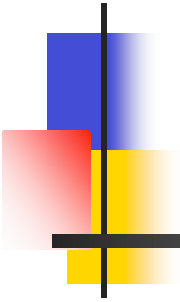


Linkage of Early Childhood Health and Education Data



Craig Mason, Ph.D.
University of Maine

June 2008

Maine Education Policy Research Institute
College of Education and Human Development
University of Maine
5766 Shibles Hall
Orono, ME 04469-5766

LINKAGE OF EARLY CHILDHOOD HEALTH AND EDUCATION DATA

Introduction

By sharing information regarding the children they are serving, health and education-based agencies could improve service delivery, as well as better plan and evaluate those services being delivered to families. Consequently, linking and sharing early childhood health and education data is increasingly seen as an important step in enhancing services to children and families and informing policy makers charged with overseeing these programs. For example, sharing newborn screening information with early intervention programs can improve the timeliness of services and reduce the number of children identified as in need of care, but lost to follow-up. Linking birth records and education records is seen as central to accurately estimating rates of developmental disabilities, as well as planning for services or identifying potential environmental risks. This is a particularly significant issue in areas such as autism, where there continues to be a debate as to whether and to what degree rates may be increasing over time. However, accomplishing this is complicated by HIPAA and FERPA, which serve to protect the privacy of health and education records, respectively. This can create a “Catch-22”, where programs with an interest in sharing data so that they can better serve the children and families in their care, are unclear as to when, how or even if they are able to do so.

Therefore, in 2007 the Maine legislature passed legislation asking the Maine Education Policy Research Institute to identify challenges to such sharing, as well as possible remedies. Specifically, Public Law 2007 chapter 450, sec B-3 states, “...the

Maine Education Policy Research Institute may conduct a review of the necessary technical and legal advances that would enable data linkage to facilitate research projects that would involve linkage of personally identifiable health and MaineCare data and personally identifiable education data in a way that would be in compliance with federal privacy law and regulation.... The report may address how to use data linkage studies to assess the effectiveness and efficiency of delivery of early intervention, education and health and human services to individuals with disabilities, including the number of case managers per child.”

For the purpose of this report, data linkage/data sharing will be distinguished based upon whether it serves one of two purposes: Clinical case management and policy/public health research. Data linkage in connection with clinical case management involves the sharing of health information for a specific child across different data systems. For example, this would refer to a specific child who has been flagged by a newborn screening program being referred to the appropriate diagnostic/early intervention services. The purpose of data linkage in such situations is to improve the timeliness of referrals and reduce loss-to-follow-up.

Data linkage for policy or public health research involves linking large-scale or population data across different systems, including both clinical cases and non-clinical children. For example, this would include linking birth certificates and education records in order to estimate state-wide rates of autism or other developmental disabilities. The purpose of data linkage in this manner is often to answer specific policy questions or inform policy makers.

Data Sharing For Clinical Case Management

Current Federal and State law allow for limited sharing of child health data with Part C or other education services for the purpose of clinical case management.

(1) Limited Referral Without Consent.

Existing State and Federal law would allow state child health programs to make limited referrals directly to Part C without requiring consent of the child's parent or guardian. In this situation, information is limited to the child and parent's names and contact information. Details regarding why the referral is being made could not be provided. Part C could report back to the referring health programs aggregate information regarding the outcome of these referrals, but information regarding individual children would not be permitted (including services received). For example, assume a newborn, James Smith, son of John and Mary Smith residing at 123 Apple Drive, Augusta, does not pass a newborn hearing screen. This type of limited referral would simply involve providing the name of the child, James Smith, the name of the parents, John and Mary Smith, and contact information, such as their address of 123 Apple Drive, Augusta, to Part C. Part C would not be notified as to why the referral is being made (i.e., \newborn hearing screening results).

Implementation of this approach would not require changes to existing law or regulations. However, memorandums of understanding (MOUs) between programs would need to be established. Information contained in the MOUs would need to address the specific limited referral data that would be provided by the screening or health

program to Part C, including the identification of individuals or entities within the program permitted to make such referrals, the mechanisms for data transfer, and acceptable security protocols.

MOUs would also need to address the nature of any aggregate data that would be reported back to the health agency, including the degree of anonymity incorporated into the aggregate data. Specifically, the degree of anonymity refers to the smallest aggregate grouping allowed in any such report. For example, if a requirement is established that the degree of anonymity be set at 15 individuals, the closest that one could get to identifying a specific person would be to know that he or she is a member of a group of at least 15 people. Continuing with this example, if 16 children (8 male, 8 female) in Penobscot County were referred to Part C in 2008 by the Newborn Hearing Screening Program, Part C would be able to later notify the Newborn Screening Program that 16 children were referred and 11 were currently enrolled in Part C services. Part C would not be able to report how many males were in services, as this would reflect only 8 children. It should be noted that establishing the actual degree of anonymity is a more complex process than is reflected in this simplistic example, and takes into consideration how various pieces of information reported separately may be combined to further reduce the actual anonymity provided in a report.

There appear to exist no technological limitations that would prevent Department of Health and Human Services (DHHS) early childhood health programs from providing limited referral information to Part C. This could be made either automatically or on-demand based on manual review of DHHS staff. As noted in the description of the MOU requirements, specific details regarding the data transfer protocol and security

requirements would need to be agreed upon by all parties. In addition, if parties agree that aggregated information is to be provided back to the health programs by Part C, staff in those programs will need to develop the appropriate queries and reports. However, this can be accomplished as a separate or later step in this process.

Summary. There currently exists no fundamental barrier to prevent DHHS early childhood health programs from providing limited referral information to Part C. There exist no technological limitations that would prevent electronic referrals, either automatically generated or based on staff review.

(2) Detailed Referrals With Consent.

With parental consent, state early childhood health programs could make limited detailed referrals directly to Part C. Details regarding the nature of the referral being made could be provided, as well as other screening and health information covered by the consent. If the original consent allowed for Part C to report information back to the health agency, sharing of such information would be permitted, assuming that such consent was not withdrawn by subsequent permissions with either of these entities. Regardless, aggregate information regarding the outcome of referrals would continue to be permitted, as described in the previous section.

Implementation of this approach would also not require changes to existing law or regulations. However, MOUs would need to address the specific referral and clinical data that would be provided to Part C, as well as the nature of any individualized or aggregate data that would be reported back to the referral source, as described previously.

Technologically, detailed referral based on parental consent would require additional changes beyond those noted under “*Limited Referral Without Consent*”. This would include systems for collecting and electronically storing individual-level consent information within the health programs. Furthermore, data transfer protocols would now have to address both limited and detailed information, based upon whether consent was received or denied, and if consent was received, the nature and extent of such consent. Similar issues would need to be addressed by Part C programs, but again, this can be accomplished in a separate or later step.

Summary. There currently exists no inherent barrier to prevent DHHS early childhood health programs from providing detailed referral information to Part C, based upon parental consent. However, this would require the development of an appropriate consent form, and its adoption across all screening programs. Mechanisms for collecting and storing this information would also need to be developed and tested. Similarly, while no inherent technological limitations exist, this would involve minor-to-modest changes to existing data systems.

(3) Screening as Part C Point of Entry.

An additional option is possible when specifically focusing on the sharing of data between state newborn screening programs (hearing and metabolic) and Part C. This would be to define screening as a point of entry into Part C. Arguably, this would be philosophically consistent with the mission of newborn screening programs in general, in that their underlying purpose is to identify children at risk for health or developmental challenges so that they can obtain appropriate diagnostic and early intervention services

in an efficient and timely manner. By defining screening as a point of entry into Part C, this would essentially recognize and acknowledge the ultimate goal of screening as a mechanism for entry into this system.

As an entry point for Part C, details regarding screening results could be provided to Part C for those infants flagged by atypical screens. This would not require formal MOUs between screening programs and Part C; however, this may require changes in state law and/or regulation.

Summary. While this may provide the most direct route to timely and complete referrals from newborn screening programs to Part C, this would require a further detailed review of state law and regulation in order to determine what, if any, changes may be required to adopt this approach.

Data Linkage for Policy or Public Health Research

In addition to case management benefits, many public policy or public health questions relevant to child health and education are best answered through broadly-based, population-level data linkage. Whereas the previous case management examples focused on sharing information regarding a small subset of specific individuals based on flagged health data (e.g., newborn screening results), data linkage for policy or public health research involves linking entire population-based data systems, including both clinical cases and non-clinical (i.e., general population) children. For example, linkage of newborn hearing screening data with future education data may help policy makers evaluate the degree to which screening and early identification leads to lower rates of

special education placement or academic difficulties for children. Similarly, linkage of birth certificates and special education data is needed to accurately estimate rates of autism, as well understand regional differences or patterns over time.

However, given privacy concerns, regulations, and state and national law, the challenge is how to enable such linkage without one party revealing sensitive health or education information to the other. Therefore, the following material summarizes a protocol through which childhood health data and education data could be electronically linked for the purpose of informing policy, without either party (health or education) releasing identifiable sensitive health or education data to the other.

1. Identification of a trusted third-party. The first step would be to identify a trusted third-party, ideally an entity that has existing independent relationships with both the Department of Education (DOE) and DHHS. The purpose of this trusted third party is to serve as an independent nexus for assigning and managing state-wide, cross-agency unique ID numbers for children in Maine.

2. Establishment of Directory-Type Information. DOE and DHHS would then provide this trusted third-party with directory-type information, such as a child's name, date of birth, location of birth, mother's name and date of birth. No clinical or other sensitive information would be included. In essence, DOE would be providing the trusted third-party simply with a list of children attending public school in Maine, but would not be providing any information regarding a child's academic status, whether a child is in special education, in remedial

programming, receiving free or reduced lunch, etc. Similarly, DHHS would be providing the third-party solely with a list of children born in Maine, but would not be indicating any medical history, screening or diagnostic findings, etc, for any of these children.

3. Creation of MaineID. The third-party then links these two very limited data sets and assign a new unique ID number (MaineID) to each child. The third-party then provides DOE with a copy of the data they submitted, with the MaineID added. DHHS is also provided a copy of the data they submitted, with the MaineID added.

Once a common ID exists across systems, it becomes possible to create linked data sets that are nevertheless de-identified. This will allow state officials and public health/policy researchers to examine important policy issues involving both health and education data, without revealing identifiable sensitive information. One possible such strategy is illustrated below.

Example of Encrypted (Blind) Linkage

For this example, assume DHHS and DOE agree to conduct a population-based study on the impact of newborn hearing screening on reducing future special education placements. For health officials, this information would be valuable documentation of the positive impact of newborn hearing screening efforts. For education officials, this

would provide valuable information as to methods for reducing costs—personal and social—associated with special education services.

Both agencies first agree on the specific fields each will provide for the linked dataset. For this example, assume DHHS provides a field indicating the results of a child's newborn hearing screen (pass/fail/no screen), while DOE provides a field indicating whether or not the child is receiving special education services (receiving services/not receiving services). They then use the same encryption algorithm and the same predetermined, user-defined encryption key to encrypt MaineID values for these children (note: the encryption algorithm would require certain properties, such as commutativity, that are not discussed here). Because the same algorithm and key were used to encrypt MaineIDs, the encrypted MaineID will be the identical for the same person in both the DOE and DHHS datasets.

Each agency then creates a new data file containing the encrypted MaineID and the additional information they are sharing. For example, DOE creates a file containing two fields: the encrypted MaineID and special education status; DHHS provides a file containing two fields: the encrypted MaineID and screening status. Agencies should also independently evaluate the degree of anonymity in their limited data sets. They provide copies of these files to the trusted third-party for linkage.

The trusted third-party matches these records based on the encrypted MaineIDs, but because the unencrypted MaineID is not included and the third-party does not know the value of the encryption key, the third-party can not identify any individuals in either of the original datasets, nor in the final linked dataset. Once the records in the two data files are matched, the third-party creates a new data file containing the matched record

information, but now deletes the encrypted MaineIDs. In other words, the file indicates each child's screening result and special education status, but does not indicate the child's MaineID, which they did not have to begin with, nor the encrypted MaineID.

Copies of this linked data file are then provided to DOE and DHHS. Because the encrypted MaineID has been removed from the linked file, neither DOE nor DHHS is able to identify any children in the linked dataset, or even determine which of their own children ultimately appear in the linked dataset. As a result, while both agencies are able to obtain a data file containing linked health-education data, the following characteristics hold true...

- DOE does not know the newborn hearing screening results for any child in the education dataset
- DHHS does not know the special education status for any child in the health dataset
- The trusted third-party does not know either the screening result or the special education status of any child in either dataset
- No party knows which children in either the education or health dataset ultimately appear in the linked dataset

The various steps in this procedure are outlined in Table 1 which follows the Conclusions below.

Conclusions

Data sharing between health and education programs can be an effective tool for improving the timeliness of service delivery, as well as providing valuable information to policy makers for evaluating program effectiveness and planning and coordinating care. In some areas, such as making limited referrals from newborn screening programs into Part C or early intervention, doing so may require no or minimal legislative action, and instead is dependent upon the relevant agencies and programs developing acceptable memorandums of understanding. Other areas, such as large-scale data linkage, may require additional collaborations or commitments, but have the potential to provide valuable information to policy makers regarding program effectiveness, as well as coordination and planning for future service needs and patterns of services. A first step would be to conduct a pilot study on the feasibility of performing the type of linkage described in the section titled “*Data Linkage for Policy or Public Health Research*”. This would allow for the evaluation of the degree to which such linkage is possible, as well as identify any technological or data issues that would need to be addressed. In this regard, it is worth noting that MEPRI is well-positioned to help in this type of work, in that it includes researchers experienced with both the health and education systems within Maine, as well as experts on informatics and data linkage.

Table 1. *Illustrative Example of Process for a Population-Based Study on the Impact of Newborn Hearing Screening on Reducing Future Special Education Placements.*

- Both agencies agree on the specific fields each will provide for the linked dataset.
 - DHHS: newborn hearing screening result
 - DOE: special education status
- Both agencies encrypt MaineID values in their dataset using the same encryption algorithm and encryption key
 - Note: The encryption algorithm would require certain properties, such as commutativity, that are not discussed here
- Each agency then creates a new data file containing the encrypted MaineID and the additional information they are sharing
 - DHHS file contains: the encrypted MaineID and screening result
 - DOE file contains: the encrypted MaineID and special education status
 - Note: Both agencies should also evaluate the degree of anonymity in their limited data sets.
- DOE and DHHS provide copies of these files to the third-party for linkage.
 - Note: Because the unencrypted MaineID is not included, the third-party can match records in the two data files, but can not identify any individuals in either.
- The third party creates a new file by linking the two files
 - Note: The trusted third party should also evaluate the degree of anonymity in the linked data set

- The third party deletes the encrypted MaineID from the new file
 - Linked file contains: hearing screening result and special education status.
- Copies of this linked data file are provided to DOE and DHHS
 - Because the encrypted MaineID has been removed from the linked file, neither DOE nor DHHS is able to identify any individuals in the linked data file.