

November 2006 CMS MIG Conference (Chicago, IL)
Workgroup #6: Working with Administrative Data

Facilitator: Su Liu (MPR)

Session Objectives

- Understand the transition from designing research question to implementing analysis plan
 - Learn what data are available in the integrated database MPR has developed, including SSA and Medicare data
 - Identify key issues to consider when working with administrative data
 - Develop table shells and analytic specifications for aggregated data analysis
-

Discussion Notes

- ❖ Even though topic of the discussion is “working with administrative data”, workshop participants agreed that many issues we discussed can be applied to data analysis in general, including survey data.
- ❖ States vary in terms of research resources and approaches: some hire own researchers within the project, some contract with universities, others rely on inter-department cooperation.
- ❖ Participants reviewed the data elements currently available in the federal integrated database MPR has developed, and discussed:
 - Certain data elements may have unstable quality and should be used with caution;
 - Because of the nature of administrative data, federal agencies may periodically update or change the structure of their data files to fit the programs’ needs;
 - MPR is in the process of merging to an updated version of the Ticket Research File (TRF.05), with some new data elements added, for example, trial work period data;
 - The possibility of using newly available Medicare Part D data CMS is developing;
 - Each of the data source included in the integrated database serves different populations; one should not expect 100% overlapping between two data sources, which may have further implication on how the results can be interpreted.
- ❖ Participants acknowledged that the turn-around time for the MIG-RATS and MPR to respond to states’ aggregated data requests may vary depending on the nature and magnitude of the request itself; and suggested a clearer guidance can be offered ahead of time on:
 - Types of requests that would take short vs. long time to respond;
 - Alternative data sources and trade-offs to answer the same question;
 - Interpretation of the results after the analysis.
- ❖ Several participants pointed out the rich data resources available at the state level: although it may take huge efforts to break the silos across agencies, the end results can

be great. MIG-RATS can help share individual states' experience among others via the website, so that states can learn from each other in achieving the goal of maximizing utility of state-level data that are housed in different agencies.

- ❖ Participants have high expectation of the MIG-RATS website, and suggested various information that can be shared there to prevent states from reinventing wheels, for example, a subject-oriented research corner on topics such as disability grouper.
- ❖ The aggregated data analysis requests do not need to be only on Medicaid Buy-In participants. Other subgroups of interest can be analyzed as well, as long as they are captured in the integrated data and personal identifiers are submitted in the finder file from states.
- ❖ The aggregated data analysis requests do not need to be at a single state level. In fact, a project with multiple states involved can often produce richer findings, and offer cross-state comparison on the same subject.
- ❖ Some participants think adding other useful variables in states' finder files may not be a bad idea, for example, names (to allow better matching rates) and disabling conditions as recorded by states during Buy-In determination.
- ❖ The series of products from MPR, including issue briefs and other reports, contain a great deal of information that can be useful to states. All are available on MPR's website at www.mathematica-mpr.com (search Medicaid Buy-In).
- ❖ Participants discussed a number of key issues researchers should pay attention to when using administrative data to conduct aggregated analysis; these issues should be carefully considered in submitting an aggregated data analysis request:
 - Maintain a sense of the "big picture";
 - Preliminary exploratory analyses help to understand the limitations of the data and unit of analysis (e.g. means, frequency distributions, etc);
 - Build in quality checkpoints in the specifications if possible; and upon getting the output, look out for outliers and check for data quality;
 - Break the analytic specifications into sequential programming steps;
 - Clearly define the definition of variables (original, intermediate, and final) and format for output data (e.g. table shells, chart configuration, etc); particularly, be specific on:
 - Data sources
 - Time periods (trends by year or all years combined)
 - Sample population (subgroups or whole population)
 - At which time point should the time-variant variables be defined
 - How many categories should be included in categorical variables
 - How to treat missing data
 - Cross check between the federal-level data and state-level data, if possible.