Frequently Asked Questions
Data Set II: Hospital

General Description

1. **How does this data differ from the information about hospitals contained in Data Set I?** In Data Set I, we included hospitalizations summarized across the experience period. This provided an easy method of analysis but eliminated many details about the length of stay, admission dates and discharge dates. Data Set IIB: Hospital offers claims-level data on hospitalization incidents, allowing for insight into these details. Additionally, this data set allows us to include more substantial information on DRGs and procedures in hospital.

2. **How will technical specifics of the data sets differ from those of Data Set I?** Like Data Set I, this data will be delivered via download from our FTP site, in the form of a text file delimited by commas. Compared to Data Set I, this file will have a relatively small number of observations of a relatively small number of attributes.

3. **How is this table intended to be used?** This detailed table is most meaningful when used in conjunction with either the Recipient or Provider table. This will require the user to have the skills and data environment to summarize the data and join the results to Recipient and/or Provider tables in Data Set I.

4. **How can my organization analyze the hospital data of our target population only?** We will deliver data based on the criteria used to generate your first data pull(s) (unless you request based on different criteria). The data release is oriented such that you can join this data to the Data Set I population you already have. This allows you to complete the same sorting steps you did previously to view the target population of interest to you in Data Set I.

5. **Does the data set have any relationship with data sets available from the Illinois Department of Public Health (IDPH)?** In creating this data set, we have noted the existence of discharge data from IDPH. We have loosely based our work on the discharge data set. Wherever possible, we have matched the data elements previously available. Exceptions to this include the fields we do not have. (Please see FAQ items below on this point.)

6. **What IDPH fields are excluded from the data set?** IDPH fields not included in our data are missing because they are either unavailable in our data warehouse, or irrelevant because of the nature of HFS data. Data collected by IDPH that are not collected by HFS include information on Do Not Resuscitate (DNR) orders, employment-related health conditions and accidents, crime victim status, and the birth weights of newborns.

   We provide information specific to a given hospital claim, including specifications of payer and insurance groups. These fields do not match IDPH data closely because our data sets use Medicaid recipient status as an inclusion criterion, prompting formatting changes with respect to data presentation. We nonetheless include a variety of information on insurance payers.

7. **Does the hospital number included in the data set match the Illinois Department of Public Health's hospital numbers?** The hospital numbers in our data set are assigned irrespective of other organizations' numbering systems. While there is a significant degree of matching between the IDPH data and ours, correlation is imperfect. We advise data users to avoid doing inner joins of HFS and IDPH data sets by hospital number, as this would result in many dropped records.

8. **Are individual providers included in the data set under their own provider ID number?** No. The data set includes the National Provider IDs (NPIs) associated with a facility, not an individual provider. These are included as hospital numbers.

Recipient and Providers

9. **Does this data set differentiate individual lives?** Discharge data often has the limitation of representing individual discharges rather than individual lives. This can result in incorrect estimates of the prevalence of various procedures and diagnoses per capita. Our hospital claims data includes many of the advantages of discharge data, including specifics of individual hospitalizations. But because it can be joined to Data Set I, this data can be assessed with respect to specific lives rather than with respect to specific discharges.

10. **What recipients are included in the data sets?** The recipient population has the same inclusion criteria as Data Set I: all recipients of full benefits during the calendar year 2010, irrespective of age and geographic location. Only those recipients with at least one hospital admission during the year have additional data available. (Exclusions are likewise the same as those in Data Set I, including the exclusion of Integrated Care Program recipients. Please see our Data Set I FAQ for further information on these exclusions.)

11. **How are Medicare recipients included in the data sets?** Medicaid recipients who also receive Medicare are included in all our data sets. (Joining Data Set IIB to Data Set I allows for selecting only those recipients who are dually eligible, using Data Set I's Dual Eligible Indicator field.) Medicare Paid Amount is a data field in Data Set IIB: Hospital. Payments are largely from Medicare Part A (hospital insurance) and Medicare Advantage (which is also known as Medicare Managed Care, and which supersedes Parts A and B combined). Medicare Part B (clinical care insurance) and Part D (pharmacy insurance) are not relevant to this data set.

12. **What provider types are included?** The provider types included in the data set are general hospitals (provider type 030), rehabilitation hospitals (provider type 032) private psychiatric hospitals (provider type 031), and state-operated psychiatric hospitals (provider type 034).

    A wide variety of providers operate in hospital settings. These providers bill services via a system separate from the hospitalization claims, and are therefore not included in this data set.

Types of Service

13. **What Types of Service are included in this data set?** In this data set, we have captured all the Types of Service that are delivered as inpatient care. To increase the specificity of data analyses in Data Set I, where the same hospital data was presented in aggregate, we divided Inpatient Services into Substance Abuse, Psychiatry, Maternity and Delivery, Newborns, and "All Other." We maintain those categories in this data set in the field labeled "Type of Service Description" in this data set. Additionally, we capture "Emergency Services for Undocumented Aliens" as a type of inpatient service, as this category uniquely captures maternity and delivery.

14. **Are encounter stays recorded in this data set?** Encounter stays, a type of inpatient care for patients enrolled in managed care, are not included in this data set. (They were included in Data Set I only with respect to Add-on Payment costs.)

Diagnoses and Procedures

15. **What does the field "DRG 1995" mean? Are up-to-date Diagnostic Related Groups (DRGs) included in the data set?** *The DRGs included in this data set are not the most current available DRGs.* This data set includes Version 12 DRGs, which were originally created in 1995. These codes are obsolete and are not used in most hospitals currently. HFS persists in using these DRGs on its hospital claims because these DRGs are defined by Illinois law as the basis of our payment system. To emphasize that the codes are outdated, we have included the year they originated, 1995. Please note that more up-to-date versions, such as MS-DRGs and APR-DRGs, are available. Mapping these DRGs onto the ones included in our data is the responsibility of the end user.

16. **How many diagnostic codes are included per claim?** We include up to 25 diagnostic codes for a hospital claim. Additional diagnostic codes, where present, are simply indicated (by a 1/0 indicator) in an accompanying column. For series bills (which are submitted over time for a single hospitalization that lasts longer than thirty days), a primary diagnosis code is carried across all the claims for that hospitalization, and all additional diagnoses are dropped after the first claim.

17. **What does "Present on Admission" (POA) mean?** Present on Admission is an additional data field included for each diagnostic code. Present on Admission, as defined by Section 5001(c) of the Deficit Reduction Act (DRA) of 2005, is an indicator applied to each diagnostic code as part of a reimbursement determination. The intention of POA indicators is to encourage the prevention of conditions acquired in hospital (nosocomial infections and preventable injuries) by reducing or eliminating payment of the conditions that are not present on admission. POA codes are folded into diagnostic groupers, and payments are made on the basis of this grouper.

18. **Why are some Diagnosis fields and "Present on Admission" fields blank?** We record up to 25 diagnoses per claim. For each of these, a POA code should be present. Many recipients will not have this many diagnoses, however, and all unused diagnostic and POA fields will be blank.

19. **What are the codes for "Present on Admission" fields 2 through 25?** They are the same as the codes for the first Present on Admission item (DiagPOAClaimCd1). The data sources are also the same.

20. **How many procedure codes are included per claim?** We include up to 25 ICD-9 procedure codes per claim. Additional procedures (>25) are indicated by a separate indicator. If the recipient had less than 25 procedures during a hospitalization, the excess fields will be blank.

21. **Why do so many procedure date fields show the date January 1, 1901?** It is possible to have fewer than 25 procedure codes, and therefore fewer than 25 corresponding procedure dates. The date 01/01/1901 is the default value for those procedure date fields that have no corresponding procedure code. We include this value in order to assist the users of software programs that do not easily accommodate "null" values in date fields.

Billing, Costs, and Payments

22. **What bills are included in the data set?** HFS recognizes a variety of different bills: those created at the beginning of a hospitalization; interim bills representing each 30-day period a recipient remains in the hospital (also known by the synonyms "series bills" or "interim bills"); discharge bills, with the specialized charges associated with the ending of a hospitalization; and bills for additional charges, which typically reflect adjustments to previous bills. This information is contained in the field Bill Type Code.

23. **If a claim includes series bills (covering a continuous hospitalization of more than 30 days), which bill is used to gather data about this hospitalization?** Most of the claims that involve series bills conclude with a discharge bill (BillTypeCd = 4). Where possible, this bill is used as the source for the patient status code, DRG1995 codes, admission diagnosis, admission source code, admission type code, pricing code, outlier day cost indicator, and Type of Service description. Additionally, diagnostic codes, diagnostic POA claim codes, procedure codes, and procedure dates (up to 25 total in each case) are pulled from this bill.

    Some hospitalizations extend past the final day of the experience period, and for this reason do not include a discharge bill in the data set. (These claims can include a discharge date after the end of the experience period, but will not include further information originating after December 31, 2010.) For these claims, we use the final bill submitted (BillTypeCd = 3) to pull the patient status code, DRG1995 codes, admission diagnosis, admission source code, admission type code, pricing code, outlier day cost indicator, and Type of Service description. We pull up to 25 diagnostic codes, diagnostic POA codes, procedure codes, and procedure dates from the first interim bill submitted (BillTypeCd = 2).

24. **What billing form is used for the claims included in the data set?** All bills included in our data are based on the UB92 billing form. No other form is included.

25. **What is included in a cost?** The cost fields in the data set are like the ones present in the data sets previously released: they reflect Medicaid's net liability for a given claim. The data set separately describes the charge the provider originally sought; copayments; sum Third Payer Liability (TPL); Medicaid charges; Medicare charges; Disproportionate Share Hospital payments, Medicaid Percentage Adjustment and High Volume Adjustments; Add On, Capital and Covered Charge payments; and a variety of other payments. It does not include static payments to hospitals that are not related to a specific hospital stay.

26. **How are hospital payments determined?** Hospital payments are determined via a calculation that incorporates multiple facts about a particular inpatient visit. Please see the downloadable DRG calculator spreadsheet on our website for more information. The payments generated by this worksheet are Medicaid's net liability including any costs that may be covered by Medicare or Third Payer liability for recipients who have such coverage.

27. **What is the DRG calculation worksheet?** The DRG calculation worksheet allows a data user to figure out the total cost associated with a particular admission. You can find this worksheet as a download on the HFS Data Releases webpage. The download includes instructions.

28. **How are Disproportionate Share Hospital payments included in the data set?** Disproportionate Share Hospital (DSH) payments are intended to support hospitals that serve a large percentage of low-income/indigent patients. These hospitals are defined by Section 148.120 of the 89 Illinois Administrative Code as serving at least one-half a standard deviation above the mean Medicaid utilization rate (MIUR), or serving low-income patients at a rate of 25% per annum. A complete listing of criteria and eligible facilities is available on the HFS website. DSH payments are calculated for each claim, and can therefore be included as a line item towards a total payment.

29. **What is a Medicaid Percentage Adjustment Add-On Payment? How does it differ from DSH payments and MHVA payments?** Medicaid Percentage Adjustment (MPA) Add-On Payments are intended to supplement income to hospitals that provide service to a relatively high rate of Medicaid patients (defined by Section 148.120 of the 89 Illinois Administrative Code as serving at least one-half a standard deviation above the mean Medicaid utilization rate, or serving low-income patients at a rate of 25% per annum, plus meeting a sufficient number of six additional criteria identified in Section 148.122 of the Code). MPA payments are paid in increments per claim, and can therefore be included as a line item towards a total payment. (A complete listing of criteria and eligible facilities is available on the HFS website.) These payments are provided incrementally, per claim. This field is recorded separately to the net liability and other payment fields.

30. **What is a Medicaid High Volume Adjustment Add-On Payment? How does it differ from DSH payments and MPA payments?** Medicaid High Volume Adjustment (MHVA) Add-On Payments are intended to help support hospitals who serve a high volume of Medicaid patients. Eligibility for these payments is determined on a hospital-by-hospital basis and is based on the hospitals' matching the criteria set forth in Section 148.120, 148.122 and 148.290(d) of the 89 Illinois Administrative Code. The last of these specifically states that MHVA-eligible must "not be a county-owned hospital…. or a hospital organized under the University of Illinois Hospital Act… in the MHVA rate period." This is the major difference between MPA and MHVA payments. MHVA payments, like MPA payments, are defined by criteria additional to those that are applied to DSH payments. A complete listing of criteria and eligible facilities is available on the HFS website.

31. **How are Per Diem payments represented in the data set?** Per Diem payments are designed to reimburse hospitals based on the total number of days of a given admission, in addition to payments derived from the specific DRG. Only University of Illinois at Chicago hospital, Cook County hospital, rehabilitation hospitals, psychiatric hospitals, children's hospitals, long-term stay hospitals and certain rural hospitals are eligible for Per Diem payments. (Out-of-state non-cost-reporting hospitals are also paid via this method, but are not included in this or any other data set.)

32. **What is the Medicaid co-pay? How is this dollar amount determined?** Under Title XIX, Medicaid recipients are assessed a co-payment for each day they are inpatients. This figure is $3.00 per day at hospitals billed under DRG methodology; $3.00 per day at hospitals with a Per Diem reimbursement rate of $325 or more; $2.00 at hospitals with a Per Diem reimbursement rate of $275-325; and $0.00 at hospitals that have a Per Diem reimbursement rate of less than $275.

33. **What is a non-covered charge?** This is the amount of a bill that is not covered by any of the benefits program available to the person receiving the healthcare service. This amount is not the

responsibility of the patient (although the patient can be liable for cost-sharing amounts, which are in separate fields).

Data Use

34. **What is an event?** An event is one hospital admission.

35. **What is a unit?** A unit is one day of a hospital stay.

36. **Is there an overall count for events, units and costs for the claims data included here?** No. To easily access this information, please use Data Set I, which provides this information for the same population in the same timeframe. Partners can also easily calculate counts for recipients and providers. Please note that counts for 2010 will balance with the counts for Data Set I, but the inclusion of hospital care beyond the pre-set experience period will result in the appearance of events, units, and costs data higher than those values in Data Set I.

37. **What is a covered day? What is a non-covered day?** On occasion, a person can become eligible or enrolled while in the hospital, or can lose Medicaid eligibility while in this hospital. On other occasions, a person may enter the hospital with an authorization for Medicaid payment for a limited number of days. These situations can results in a number of days for which Medicaid will reimburse costs ("covered days"), and a number of days for which Medicaid does not pay ("non-covered days"). This data field contains these counts for those recipients to which it applies. Each day is counted as one unit irrespective of whether or not the recipient was covered by Medicaid on that day.

38. **How are inpatient hospitalization days and Length of Stay (LOS) counted in the data set?** HFS uses the traditional health insurance calculation for total inpatient days: Discharge Date – Admission Date = total number of inpatient days (Length of Stay). (Equivalently, one may consider this as counting only the nights of a stay, or as disregarding the day of discharge.) Under this formula, a person who is admitted to the hospital on a Monday and discharged the next Monday has a seven-day (not eight-day) inpatient stay. Length of stay (LOS) in the data set is calculated this way. This figure should equal the sum of covered and non-covered days.

    An exception occurs for interim (series) bills, which use the formula Discharge Date – First Date of Bill +1 to ensure proper payment of all the days in the bill. This can include the first bill in a series of bills for a hospital stay, and is therefore relevant to Bill Type 2 (first bill in a series) and Bill Type 3 (continuing bills in a series). The first date of the bill will be the admission data for Bill Type 2, and will be the first date of the bill (typically the first day of the month) for Bill Type 3.

    Please note the data set will report an inpatient stay of one day for a person who checked into and out of the hospital on the same day, even though the formula would generate an inpatient stay of zero days.

39. **What does the indicator "When Earliest ≠ Admit" mean?** It is possible for a patient to be admitted at the point prior to the start of the experience period (January 1, 2010). This indicator = 1 for patients who were admitted to the hospital before the turn of the calendar year. Other data fields

provide additional information as to their length of stay, covered days, etc., both inside and outside of the experience period.

40. **What does the indicator "When Latest ≠ Discharge" mean?** It is possible for a patient to remain in the hospital past the end of the experience period (December 31, 2010). This indicator = 1 for patients who were not discharged from the hospital as of the end of the calendar year. Other data fields provide additional information as to their length of stay, covered days, etc., both inside and outside of the experience period.

41. **Does this data set include information on mortality?** Death during hospitalization is not included as a separate data field in Data Set IIB: Hospital. However, Patient Status Code includes several classifications of "Expired," which is synonymous with deceased. Please consult this field for information on death. Further, Death Date was included in Data Set I and can be used after a join is completed to specify deaths within a certain span or a specific patient population. Please note that no element of any of the data sets includes specifics as to the cause of death.

42. **Does the data set include information on a patient's destination after discharge?** Yes. The data field called "Patient Status Code" conveys the patients' immediate destination after a discharge, including codes such as "left against medical advice/discontinued care" and "expired."

43. **Does the data set include a readmission indicator? Is it possible to calculate readmission rates using the data?** There is no readmission indicator in the data set. However, the hospital claims data documents multiple admissions over the experience period, and it is therefore possible to calculate the rate of readmission within 7, 30, or any other number of days.

44. **Is comparing the data or joining the data to an outside source acceptable**? This data set is deliberately arranged to mimic typical discharge data available from Illinois Department of Public Health to the largest extent possible. This is intended to allow data-savvy organizations to work with a familiar format, but also to allow for the insights that can arise from the comparison of a Medicaid-only data set and a data set that includes all payers. The feasibility of joining the data table to a table from another source depends on the specific fields of interest and the contents of the outside source.

    However, please note that joining this data to any other data in a way that reveals the identity of any person represented in the data set is a violation of HIPAA law. This is grounds for termination of the Data Use Agreement and therefore a termination of your organization's data access.

45. **Can the data in this data release be used to create baselines for quality measures defined by HFS as important to care coordination?** We know that our partner organizations have an interest in the use of our data to analyze recipients' use of specific services important to quality measures. We recognize some hospital claims data could possibly be useful towards this end. In particular, this data set is useful for assessing 30-day readmission rates. However, please note that the CCIP hospital claims data table has not been designed expressly for this function. We urge all partner organizations to carefully review the requirements of quality measures as they use relevant data from any source. We also urge partners to note that Section 3A of the Data Use Agreement specifies that this data may only be used for the purposes of Care Coordination Innovations Project

proposal preparation. Any other use, including for assessment of quality measures, is a violation of this and can impel HFS to terminate a Data Use Agreement and/or otherwise seek remedies.

46. **What is the "Data Issue Indicator"?** For the sake of expediency, we are releasing this data before completing all the changes necessary to correct errors. The remaining errors affect a very small number of the claims in this data set (less than 100 out of 50,000 claims) and occur only in certain categories where our claims-level data is recorded in slightly unorthodox ways, largely with respect to payment coding. Those affected claims are marked with our "Data Issue Indicator." Data users may wish to eliminate them from their considerations or to consider analysis of costs carefully. Inquiries are welcome.