



Concordance between survey reported childhood asthma and linked Medicaid administrative records

Benjamin Zablotsky & Lindsey I. Black

To cite this article: Benjamin Zablotsky & Lindsey I. Black (2018): Concordance between survey reported childhood asthma and linked Medicaid administrative records, Journal of Asthma, DOI: 10.1080/02770903.2018.1455854

To link to this article: <https://doi.org/10.1080/02770903.2018.1455854>



Published online: 17 May 2018.



Submit your article to this journal [↗](#)



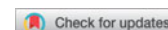
Article views: 39



View related articles [↗](#)



View Crossmark data [↗](#)



Concordance between survey reported childhood asthma and linked Medicaid administrative records

Benjamin Zablotzky and Lindsey I. Black

National Center for Health Statistics, Hyattsville, MD, USA

ABSTRACT

Objective: Agreement between administrative and survey data has been shown to vary by the condition of interest and there is limited research dedicated to parental report of asthma among children. The current study assesses the concordance between parent-reported asthma from the National Health Interview Survey (NHIS) with Medicaid administrative claims data among linkage eligible children from the NHIS. **Methods:** Medicaid Analytic eXtract (MAX) files from the Centers for Medicare & Medicaid Services (CMS) (years 2000–2005) were linked to participants of the NHIS (years 2001–2005). Concordance measures were calculated to assess overall agreement between a claims-based asthma diagnosis and a survey-based asthma diagnosis. Structural equation modeling was used to assess the association between demographic, service utilization, and co-occurring conditions factors and agreement. **Results:** Percent agreement between the two data sources was high (90%) with a prevalence-adjusted bias-adjusted kappa of 0.80 and Cohen's kappa of 0.55. Agreement varied by demographic characteristics, service utilization characteristics, and the presence of allergies and other health conditions. Structural equation modeling results found the presence of a series of co-occurring conditions, namely allergies, resulted in significantly lower agreement after controlling for demographics and service utilization. **Conclusions:** There was general agreement between asthma diagnoses reported in the NHIS when compared to medical claims. Discordance was greatest among children with co-occurring conditions.

ARTICLE HISTORY

Received 14 November 2017
Revised 27 February 2018
Accepted 19 March 2018

KEYWORDS

Asthma; children; diagnosis;
national survey;
parent-report; agreement

Introduction

Asthma is a prevalent chronic condition that is characterized by recurring periods of wheezing, chest tightness, shortness of breath, and coughing (1). Asthma can affect persons of all ages, though it usually is first diagnosed during childhood. As children grow, it is possible that their asthma can be asymptomatic as either their allergen sensitivities change, or their airway passages expand and inflammation is reduced. Recent parent-reported survey data suggest that 13% of children have ever had a diagnosis of asthma from a health care provider, whereas only 8% of children still have asthma (2). Although asthma can potentially be self-managed or controlled with the use of medications and avoiding exposure to allergens, each year approximately 10% of children with asthma have emergency department visits due to symptoms or attacks (3). Accurate estimates of the prevalence of asthma among children in the United States are important for planning adequate health promotion programs and interventions that may improve long-term outcomes for children with the condition.

The National Health Interview Survey and the National Health and Nutrition Examination Survey are the primary data sources used to estimate the prevalence of asthma in the United States. Health surveys such as these have the ability to capture a wide range of measures that can be used to examine factors associated with asthma, as well as to estimate the prevalence of diagnosed asthma and other conditions among population subgroups. However, surveys are limited in their reliance on self-report, or in the case of children, parent-report for collecting important indicators and conditions, which are subject to recall bias. Alternatively, administrative data such as medical claims or electronic patient records are another way to measure the prevalence of conditions such as asthma. Medical records may be more accurate as they are not subject to the respondent's ability to recall being diagnosed with a given condition. However, these sources are limited in that they can only describe the population that receive services, are dependent on accurate notation in the medical record, and can only capture services that are billable (4, 5). Administrative

CONTACT Benjamin Zablotzky, PhD  bzablotzky@cdc.gov  National Center for Health Statistics, Centers for Disease Control and Prevention, 3311 Toled Road, Hyattsville, MD, 20782, 301-458-4621, USA.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ijas.

 Supplemental data for this article can be accessed on the [publisher's website](#).

© 2018 Taylor & Francis Group, LLC

data sources that capture diagnoses and procedures, such as medical records, are also limited in that they typically do not include individual characteristics that are useful in describing subpopulations.

Studies that utilize linkage of multiple data sources are particularly valuable to inform public health practitioners as they are able to maximize the value of each data source by augmenting additional information (e.g., a combined dataset of survey responses with administrative claims (6)). Such a linkage could allow one to explore questions regarding the reliability of self- or parent-report of a diagnosis of a medical condition by examining concordance between the two data sources. Moreover, through the incorporation of survey-level sociodemographic characteristics and other health conditions, it would be possible to examine what factors may lead to lower or higher rates of agreement, and whether particular subgroups may be less reliable in their reporting.

Previous agreement studies in adults have demonstrated that individual demographics (7–11), the severity of an individual's condition (12–14), including the presence of physical symptoms (15), and the location and types of services received (13, 14) can impact concordance when comparing survey responses to administrative data. Demographics, service utilization, and the presence of co-occurring conditions represent important factors to consider when evaluating concordance.

Agreement studies within a child population are limited, likely a consequence of the dependence on parent-recall rather than self-report. However, parental recall has been found to be accurate on a series of items, including a child's birthweight, health care utilization (including for asthma health services), and atopic disease and other illnesses (16–19). Additional research may be useful to better understand how parental report of conditions impact concordance, particularly asthma, a condition that is often comorbid with other developmental, physical, and behavioral conditions (20–22).

The objective of the current study is to assess the concordance between parent-reported asthma from the National Health Interview Survey with administrative data among respondents who were linkage-eligible. In addition, the association between agreement of survey and administrative data and child demographics, presence of co-occurring conditions, and service utilization characteristics will be examined.

Methods

Data source

The current study used Medicaid administrative claims data from the Centers for Medicare & Medicaid Services

(CMS) between 2000 and 2005 linked to participants of the 2001–2005 National Health Interview Survey (NHIS), utilizing Medicaid Analytic eXtract (MAX) files. These data are restricted use data and can only be accessed via the National Center for Health Statistics (NCHS) Research Data Center (www.cdc.gov/nchs/data-linkage/Medicaid-restricted.htm). The MAX files are administrative data consisting of calendar-year files that contain utilization and expenditure information for individuals enrolled in Medicaid/CHIP (6). More specifically, MAX files consist of administrative claims for Medicaid enrollees resulting from: (1) inpatient hospital services; (2) institutional long-term care records for services provided by four types of long-term care facilities: mental hospitals for the aged, inpatient psychiatric facilities for persons under age 21, intermediate care facilities for the mentally disabled, and nursing facilities; (3) filled prescription drugs, over-the-counter drugs, and other items dispensed by a free-standing pharmacy (non-hospital based); (4) physician and professional services, outpatient and clinic services, durable medical equipment, hospice, home health, and laboratory and x-ray results. Data for this analysis come from each of these four types of administrative claims.

The NHIS is a nationally representative survey of the civilian noninstitutionalized U.S. population. It is conducted continuously throughout the year by NCHS. The NHIS is an in-person interview conducted in the respondent's home. In some instances, follow-up to complete the interview is via telephone. The survey consists of (a) a family core component, with questions asked about all members of the family; (b) a sample adult component, which collects additional information from one randomly selected adult per family; and (c) a sample child component, which collects additional information about one randomly selected child per family. The sample child component is completed by a knowledgeable adult respondent, usually the parent. Data for this analysis come from the sample child and family core components of the NHIS. For more information about the NHIS, visit <https://www.cdc.gov/nchs/nhis.htm>. Response rates for the 2001–2005 Sample Child component of the NHIS ranged from 78%–81%.

Details of the linkage methodologies and the linked data files have been published elsewhere (6, 23). In short, linkage eligibility was defined as children for whom an adult respondent, acting on behalf of their child, had not refused to provide their Social Security Number (SSN) or health insurance claim numbers, had provided sufficient personal identifier information about their child, and had an SSN verified by the Social Security Administration Enumeration Verification System. Linkage eligibility for children in the 2001–2005 NHIS ranged from 42% to 62%

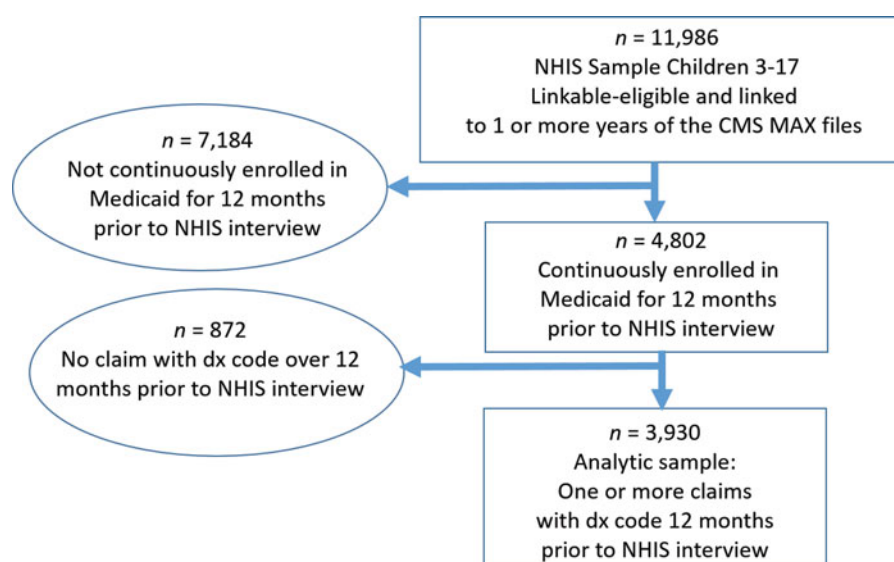


Figure 1. Flow chart of selection into the analytic sample.

and successful linkage rates of children ranged from 13% to 20%.

Sample

There were 11,986 linkage eligible children aged 3–17 who had linked NHIS-MAX observations for at least one calendar year, representing 23% of the total NHIS sample from 2001 to 2005 (see Figure 1). The age range 3–17 years was selected due to the uncertainty of an accurate asthma diagnosis among very young children (24). The analytic sample was restricted to children with Medicaid who had nonduplicative (i.e., records in only one state in a given year) continuous enrollment status for 12 months prior to the date of their NHIS interview in order to increase the likelihood that all services and prescriptions received would be documented, removing 7184 children from the analytic sample. An additional 872 children were removed from the analytic sample because they had no claims with a valid diagnostic code in the 12 months prior to the date of their NHIS interview. The final eligible analytic sample consisted of 3930 children (8% of the total NHIS sample from 2001 to 2005). The analytic sample contained children from 49 states and the District of Columbia.

Outcomes

MAX asthma diagnosis: Using the MAX files, asthma cases were defined based on an ICD-9 diagnostic code of asthma (493.xx) in any claim (e.g. outpatient, hospitalization, or long-term care) in the 12 months prior to the NHIS interview date or a prescription filled 12 months prior to the NHIS interview date using the National Committee for Quality Assurance's (NCQA) National

Drug Code (NDC) list for asthma medications (25). This derived diagnosis will herein be referred to as the claims-based asthma diagnosis, and was intended to maximize the sample yield (26, 27) thus sacrificing specificity to be inclusive of children who may have any asthma diagnosis claim, but infrequent visits to the doctor (28).

NHIS asthma diagnosis: Using the NHIS, asthma cases were defined based on affirmative answers to both “Has a doctor or other health professional ever told you that [child’s name] had asthma?” and “Does [child’s name] still have asthma?” by the sample child respondent. This diagnosis will herein be referred to as the survey-based asthma diagnosis.

Covariates

Sociodemographic characteristics: Using the NHIS, parent-reported sociodemographic characteristics examined included the child’s sex, age, race/ethnicity (non-Hispanic white, non-Hispanic black, non-Hispanic other, and Hispanic), family income recorded as a percentage of the federal poverty level (<200%, ≥200%), educational attainment of the highest educated household member (less than high school education, high school education, and greater than high school education), geographic region of residence (Northeast, Midwest, West, and South), metropolitan statistical area (MSA) status of residence (not in an MSA, small MSA, and large MSA), family structure (two-parent family versus all other family types), and respondent’s relationship to sample child (parent or other guardian).

Service utilization characteristics: Using the NHIS, parent-reported service utilization measures were examined. Respondents were asked a series of questions about

what type of providers their child had seen in the past 12 months. The providers included a generalist (pediatrician, family medicine doctor, internal medicine doctor), a specialist, a mental health professional (psychiatrist, psychologist, psychiatric nurse, or clinical social worker), or a therapist (physical therapist, speech therapist, respiratory therapist, audiologist, or occupational therapist). Separately, questions on locations the child had received services were asked. The locations of service examined included a doctor's office or clinic visit, an emergency department visit, or a home visit.

Co-occurring conditions: The NHIS include a series of questions regarding health conditions experienced by the child in the past 12 months. Along with being asked about an asthma diagnosis, respondents were also asked if "During the past 12 months, has [child's name] had any of the following conditions ... [condition name]". These conditions included hay fever, any kind of respiratory allergy, any kind of food or digestive allergy, eczema or any kind of skin allergy, frequent or repeated diarrhea or colitis, anemia, frequent or severe headaches (including migraines), three or more ear infections, and seizures. Allergies and the other health conditions included in the current analysis have been found to be commonly associated with childhood asthma (29–32).

Statistical analysis

Demographic and clinical characteristic differences between children in the analytic sample and those who were linkage eligible, but were not selected were compared utilizing bivariate logistic regressions that were weighted and accounted for the complex survey design of the NHIS (see Table 1).

Concordance measures among children in the analytic sample, which were also weighted and incorporated the complex survey design of the NHIS, included overall agreement between the survey-based asthma diagnosis and the claims-based asthma diagnosis. A prevalence-adjusted, bias-adjusted kappa (PABAK) (33), was presented alongside Cohen's kappa, as Cohen's kappa may be skewed when the prevalence of the condition of interest is small within the population. Cohen's kappa may also not fully account for biases that are introduced when differing methodologies are used for ascertainment, such as when comparing survey-based methods and claims-based methods (34). Differences between subgroups of children on percent agreement were calculated using logistic regressions and differences in the proportion of survey-based and claims-based asthma diagnoses among subgroups of children were calculated using McNemar's test via adjusted χ^2 tests.

In order to properly account for correlations that may exist between the factors of co-occurring conditions,

service utilization, and demographics, structural equation modeling (SEM) was employed. SEM has some additional capabilities compared to other types of modeling (e.g., linear regressions), namely the ability to concurrently analyze both the association between multiple dependent factors while also exploring the loadings of observed items on these dependent factors, all while simultaneously accounting for the possibility of measurement error within items (35).

A three-factor model of the covariates of demographics (child's sex, age, race/ethnicity, family income recorded as a percentage of the federal poverty level, educational attainment of the highest educated household member, geographic region of residence, MSA status of residence, family structure, and respondent's relationship to sample child), co-occurring conditions (hay fever, any kind of respiratory allergy, any kind of food or digestive allergy, eczema or any kind of skin allergy, frequent or repeated diarrhea or colitis, anemia, frequent or severe headaches (including migraines), three or more ear infections, and seizures), and service utilization (saw generalist, saw specialist, saw mental health professional, saw therapist, office visit, emergency department visit, and home visit in the past 12 months) were modeled directly onto agreement and allowed to correlate with each other. All covariates were modeled as binary, with a series of dummy variables for demographic variables. All SEM modeling was conducted within STATA 14.0 (36), utilizing a maximum likelihood estimator which could account for missingness on a given covariate, and were weighted and incorporated the complex survey design of the NHIS using subpopulation estimation for the analytic sample.

A sensitivity analysis utilized a weighted least squares estimator (asymptotic distribution free) for the aforementioned model to determine how missingness on various covariates may impact the SEM, and a secondary model with demographics modeled onto conditions and services, instead of directly onto agreement was also tested. For all models, a sub-analysis was conducted for any significant coefficients to further explore individual items. Finally, a secondary sensitivity analysis was conducted to assess the effect of excluding children with conditions known to use asthma medications for treatment (cystic fibrosis, bronchopulmonary dysplasia, and trachelomalacia) (37).

Results

Sample differences

Table 1 presents demographic characteristics, service utilization, and co-occurring conditions as measured through data from the NHIS of linkage eligible children stratified by inclusion ($n = 3930$) or exclusion ($n = 8056$)

Table 1. Demographic characteristics, service utilization, and health conditions among linkage eligible Medicaid children, by inclusion in the analytic sample, National Health Interview Survey 2001–2005.

	Total linkage eligible sample (<i>n</i> = 11,986) Weighted % (SE)	Analytic sample (<i>n</i> = 3930) Weighted % (SE)	Remaining sample (<i>n</i> = 8056) Weighted % (SE)
Sex			
Boys	50.2 (0.5)	49.6 (1.0)	50.4 (0.6)
Girls	49.8 (0.5)	50.4 (1.0)	49.6 (0.6)
Age group			
3–10	59.4 (0.6)	60.8 (1.0)	58.7 (0.7)
11–17	40.6 (0.6)	39.2 (1.0)	41.3 (0.7)
Race/Ethnicity			
Non-Hispanic white	49.9 (2.7)	45.7*** (3.1)	52.1 (2.7)
Non-Hispanic black	24.3 (1.5)	28.9*** (1.9)	22.0 (1.5)
Non-Hispanic other	4.1 (0.4)	4.1 (0.7)	4.1 (0.4)
Hispanic	21.7 (2.8)	21.4 (3.4)	21.8 (2.6)
Federal Poverty Level			
≥200%	30.2 (0.9)	15.4*** (0.9)	37.2 (1.1)
<200%	69.8 (0.9)	84.6*** (0.9)	62.8 (1.1)
Highest education level in household			
Less than high school	19.0 (1.2)	23.8*** (1.6)	16.6 (1.1)
High school	33.9 (0.7)	35.9* (1.1)	32.9 (0.8)
More than high school	47.2 (1.0)	40.3*** (1.5)	50.5 (1.1)
Household structure			
Two parent	46.5 (0.9)	38.7*** (1.1)	49.5 (0.9)
Other or unknown	53.5 (0.9)	61.3*** (1.1)	50.5 (0.9)
Region			
Northeast	15.5 (4.9)	16.1 (5.7)	15.2 (4.6)
Midwest	23.8 (3.7)	24.0 (3.8)	23.8 (3.9)
South	42.6 (4.4)	40.5 (4.4)	43.6 (4.6)
West	18.1 (4.3)	19.5 (5.0)	17.4 (4.0)
MSA			
Large urban	33.5 (2.0)	36.7** (2.9)	31.9 (1.8)
Small urban	40.2 (1.7)	35.0*** (2.0)	42.7 (1.8)
Not in MSA	26.4 (2.6)	28.3 (3.1)	25.4 (2.6)
Respondent			
Parent	90.1 (0.4)	87.1*** (0.7)	91.6 (0.4)
Other adult	10.9 (0.4)	12.9*** (0.7)	8.4 (0.4)
Providers seen in past year			
Generalist	78.8 (0.7)	84.8*** (1.0)	75.9 (0.9)
Specialist	12.5 (0.5)	14.8*** (0.8)	11.3 (0.5)
Mental health professional	10.0 (0.4)	13.6*** (0.8)	8.2 (0.4)
Therapist	7.8 (0.4)	10.9*** (0.7)	6.2 (0.3)
Any provider	81.6 (0.7)	87.6*** (0.9)	78.7 (0.8)
Location received care in past year			
Office	87.1 (0.7)	93.3*** (0.6)	84.1 (0.8)
Emergency room	26.3 (0.6)	31.6*** (1.1)	23.7 (0.7)
Home	0.6 (0.1)	1.1*** (0.2)	4.1 (0.1)
Any location	88.5 (0.6)	94.3*** (0.6)	85.7 (0.7)
Allergies			
Hay fever or seasonal allergy	10.8 (0.5)	11.6 (0.7)	10.4 (0.5)
Respiratory allergy	13.7 (0.6)	15.1* (0.9)	13.0 (0.6)
Digestive or food allergy	3.9 (0.2)	4.0 (0.3)	3.8 (0.3)
Eczema or skin allergy	10.5 (0.4)	12.0** (0.7)	9.8 (0.5)
Any allergy	27.0 (0.7)	29.5** (1.1)	25.8 (0.8)
Other health conditions			
Diarrhea or colitis	1.7 (0.2)	2.2** (0.3)	1.4 (0.2)
Anemia	1.8 (0.2)	2.1 (0.3)	1.6 (0.2)
Headaches or migraines	9.0 (0.4)	10.6*** (0.7)	8.2 (0.5)
Frequent ear infections	6.6 (0.2)	8.2*** (0.5)	5.9 (0.3)
Seizures	1.2 (0.1)	2.1*** (0.3)	0.8 (0.1)
Any other health condition	17.3 (0.5)	20.3*** (0.8)	15.8 (0.6)

Notes: SE, standard error; MSA, metropolitan statistical area.

Differences between subgroups were tested with bivariate logistic regressions.

* $p < .05$ ** $p < .01$ *** $p < .001$.

to the analytic sample and the total linkage eligible sample overall ($n = 11,986$). Children included in the analytic sample were 49.6% boys and 60.8% were aged 3–10 years. About 46% of the children were non-Hispanic white, 29% were non-Hispanic black, 21% were Hispanic,

and 4% were non-Hispanic other. Approximately two in five children lived in two-parent families (38.7%), and 84.6% of children lived in households with family incomes <200% of the federal poverty level. All children lived in the United States with the greatest portion living in

the South, (40.5%), followed by the Midwest (24.0%), West (19.5%), and Northeast (16.1%). Most children had an office visit in the past 12 months (93.3%), or saw a provider in the past 12 months (94.3%). Approximately 3 in 10 children had some form of allergy (29.5%), while approximately 1 in 5 children had some other health condition (20.3%).

Children in the analytic sample were more likely to be non-Hispanic black, living in households below the 200% federal poverty level, without two parents, and in a large urban metropolitan statistical area when compared to children who were linkage-eligible, but not in the analytic sample. Children in the analytic sample were more likely to have a parent as the respondent for the NHIS interview than children not in the analytic sample. Additionally, children in the analytic sample were also more likely to have seen any provider, received care at any location, and been diagnosed with either an allergy or other health condition in the past 12 months.

Concordance between asthma case definitions

For approximately 90% of children in the analytic sample (89.9%), the diagnostic status for asthma matched when comparing the survey-based definition to the claims-based definition (see Table 2). The prevalence-adjusted bias-adjusted kappa was 0.80 while Cohen's kappa was 0.55. The prevalence of survey-based asthma diagnosis was 14.6% which was significantly higher than the claims-based asthma diagnosis of 10.6% ($p < .001$), a finding true of most comparisons across subpopulations. Agreement was higher for girls than boys ($p < .05$), while children who had received care in the past year had lower agreement than children who had not received care in the past year ($p < .01$), and children with allergies and other health conditions in the past 12 months had lower agreement than children without these conditions ($p < .001$).

Structural equation modeling

A three-factor model (see Figure 2) was constructed to evaluate the relationship among a child's demographics, service utilization, and co-occurring conditions and the concordance (modeled as agreement) of a child having a claims-based asthma diagnosis and a survey-based asthma diagnosis.

After controlling for demographics and services, co-occurring conditions had a significant negative effect on agreement ($\beta = -0.25$, $p < .001$). However, after controlling for demographics and co-occurring conditions, service utilization had no significant effect on agreement nor did demographics after controlling for co-occurring conditions and services. There was a significant positive correlation found between the factors of service utilization

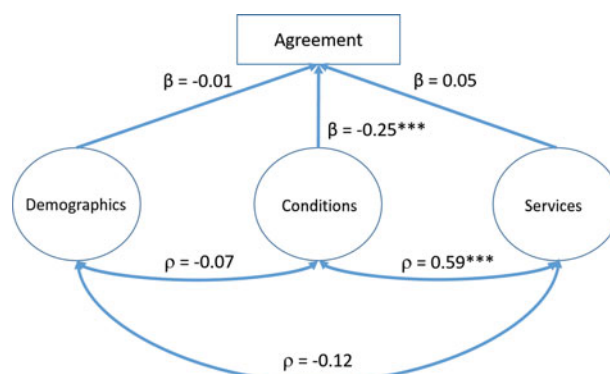


Figure 2. Structural equation modeling of agreement by demographics, health conditions and services. Notes: The following variables were used in creating the latent variables. *Demographics*: child's sex, age, race/ethnicity, family income recorded as a percentage of the federal poverty level, educational attainment of the highest educated household member, geographic region of residence, metropolitan statistical area (MSA) status of residence, family structure, respondent's relationship to sample child. *Conditions*: hay fever, any kind of respiratory allergy, any kind of food or digestive allergy, eczema or any kind of skin allergy, frequent or repeated diarrhea or colitis, anemia, frequent or severe headaches, including migraines, three or more ear infections, seizures. *Services*: In the past 12 months, saw generalist, specialist, mental health professional, and therapist, had an office visit, emergency department visit, home visit. *** $p < .001$. **Source:** 2000–2005 CMS Medicaid Analytic eXtract (MAX) data, linked to 2001–2005 National Health Interview Survey data.

and co-occurring conditions ($\rho = 0.59$, $p < .001$), but there was not a significant correlation found between service utilization and demographics, nor co-occurring conditions and demographics.

A sensitivity analysis was performed to see what impact using a different estimator (weighted least square), which did not have the ability to account for missingness on a given covariate, would have on the main findings. This analysis revealed comparable results, but with less power. A secondary model of demographics modeled directly onto co-occurring conditions and services produced nearly identical coefficients—a likely product of the nonsignificant direct pathway between agreement and demographics [see Online Supplemental Figure].

Finally, the secondary sensitivity analysis conducted assessed the effect of excluding children with conditions known to use asthma medications for treatment (cystic fibrosis, bronchopulmonary dysplasia, and trachelomalacia ($n = 6$)). The main results did not change given this degree of misclassification.

Co-occurring conditions subanalysis

The significant correlation between co-occurring conditions and agreement provided justification for a secondary analysis to further explore the impact of

Table 2. Weighted measures of concordance between survey-based asthma cases and claims-based asthma cases.

	Percent agreement (%)	Kappa	Prevalence adjusted, bias-adjusted kappa	Survey-based prevalence (%)	Claims-based prevalence (%)	Difference between prevalence estimates
Overall	89.9	0.55	0.80	14.6	10.6	4.0***
<i>Race/Ethnicity</i>						
Non-Hispanic White	89.7	0.52	0.79	14.2	10.1	4.1***
Non-Hispanic Black	90.0	0.59	0.80	17.6	10.7	6.9***
Non-Hispanic other (reference)	95.4	0.74	0.91	10.7	8.8	1.9
Hispanic	89.2	0.49	0.78	12.3	11.8	1.5
<i>Sex</i>						
Boys (reference)	88.6	0.53	0.77	16.4	11.9	4.5***
Girls	91.3 ^a	0.55	0.83	12.9	9.3	3.6**
<i>Age</i>						
3–10 years (reference)	89.4	0.54	0.79	14.8	11.7	3.1***
11–17 years	89.8	0.56	0.82	14.5	8.9	5.6***
<i>Saw Selected Provider</i>						
Seen in past 12 months (reference)	89.4	0.55	0.79	15.7	11.4	4.3***
Not seen in past 12 months	93.5 ^a	0.42	0.87	7.4	4.5	2.9**
<i>Received Care at Selected Location</i>						
Care received in past year (reference)	89.5	0.54	0.79	15.2	11.0	4.2***
No care received in past year	96.3 ^b	0.42	0.93	4.2	2.6	1.6
<i>Allergies</i>						
Present in past 12 months (reference)	84.2	0.59	0.68	29.7	21.3	8.4***
Not present in past 12 months	92.6 ^c	0.44	0.85	8.2	6.1	2.1**
<i>Other Health Conditions</i>						
Present in past 12 months (reference)	84.8	0.55	0.70	25.4	17.0	8.4***
Not present in past 12 months	91.3 ^c	0.54	0.83	11.8	9.0	2.8**
<i>Child Respondent</i>						
Parent	90.0	0.55	0.80	14.7	11.0	3.7***
Other adult	89.6	0.48	0.79	14.2	8.1	6.2***

Notes: Selected providers include: generalist (pediatrician, family medicine doctor, or internal medicine doctor), specialist, mental health professional (psychiatrist, psychologist, psychiatric nurse, or clinical social worker), therapist (physical therapist, speech therapist, respiratory therapist, audiologist, or occupational therapist); selected locations include: doctor's office or clinics, hospital emergency room, home; selected allergies include: hay fever, any kind of respiratory allergy, any kind of food or digestive allergy, eczema, or skin allergy; selected other health conditions include frequent diarrhea or colitis, anemia, three or more ear infections, seizures, frequent headaches, or migraines.

^aDifference in agreement between subpopulations is significant at $p < .05$.

^bDifference in agreement between subpopulations is significant at $p < .01$.

^cDifference in agreement between subpopulations is significant at $p < .001$.

* $p < .05$ ** $p < .01$ *** $p < .001$.

Source: 2000–2005 CMS Medicaid Analytic eXtract (MAX) data, linked to 2001–2005 National Health Interview Survey data.

selected conditions on the concordance between asthma definitions (see Figure 3). An additional structural equation model with two factors was developed which included: allergies (hay fever or seasonal allergies, respiratory allergies, food or digestive allergies, and eczema or skin allergies), and other health conditions: (diarrhea or colitis, anemia, frequent or severe headaches (including migraines), frequent ear infections, and seizures). After controlling for other health conditions, allergies had a significant negative effect on agreement ($\beta = -0.23$, $p < .001$), but other health conditions did not have an effect on agreement after controlling for allergies. There was a significant positive correlation found between allergies and other health conditions ($\rho = 0.60$, $p < .001$).

Discussion

There was a high level of agreement (approximately 90%) found between a claims-based asthma diagnoses definition and a survey-based asthma diagnosis definition within the Medicaid Analytic eXtract data from the

Centers for Medicare & Medicaid Services linked with the National Health Interview Survey. Analysis of a series of factors examined concurrently, which included child demographics, service utilization, and co-occurring conditions, revealed that demographic characteristics did not affect the concordance between asthma diagnoses, nor did the frequency and type of services received. Instead, the presence of a series of co-occurring conditions, namely allergies, resulted in significantly lower agreement after controlling for the other two factors. It is possible that higher prevalence of asthma typically seen among children who also have allergies (38, 39) may result in poorer recall when a child has received multiple diagnoses.

The prevalence of survey-based asthma was frequently significantly higher than the prevalence of claims-based asthma across a series of subpopulations. This could be a reflection of a systematic bias based on the different definitions of asthma in the NHIS and MAX. It is possible that there were missed cases and underreporting within the MAX file, where doctors failed to include asthma as one of the diagnostic codes if the condition was not

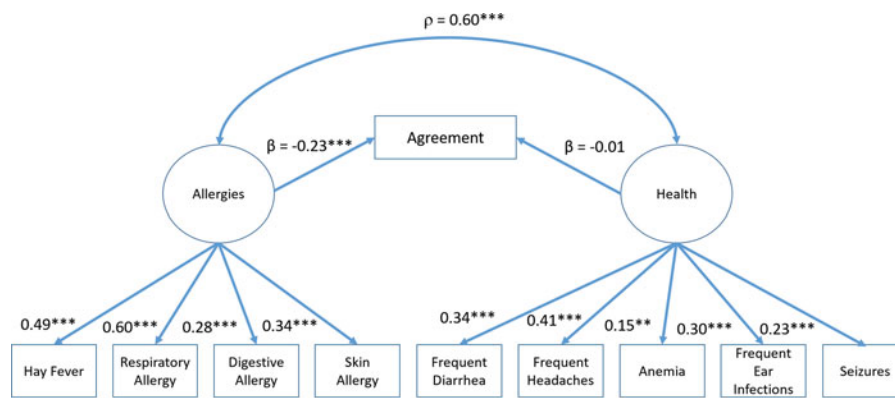


Figure 3. Structural equation modeling of agreement by physical health conditions and allergies. ** $p < .01$ *** $p < .001$. **Source:** 2000–2005 CMS Medicaid Analytic eXtract (MAX) data, linked to 2001–2005 National Health Interview Survey data.

the primary need of the visit itself, or it did not impact the length of stay or treatment for the visit. As a result, the asthma cases seen within the MAX file may be more severe than those cases captured within the NHIS, as children with more severe asthma are more likely to have been seen by a doctor in the past year (26). Additionally, the MAX file may not capture children with milder cases who have not needed recent treatment, have not filled asthma medications, or those who have received treatment outside of Medicaid. Previous work has identified that it is possible that Medicaid recipients receive asthma services exclusively through school or another health insurance plan (40–42).

The NHIS-MAX linkage used in this study has the advantage of combining self-reported health status and socio-demographic information from a survey with claims from the Medicaid program, resulting in unique population-based information that can be used for an array of epidemiological and health services research. This linkage is also unique as it is not regionally limited and represents a wide range of children that received Medicaid. Further, this linkage provides the ability to compare two sources of information for an individual while taking into account a range of health conditions and demographic differences.

Limitations

Although the analytic sample encompasses a nationwide cohort of children within Medicaid with continuous enrollment over a 12-month period, this study inclusion restriction impacted the generalizability of the sample given typical breaks in coverage seen among children receiving Medicaid (43, 44). It should be noted that children who receive medical services through Medicaid who have been diagnosed with asthma are also not representative of the greater population of children

with asthma. In fact, recent research has found significant differences in asthma management between children enrolled in a public or private health insurance plan, with children with public insurance being more likely to use the hospital emergency department and to discontinue preventive medications (45). Nonetheless, children with Medicaid or some other public insurance account for almost half of children in the United States diagnosed with asthma (2). Finally, it is important to note that reporting standards, cost-sharing practices, and managed care plans may differ from state to state, particularly over a 5-year period, resulting in discrepant proportions of cases represented in the final analytic sample. However, it is worth noting that 49 states and the District of Columbia contributed cases to the sample.

The definitions of both claims-based asthma and survey-based asthma in this study have limitations. While claims are generally an accurate reflection of diagnoses, it is possible that a diagnosis of interest may be missed either because the associated claim occurs beyond the scope of the study period (e.g., the child did not receive services for their asthma in a given year) (28) or the claim may never appear in the records because the child received services through a non-Medicaid provider (46). Furthermore, MAX prescription claims are based on prescriptions being filled, and it is known that children with Medicaid are less likely to fill asthma prescriptions compared with children that do not have Medicaid (47).

Finally, for claims-based cases there was the potential for children to have asthma medication claims for treatment of non-asthma conditions. Results from a sensitivity analysis suggest that the main results are unchanged given this degree of misclassification. Survey-based asthma cases are also subject to misclassification due to the reliance on parent report that are subject to recall bias and social desirability.

Conclusions

Results from this study demonstrate general agreement between asthma diagnoses reported in survey and Medicaid claims data among a sample of linkage-eligible children in the NHIS. Potential limitations exist in using either method to describe the prevalence of childhood asthma. Children with co-occurring conditions had greater discordance, although neither the child's demographics nor their service utilization was correlated with discordance. Findings from this study demonstrate the feasibility of evaluating agreement of a condition while incorporating socio-demographic characteristics and co-occurring conditions in Medicaid covered children. Furthermore, it highlights the practical value of data linkage for researchers interested in examining the reliability and factors impacting the reliability of parent-report on a survey of children's health.

Acknowledgments

The authors would like to thank Lara Akinbami, Lisa Mirel, and Cordell Golden for their guidance and expert knowledge as it relates to childhood asthma and data linkage.

Conflicts of interest and source of funding

The authors have no financial relationships relevant to this article to disclose. The authors have no conflicts of interest to disclose.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Supplementary material

Supplemental data for this article can be accessed on the publisher's website.

References

1. National Asthma Education and Prevention Program. Expert Panel Report 3 (EPR-3): Guidelines for the diagnosis and management of asthma-summary report 2007. *J Allergy Clin Immunol*. 2011;120(S5):S94-S138.
2. Bloom B, Simpson JL. Tables of Summary Health Statistics for U.S. Children: 2015 National Health Interview Survey. [Cited 2016 Dec 7] Available from: <http://www.cdc.gov/nchs/nhis/SHS/tables.htm>.
3. Akinbami LJ, Moorman JE, Bailey C, Zahran HS, King M, Johnson CA, et al. Trends in asthma prevalence, health care use, and mortality in the United States, 2001–2010. *NCHS Data Brief*. 2012;94:1–8.
4. Roos LL, Nicol JP, Cageorge SM. Using administrative data for longitudinal research: comparisons with primary data collection. *J Chronic Dis*. 1987;40(1):41–9. doi:10.1016/0021-9681(87)90095-6.
5. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol*. 2004;57(2):131–41. doi:10.1016/S0895-4356(03)00246-4.
6. Golden C, Discroll AK, Simon AE, Judson DH, Miller EA, Parker JD. Linkage of NCHS population health surveys to administrative records from Social Security Administration and centers for Medicare Medicaid services. *Vital Health Stat 1*. 2015;58:1–53.
7. Hunger M, Schwarzkopf L, Heier M, Peters A, Holle R. Official statistics and claims data records indicate non-response and recall bias within survey-based estimates of health care utilization in the older population. *BMC Health Serv Res*. 2013;13(1):1–11. doi:10.1186/1472-6963-13-1.
8. Wallihan, DB, Stump TE, Callahan CM. Accuracy of self-reported health services use and patterns of care among urban older adults. *Med Care*. 1999;37(7):662–70. doi:10.1097/00005650-199907000-00006.
9. Marshall SF, Deapen D, Allen M, Anton-Culver H, Bernstein L, Horn-Ross PL, et al. Validating California teachers study self-reports of recent hospitalization: comparison with California hospital discharge data. *Am J Epidemiol*. 2003;158(10):1012–20. doi:10.1093/aje/kwg256.
10. Merkin SS, Cavanaugh K, Longenecker JC, Fink NE, Levey AS, Powe NR. Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients. *J Clin Epidemiol*. 2007;60(6):634–42. doi:10.1016/j.jclinepi.2006.09.003.
11. Norrish A, North D, Kirkman P, Jackson R. Validity of self-reported hospital admission in a prospective study. *Am J Epidemiol*. 1994;140(10):938–42. doi:10.1093/oxfordjournals.aje.a117182.
12. Muggah E, Graves E, Bennett C, Manuel DG. Ascertainment of chronic diseases using population health data: a comparison of health administrative data and patient self-report. *BMC Public Health*. 2013;13(16):1–8.
13. Short ME, Goetzel RZ, Pei X, Tabrizi MJ, Ozminkowski RJ, Gibson TB, Dejoy DM, et al. How accurate are self-reports? An analysis of self-reported healthcare utilization and absence when compared to administrative data. *J Occup Environ Med*. 2009;51(7):786–96. doi:10.1097/JOM.0b013e3181a86671.
14. Tisnado DM, Adams JL, Liu H, Damberg CL, Chen WP, Hu FA, et al. What is the concordance between the medical record and patient self-report as data sources for ambulatory care? *Med Care*. 2006;44(2):132–40. doi:10.1097/01.mlr.0000196952.15921.bf.
15. Cornish R, Henderson J, Boyd AW, Granell R, Van Staa T, Macleod J. Validating childhood asthma in an epidemiological study using linked electronic patient records. *BMJ Open*. 2014;4(4):e005345. doi:10.1136/bmjopen-2014-005345.
16. Adegbeye ARA, Heitmann BL. Accuracy and correlates of maternal recall of birthweight and gestational age. *BJOG*. 2008;111(7):886–93. doi:10.1111/j.1471-0528.2008.01717.x.
17. D'Souza-Vazirani D, Minkovitz CS, Strobino DM. Validity of maternal report of acute health care use for

- children younger than 3 years. *Arch Pediatr Adolesc Med*. 2005;159(2):167–72. doi:10.1001/archpedi.159.2.167.
18. O'Sullivan JJ, Pearce MS, Parker L. Parental recall of birth weight: how accurate is it? *Arch Dis Child*. 2000;82(3):202–3. doi:10.1136/adc.82.3.202.
19. Ungar WJ, Davidson-Grimwood SR, Cousins M. Parents were accurate proxy reporters of urgent pediatric asthma health services—a retrospective agreement analysis. *J Clin Epidemiol*. 2007;60(11):1176–83. doi:10.1016/j.jclinepi.2007.02.002.
20. Blackman JA, Gurka MJ. Developmental and behavioral comorbidities of asthma in children. *J Dev Behav Pediatr*. 2007;33(1):24–31. doi:10.1097/DBP.0b013e3182396895.
21. Grupp-Phelan J, Lozano P, Fishman P. Health care utilization and cost in children with asthma and selected comorbidities. *J Asthma*. 2001;38(4):363–73. doi:10.1081/JAS-100001492.
22. Boulet LP, Boulay MÈ. Asthma-related comorbidities. *Expert Rev Respir Med*. 2011;5(3):377–93. doi:10.1586/ers.11.34.
23. Simon AE, Driscoll AK, Golden C, Tandon R, Duran CR, Miller EA, Parker JD. Documentation and analytic guidelines for NCHS surveys linked to Medicaid Analytic eXtract (MAX) files. [Cited 2016 Oct 12] Available from: http://www.cdc.gov/nchs/data/datalinkage/documentation_and_analytic_guidelines_nchs_survey_max_linked_data.pdf
24. Martinez FD, Wright AL, Taussig LM, Holberg CJ, Halonen M, Morgan WJ. Asthma and wheezing in the first six years of life. *N Engl J Med*. 1995;332(3):133–8. doi:10.1056/NEJM199501193320301.
25. National Committee for Quality Assurance (NCQA). (2015). HEDIS 2015 NDC lists Asthma Medication Ration (AMR) Table AMR-A: asthma controller and reliever medications. [Cited 2016 Oct 6] Available from: <http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2015/hedis-2015-ndc-license/hedis-2015-final-ndc-lists>.
26. Huzel L, Roos LL, Anthonisen NR, Manfreda J. Diagnosing asthma: the fit between survey and administrative database. *Can Respir J*. 2002;9(6):407–12. doi:10.1155/2002/921497.
27. Lieu TA, Lozano P, Finkelstein JA, Chi FW, Jensvold NG, Capra AM, et al. Racial/ethnic variation in asthma status and management practices among children in managed Medicaid. *Pediatrics*. 2002;109(5):857–65. doi:10.1542/peds.109.5.857.
28. Maziak W, von Mutius E, Keil U, Hirsch T, Leupold W, Rzehak P, et al. Predictors of health care utilization of children with asthma in the community. *Pediatr Allergy Immunol*. 2004;15(2):166–71. doi:10.1046/j.1399-3038.2003.00105.x.
29. Austin JK, Smith MS, Risinger MW, McNelis AM. Childhood epilepsy and asthma: comparison of quality of life. *Epilepsia*. 1994;35(3):608–15. doi:10.1111/j.1528-1157.1994.tb02481.x.
30. Leynaert B, Neukirch F, Demoly P, Bousquet J. Epidemiologic evidence for asthma and rhinitis comorbidity. *J Allergy Clin Immunol*. 2000;106(5S):S201–5. doi:10.1067/mai.2000.110151.
31. Mortimer MJ, Kay J, Gawkrödger DJ, Jaron A, Barker DC. The prevalence of headache and migraine in atopic children: an epidemiological study in general practice. *Headache*. 1993;33(8):427–41. doi:10.1111/j.1526-4610.1993.hed3308427.x.
32. Thomas EM. Recent trends in upper respiratory infections, ear infections and asthma among young Canadian children. *Health Rep*. 2010;21(4):47–52.
33. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423–9. doi:10.1016/0895-4356(93)90018-V.
34. Chen G, Faris P, Hemmelgarn B, Walker RL, Quan H. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. *BMC Med Res Methodol*. 2009;9(1):5–12. doi:10.1186/1471-2288-9-5.
35. Gefen, D., Straub, D., Boudreau, M. C. (2000). Structural equation modeling and regression: guidelines for research practice. *Commun Assoc Inf Syst*. 2000;4:1–78.
36. StataCorp. Version 14. StataCorp LP. TX: , College Station; 2015.
37. Capo-Ramos DE, Duran C, Simon AE, Akinbami LJ, Schoendorf KC. Preventive asthma medication discontinuation among children enrolled in fee-for-service Medicaid. *J Asthma*. 2014;51(6):618–26. doi:10.3109/02770903.2014.895010.
38. Bröms K, Norbäck D, Eriksson M, Sundelin C, Svärdsudd, K. Prevalence and co-occurrence of parentally reported possible asthma and allergic manifestations in pre-school children. *BMC Public Health*. 2013;13(1):764–71 doi:10.1186/1471-2458-13-764.
39. Garcia-Aymerich J, Benet M, Saeys Y, Pinart M, Basagana X, Smit HA, et al. Phenotyping asthma, rhinitis and eczema in MeDALL population-based birth cohorts: an allergic comorbidity cluster. *Allergy*. 2015;70(8):973–84. doi:10.1111/all.12640.
40. Clark NM, Brown R, Joseph CL, Anderson EW, Liu M, Valerio MA. Effects of a comprehensive school-based asthma program on symptoms, parent management, grades, and absenteeism. *Chest*. 2004;125(5):1674–9. doi:10.1378/chest.125.5.1674.
41. Halterman JS, Szilagyi PG, Fisher SG, Fagnano M, Tremblay P, Conn KM, Wang H, Borrelli B. Randomized controlled trial to improve care for urban children with asthma: results of the school-based asthma therapy trial. *Arch Pediatr Adolesc Med*. 2011;165(3):262–8. doi:10.1001/archpediatrics.2011.1.
42. Tinkelman D, Schwartz A. School-based asthma disease management. *J Asthma*. 2004;41(4):455–62. doi:10.1081/JAS-120033988.
43. Simon AE, Driscoll A, Gorina Y, Parker JD, Schoendorf KC. A longitudinal view of child enrollment in Medicaid. *Pediatrics*. 2013;132:656–62. doi:10.1542/peds.2013-1544.
44. Simon AE, Schoendorf KC. Medicaid enrollment gap length and number of Medicaid enrollment periods among US children. *AJPH*. 2014;104(9):e55–61. doi:10.2105/AJPH.2014.301976.
45. Chang J, Freed GL, Prosser LA, Patel I, Erickson SR, Bagozzi RP, Balkrishnan, R. Comparisons of health care utilization outcomes in children with asthma enrolled in private insurance plans versus Medicaid. *J Pediatr Health Care*. 2014;28(1):71–79. doi:10.1016/j.pedhc.2012.11.001.

46. Bronstein JM, Santer L, Johnson V. The use of Medicaid claims as a supplementary source of information on quality of asthma care. *J Healthc Qual.* 200;22(6):13–18. doi:[10.1111/j.1945-1474.2000.tb00160.x](https://doi.org/10.1111/j.1945-1474.2000.tb00160.x).
47. Finkelstein JA, Barton MB, Donahue JG, Algatt-Bergstrom P, Markson LE, Platt R. Comparing asthma care for Medicaid and non-Medicaid children in a health maintenance organization. *Arch Pediatr Adolesc Med.* 2000;154:563–8. doi:[10.1001/archpedi.154.6.563](https://doi.org/10.1001/archpedi.154.6.563).