SECTION 2: MEASUREMENT

Administrative Data for Quality Improvement

Rachel M. Schwartz, MPH*; David E. Gagnon, MPH‡; Janet H. Muri, MBA*; Q. Rose Zhao, MS§; and Russell Kellogg, BA$

ABSTRACT. This article discusses the use of administrative data for quality improvement in perinatal and neonatal medicine. We review the nature of administrative data and focus on hospital discharge data as the primary source of hospital- and community-based assessments. Although discharge abstract data lack the richness of primary data, these data are the most accessible comparative data source for examining all patients admitted to a hospital. When aggregated to the state level as occurs in more than 30 states, hospital discharge data reflects hospital utilization and outcomes for an entire geographic population at the state and community level.

This article reviews some of the weaknesses of administrative data and then focuses how these data can be used for hospital- and community-based assessment of perinatal care citing as examples the measures of perinatal process and outcome used by the National Perinatal Information Center in its Quality/Efficiency Reports for member hospitals and a study of perinatal high-risk care in the State of Florida. The use of discharge abstract data for performance measurement at either the hospital or the system level requires a thorough understanding of how to select a patient group, its characteristics, the intervention, and the outcomes relevant to that patient group. In the perinatal arena, the National Perinatal Information Center has selected and presents those measures that rely on data items shown to be the most reliable based on validity studies and clinician opinion, delineation of the intervention, and the measurement of what occurred.

As hospitals respond to the recent pressures of the Joint Commission on Accreditation of Healthcare Organizations and other quality assurance entities, the accuracy of the discharge data will improve. With accepted caution, these data sets are invaluable to researchers studying comparative populations over time or across large geographic areas. Pediatrics 1999;103:291–301; administrative data, claims data, hospital discharge data, secondary data, process measurement, outcome measurement, population studies.

WHAT IS ADMINISTRATIVE DATA?

Administrative data are data sets that have been created in the health services area for purposes that are usually related to billing for services. They are also called secondary data sets because they are generated as a by-product of a nonresearch activity. An example of primary data is data collected for a clinical trial or medical record data. Administrative data, like primary data, exist at the patient level and can be aggregated to the provider, the community, or the state. There are two major types of administrative databases that are used in health services research: claims data and hospital discharge abstract data. They both include information on the reason the patient sought services in the form of a diagnosis and information on the charges for those services. In the case of inpatient care they include the duration of the care. Although this article focuses only on hospital discharge abstract data, in this section we will provide some basic information about each type.

Claims data files are created by payers from bills

T

He purpose of this article is to discuss the uses of administrative data, specifically hospital discharge abstract data, for quality improvement in neonatal and perinatal medicine. Because clinicians may not be familiar with the nature of this type of data, the article begins with a discussion of the definition of administrative data (also known as secondary data) including the source of the data and the items that are most often included. We then discuss the problems with administrative data so that all readers are familiar with its limitations. Finally, we focus on how these data can be used from two perspectives—that of the hospital and that of the population. In the former case the users are the hospitals’ staff including administrative and clinical staff; and in the latter the users include: insurers, state public health officials, and a wide ranges of other parties interested in issues of quality of care. We highlight how, despite the many limitations of this data, it can be used to examine a range of issues of interest to neonatal and perinatal clinicians, administrators, and policy makers.

From the National Perinatal Information Center, Providence, Rhode Island. *Vice President(s); ‡President; and §Research Associate(s), National Perinatal Information Center. Received for publication Sep 9, 1998; accepted Sep 9, 1998.

ABBREVIATIONS. UB, uniform bill; HCFA, Health Care Financing Administration; UHDDS, Uniform Hospital Discharge Data Set; NAHDO, National Association of Health Data Organizations; HCUP, Health Care Utilization Project; AHCPR, Agency for Health Care Policy Research; NPIC, National Perinatal Information Center; ICD-9-CM, International Classification of Diseases, 9th Revision Clinical Modification; VBAC, vaginal birth after cesarean section; QA, quality assurance; IVH, intraventricular hemorrhage; JCAHO, Joint Commission on Accreditation of Healthcare Organizations; IUGR, intrauterine growth retardation.

ABSTRACT. This article discusses the use of administrative data for quality improvement in perinatal and neonatal medicine. We review the nature of administrative data and focus on hospital discharge data as the primary source of hospital- and community-based assessments. Although discharge abstract data lack the richness of primary data, these data are the most accessible comparative data source for examining all patients admitted to a hospital. When aggregated to the state level as occurs in more than 30 states, hospital discharge data reflects hospital utilization and outcomes for an entire geographic population at the state and community level.

This article reviews some of the weaknesses of administrative data and then focuses how these data can be used for hospital- and community-based assessment of perinatal care citing as examples the measures of perinatal process and outcome used by the National Perinatal Information Center in its Quality/Efficiency Reports for member hospitals and a study of perinatal high-risk care in the State of Florida. The use of discharge abstract data for performance measurement at either the hospital or the system level requires a thorough understanding of how to select a patient group, its characteristics, the intervention, and the outcomes relevant to that patient group. In the perinatal arena, the National Perinatal Information Center has selected and presents those measures that rely on data items shown to be the most reliable based on validity studies and clinician opinion, delineation of the intervention, and the measurement of what occurred.

As hospitals respond to the recent pressures of the Joint Commission on Accreditation of Healthcare Organizations and other quality assurance entities, the accuracy of the discharge data will improve. With accepted caution, these data sets are invaluable to researchers studying comparative populations over time or across large geographic areas. Pediatrics 1999;103:291–301; administrative data, claims data, hospital discharge data, secondary data, process measurement, outcome measurement, population studies.

WHAT IS ADMINISTRATIVE DATA?

Administrative data are data sets that have been created in the health services area for purposes that are usually related to billing for services. They are also called secondary data sets because they are generated as a by-product of a nonresearch activity. An example of primary data is data collected for a clinical trial or medical record data. Administrative data, like primary data, exist at the patient level and can be aggregated to the provider, the community, or the state. There are two major types of administrative databases that are used in health services research: claims data and hospital discharge abstract data. They both include information on the reason the patient sought services in the form of a diagnosis and information on the charges for those services. In the case of inpatient care they include the duration of the care. Although this article focuses only on hospital discharge abstract data, in this section we will provide some basic information about each type.

Claims data files are created by payers from bills
generated by providers seeking payment for services rendered. These data include a wide range of patient utilization information including information on all types of inpatient and outpatient services and basic characteristics of the patients such as age, sex, and diagnosis. For inpatient hospital-based care the claims are created using the Uniform Bill (UB)-82 (1982 UB later updated to the UB-92, an electronic version), a federal government required format for all payers. The UB-92 is based on standard revenue codes used to create the patient’s bill. There is a charge master that provides a revenue code for every billable health care activity or object including all visits, procedures, inpatient stays, medical visits and pharmaceuticals. An example of the UB-92 format is provided in Fig 1.

Claims data can be used to examine patients’ utilization and outcome of services over time across a wide range of providers when the patient remains enrolled with the same payer. Patients can be characterized by age, sex, diagnosis, and procedures. Subgroups can also be studied. However, the analysis of claims data are limited to what is billed to the specific payer and is limited to enrollees within that payer. For example, if a Medicaid enrollee becomes enrolled during the intrapartum period the claims file would not include information about services received before enrollment.

The claims data sets are created by the Health Care Financing Administration ([HCFA], the federal agency responsible at the national level for Medicare and Medicaid—www.hcfa.gov) for Medicare and by the states for Medicaid. Such claims files exist for commercial payers but are largely unavailable publicly. Medicare and selected states’ Medicaid claims files can be obtained through HCFA for research. In addition, often states will work directly with researchers in providing access to such data files. Confidentiality is protected by HCFA and the states through the use of signed confidentiality agreements and limitations on access to information that could identify an individual person. In summary, a claims data file can provide information on the users of services through time classified by diagnoses or characteristics across a range of inpatient and outpatient providers as long as the person remains enrolled.

Hospital discharge abstract data sets include only the hospital inpatient component of care. The core of the data are from the Uniform Hospital Discharge Data Set (UHDDS) that is abstracted from the patient’s medical record after discharge. The UHDDS abstract mandated for Medicare and Medicaid by the federal government in 1974 is used by all hospitals for all inpatients in the United States. The elements included in the UHDDS form the core of the inpatient clinical information available on all hospital discharge data sets: hospital identifier, patient identifier, date of birth, sex, race, residence, admission date, discharge date, attending physician identifier, operating physician identifier, diagnoses, procedures with dates, discharge disposition of the patient, and expected source of payment. The discharge abstract databases prepared by states for all hospitals usually include summarized data from the UB-92 billing data. Some states have added additional detail to the files. Hospital discharge abstract data sets describe hospital-based services by providing a snapshot of all patients discharged from a hospital in a given period. This is an important difference between claims files and discharge abstract databases. Unlike claims databases, discharge abstract databases can examine patients through time only when the patient exclusively uses the same hospital for services. In addition, although all hospitals collect and store data on each patient, not every state has a repository for all the hospitals’ data in the state. The National Association of Health Data Organizations (NAHDO—www.nahdo.org) published a Data Resource Guide that is being updated this year. They estimate that there are almost 30 states with state-mandated data systems. There are another half dozen states with systems that are not controlled by the state government but by the hospital association.

Although claims data sets have the advantage of including both inpatient and outpatient care and can be linked through time for enrollees, hospital discharge abstract data sets cover all patients regardless of payer and reflects the hospital utilization and outcome experience of an entire geographic population. It allows for the comparison among populations across payers and across populations over time.

States usually will make statewide public use data files available to researchers for a fee. Some states have strict confidentiality rules but will nonetheless create special research data files in addition to their public use files. Often hospital associations will not release the data because its collection is not a state mandated function but rather a function of the hospital association. In these cases, special permission is required; and even then, direct access to patient level data may not be provided. In addition to state databases, there are other databases that have been created using hospital discharge abstract data. One created by merging data from multiple states is the Health Care Utilization Project (HCUP)-3 data—a National Inpatient Sample, which contains all discharge records from a 20% sample of US community hospitals from 17 states created by the Agency for Health Care Policy Research ([AHCPR]—www.ahrpr.gov) to study trends in cost and utilization across the United States. HCUP-3 data sets, available publicly on CD-ROM, extend back to 1988. These early data sets, however, are not as complete because the participation rate was not as high in the earlier years. AHCPR also has a nonsample file from 12 states that extends back to 1988; however, this data set is not available publicly, but rather is used by AHCPR for intramural research.

There are also proprietary reporting systems that rely on individual hospital’s submission of data on a voluntary basis usually for the purpose of comparative performance measurement. The longest standing system of this sort was initiated in the early 1970s by Commission on Professional and Hospital Activities and is based on the data from the UHDDS. This organization was merged into the HCIA, Inc, a proprietary data center in Baltimore, MD. Hospitals de-
cide on their own to participate and submit patient level hospital data files directly to the reporting system. Another example is the data set of the National Perinatal Information Center ([NPIC]—www.npic.org) initiated in 1985. NPIC relies on the submission of case level discharge abstract and UB-92 data for all patients admitted to each member hospital. As noted there other systems like that of NPIC; however, the NPIC system is the only one that includes features that are designed to allow a special focus on perina-

Fig 1. An example of the UB-92 format.
tial care. First, the NPIC Perinatal Center Database requests specific perinatal items including: birth weight, gestational age, Apgar, days of ventilation, linkage between the mother and newborn record, and admission to the neonatal intensive care—all which augment the analysis of perinatal issues. Second, validation of the data after submission by the hospital is used to improve the quality of the data. Data validation involves a preliminary analysis of the quality and validity of data and the return of questionable data to the hospital for scrutiny. For example, birth weight information and discharge disposition are two data elements checked for accuracy during the data validation process. The importance of these is highlighted further below. In this article we will give examples from our work with state data sets and the NPIC data to demonstrate how these data can be used in quality improvement.

ADMINISTRATIVE DATA CRITIQUE

In the measurement of quality there are two types of measures—those that measure the process of care and those that measure the outcome. An oversimplification would say: process refers to the type of treatments provided; and the outcome refers to whether or not the patient got better, sicker, or died. However, to measure these with any data set, a number of factors must be true about the data. First, the patients must represent a class of patients with the characteristics of the group we are describing; and they should represent a fair comparison to each other. More specifically, researchers select patients to study based on similar characteristics such as age group, sex, race/ethnic group, or diagnosis. Therefore, proper definition of the group, ie, who is in or out of the data analysis, depends on the accuracy of the recording of those items. In addition, when patients are selected on one set of parameters such as diagnosis, it may be necessary to risk-adjust based on other characteristics like age or whether or not the patient has a specific comorbid condition or group of conditions. Second, the treatments or process of care must reflect what actually occurred. In other words, a patient who received a specific surgery or other procedure must have a postoperative complication had the complication recorded.

For each of these three elements of analysis the data items must be accurate or at least accurate enough so that the analysis reflects reality. In a secondary data set where we do not have the luxury of going back to the chart for every analysis, we need to have a good understanding of the limitations of the data—what it can and can not do. In this section we discuss some of the general critiques of administrative data and some specific problems we have identified in our experience in the perinatal area. For both state and NPIC data, we will review key data items present in administrative data sets and their limitations for patient characterization, risk adjustment, process of care, and outcome assessment.

Age

In the use of administrative data for quality measurement, the date of birth variable is absolutely critical. It allows the computation of age for any patients as well as the identification of newborns, infants readmitted to the hospital within 28 days of birth, and identification of the age of the newborn at death. In administrative data, date of birth is a requested item; and generally in these data sets, age is accepted as being accurate based on patients report of date of birth. Age is a simple, straightforward variable with good face validity, which means it measures what it claims to measure. Comparisons of medical records and discharge abstract data for NPIC member hospitals has shown high concordance between the two (unpublished analysis of 5,283 abstracted medical records with NPIC linked discharge abstract information data base funded by AHCPR for the study entitled, Antenatal Corticosteroid Dissemination Study). Some public use state data sets, however, do not include date of birth but rather a variable called age, which is used to protect patient confidentiality. In such cases, the researcher or person requesting the data must be careful to ensure that the data file actually provides a variable for neonates separately; otherwise, all infants <1 year of age will be the finest distinction possible. If available, date of birth is the least problematic variable and can be used as a method of subdividing cases or risk-adjusting in the maternal population as well.

Diagnosis and Procedures

Patient Selection

Administrative data sets use the International Classification of Diseases, 9th Revision Clinical Modification (ICD-9-CM) coding scheme to describe patients in great detail. This method uses internationally agreed coding conventions. These are updated periodically and provide a framework for examining the diseases and conditions for which a person enters the hospital. There are thousands of codes that are ordered by the organ system affected with up to 5 digits to describe it.

The explication of all the codes is provided in three volumes. Volume 1 includes the diagnoses as a tabular list by code number organized by 17 body systems and condition. For example, codes between 630 and 677 include all the diagnoses related to pregnancy under a heading entitled Complications of Pregnancy, Childbirth and the Puerperium while codes 760 to 779 include all the codes used for neonates entitled Conditions Originating in the Perinatal Period. Congenital anomalies are organized together as one group independent by organ system. Volume 1 details each code concerning what is and is not captured within a specific diagnosis. Volume 2 lists diagnoses alphabetically, and Volume 3 lists the procedures numerically and alphabetically.

Hospital discharge abstract data sets from the states usually have anywhere from 5 to 25 ICD-9-CM diagnostic fields. The NPIC data specifications allows up to 15 codes. The first code is always the
principal diagnoses and reflects the dominate reason the patient was treated in the hospital. In the perinatal arena, pregnancy is straightforward enough so that selection of pregnant women admitted for antepartum, intrapartum, or postpartum care can be accomplished accurately and easily using these ICD-9-CM codes.

Diagnosed medical problem or condition can be used to select patients to be studied. Patients can also be selected based on both diagnosis and procedure codes or procedure alone. Patients who have a cesarean section or a forceps delivery are identified using procedure codes. Here, again, identifying these women is extremely reliable because the condition is one that is obvious to both the patient and the clinicians. As noted above, newborns are selected based on age and there is virtually no problem in the identification of newborns. Newborns can be further described as transfers from another hospital, readmissions, or multiple births using diagnostic codes.

Although selection of cases is generally not a problem in the perinatal area, it can be more of a generic problem when using administrative data sets. Often administrative data are weak in the identification of cases using diagnosis codes for a number of reasons: physician unfamiliarity with the use of ICD-9-CM coding and therefore a lack of congruity between the charting by the physician and the abstraction of the principal diagnosis by medical records staff; ICD-9-CM codes often do not allow for adequate discrimination among cases with the same disease because the diagnosis alone does not capture the severity; physician listing diagnoses on the chart that are suspected but not necessarily final; and upcoding of severity to enhance payment. Finally, a diagnosis in the discharge abstract data does not include the timing of the problem so patients admitted with certain diagnoses can not be distinguished from those who become sick in the hospital. The Institute of Medicine studied the reliability of discharge abstract data in the late 1970s and found that, while information related to dates of admission, discharge, and age were accurate as were payer and gender, the reliability of diagnoses was inconsistent with only 65.2% agreement on the principal diagnosis. However, they also determined that most discrepancies resulted from differences in interpretation of the medical record for the purposes of coding, ie, they were related to placement of the diagnosis (principal or secondary) rather than whether or not it was present. The result of such a disagreement can be the miscoding of the patient if selection were to be based on principal diagnosis alone. A similar finding was present in a study of calibration discharge abstract data reported by Romano. In this study where discharge abstract data were compared with abstracted medical record data the researchers found that in seven of eight comorbid categories, sensitivity exceeded 85%.

Diagnosis for Process and Outcome Measurement

The problems related to the under and/or overidentification of patients are the same reasons great care must be taken when using administrative data for process and outcome measures. The selection of diagnoses where coding and physician recognition of disease are most congruent represent the diagnoses that can best be used for process and outcome measures. In the examination of process, procedure data such as the presence of a cesarean section, repeat cesarean section, vaginal birth after cesarean section (VBAC), the use of forceps, or episiotomy that rely on procedure codes to describe the process of obstetric care are the least problematic because there is little judgment involved in whether or not the procedure occurred and is recorded on the chart. Hsia found that procedure code is extremely accurate (0.88 to 0.95 correspondence for 10 of 15 procedure codes examined) when the order of the procedure on the abstract is not considered. We believe selection of the rate of cesarean section and forceps delivery as measure of process of care reflect the use of data items on which there is little disagreement. This same issue is true with regard to neonatal surgery; surgical intervention as a process of care measure is a major intervention where the coding can be accepted as accurate because all newborns who had the surgery are coded and that those who did not have the surgery are not coded.

Diagnosis data, unlike procedure data, are fraught with problems. Examples in obstetrics where diagnosis data are useful include intrapartum complications such as laceration where there is little debate among clinicians. Postdelivery infection, however, is an example of the use of diagnosis codes to develop an obstetric outcome measure that could be problematic especially when comparing hospitals. For example, there are multiple codes for coding infection; in addition, infection may be coded when it is suspected but not present. The ICD-9-CM coding scheme does not include clinical requirements for coding-specific diagnoses. Hospitals with high infection rates may be those where the quality assurance (QA) program is extremely vigilant, reflecting very aggressive coding and charting of infections while other hospitals in the comparison group may be more lax. Iezzoni found that hospitals with higher complication rates code more diagnoses per case. This is not to say postdelivery infection rates should be ignored; rather, they may be used for monitoring by hospitals and internal assessment over time; but within comparative data sets, they must be interpreted with caution.

In examining newborns, the use of diagnosis codes for measuring outcomes is problematic for almost every diagnosis we might wish to select. As noted above, diagnoses of interest may not occur as the principal diagnosis. To improve the accuracy of analyses, NPIC has written many of its analytic algorithms by focusing on specific conditions to identify a patient regardless of where the diagnosis is mentioned. As a result, the placement of the diagnosis becomes unimportant in the examination of rates of occurrence. Thus, to study rates of patients with intaventricular hemorrhage, all cases where this diagnosis appears will be selected. Even so, this approach does not eliminate a frequent problem; ICD-9-CM coding does not capture the gradations of
severity of intraventricular hemorrhage (IVH) or the possibility that overidentification may be present at an indeterminate level because patients with suspected IVH may be coded the same way as those with documented IVH.

Others diagnoses have similar problems. For example, hospitals vary in their use of the diagnosis of respiratory distress syndrome versus hyaline membrane disease and respiratory problems not otherwise specified. In other words, across hospitals and even across physicians, acceptance of a very specific diagnosis varies. The more the acceptance varies, the more difficult the diagnosis is to use as a measure of outcome. NPIC does use measures based on diagnosis; but because of problems with acceptance of ICD-9-CM coding of diagnoses as accurate reflections of patients’ illness(es), NPIC hospital members are told by staff at NPIC that these analyses are guides to help them understand their case-mix, their own trends over time, and how they compare with similar perinatal centers. We often suggest if hospitals have high rates on performance measures based on diagnoses, they should request a listing of the medical record numbers of these cases for more careful scrutiny (see below). In her analysis of complication rates for five groups of medical and surgical patients Iezzoni13 draws a similar conclusion. As performance measurement becomes increasingly popular and evidenced-based medicine becomes increasingly the focus of research, the use of administrative data-derived performance measures by hospitals as a mechanism for understanding their own data, the coding, and the problems with measurement will become an important part of the entire QA and quality improvement process.

Birth Weight

Birth weight can be used to select groups of high-risk newborns and used as a mechanism of case-mix adjustment because it is a proxy for prematurity. Birth weight is recorded on the birth certificate and is always in the newborn chart. A unique aspect of the NPIC hospital discharge data are the presence of actual birth weight data in grams. We have shown that birth weight alone explains 21% of the variance in length of stay and birth weight in combination with transfer status describes resource use almost as well as case-mix adjustment systems far more complicated.14 Even when actual birth weight is not available, it can be found in most data sets because it is the 5th digit of the ICD-9-CM code for the diagnoses of Slow Uterine Growth and Fetal Malnutrition (764) and Extreme Prematurity (765). Table 1 shows the percent agreement between birth weight and ICD-9-CM 5th digit coding. Thus, researchers can feel confident that state administrative data sets that have reasonably good diagnosis coding can be used to identify low birth weight infants.

Disposition

Disposition refers to the discharge disposition of the patient. There are three main categories: death, home, transfer. Transfer can be to another acute facility, home care, or to a skilled nursing facility. Death is regarded as extremely accurate when the infant weight is above 500 g. NPIC validates mortality data with the hospitals, and the problems with this variable are related to newborns <500 g. More specifically, there are hospitals that will have cases who die in the delivery room below this 500-g category who are not counted as livebirths; and no newborn discharge abstract record is created for them. The policy with regard to these cases varies from hospital to hospital; and while they may be treated uniformly for the purposes of vital statistics, they are not necessarily treated uniformly from the perspective of the hospital discharge abstract and billing system. As a result, all the analysis of neonatal outcome performed by NPIC excludes newborns <500 g. Thus, neonatal in-hospital mortality for infants >500 g can be used as a reliable data item for outcome measurement.

Race

Race and ethnicity is requested in six or more categories in most administrative data sets and usually include: white, black or African-American, Hispanic, Asian, Native American, and other. There is extensive literature on the issue of meaningfulness and accuracy of race in health services research.15–18 Here we will highlight some of the basic issues of the problem that arise from a variety of sources. First, there is the conceptual problem—we do not really understand or even know the difference between race and ethnicity. For example, when we ask if a person is Hispanic are we asking about whether their origin is from a Spanish-speaking country in Latin America even if they are of European origin or from a native mountain people living in the high Andes? The second problem arises from the structure of the administrative databases of the sort we are discussing here. The data set allows only one field to accurately reflect this characteristic of the patient. Thus, a person of mixed race often has her most dominate

<table>
<thead>
<tr>
<th>Actual Birth Weight</th>
<th>Birth Weight Code From DX 764–765’s 5th Digits</th>
</tr>
</thead>
<tbody>
<tr>
<td>500–749 g</td>
<td>91.19%</td>
</tr>
<tr>
<td>750–999 g</td>
<td>92.06%</td>
</tr>
<tr>
<td>1000–1249 g</td>
<td>91.07%</td>
</tr>
<tr>
<td>1250–1499 g</td>
<td>92.52%</td>
</tr>
<tr>
<td>1500–1749 g</td>
<td>91.09%</td>
</tr>
<tr>
<td>1750–1999 g</td>
<td>91.63%</td>
</tr>
<tr>
<td>2000–2499 g</td>
<td>84.80%</td>
</tr>
</tbody>
</table>
characteristic coded as the race because much of the information is coded based on observation rather than self-report. However, even with self-report a person is being asked to characterize who they are in one word. Third, with regard to infants—the patients with whom we are most concerned here—the problem is confounded in administrative data because the infant’s race is dominated by observation of the mother, which is what is recorded in the medical record. Thus, in a case where the mother is black and the father is white, the child will be coded black. The birth certificate, however, allows separate coding of the parents based on self-report so analysis can be done in a somewhat more sophisticated manner by taking into account both parents. Finally, there is the issue of whether what is coded in the administrative databases can be corroborated from other data sources. In a study where NPIC went back to the chart, (unpublished analysis of 5,283 abstracted medical records with NPIC linked discharge abstract information data base funded by AHCPR for the study entitled, Antenatal Corticosteroid Dissemination Study), we have found 91% agreement with the coding on race for black and 84% for white with what is in the medical record; however, unless the birth certificate is with the chart, the race of the child is based solely on observation or self-report of the mother. Despite these problems, much of the US population of whites and blacks considers itself to be either one or the other and, as discussed below, broad distinctions are important to keep in mind.

When race data has so many issues associated with it, why use this information at all? In the literature outcome has varied with race. When examining newborns using vital statistics data or administrative data, non-white or black babies in lower gram weight groups actually have had lower mortality rates than whites. This means that, when examining morbidity and mortality within a hospital or across an entire community or communities, it is important or adjust for this fact. As an example, a QA manager and the director of neonatology wish to compare the mortality of infants 500-g to 750-g among a group of hospitals. One hospital is in an inner-city while another is in a suburb. As a result, merely comparing the two hospitals without adjusting for race (however imperfect that adjustment is) would make the hospital in the suburbs appear as though it had a higher mortality because its underlying population is white. There has been significant research speculating about why this difference is present—biological, genetic, and social hypotheses have been presented. Interestingly, we have found that the difference that was so dramatic in the past work has become less important more recently. This may be because technologies such as surfactants are improving neonatal mortality more for whites than blacks. Indeed, in the models presented in this article, race is not a significant predictor of mortality across the data set; however, because of its historical importance, NPIC continues to use it in risk adjustment as do many others performing measurement in newborn care.

QUALITY MEASUREMENT USING DISCHARGE DATABASE

Hospital Level

Obstetrics

The NPIC Quality/Efficiency Report of the last 5 years has reported on eight obstetric measures of quality. The state of the art with regard to diagnosis and procedure data, leads us to recommend that the best measures for obstetrics are those that focus on the process of care. These include cesarean section rate, VBAC rates, repeat cesarean section rates, and obstetric readmission rates. Outcome measures are limited to complications of delivery that are subject to minimal interpretation, and would include those involving significant findings such as laceration by degree and wound infection. Anesthesia complication is a softer measure, but because of its importance in hospital quality assurance programs, it is reported by NPIC. These measures are examined every year by all the hospitals and input is sought regarding their reliability.

The quality obstetric measures are not risk-adjusted by age or other factors and are shown in Table 2.

In these reports hospitals are given their own rates and compared with all perinatal centers reporting. All the hospitals are shown graphically for each measure so that the range of variation is clear (see Fig 2 and 3) (Analysis: NPIC NICU admission data in 30 NPIC member hospitals in 1996). The mean for the group is shown as a dotted line and the confidence interval around the mean is also shown. The confidence interval is based on the assumption that the group of participating hospitals is a random sample of all perinatal centers. Of most interest to members is whether or not the hospital falls above or below the confidence interval of the measure. Hospitals can request a listing of the medical record numbers of cases that fall in the category being studied. Alternatively, if they wish to raise their VBAC rate, they request an analysis of the cesarean section rate by physician identified by code. The reports are used as an integral part of the quality improvement activities within the hospital. This is especially important in the context of the extensive monitoring that will be occurring through the Joint Commission on Accreditation of Healthcare Organizations (JCAHO)–www.jcaho.org in the future under its new ORYX program which is designed to monitor hospital performance (see below).

Neonatal

For newborns the Quality/Efficiency Report also includes eight measures of quality. Of these eight, one is a process measure and seven are outcome measures. Three of the seven outcome measures focus on in-hospital mortality, and four focus on diagnosis. These measures were selected by working with neonatologists in NPIC member hospitals using data elements that both the NPIC and hospital staff believe to be most reliable. Mortality measurement uses key patient characteristics available on the dis-
The first measure is **Outborn Mortality Rate: Neo-nates 500 to 1499 g Transferred In**. The rate of unadjusted outborn newborn mortality is analyzed separately from inborn because these cases receive both...
OB and immediate newborn care from locations outside the control of the hospital being studied. Hospitals with very high transfer in rates tend to have higher neonatal mortality.

Although outborn mortality is presented as a raw rate, both a raw rate and adjusted actual versus expected deaths are calculated for the next two measures. In-hospital Deaths: Inborn Singletons 500 to 1499 g and In-hospital Deaths: Inborn Singletons 500 to 2499. Actual versus expected deaths are based on a logistic regression model. The model is used to adjust for case-mix differences present in the hospital that go beyond differences in the birth weight distribution. The original model presented here was published in the New England Journal of Medicine in 199424 where it was used to show that mortality declined after surfactant became available. The model has been refined and tested over a period. The newest version presented in Table 3 shows the coefficients of the model for the variables on which we adjust: sex, race, presence or absence of a congenital anomaly (ICD-9-CM codes 740–759), presence or absence of intrauterine growth retardation (IUGR–ICD-9-CM code 764), and birth weight category. Anomaly and IUGR are the only two types of diagnoses used to adjust because these are diagnoses that can not be affected by the newborn care provided. They represent conditions that the neonatologists find unalterable but are very important to outcome. As can be seen, neonates with an anomaly are 2.9 times more likely to die compared with those without an anomaly. Those with IUGR are half as likely to die. Although race is insignificant in this presentation, the original models showed race to be highly significant. When further exploration is requested by hospital, subcategories of patients are analyzed separately. This model is highly predictive with a c-Statistic of 0.85 for newborns >500 g and 0.79 for newborns 500 to 1499 g.

The NPIC outcome measures that focus on diagnoses are as follows:

- Percent of Inborns With Diagnosis of Significant Birth Trauma. This measure is the percent of all inborns (including multiple gestation and those without valid birth weight) with a discharge diagnosis of significant birth trauma. Inborn cases with significant birth trauma are divided by all inborns. Significant birth trauma is defined as cases with an ICD-9-CM code of 767.0 or 767.2 to 767.8.

---

**TABLE 3.** NPIC Model for Predicting Mortality by Birth Weight Category for Inborn Neonatal Intensive Care Unit Admissions With Birth Weight >500 Grams

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Coefficient</th>
<th>P Value</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>1.3256</td>
<td>.004</td>
<td>1.096–1.603</td>
</tr>
<tr>
<td>White</td>
<td>0.8727</td>
<td>.162</td>
<td>0.721–1.056</td>
</tr>
<tr>
<td>Anomaly</td>
<td>2.8505</td>
<td>.000</td>
<td>2.339–3.374</td>
</tr>
<tr>
<td>IUGR</td>
<td>0.589</td>
<td>.050</td>
<td>0.347–0.999</td>
</tr>
<tr>
<td>Birth weight category</td>
<td>0.588</td>
<td>.000</td>
<td>0.563–0.613</td>
</tr>
</tbody>
</table>

---

Fig 3. Third and fourth degree perineal laceration rate, without forceps or vacuum extraction.
Bronchopulmonary Dysplasia Rates: Surviving Neonates 500 to 1499 g. Bronchopulmonary dysplasia (code 770.7) for all surviving neonates weighing 500 to 1499 g at birth.

Meconium Aspiration Syndrome Rate: Inborns. The rate of neonatal meconium aspiration syndrome (ICD-9-CM code 770.1) for inborn neonates.

Meconium Aspiration Syndrome Rates: Outborns. The rate of neonatal meconium aspiration syndrome for neonates transferred to your hospital.

One process measure is used for newborn quality assessment:

Inborn Readmission Rate: Readmissions for Normal Newborns Discharged Within Two Days. The rate is normal newborns readmitted within 28 days after discharge divided by all normal newborns discharged in 2 days. This rate does not include infants first readmitted to another facility.

These measures are used by NPIC member hospitals to compare themselves over time and to a group of other perinatal centers. Many of these hospitals have been working on improving their discharge abstract coding systems for 13 years. The hospitals review their reports in the data validation stage and the reporting stage and provide feedback on whether they feel the measures reflect reality. State data systems, unfortunately, do not always report back on measures to individual hospitals. If they do provide feedback, the analysis is usually limited and not focused on perinatal care with the possible exceptions of cesarean section and VBAC rates. This is changing, however, as hospitals are being required to report on performance to the JCAHO under the ORYX initiative.

Beginning July 1, 1998, the ORYX initiative required all hospitals to give patient-specific data to a measurement system meeting the JCAHO criteria. The data must be given to the system quarterly and the system, in turn, must process the data and forward 2 to 5 measures of quality chosen by the hospital from the system’s approved measures list to the JCAHO. For the first year, the measures must represent at least 20% of the hospital’s total discharges. Each year the number of measures and percent of discharges covered increases. A major concern of data analysis specialists, like those at NPIC who have been performing data collection and analysis for many years, is the variation in data quality and validation on which comparative measures are being created. Newer measurement systems may not be aware of the pitfalls and problems of the data outlined here.

Geographic or Population-based Use of Discharge Abstract Data

So far we have focused on the measures that can be used by hospitals for comparison over time and comparison with other similar hospitals. These same data sets, when available on the statewide level, can be used to assess whether the system of perinatal care in the state or region is functioning optimally.

In regionalized perinatal care one major focus is whether or not high-risk mothers and newborns receive care in the most optimal location. Generally, it is accepted that only those hospitals considered tertiary referral centers should care for high-risk newborns usually defined as those who are extremely premature or <1500 g; however, there is no formal national cutoff and there has been widespread expansion of level II units. Monitoring where high-risk newborns are born can be both a useful and meaningful use of state hospital discharge abstract data.

Monitoring the provision of high-risk care is easily accomplished in states like Florida, Massachusetts, New York, and New Jersey where the state collects information on the levels of care of all the hospitals who provide perinatal services. In areas where this is established, the states can create analyses of where newborns are born by gram weight. They can create these tables, by region, race, payer, and then over time identify if there is a shift in the location of care for the population of high-risk newborns and whether this shift is away from the highest level hospitals. Although this analysis can be improved if the data sets include both vital statistics and discharge abstract data, very simple analyses by region and hospital can provide a realistic presentation of which hospitals are caring for the higher risk patients. In a study performed by these authors for Children’s Medical Services in Florida, a process and outcome analysis of regionalized care was performed using the Florida discharge abstract data for 1993. The process involved examining where newborns were delivered by birth weight category and whether or not they were transferred in a timely manner if born outside a perinatal center. The process analysis revealed that the Florida perinatal guidelines were being followed to a significant degree. More than 75% of newborns <1000 g were delivered in a tertiary center. Among the remaining 24% born at a level II hospital (250 newborns), 30 were not transferred at all and 26 were held more than 1 day before transfer. Thus, a total of 8.5% of all newborns in this weight category could have been managed more effectively strictly from the point of view of the Florida guidelines that do not allow level II facilities to care for newborns <1000 g. We concluded that a selected group of hospitals (those with higher levels of technology within the level II group) were not following state guidelines for transfers but that for the most part, the regionalized system was functioning well.

This study also examined mortality by level of care and gram weight. Discharge abstract data are imperfect for this purpose because transferred newborns who die cannot be allocated back to the actual hospital of birth. To compensate for this we estimated the number of deaths within level III hospitals that could be attributed to the Level II transferred newborns and then allocated these deaths back to the level II hospitals as a group. Mortality for newborns <1000 g was 34.5% in level III compared with 37.1% in the level II group (P = .19). There was no statistically significant elevation in mortality in the level II group.

In cases where systems appear to function poorly, more detailed assessment of mortality and morbidity.
can be accomplished using linked vital statistics and discharge abstract data sets. We recommended organizing such a monitoring activity at the state level in Florida. Such efforts at the state level will become increasingly important as level IIs compete with perinatal centers and payers begin to allocate contracts to less expensive providers. Such a shift could result in increases in high-risk cases among hospitals with less technological sophistication30 (as found by Phibbs) resulting in newborn outcomes that are less than optimal.

The role of states and research groups in monitoring the delivery of services will become more important as managed care plays an increasing role in the provision of services. And while there are incentives for managed care providers to provide preventive services, there is little incentive to provide expensive high-technology care. And while there may be only limited evidence of adverse outcomes by payer,30 it is important to use the data sets and resources readily available in an optimal manner. Although these data sets have weaknesses, they also have enormous strengths. It is the role of providers, clinicians, and administrators to keep a careful tab on the systems of perinatal care as they evolve. State level administrative data sets are such a mechanism.

CONCLUSION

This article explored only some of the issues with regard to administrative data types, quality, and uses. The context for the work of NIPIC has mainly focused at the hospital level, but the link between the hospital level and the community level is one that can be made for more and better studies as the data are accepted as being more accurate. Accuracy, however, can only be improved at the source. To do this, hospitals need a vested interest in accurate reporting. As the work of JCAHO through ORYX becomes more prominent, hospitals’ attention to data will sharpen. Over time each hospital will work toward better coding, and in the future statewide data sets will be increasingly usable for comparative population studies.

REFERENCES

12. Ibid
13. Ibid
15. LaViest TA. Beyond dummy variables and sample selection: what health services researchers ought to know about race as a variable. Health Serv Res. 1994;29:1–16
Administrative Data for Quality Improvement
Rachel M. Schwartz, David E. Gagnon, Janet H. Muri, Q. Rose Zhao and Russell Kellogg

*Pediatrics* 1999;103;291

Updated Information & Services
including high resolution figures, can be found at:
http://pediatrics.aappublications.org/content/103/Supplement_E1/291

References
This article cites 17 articles, 1 of which you can access for free at:
http://pediatrics.aappublications.org/content/103/Supplement_E1/291.
full#ref-list-1

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Administration/Practice Management
http://classic.pediatrics.aappublications.org/cgi/collection/administration/practice_management_sub
Quality Improvement
http://classic.pediatrics.aappublications.org/cgi/collection/quality_improvement_sub
Fetus/Newborn Infant
http://classic.pediatrics.aappublications.org/cgi/collection/fetus:newborn_infant_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
https://shop.aap.org/licensing-permissions/

Reprints
Information about ordering reprints can be found online:
http://classic.pediatrics.aappublications.org/content/reprints
Administrative Data for Quality Improvement
Rachel M. Schwartz, David E. Gagnon, Janet H. Muri, Q. Rose Zhao and Russell Kellogg

*Pediatrics* 1999;103;291

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/103/Supplement_E1/291