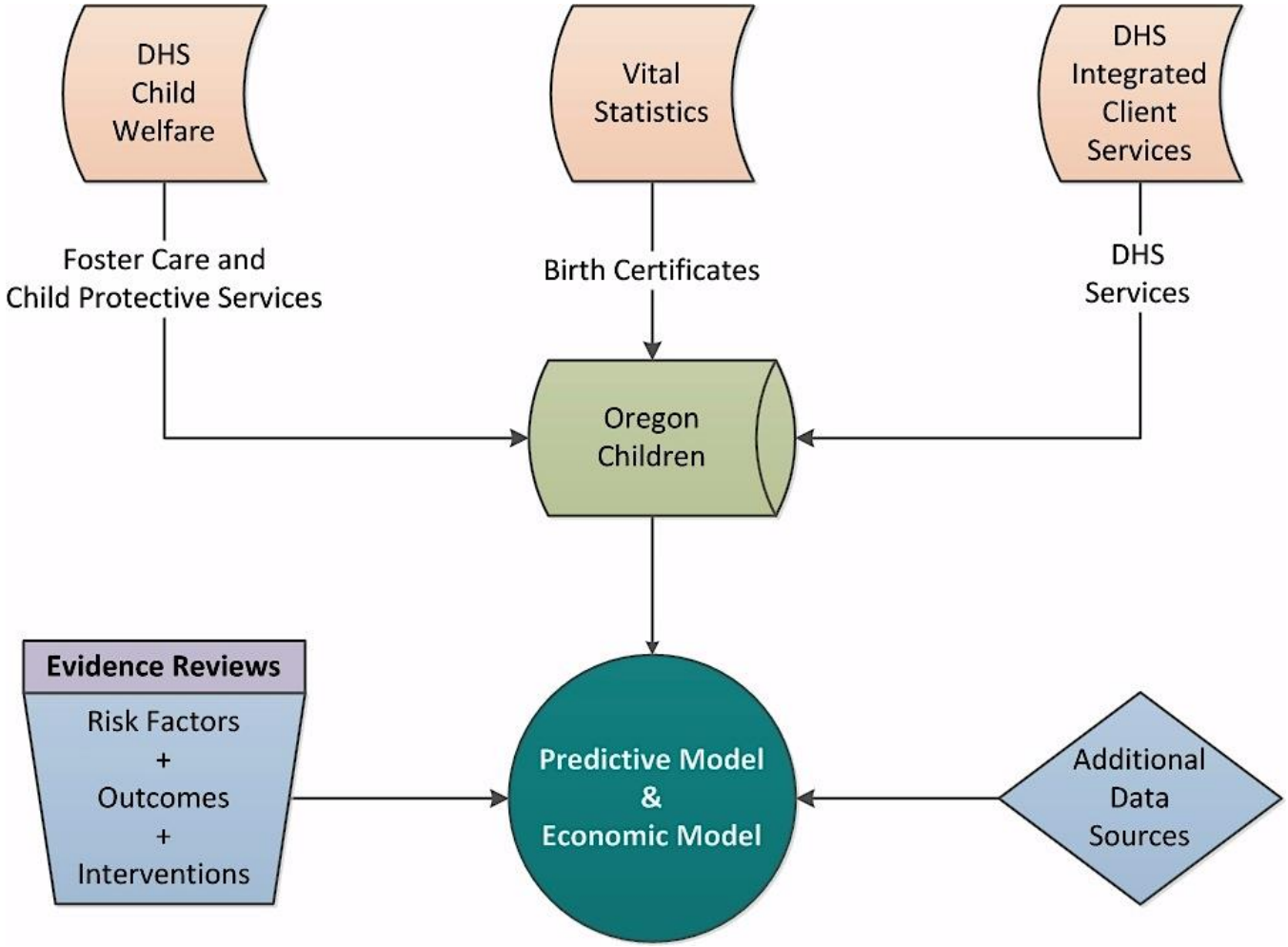
- What are the on-the-ground realities of collecting certain data?
- Is there data that the agency doesn't trust?
- How do we interpret missing or duplicate data?
- Have fields evolved over time?

Step 2.2: Match Identifiers

DHS Integrated Client Services (ICS)

- Agencies send data to ICS
- ICS matches individuals and deidentified data
- ICS assigns unique Study IDs
- Project receives a unified dataset

Mapping the Integration



CENTER FOR EVIDENCE-BASED POLICY

Step 3.1: Analyze the Data

Iterative Process of Statistical Analysis and Modeling

- Begins at the very start of the project with research design
- Is refined by (and influences) data collection and integration
- Is as much a means of auditing the data integrity and reliability, as it is a tool for informing decision-makers
- At its best… raised more questions than it answers

Step 3.2: Capture Lessons

Build a Business Case for a Full IDS

- Answer the original research questions
- Demonstrate utility of integrated data to decision-making
- Capture lessons learned from the Use Case project
- Identify the costs and benefits of a fully-realized IDS

Lessons Learned

- Framing… Pathway presentation fits in between IDS Intro and State System presentation
- Human Engineering… Center took time to build relationships in order to secure data sharing agreements
- Would have gone faster if we started with vital statistics (denominator) and then capture DHS data.
- Use Case pathway is about learning from mistakes at a manageable scale.