1. Can you provide some approximate size information for each of the databases being constructed - # of records, # variables/bytes per record.

The 1997 Nationwide Inpatient Sample, Release 6, contains over 7 million observations from over 1,000 hospitals with approximately 123 variables. The NIS data set is distributed in a set of six CD-ROMs. Five of the CDs contains data, the sixth consists of documentation. More information may be found at the HCUP web site, http://www.ahrq.gov/data/hcup/nisintro.htm

Information for the1997 State Inpatient Databases is available in the SID Application Kit found in the HCUP RFP Reference Materials or available for downloading at http://www.ahrq.gov/data/hcup/hcupsid.htm#head4 . Specifically, page 6 in the Application Kit lists number of discharges and hospitals per state, per year in the SID for the 13 state partners participating in the HCUP Central Distributor.

For 1997 data, Ambulatory Surgery (AS) databases were constructed from nine states. Sizes of these databases vary. For example, the Colorado and Connecticut AS databases both contained over 200,000 observations with approximately 90 variables. Both CO and CT had approximately 400,000 inpatient discharges in their respective SID that year.

2. In the process of coding and editing the data from each state, it is certain that edit failures will occur and data inconsistencies will be found. Examples:

ICD-9-CM code on the state record is not found on the valid list of ICD-9-CM codes.

ICD-9-CM code on the state record is not consistent with gender of discharge compared to valid gender-specific ICD-9-CM codes.

In application of the CCS, diagnosis data from the state record does not map into a CCS category.

Out-of-range values, invalid codes, or coding inconsistent with procedures are found in the process of data editing.

Based on past experience, can you quantify the edit failure rate that should be expected with this data (e.g. edit failures / 1000 records). Also, is there an expectation on how the edit failures will be remediated? (Possible solutions could range from human intervention and follow-up with the hospital in question, to some sort of automated replacement of the item in question that would ensure consistency).

Edits and failures occur relatively infrequently and vary across data elements with virtually no failures for many edit checks and 2-3 % failures for some edit checks. We can provide the following guidance. A basic verification of the entire data file is conducted after the data are received from the state data source to determine if the data are complete and to verify that the data are readable and match those requested. If any discrepancies are found at that

time, HCUP contractor staff contact the data organization that provided the data to resolve any problems. Offerors are referred to the RFP reference document, "Final Report on Development and Analysis of HCUP - Summary Version", pages 11-12 for further detail on the current procedures.

At the beginning of data processing, edit checks are performed on many data elements. As described, past experience indicates that there are relatively few errors found in the individual data records. In the majority of cases, records are only "flagged" to indicate discrepancies. In rare circumstances, such as if a high rate of missing or anomalous values is observed, the contractor might also re-contact the state data organization to resolve the discrepancy during data processing or to obtain replacement data. The hospital or outpatient facility is never contacted for follow-up.

In response to this technical question, we have added the following document to the HCUP Reference Materials site at http://198.179.0.100:8040/hcupref , "NIS Technical Supplement 2, Quality Control in HCUP Data Processing." This supplement describes the quality review guidelines and edit checks performed on HCUP data.

3. In section C.3.1, a requirement for transferring HCUP data from the previous contractor is described. Can you provide greater detail about the types and magnitude of this data, and what the new contractor is required to do with it (make data available on-line, archive the data, etc.) It is clear that the 1998 and 1999 inpatient data are to be processed as rapidly as possible, but it is not clear what is to be done with the inpatient data prior to 1998 (data used to create the SID and the NIS).

The current contractor maintains a variety of project files. These include, but are not limited to, the SID, NIS, Ambulatory Surgery and Emergency Department data files for HCUP partner states. For example, for the SID there are files for eight states for data year 1988, 11 states for 1989 to 1992, 17 states for 1993 to 1994, 19 states for 1995 to 1996, and 22 states for 1997. There are 6 releases of the NIS, each year of data is held on approximately 4-6 CDs. We have Ambulatory Surgery data for 3 states from 1988-1989, 4 states for 1990-1992, 5 states for 1993-1995, and 9 states for 1996 - 1997. They also have the master CDs for the Central Distributor files for 13 partner states for 1995 through 1997. Other files consist of documentation/codebooks, SAS programming files, American Hospital Association files, and project reports for each database and for each year. For some earlier years of data, the files are maintained on tape at the NIH computer center where they will remain. AHRQ maintains its own copies of all deliverable files.

The new contractor is required to maintain a library of these data files to 1) serve as a historical archive of project products; 2) serve as reference material in developing new years of HCUP data; 3) prepare special analyses upon request of the project officer; 4) support technical assistance; and 5) with respect to the Central Distributor, to create copies of the SID public release files for purchase by researchers. The new contractor will determine the best approach for how to maintain these files.

4. For the crosswalk files described on page 31 - can you provide any information as to how this was created in the past? Are there common variables in each source data set

that can be linked one to one? Or is this done by some pattern matching on character data (e.g. hospital name) combined with other fields (e.g. zip) that indicate a "match".

The AHA crosswalk files are created and documented for each HCUP data year. The reference documents, "Guide to the HCUP 3 Database" and "Final Report on Development and Analysis of HCUP - Summary Version" each provide basic descriptions of the process. In response to this technical question, we have added the following document to the HCUP Reference Materials site at http://198.179.0.100:8040/hcupref , "NIS Technical Supplement 3, Mapping Source-Specific Hospital Identifiers to AHA Hospital Identifiers." This supplement describes the linking of hospital identifiers on HCUP data.

Generally, however the crosswalk was created using the name and location of the hospital in the AHA file and matching it to the name and location of the hospital in the state data. This has resulted in a historical file that is used from year to year to create the most current crosswalk and which provides the identification for the majority of hospitals in our database. Typically, in any given year, approximately 5% of the hospitals change due to mergers, de-mergers, closings or openings. Those 5% of hospitals are usually hand-matched using a variety of available data.

5. Presumably, the data from each state are collected by the state from all or most of the hospitals in the state. Can it be assumed that all of the records contained on the state file have been formatted into a common format, even if different formats are used by each hospital?

Yes, each state delivers its data in a uniform format. However, formats vary across the states.

6. On what type of media is the data available from the states? (CD's, 9 track tapes, etc.)

Data from the states comes in all forms of media. The most typical are CD's and 9 track tapes.

7. On what type media will the data be received from the prior contractor? (CD's, 9 track tapes, etc.)

Data will primarily come on CD's.

8. On what type of media does the data need to be returned to the states and to AHRQ? (CD's, 9 track tapes, etc)

RFP Sections C.4.2.1.5 and C.5.3 and the delivery schedule indicate that data need to be returned on CDs. However, as noted in Section C.3, "It is possible that future years of this contract will require other formats and/or media based on emerging or existing technologies, for example, data might be received through Internet transfer to a central AHRQ data warehouse."

9. In the last 2 - 3 years of the contract period, there are several mentions of ideas that AHRQ would like to try. There are also plans to create more databases on outpatient

and freestanding care. How should we budget these still undefined ideas in the out years?

As stated in the Delivery Schedule (Section F.3), offerors should propose the timing and technical approach to the tasks that are not pre-determined. Any assumptions should be documented. Information is usually provided that indicates the likely start date. For example, the RFP states in Section 8, Process and Develop Ambulatory Care / Office / Clinic Visit databases of the SOD, that the task will begin in the second year of the project. The budget should correspond to the proposed approach and timing.

10. Recruitment and retention: The RFA suggests that the recruitment and retention process is a minimal part of the overall project, when in fact, it may be a substantial part of the project. Marketing HCUP and demonstrating HCUP's value to state health data agencies should be an annual concern and in fact, may be increasingly more important than the technical aspects, which we would expect to become routine.

Should bidders be advised that recruitment and retention may be a substantial part of the project and that these "non-technical" components of the project may be as challenging as the data processing activities?

In what way will the grant award process take the import of the recruitment and retention activities into account?

The critical importance of recruiting and retaining data partners is highlighted in the second paragraph of the first page of the statement of work,

> "The participation of state partners is essential for success of the HCUP project and is based on cooperative, detailed agreements made between AHRQ and each state Partner."

In addition, establishing the value of HCUP partnership requires a non-routine, individualized approach with each existing and potential partner. The statement of work describes a variety of tasks that must be accomplished in order to support the data partnership. It is incumbent upon offerors to demonstrate their understanding of the requirements of this task by describing their approach and the level of effort required to achieve these goals.

The contract award process requires that the external Technical Review Panel evaluate proposals based on the technical evaluation criteria in the RFP. Specifically, qualifications that relate to the recruitment and retention task are detailed in Section L.9 under A. Proposed Technical Approach, B. Organizational/Corporate Experience, and C. Qualifications of Proposed Staff.

11. Data submittal: Timeliness is an important factor in compiling nationwide data. With HIPAA Administrative Simplification provisions, states and providers will be converting their submission specifications to enable electronic submission of standardized data. The RFP does not address alternatives for state submission of data files. If a bidder proposes an option for electronic submission of the state data or a pilot trial for select

states, would this be considered favorably by AHRQ, provided the proper security provisions were in place?

As stated in RFP section L.9 Technical Proposal Instructions,

> "Methodologies followed by the previous contractor are provided in this RFP with varying levels of detail to provide insight into the complexity of the HCUP project. Because HCUP is a continuing project in which future activities will need to be compatible with previous approaches, a fair amount of detail is provided to illustrate the approaches and steps undertaken. However, unless specifically stated, the methodologies described are primarily provided as examples. **Offerors are strongly encouraged to propose alternative, technologically and cost-efficient methods for achieving HCUP goals when appropriate.** Ideally, proposed approaches will maintain compatibility with previous HCUP data years while identifying methodologies that take maximum advantage of state-of-the-art data processing techniques that will carry the project forward over the next five years. "

Additionally, RFP Section 15.4.4.3 Other technological innovations states,

> "In consultation with the Project Officer, the contractor will explore and implement other technological innovations as appropriate to improve the efficiency and effectiveness of HCUP processes. For example, should it become feasible to obtain data from contributing Partners via the Internet, such a process will be developed upon the direction of the HCUP Project Officer."

12. SID Central Distributorship activities: When a vendor who purchases and sells state data commercially also serves as a distributor of the SID for the HCUP project, how does AHRQ determine that there is no conflict of interest between these functions? Should the bidder include in the RFP response a proposed mechanism to identify potential conflicts of interest and define specific measures to control for these conflicts? If a bidder does include statements of potential conflict of interest and related policies to account for this concern, how are such disclosures and policies verified and by whom?

The Federal Government defines conflict of interest in the Federal Acquisition Regulations (FAR) Part 9.5. Based on that definition and the question as posed, the situation described above does not appear to constitute an organizational conflict of interest. Bidders need not address conflicts of interest in their proposal.

AHRQ contracts are closely monitored by Program, Contract and Legal Staff to ensure compliance with all Federal rules and regulations including such issues as any potential conflict of interest.

13. Please clarify the recommended small business goals of the RFP. Currently, the RFP indicates 23% in Section A - Solicitation Form, paragraph 3, and 30% in Section L.12.B.(g).

Section L.12.B(g) is in error. The recommended small business goals of this RFP is 23%.

14. Please provide the name of the incumbent contractor and the total contract value of the previous award.

The incumbent contractor is SysteMetrics, Inc., A Medstat Division. The total contract value including all modifications is $15,359,241.

15. Please clarify whether or not the incumbent contractor was provided government-furnished property, facilities, equipment, and/or supplies under the previous award. If so, please confirm that offerors must propose these items in their proposals or identify which items will be provided by the government.

The incumbent contractor was **not** provided government-furnished property, facilities, equipment, and/or supplies under the previous award.