# Frontiers in Child Care Research
## Annual Meeting of the Child Care Policy Research Consortium
### Radisson Lord Baltimore Hotel
### March 8–11, 2005

### SESSION 11

**Linking Data from Different State Systems**

*What are we learning from efforts to use a common identifier across different administrative datasets—including data from child care subsidy, licensing, and registry systems, as well as TANF, Unemployment Insurance, and Census data?*

Facilitator:     Rod Southwick, Massachusetts Office of Child Care Services

Discussants:    Mairéad Reidy, Chapin Hall Center for Children, University of Chicago

                  Alan Sweet (retired), Wisconsin Department of Workforce Development

                  W. David Patterson, Office of Research and Statistics, South Carolina Budget and Control Board

Scribe:            Shawn Marie Pelak, Child Care and Early Education *Research Connections*, Inter-university Consortium for Political and Social Research, University of Michigan

                  Rosemary Kendall, National Child Care Information Center


## W. David Patterson
- South Carolina Budget and Control Board is a service agency
- Neutral, no programmatic responsibilities
- Statistical and research focused
- Experienced in gathering, integrating, and disseminating data
- Build off of existing systems (legacy systems from state agencies and private sector)
- Create a unique ID (not related to any other number)
- Identifiers pulled off of the statistical data
- Data always "owned" by the originating agency. Must have permission to use or link any data

Using data in the Data Warehouse, agencies can:
- Evaluate programs
- Look at outcomes
- Better understand how their programs interact with other agency programs

- Study health, human service, education, and law enforcement issues
- Analyze statistical (aggregate) information
- Access analytic data cubes
- Partner in the development of customized software applications

How is ORS involved with early care and education?
- Immediate assistance in the migration of all SSBG-funded activity to DSS
- Support in finding effective interim solutions to fragmented and/or outdated child care data systems
- Development of an information system infrastructure capable of fulfilling the DSS Strategic Plan for Child Care Quality (July 2004):
    - Facilitating data exchange and common standards of evaluation among agencies responsible for both public and nonpublic early care and education facilities
    - Compiling and tracking information on facilities, staff, and children
    - Automating on site visits by DSS Licensing, Child Care Quality, and Department of Education monitors and evaluators
    - Allowing linkage to the data warehouse for expanded analyses and "value added" uses
    - Providing useful information and "tools" to child care providers and parents via the Internet
- What about cost?
    - 15 percent appropriated cost
    - Agencies provide match
    - $150,000 for Stars project
- What about getting everyone onboard?
    - Section leader creates trust over 30 years
    - Long history of always staying neutral
    - Agencies always in full control of their data
    - Eventually the agencies get to see the benefit of having data quickly
- Can know if a child is both in subsidy and in a school if the agencies ask for it
- Child protective services is included and can be linked to other data
- Unique identifier system
    - Done in SASS, uses Social Security, Medicaid, race, gender, age
    - Data as acquired run through every month
    - Error rate 0.43 percent
- South Carolina can use database to address fraud and abuse
- Could follow children over time to see how they use various education systems.
- Most of the data go back to 1993; some records are older
- Database holds some wage information

<u>**Alan Sweet**</u>

**Handout: Wisconsin Data Sharing Project**

- Works with University of Wisconsin Extension and Department of Workforce Development
- Worked with funding provided under the child care data capacity grant
- Moving in direction of quality rating system and tiered reimbursement
- Jason Bierbrauer, UW-Extension, is the technical person at the meeting
- Experience in building merged data system—near completion

Central data base comprised of records from two regulatory systems and their respective records:
- Licensing (group center and family programs) – Licensing administered in different state agency. Merging their data first required building a partnership and trust. A further challenge then arose due to their system being quite outdated.
- Certification (license-exempt family care) – Relatively new, web-based system already existed. Used existing data base as core structure.
- Accreditation
- Wisconsin Shares (child care subsidy system)
- Child and Adult Care Food Program system
- Child Care Data Warehouse
    - o Use for regular reporting and ad hoc reporting (e.g., legislature, media)
    - o Market rate survey data analysis annually
- Interface with child care registry
    - o Wisconsin is the leading state for *The Registry*
    - o This will help Wisconsin move forward on quality rating system
- Working with NACCRRA on daily upload of data to update records in NACCRRAware
- Child Care Web mapping (out of UW-Extension office)
    - o Public access Web site
    - o Information relating to all child care providers in Wisconsin
        - Query to get map or information
        - Plan to add Head Start and Pre-Kindergarten
        - Ideal infrastructure to add public education in the future
- Plan to develop public access web site to see providers' star ratings
    - o Would not replace CCR&Rs
- All built using a single common identifier—had to negotiate with agencies
    - o Individual systems can keep their own system IDs and also use common ID (10-digit provider number plus 3-digit location number)

**Lessons learned:**
- Wisconsin
    - o It is difficult to merge systems
    - o Systems of each agency depend on technology and can vary widely
    - o Funding is always a concern

- o Providers have concerns regarding privacy
  - o There is always a new frontier—issues are never settled
  - o Have to communicate so work is not reinvented
- Rhode Island
  - o The quality of weakest link prevails
- Connecticut
  - o State has unused data
- Maine
  - o It is difficult to work across agencies

**Questions**
- Are there collected data that have never been used?
  - o Some of the 801 data
  - o Process of defining requirements involved (sifting through what was gathered); some of those data may have been unnecessary
- Common identifier clarified?
  - o Somewhat random—in South Carolina for providers
  - o For providers—there is more sense in assignment
- Which was worse, working out the business process or the technical part?
  - o The business process—then technical people could start coding
- Who determines what the outputs are?
  - o South Carolina: on a case by case basis—committee
  - o Wisconsin: blend of formal reporting requirements plus individual agency needs plus ad hoc requests
- Do you use some of the SSN numbers, and then people are assigned a different level?
  - o External number that shows up is the 13-digit number that is generated
  - o Do not use SSN for clients except for determining eligibility coordinate, clear for past participation
  - o South Carolina—SSN is used to give number and it is later stripped and put in "lock box"
- Massachusetts—use SSN to link data from child care to DSS to foster care reimbursement
  - o Used to match informal providers with TANF
  - o All very secure—(e.g., on common main frame)
- Is the fact that providing SSN is voluntary a barrier?
  - o Massachusetts—yes
  - o Wisconsin—no
  - o South Carolina—can match several ways

**Mairéad Reidy**

**Child Care Research Questions Addressed at Chapin Hall Using linked individual-level State Administrative Data**

- Chapin Hall has been involved in the development of linked administrative datasets for more than 20 years.
- We typically link data across systems using probabilistic matching when data systems do not share a common identifier.
-  Key research interest at Chapin Hall is in patterns of take-up of means tested programs including food stamps, WIC, and Child Care subsidies. We are interested in the prevalence of nonparticipation among those who are eligible; and the relationship between take-up of programs and family and child outcomes.
- Our interest in the relationship between childcare subsidy take up and workforce issues/employment stability culminated a few years back in collaboration between Illinois, Maryland, and Massachusetts to examine child care subsidy use and employment duration of TANF mothers during the early years of welfare reform. This 3-State study was funded by the Child Care Bureau.
- Study follows a series of TANF entry cohorts over time. It tracks child care subsidy eligibility and take-up, and the relationship between subsidy take-up and employment duration over-time.
    - Linked UI wage reporting data is used to identify those eligible for the child care subsidy because they have found employment with wages below the child care subsidy state threshold.
    - Linked child care subsidy use data is used to distinguish between those who take-up the subsidy and those who do not.
    - Linked UI data is used to examine both groups' subsequent employment patterns.
- Results—there is a strong correlation between take up and employment in all three states.
    - Taking up the child care subsidy decreases the probability of ending employment by 43 percent in Illinois, by 31 percent in Maryland, and by 25 percent in Massachusetts.
- These results provide critical evidence that the child care subsidy system plays an important role in supporting the self-sufficiency of families.
- Much remains to be learned however. We have not yet been able to examine the relationship between subsidy tale-up among other low-income families who have no history of TANF use (i.e. the working poor).
- With funding from the Child Care Bureau in 2004, we now have a new unique collaboration with the U.S. Census Bureau to examine child care subsidy take-up patterns among all income eligible populations.
    - Participating states are Illinois, Maryland, and Texas.
    - Administrative data will be blended with individual census data
    - 2001 Census data will be used to identify who is eligible for the child care subsidy (taking account of different State eligibility requirements)
    - Blended child care subsidy data will be used to examine take up rates.

- o   Blended UI wages will be used to examine employment outcomes and continued eligibility.
        - o   We hope to examine more employment and welfare outcomes than have been analyzed to date.
        - o   This year our work is to put confidentiality agreements in place to enable sharing of administrative data with the Census Bureau.
        - o   Putting in place the agreements is vital work and it takes significant commitment.
- Matching process is complicated.
        - o   No single identifier can be relied upon with complete reliability.
        - o   An individual can have multiple IDs within one data system.
        - o   SSN can be used across systems in Maryland.
        - o   In Illinois, most reliable method is "probabilistic record matching."
                - Once a match has been determined, individual given a unique id
                - This unique ID is then used in all records
                - Probabilistic record matching has been used for 15–20 years
                - This system may soon become expensive
- Confidentially is a key issue—key strategy to keep confidentiality:
        - o   Findings are never reported at the individual level—always in aggregate
- Cooperative relationships are essential to:
        - o   Scrupulously maintain confidentiality
        - o   Share results with state agencies
        - o   Add to the policy relevance of the work through opportunity to review and comment
- Having linked data has allowed us to answer many more questions than individual data systems can answer; linked data provides powerful information.
- Linked individual-level data can be expensive to put in place initially.
- Data were not gathered typically to answer research questions, and so we are limited to the socioeconomic and demographic variables contained in these datasets.

**Questions and Comments from the Audience**

What is your web site?
   www.chapinhall.org

When do you expect to have answers?

   We do have information from our first 3-state study. The report Child Care Subsidy Use and Employment Duration of TANF mothers during the early years of Welfare reform can be found on our Web site. We have found for example that:
   - o   Roughly half of our sample of TANF entrants become eligible for the child care subsidy through employment during the study period;
   - o   Subsidy take-up does not exceed 35 percent of the eligible population in any of the three states; and

- o Take-up of the child care subsidy program is strongly correlated with employment retention.

We expect to have results from the collaboration with the Census Bureau in 2007.

We do not yet know "why they do not take subsidy up," because we cannot get that information from the administrative data. It will require a survey.

Question to Jan Staveley from the Jacob France Institute in Maryland—What would be the impact of moving subsidy from Human Services to Education?

That would have an impact on follow-up.

They would have to renegotiate with new agencies.

Relationships are extremely important.

It's a lot of work to keep demonstrating the value of data that can cut across "regime changes"