



HEALTH CARE AND HUMAN SERVICES POLICY, RESEARCH, AND CONSULTING—WITH REAL-WORLD PERSPECTIVE.

Study of Health Outcomes in Children with Autism and their Families

Task A: Chart Study

Final Report

Prepared for: National Institute of Mental Health

Submitted by: The Lewin Group, Inc.

1 March 2012

Study of Health Outcomes in Children with Autism and Their Families

Task A: Chart Study

Final Report

Prepared for:

National Institute of Mental Health

Submitted by:

The Lewin Group, Inc.

1 March 2012

Executive Summary

Introduction and Study Objectives

The National Institute of Mental Health (NIMH) contracted with The Lewin Group to conduct a two-year study from September 2010 to September 2012 entitled “The Study of Health Outcomes in Children with Autism and their Families.” This study seeks to address a significant gap in the empirical knowledge base about the trajectories of health outcomes and utilization of health care services among children with autism spectrum disorders (ASD), their siblings, and their parents.

Task A has two main subtasks:

- A baseline claims analysis to identify and describe children with ASD and the control cohorts from the large administrative dataset; and
- A medical chart review to validate the claims-based identification of children with ASD in the study population, or the “chart study.” Its purpose is to evaluate our ability to identify individuals with ASD within the research claims databases by comparing claims-based ASD case* identification to ASD status as documented in medical charts.

The purpose of this report is to describe the results of the Task A Chart Study, the methodology which lead to those results, and the implications for future analyses in Tasks B, C, and D.

Goals of the Task A Chart Study are to:

- Acquire 400 medical charts for a sample of children with ‘Likely’ or ‘Possible’ ASD and also acquire medical charts for an ‘Enriched control’ group of children without ASD codes in their available claims but with other characteristics postulated to relate to being false negative cases.
- Abstract relevant information from medical charts and conduct a clinical review to assign a final ASD case categorization to each subject (the ‘gold standard’ for ASD case definition in our study).
- Calculate the Positive Predictive Values for the claims-based ASD case algorithms relative to the defined ‘gold standard’.

Study Design and Analytic Strategy

The primary data source was OptumInsight’s research database containing claims from the large health plan affiliated with OptumInsight. This has information on commercially insured individuals from 01 January 2001 to 31 December 2009. All sample members selected for the study were required to have a minimum of 6 months of continuous enrollment with simultaneous medical, pharmacy, and behavior health coverage between 2001 and 2009. This study involved the selection of three cohorts from the larger Task A study sample. This included the identification of Likely ASD and Possible ASD samples based on our claims-based ASD case

* The term “case” is used in this report to refer to the result of the application of the chart-based ‘gold standard’ to the information about an individual child with ASD, but not to the child him/herself.

algorithms and the identification of subjects who have no indication of ASD but are at risk of being false negatives based on the Likely and Possible definitions ('Enriched controls').

Medical charts for a random sample of the children identified in the claims data were then procured for the study. The medical charts abstraction and review process was adapted from the approach used by the Centers for Disease Control's (CDC's) Autism and Developmental Disabilities Monitoring (ADDM) Network. The abstracted information was then reviewed by an experienced clinician who classified children as Level 1 confirmed ASD case (i.e., met ADDM's criteria for ASD), Level 2 confirmed ASD case (i.e., did not meet ADDM's criteria for ASD, but had evidence such as a clinician-documented ASD diagnosis), ASD ruled-out, or unconfirmed. This classification was considered the 'gold standard' for ASD diagnosis and identification for this study. The positive predictive value of our claims-based ASD case algorithms was then calculated relative to the 'gold standard'.

Results

The key findings regarding sampling and the positive and negative predictive values include:

- To reach the goal of 400 charts across the three cohorts (175 Likely, 175 Possible, 50 Enriched) 2,400 subjects were randomly sampled from the 23,004 eligible subjects. This goal was reached, as 418 charts (180 Likely, 180 Possible, 58 Enriched) were abstracted.
- For the Level 1 ASD confirmation criteria the unweighted PPV was 60.1 in the Likely ASD cohort and 43.3 when the Likely and Possible cohorts were combined. The weighted PPV was 60.9 in the Likely ASD cohort and 45.0 when the Likely and Possible cohorts were combined.
- Using the Level 1 or Level 2 criteria, the unweighted PPV was 87.4 in the Likely ASD cohort and 74.2 when the Likely and Possible cohorts were combined. The weighted PPV was 87.3 in the Likely ASD cohort and 76.5 when the Likely and Possible cohorts were combined.
- Only 1 subject of the 60 Enriched controls met gold-standard criteria for ASD. This suggests that negative predictive value for claims-based case finding criteria will indeed be very high.

Implications and Recommendations

This study undertook what is, to our knowledge, the first validation study of claims-based ASD case algorithms in a large sample. Our study offers a method for using medical chart review as the 'gold standard' to validate claims-based ASD case algorithms. Overall, the chart study results endorse the ability to use claims data for research about ASD in children and associated health outcomes and utilization. Claims data is able to identify children who have actually been validly diagnosed with ASD, which may also prove useful for research about the etiology of ASD and the role of claims-based risk factors.

Implications for Tasks B, C, and D include:

- We will revise the Likely ASD criteria to only include children with two or more claims with an ASD diagnostic code. The presence of a risperidone prescription with one claim with an ASD diagnostic code will no longer be considered in the Likely criteria. Possible

ASD criterion will remain one claim with an ASD diagnostic code. As a result of this change, those children who were categorized as Likely ASD using the criterion of one ASD diagnostic code and a risperidone prescription (3.4% of Likely ASD cohort; 1,189 of 34,754 study subjects) will be re-categorized as Possible ASD. The total number of children identified as having ASD would be unchanged.

- We would discourage future claims-based ASD case algorithms from incorporating risperidone unless further work is done on the ability of risperidone use to identify an ASD diagnosis.
- The PPV increases from 74.2% to 87.4% when the Possible ASD cohort is not included in the case definition. Consequently, in the remaining analytic Tasks for this project we will primarily use the Likely ASD cohort for analyses. The Possible ASD cohort will be held in reserve and may be used to supplement subgroup analyses that have small sample sizes.

In summary, we will primarily use the two-claim (Likely ASD) claims-based case algorithm based on the presence of ICD-9 codes to identify children with ASD and their family members in Tasks B, C, and D. This is because of the differences in the PPV of the two algorithms as well as the differences in demographic and clinical conditions reported in Task A Claims Study Report delivered to NIMH on October 17, 2011.

Table of Contents

EXECUTIVE SUMMARY	i
Introduction and Study Objectives.....	i
Study Design and Analytic Strategy	i
Results.....	ii
Implications and Recommendations.....	ii
I. INTRODUCTION AND BACKGROUND.....	1
A. Overview of Study.....	1
B. Claims-based ASD Case Algorithms	2
II. STUDY OBJECTIVES	4
III. STUDY DESIGN	5
A. Claims Data Sources.....	5
B. Study Reviews.....	6
1. <i>Health Plan Approval</i>	6
2. <i>Institutional Review Board (IRB) Review</i>	6
3. <i>Health Plan Medical Director Notification</i>	7
4. <i>Confidentiality</i>	7
C. Study Sample.....	7
1. <i>Subject Eligibility Criteria</i>	7
2. <i>Sampling strategy</i>	10
3. <i>Provider Selection</i>	11
4. <i>Identification and Observation Periods</i>	12
D. Medical Chart Abstraction Implementation.....	13
1. <i>Professional Chart Abstraction Firm</i>	13
2. <i>Abstraction Training</i>	14
3. <i>Provider Identification and Participation</i>	14
4. <i>Chart Screening Criteria</i>	15
5. <i>Chart Data Abstraction</i>	16
6. <i>Assignment of Final Case Categorization of ASD ('gold standard' for this study)</i>	16
7. <i>Quality Assurance</i>	17

E.	Description of Claims and Chart Variables	19
1.	<i>Chart Data Variables</i>	19
2.	<i>Variables Determined by Clinical Review of Abstracted Chart Data</i>	20
3.	<i>Claims Variables</i>	20
F.	Analytic Strategy.....	22
1.	<i>Descriptive Analyses</i>	22
2.	<i>Estimating Predictive Values</i>	22
IV.	RESULTS	24
A.	Sampling.....	24
B.	Informative Chart Findings	27
C.	Final Assessment of ASD ('gold standard')	28
D.	Estimated Predictive Values of Alternate Claims-Based ASD Case Algorithms.....	30
1.	<i>Positive Predictive Value</i>	30
2.	<i>Negative Predictive Value in Subjects with ASD-associated Conditions</i>	34
E.	Comparison of True and False Positives	34
F.	Generalizability	39
V.	DISCUSSION AND CONCLUSION	43
A.	Comparison to the Literature	43
B.	Summary of Results	43
1.	<i>Positive Predictive Value</i>	43
2.	<i>Negative Predictive Value</i>	45
3.	<i>Potential to Augment Claims-based ASD Case Algorithm</i>	45
4.	<i>Generalizability</i>	45
C.	Strengths and Limitations	46
1.	<i>Strengths</i>	46
2.	<i>Limitations</i>	47
D.	Implications and Recommendations.....	48
VI.	REFERENCES	50
	APPENDIX A: CHART ABSTRACTION FORM	A-1
	APPENDIX B: RESULTS OF PILOT STUDY	B-1
	APPENDIX C: CLAIMS-BASED VARIABLES DEFINITIONS	C-1

APPENDIX D: SAMPLE SELECTION AND ATTRITIOND-1

APPENDIX E: ADDITIONAL CHART DATA TABLES E-1

APPENDIX F: REPRESENTATIVENESS OF SOURCE POPULATIONF-1

List of Figures

Figure 1. Identification and Observation Periods.....	13
Figure 2. Sample Selection.....	26

List of Tables

Table 1. ASD-Associated Conditions.....	10
Table 2. Targeted Sampling Strata for Chart Population.....	11
Table 3. Provider Tiers.....	12
Table 4: Comparison of Final Case Categorization of ASD Inter-rater Agreement.....	19
Table 5: Positive Predictive Values calculated based on two claims-based ASD algorithms and two 'gold standard' confirmed case definitions.....	23
Table 6. Eligible Study Subjects by Cohort and Sampling Strata.....	24
Table 7. Distribution of a 6 to 1 Random Sample of Eligible Children Available for Chart Selection.....	25
Table 8. Proportion of Targeted Sample Size Achieved by Cohort and Strata.....	26
Table 9. Chart Variables by Cohort (Chart Data).....	27
Table 10. Other Conditions/Health Issues by Cohort (Chart Data).....	28
Table 11. Final Case Categorization of ASD by Cohort (Based on Chart Data).....	29
Table 12. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions in full sample*.....	31
Table 13. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions, by age group*.....	32
Table 14. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions stratified by enrollment categories*.....	33
Table 15. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions stratified by provider tier*.....	34
Table 16. Behavioral Health Comorbidities by Case Classification Status (Claims Data) - Likely ASD.....	35

Table 17. Behavioral Health Comorbidities by Case Classification Status (Claims Data) - Likely or Possible ASD36

Table 18. Post-Index Psychotropic Medications by Case Classification Status (Claims Data) - Likely ASD36

Table 19. Post-Index Psychotropic Medications by Case Classification Status (Claims Data) - Likely and Possible ASD37

Table 20. Top 30 Most Frequent Medications by Case Classification Status (Claims Data) - Likely ASD38

Table 21. Top 30 Most Frequent Medications by Case Classification Status (Claims Data) - Likely and Possible ASD39

Table 22. Subject Demographic Characteristics by Abstraction Status (Claims data) - Likely ASD40

Table 23. Subject Demographic Characteristics by Abstraction Status (Claims data) - Possible ASD.....41

Table 24. Subject Demographic Characteristics by Abstraction Status (Claims data) - Enriched Controls.....42

I. Introduction and Background

A. Overview of Study

The National Institute of Mental Health (NIMH) contracted with The Lewin Group to conduct a two-year study from September 2010 to September 2012 entitled “The Study of Health Outcomes in Children with Autism and their Families.” The Lewin Group’s study team is a collaboration of organizations reflecting expertise in the epidemiology of autism spectrum disorders (ASD), health services research, and the clinical care of children and families.

This study seeks to address a significant gap in the empirical knowledge base about the trajectories of health outcomes and utilization of health care services among children with ASD, their siblings, and their parents. The project employs large administrative health care claims databases to fulfill four distinct aims:

- Task A: To identify a large and diverse number of children with ASD and a general population comparison group, along with their families, and describe factors related to each of these populations including age and gender distribution, geographic distribution, and socioeconomic characteristics.
- Task B: To describe and compare the health trajectories of children with ASD and their families to similar families without a child with ASD.
- Task C: To describe and compare the use of health services by children with ASD and their families to similar families without a child with ASD.
- Task D: To propose an approach for using administrative data to identify potential risk factors for ASD for future research.

Task A is comprised of two main subtasks: 1) a baseline claims analyses in order to identify and describe children with ASD, their siblings and parents and their respective comparison groups from the large administrative dataset; and 2) a concurrent medical chart study to validate the claims-based ASD case† algorithms in the study population. The purpose of the chart study was to evaluate our ability to accurately identify children with ASD within the research claims databases by comparing the probability that a subject identified by our claims-based ASD case algorithms actually has ASD as confirmed by the clinical review of medical charts (‘gold standard’). The focus of this report is to present the methodology, approach, and results of the Task A medical chart study. The methodology and results of the baseline claims analyses are detailed in a companion report that was delivered to NIMH October 17, 2011. Confirmation of the claims-based ASD case algorithm will reinforce the validity of the findings of the overall study and the other tasks, B, C, and D.

While much research is underway to examine the prevalence and consequences of ASD, to identify the risk factors and potential causes of ASD, and to explore potential treatments, fewer efforts have been directed toward understanding the overall health status and health care use of a

† The term “case” is used in this report to refer to the result of the application of the chart-based ‘gold standard’ to the information about an individual child with ASD, but not to the child him/herself.

very large group of children with ASD and their families.† To date, few studies have used large administrative claims databases to examine health outcomes in children with ASD, especially over an extended period of time.§ In addition, as most studies are clinical studies with small sample sizes that are not representative of the US population of children or children with ASD generally, a larger, more representative study drawn from large existing electronic datasets, without the additional burden to individuals, families, clinicians or researchers of prospective data collection can help advance the research for children with ASD and their families. Finally, longitudinal data for family members of children with ASD will inform research on how ASD impacts whole families in addition to the effects on the individual with ASD over time.

B. Claims-based ASD Case Algorithms

Claims-based studies often use an algorithm to define a diagnosis or case identification. Such algorithms involve criteria which must be met before a subject is flagged as having a particular diagnosis. These criteria typically involve aspects of the medical claim on which a diagnosis is present (e.g., inpatient vs. outpatient), the specific diagnosis and/or other associated codes (e.g., procedures) on claims, the position of the diagnosis on the claim (i.e., primary vs. secondary), the number of occurrences over time of a particular diagnostic code (e.g., two appearances on different days), and/or the occurrence of a filled prescription for the subject of a medication highly correlated to the condition of interest.

We conducted a literature review to inform our claims-based ASD algorithms. We found a common approach among other ASD reviewed studies was to define ASD cases as a child with ASD diagnostic codes in primary or secondary positions.^{1 2 3} One study, for example, based on Medicaid claims data from one large county in Pennsylvania required two separate ICD-9 primary diagnostic codes of 299.xx for ASD case identification.⁴ Another study of Medicaid claims from 42 states identified children with ASD based on the appearance of the diagnostic codes on either one inpatient claim or two outpatient claims that did not occur on the same day.⁵ A Canadian study empirically tested the performance of seven claims-based algorithms relative to a ‘gold standard’ (in this study defined as the diagnosis made by a team of trained clinicians). The seven algorithms used combinations of single or multiple ASD diagnostic codes for children using three administrative databases (i.e., hospital data, physician billing data, and outpatient mental health data). This study concluded that defining children with ASD as those children with a single appearance of an ASD diagnostic code in any of the three databases resulted in the best overall sensitivity.

We considered this literature when constructing our claims-based ASD case algorithms:

1. Likely ASD, defined as subjects with *at least* 2 medical claims with an ASD diagnostic code [Autistic Disorder (ICD-9-CM 299.0x), other specified PDD (including Asperger’s

† See the National Institute of Mental Health web page on autism: <http://www.nimh.nih.gov/health/topics/autism-spectrum-disorders-pervasive-developmental-disorders/index.shtml#>, the link there to clinical trials regarding autism, and also the research summary by the Interagency Autism Coordinating Committee at <http://iacc.hhs.gov/summary-advances/2010/>.

§ The Request for Proposal for this study, HHS-NIH-MH-2010-018 at Attachment 3 page 2 of 12 references the “significant gaps” in this area.

Disorder) or unspecified PDD (or PDD-NOS) (299.8x and 299.9x)] in any position ** during the identification period.††

2. Possible ASD, defined as subjects with only 1 medical claim with an ASD diagnostic code in any position during the identification period.

Although our claims-based ASD algorithms were informed by relevant studies in the literature, the possibility of errors remains. We recognized that subjects identified through claims-based algorithms could exhibit either Type I errors (inclusion of non-ASD cases as ASD – false positives) or Type II errors (exclusion of true ASD cases – false negatives). Such Type I or Type II errors could occur due to underlying diagnostic uncertainties (such as misdiagnosis of ASD due to lack of expertise by inexperienced or untrained clinicians); the reporting of a diagnosis in a rule-out context; a coding error; data gaps due to discontinuity in health insurance coverage or subjects having other primary insurance coverage; lack of complete data in claims; or other drivers.

Given the potential for Type I and Type II errors using claims-based ASD case algorithms, this study tests the validity of our claims-based ASD definitions by comparing the results of our claims-based algorithms to clinical diagnoses of ASD as confirmed by the clinical review of medical charts ('gold standard') for a sample of our study population. Both Type I and Type II errors are investigated. In this report we describe our detailed objectives for this study, study methods, and results, specifically the estimated predictive value associated with our two claims-based algorithms.

** Up to 4 diagnosis codes are recorded on provider claims and up to 9 diagnosis codes are recorded on facility claims. Primary position refers to the first diagnosis code listed; secondary position refers to any diagnosis position after the first diagnosis position.

†† The chart study is validating our claims-based ASD case algorithms based on ICD-9 codes only. The alternative definition of Likely ASD (one medical claim and a prescription for risperidone) included in the Task A claims study is not included in the definition of 'Likely' for this portion of the study. By excluding this criterion, we are still identifying these children with ASD in the Possible group. This allows the chart study to focus on the validity of the diagnostic code portion of the algorithms only. In addition, the proportion of children with risperidone and one claim with an ASD diagnostic code was small which would cause oversampling of this group within the chart study.

II. Study Objectives

The main objective of the Task A Chart Study was to validate our claims-based ASD case algorithms relative to ASD confirmation based on the clinical review of medical charts ('gold standard') for a portion of our study sample. Specifically, the objectives for the chart study were to:

- Acquire medical charts for a sample of children with Likely or Possible ASD (as defined above) and also acquire medical charts for an Enriched control group (defined below) of children without ASD diagnostic codes in their available claims but with other characteristics postulated to relate to being false negative cases.
- Abstract relevant information from medical charts and conduct a clinical review to assign a final ASD case categorization to each subject (the 'gold standard' for ASD case definition in our study).
- Calculate the Predictive Values for the claims-based ASD case algorithms relative to the defined 'gold standard'.
- Compare certain claims-based data elements between true and false positive ASD cases to determine if other aspects or dimensions of claims-based variables could be used to further refine the claims-based algorithms.
- Assess the generalizability of the findings from the chart study sample to the broader study population.

The remainder of this report describes the data, methods, and results of the chart study and discusses implications for the claims-based ASD case algorithms we will use for the research to be conducted in Tasks B, C, and D.

III. Study Design

This retrospective chart and claims data study used medical data, pharmacy data, enrollment information and linked abstracted chart data to assess the validity of claims-based ASD case algorithms. Data from 01 January 2001 through 31 December 2009 was used in the study. Study subjects were commercial health plan enrollees with and without diagnostic codes for ASD in available claims. This included the identification of Likely ASD and Possible ASD samples based on our claims-based ASD case algorithms and the identification of subjects who have no indication of ASD but are at risk of being false negatives (Enriched controls). Medical charts were procured for a sample of the subjects eligible for the study. Relevant clinical elements were abstracted from the charts to enable an informed clinical review that resulted in a final case categorization of ASD (the 'gold standard' for this study) to each subject. The medical charts abstraction and review was adapted from the approach used by the Centers for Disease Control's (CDC's) Autism and Developmental Disabilities Monitoring (ADDM) Network. The adaptation was done with substantive input from project team member Craig Newschaffer, PhD and project consultant Marygrace Kaiser, PhD, both former ADDM Network Site principal investigators and with the permission of the CDC ADDM project. The predictive value of our claims-based ASD case algorithms was then calculated relative to the chart-based 'gold standard'.

This section outlines the details of our study design. This includes: a) an overview of the database that was the source of the claims-based analyses and the source of sample selection for the chart review; b) the study reviews that were required for approval of the study to be in compliance with privacy and ethical policies; c) a description of the study sample, including subject eligibility criteria, sampling strategy, provider selection and identification and observation periods; d) details about the implementation of medical chart abstraction including the role of the abstraction firm, abstraction training, chart screening criteria, chart data abstraction, and quality assurance; e) a description of the claims and chart variables constructed for the study; and f) our analytical strategy.

A. Claims Data Sources

The data sources for the chart study includes both claims data and chart data. In this section we provide an overview of the claims database that was used to identify children with ASD based on our claims-based ASD case algorithms and to children that were considered as the Enriched control group based on claims. The medical charts for a random sample of the children identified in the claims data were then procured and abstracted.

OptumInsight has access to a proprietary research database ("OptumInsight Research Database") containing medical and pharmacy claims with linked enrollment information covering the period from 1993 to 2010. For 2009, for example, data relating to approximately 13.3 million individuals with both medical and pharmacy benefit coverage are available. The underlying population is geographically diverse across the US and reasonably representative of the privately insured US population.

- **Medical Claims**

Medical claims or encounter data are collected from all available health care sites (inpatient hospital, outpatient hospital, emergency room, outpatient office, surgery center, etc.) for all types of covered services, including specialty, preventive and office-

based treatments. Medical claims and coding conform to insurance industry standards. Claims for ambulatory services submitted by individual providers (e.g., physicians) use the HCFA-1500 or CMS-1500 format.⁶ Claims for facility services submitted by institutions (e.g., hospitals) use the UB-82, or UB-92, or UB-04 format.^{7 8} Medical claims include: diagnostic codes recorded with the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) diagnosis codes; procedures recorded with ICD-9-CM procedure codes, Current Procedural Terminology (CPT), or Health care Common Procedure Coding System (HCPCS) codes; site of service codes; provider specialty codes; revenue codes (for facilities); paid amounts; and other information. Typically, facility claims do not include complete information about drugs administered within a hospital. Approximately 6 months following the delivery of services is required for complete medical data due to lags in claims submissions and final claims processing. In this report, the term “medical claims” is used to refer to both claims for both physical health care and behavioral health care submitted and processed for reimbursement. Health care not processed as a medical claim (e.g., care provided as part of a wellness program or as an Employee Assistance Program - EAP) is not included.

- **Pharmacy Claims**

Claims for pharmacy services are typically submitted electronically by the pharmacy at the time prescriptions are filled. The pharmacy claims history is a profile of all outpatient prescription pharmacy services provided and covered by the health plan (including hospital discharge pharmacy fills). Pharmacy claims data include drug name, dosage form, drug strength, fill date, days of supply, financial information, and de-identified subject and prescriber codes, allowing for longitudinal tracking of medication refill patterns and changes in medications. Pharmacy claims are typically added to the research database within 6 weeks of medication dispensing.

B. Study Reviews

1. Health Plan Approval

Abstraction of individual medical charts of subjects identified directly in the OptumInsight claims database required that the study concept and any communication with providers be approved by the large US health plan affiliated with OptumInsight. Approval was received on January 10, 2011.

2. Institutional Review Board (IRB) Review

Following health plan approval, an application was submitted to the New England IRB and its affiliated privacy board, a commercial non-academically affiliated IRB, for approval of the study and the medical chart abstraction process and documents.

Initial approval from the New England IRB was given on 01 February 2011. Administrative changes and minor amendments were approved on 08 April 2011, 01 June 2011 and 22 September 2011.

This study was undertaken only after the study protocol and study documents were approved and OptumInsight was granted a Waiver of Authorization by the privacy board and a waiver of the informed consent requirement by the IRB. Upon receipt of the waivers from the IRB and privacy

board, OptumInsight provided a copy of the waiver documents and general study information to the relevant data sources for final approval to utilize those source's data in the study.

The study protocol and study documents were also submitted to Drexel University IRB for expedited review in February 2011. Approval from Drexel IRB was received in April 2011 under protocol #19676.

3. Health Plan Medical Director Notification

Following health plan, IRB, and privacy board approvals, OptumInsight sent a research study notification letter to the health plan market medical directors (MMDs) to alert them of a medical chart review study being conducted in the field. The MMDs each received an email containing a study synopsis, a copy of the network provider abstraction request letter, and a list of providers that were to be contacted for participation in the study from the MMDs' market region.

4. Confidentiality

No child's identity or medical charts were disclosed for the purposes of this study except in compliance with applicable law.

C. Study Sample

This section describes the process of identifying the study sample used to validate our claims-based ASD case algorithms. The final sample for the study included subjects who met the eligibility criteria for one of three cohorts, were sampled for inclusion in the study, and whose chart was received from the provider who was selected, contacted and agreed to participate.

1. Subject Eligibility Criteria

This study involved the selection of three cohorts from the larger Task A study sample. This included the identification of Likely ASD and Possible ASD samples based on our claims-based ASD case algorithms and the identification of subjects who had no indication of ASD via diagnostic codes in claims but are at increased risk of being false negatives due to the presence of diagnostic codes that are associated with ASD.

Eligibility criteria for these three cohorts are described below:

► Subjects Meeting Claims-Based Criteria for Likely ASD

Inclusion criteria:

- Commercial health plan enrolled individual with medical, pharmacy, and behavioral health coverage^{##}
- Evidence of ASD defined as ≥ 2 medical claims with a diagnostic code of ASD [Autistic Disorder (ICD-9-CM 299.0x), other specified PDD (including

^{##} Continuous enrollment was based on simultaneous medical, pharmacy, and behavioral health coverage. Gaps in enrollment of ≤ 32 days were bridged and included in calculation of continuous enrollment duration.

Asperger's Disorder) or unspecified PDD (or PDD-NOS) (299.8x and 299.9x)] in any position^{§§} between 01 July 2001 and 31 June 2009 (identification period)

- Age ≥ 2 years and ≤ 20 years as of year of index date (date of first medical claim for ASD after age 2)^{***}
- ≥ 6 months of pre-index date continuous enrollment (pre-index period)
- ≥ 6 months of post-index date continuous enrollment (post-index period)
- Insured under a fully insured health plan approved for studies involving personal health information

Exclusion criteria:

- Evidence of ASD (≥ 1 medical claim with an ICD-9-CM code for ASD) before the index date (if claim is at age 2 or younger). Due to the unreliability of diagnosis prior to age 2, claims for ASD prior to age 2 were not considered valid and were not accepted as part of the definition. If the child had a claim for ASD prior to age 2 and had at least one more after age 2, they would still be included.
- At least one medical claim with a diagnosis of childhood disintegrative disorder (ICD-9-CM code 299.1x) or Rett Syndrome (ICD-9-CM code 330.8x)^{###} in any position between 01 January 2001 and 31 December 2009^{##}

The Likely ASD definition for the chart study differs slightly from the Task A claims definition. The alternative definition of Likely ASD (one medical claim and a prescription for risperidone) included in Task A claims study is not included in the definition of Likely for this portion of the study. This is because the chart study is validating our claims-based ASD case algorithms based on ICD-9 codes only. With excluding this criterion, we are still identifying these children with ASD, but now they will appear in the Possible cohort (based on one diagnosis for ASD). In addition, the proportion of children with risperidone was small and we did not want to try to oversample on this characteristic.

► **Subjects Meeting Claims-based Criteria for Possible ASD**

Inclusion criteria:

- Commercial health plan enrolled individual with medical, pharmacy, and behavioral health coverage
- Only 1 medical claim with a diagnostic code for ASD in any position between 01 July 2001 and 31 June 2009; and

^{§§} Up to 4 diagnosis codes are recorded on provider claims and up to 9 diagnosis codes are recorded on facility claims. Primary position refers to the first diagnosis code listed; secondary position refers to any diagnosis position after the first diagnosis position.

^{***} Note that this index date differs from the index date identified in the baseline claims analyses for Task A, where the date of the beginning of the first CE period was identified as the index date.

^{##} While Rett syndrome and CDD are also considered types of pervasive development disorders similar to ASD, subjects with evidence of these disorders were excluded because these two disorders have different etiologies, disease progression and prognoses than Autistic Disorder, other specified PDD and unspecified PDD, subjects with evidence of these disorders were excluded because these two disorders have different etiologies, disease progression and prognoses than Autistic Disorder, other specified PDD and unspecified PDD.

- Age ≥ 2 years and ≤ 20 years as of year of index date (date of first medical claim for ASD after age 2)
- ≥ 6 months of pre-index continuous enrollment (pre-index period)
- ≥ 6 months of post-index continuous enrollment (post-index period)
- Insured under a fully insured health plan approved for studies involving personal health information

Exclusion criteria:

- Evidence of ASD (>1 medical claim with an ICD-9-CM code for ASD) before the index date (if claim is at age 2 or younger); or
- At least one medical claim with a diagnosis of childhood disintegrative disorder or Rett Syndrome in any position between 01 January 2001 and 31 December 2009

This last sample, which, for simplicity, we refer to as Enriched controls, was selected to test the expectation that negative predictive value for claims-based ASD case algorithms will be acceptably high. Subjects selected here were children without ASD codes on any available claims, but who did have codes on claims for other conditions that are associated with ASD and/or used in the rule-in, rule-out process of making an ASD diagnosis. Consequently, this sample includes subjects who would not be identified as ASD cases in purely claims-based research, but who are potentially more likely (than a random sample of subjects without ASD diagnostic codes) to include false negatives. That is, our strategy was to select a target group we believed likely to contain false negative cases and thus test and quantify the claims-based algorithm's tendency to exclude false negatives.

► **Subjects with No Claims-Based Indication of ASD at Enriched Risk of Being False Negatives (Enriched controls)**

Inclusion criteria:

- Commercial health plan enrolled individual with medical, pharmacy, and behavioral health coverage
- ≥ 1 medical claim with a diagnosis code for an ASD-associated condition (Table 1) in any position between 01 July 2001 and 31 June 2009
- Age ≥ 2 years and ≤ 20 years as of year of index date (date of first medical claim for an ASD-associated condition)
- ≥ 6 months of pre-index date continuous enrollment (pre-index period)
- ≥ 6 months of post-index date continuous enrollment (post-index period)
- Insured under a fully insured health plan approved for studies involving personal health information

Exclusion criteria:

- At least one medical claim with a diagnostic code for ASD in any position between 01 January 2001 and 31 December 2009; or
- At least one medical claim with a diagnosis of childhood disintegrative disorder or Rett Syndrome in any position between 01 January 2001 and 31 December 2009

Table 1. ASD-Associated Conditions

Code Category	ICD-9-CM CODES	Condition
Intellectual Disabilities	317.00	Mild intellectual disability
	318.00	Moderate intellectual disability
	318.10	Severe intellectual disability
	318.20	Profound intellectual disability
	319.00	Unspecified intellectual disability
Specific Conditions often Associated with ASD	759.50	Tuberous sclerosis
	759.83	Fragile X syndrome
	771.00	Congenital rubella
General Codes often Associated with ASD	348.30	Encephalopathy, not elsewhere classified
	348.80	Other conditions of the brain
	348.90	Unspecified condition of brain
	783.42	Delayed milestones
	V79.30	Screening, dev handicaps in early childhood
	V79.80	Screening, other specified mental disorders & dev handicaps
Specific Developmental Delays Associated with ASD	315.30	Dev speech or language disorder
	315.31	Expressive language disorder
	313.32	Mixed receptive-expressive language disorder
	315.40	Dev coordination disorder
	315.50	Mixed developmental disorder
	315.80	Other specified delays in development
	315.90	Unspecified delay in development

2. Sampling strategy

Given the time and labor-intensive nature of a chart study, it was not feasible to conduct a chart review for all subjects who met the inclusion criteria outlined above. Therefore, we set a target of 400 valid medical charts to be abstracted across the 3 cohorts: 1) Subjects with Likely ASD (n=175); 2) Subjects with Possible ASD (n=175); and 3) Enriched controls (n=50). Because of the relative importance of exploring positive as opposed to negative predictive value, the bulk of the sampling was devoted to the two claims-based ASD groups (further discussed in Section III.F Analytic Strategy) rather than the Enriched controls. The overall sample size of 400 was sufficient to achieve adequate levels of statistical power and also limited by resource constraints.^{sss} We recognize that this stratified sampling strategy could raise concerns about the generalizability of our results ; thus in the Results (Section IV.F) we compare the generalizability of these three cohorts to the larger sample from which they were drawn.

Targeted sampling fractions for each cohort were based on age of the child, (<8 years and ≥8 years), length of enrollment (<18 months and ≥18 months), and "richness" of claims (defined as having above the 75th percentile of total number of available claims for the entire ASD population

^{sss} In the proposal to NIMH for this study, we included detailed power calculations for alternative sample sizes. For example, with an assumption of 50% of ASD diagnosed cases in the sample and a .10 precision interval, 384 sampled cases would achieve 95% confidence.

(Likely and Possible combined). The targets for the distribution of abstracted charts are shown in Table 2.

Table 2. Targeted Sampling Strata for Chart Population

	Younger (<age 8)			Older (≥age 8)		
	Shorter enrolled (<18 mo.)	Longer enrolled (≥18 mo.)	Longer enrolled – rich claims	Shorter enrolled (<18 mo.)	Longer enrolled (≥18 mo.)	Longer enrolled – rich claims
Likely ASD	35	35	70	9	9	18
Possible ASD	35	35	70	8	9	17
Enriched Controls	10	10	20	2	3	5
Sampling Fraction	.2	.2	.4	.05	.05	.1

Our focus was on younger children (targeted to be 80% of the sample) because claims-based ASD case algorithms among younger children with ASD are expected to be more accurate since older children are likely to be, on average, further from the time of diagnosis and are also expected to have fewer medical encounters overall or related to ASD (though this will be explored empirically in other tasks of this project). Similarly, we focused on longer-enrolled subjects (75% of the sample) because a longer enrollment and thus longer monitoring is more likely to reflect a child's entire medical course since we often have incomplete information on the study subjects. Within the longer-enrolled group we paid particular attention to those subjects with "rich" claims so that we could explore whether adding claims based variables other than ASD diagnostic codes has potential for improving the validity of the claims-based ASD case algorithms – this would not be possible without a sufficient number of claims. The positive predictive value calculated as part of our results is weighted to account for this sampling strategy (discussed further in Section III.F). Enriched controls were selected so as to include roughly equal numbers from the four ASD-associated condition categories in Table 1.

An over-sample of 2,400 subjects was initially selected from the eligible sample, based on the sampling strata presented in Table 2 to provide an adequate pool to achieve a final sample of 400 abstracted charts. Oversampling was necessary to account for the multiple reasons why a chart may not be received from a provider (i.e., provider nonresponse, chart unavailable, child not seen during the date range, chart purged or destroyed, or provider requested subject approval). In addition, some charts once abstracted may not be informative. Further explanation of these reasons and the resulting final chart study sample is provided later in the report.

3. Provider Selection

Due to the logistics of identifying and procuring medical charts from many locations, only one chart could be reviewed from a single provider for each study subject limited to the duration of the enrollment period. For each sampled subject, a treating provider associated with the index claim (first ASD claim for Likely and Possible ASD or first claim for an ASD-associated condition for the Enriched controls) was identified. Whenever possible, the provider who submitted the index claim was selected (index provider). In cases where there was only one claim from the index provider or when the index provider was of the wrong tier (in order to meet the proportion of tier 1/non-tier 1 providers (See Table 3 below), another provider was selected.

Provider selection was based on the following criteria:

- **Tier of Provider.** Index providers were separated by tiers (specialties) in order to ensure that there would be a sufficient number of specialist charts included in the study (rather than generalists) considering that certain specialists were probably more likely to formally assess and diagnose ASD as well as provide written documentation. Approximately half of the charts were targeted to be obtained from a tier 1 (specialist) provider while the remaining 50% were targeted to be from tier 2 and 3 providers. With the assumption that most index providers were generalists (non-specialists), to ensure an adequate number of tier 1 providers, all providers for subjects within 6 months before and after the index date were identified and their specialty determined. When necessary, non-index providers were selected.
- **Frequency of claims.** Providers with ≥ 2 separate medical claims on different days were preferentially selected with the intention of avoiding procuring the chart from a provider who did not have an ongoing relationship with the child.

Table 3. Provider Tiers

Tier	Provider Type
1st	Developmental Pediatric Specialist
	Child Psychiatrist
	Child & Adolescent Psychiatrist
	Child Psychologist
	Child & Adolescent Psychologist
	Child Neurologist
2nd	Other Psychiatrists and Psychologists
	Speech Language Pathologist
	Pediatrician
3rd	Family/General Practice
	Internal Medicine

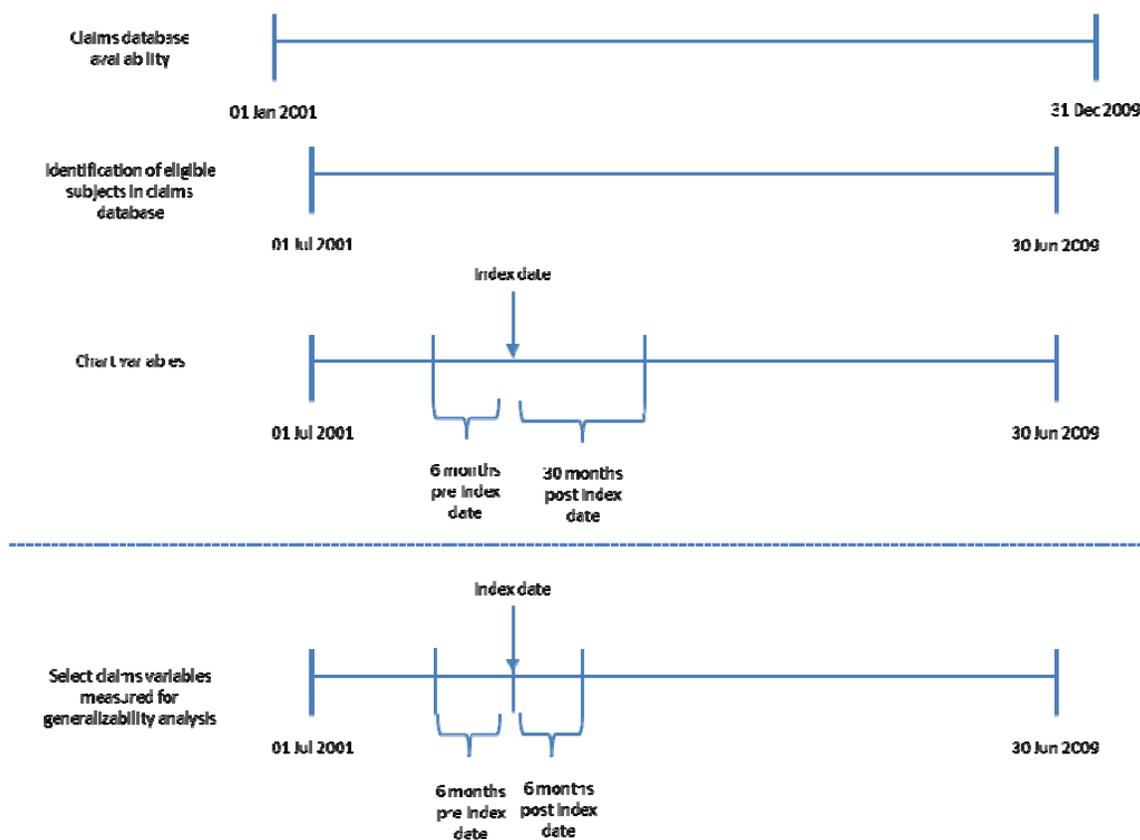
A further consideration was that our Health Plan associates requested that we exclude from the selection any providers who participated in another recent chart-based study to support the Healthcare Effectiveness Data and Information Set (HEDIS) data collection effort. Impact of this additional provision is also reported below.

4. Identification and Observation Periods

The figure below summarizes the identification and observation periods for the sampled subjects in the three cohorts (Figure 1). Our research database has data from 2001 to 2009. As mentioned above in the inclusion criteria regarding the pre- and post-index continuous enrollment of ≥ 6 months, study subjects were identified between July 2001 and June 2009. The date of the first medical claim for ASD (Likely or Possible ASD) or ASD-associated condition (Enriched controls) was set as the index date. Chart variables were assessed up to 6 months prior to and up to 30 months after the index date to allow sufficient time to accumulate and identify evidence of a diagnosis of ASD. Finally, as mentioned in the objectives above, claims-based variables were

described for the final sample and the populations from which they were drawn to assess generalizability and representativeness of the chart study sample to the eligible population. The presence of clinical conditions was assessed during the 6 month pre-index period. Health care utilization and medication use were assessed during the 6 month post-index period.

Figure 1. Identification and Observation Periods



D. Medical Chart Abstraction Implementation

Once the study sample was defined, the medical chart abstraction began. The medical chart abstraction, described in this section, follows a similar process to one described by Gearing and colleagues that outlined a methodology for conducting retrospective chart reviews in child and adolescent psychiatry.⁹ Key steps included designing a data abstraction instrument, defining guidelines for abstraction, abstracting data from charts, obtaining appropriate approvals (described above), and conducting a pilot study.

1. Professional Chart Abstraction Firm

An abstraction firm external to the Project team was selected to implement and manage all aspects of data abstraction. The chart abstraction firm was required to comply with the established practice of OptumInsight and the Privacy Rule as specified in their contract. Procedures relating to medical chart data abstraction, data handling and data transfers, ID encryption, confidentiality, data quality management and documentation were specified by the Project team.

The chart abstraction firm was responsible for the following:

- provide credentialed and/or registered abstraction staff to manage and perform abstraction
- assist in the completion and testing of medical chart abstraction guidelines
- procure medical charts from providers for abstraction
- review the medical charts received and record study data into the abstraction database
- assist in a pilot study to clarify study parameters, abstraction data elements and abstraction procedures
- utilize an electronic tracking database to track the status of procurement and abstraction at each step of the process
- conduct technical data audits of the abstraction database
- perform a content audit of the abstraction database using the inter-rater reliability quality assurance method
- comply with HIPAA confidentiality and security procedures
- organize and transfer the source documentation (i.e., hard paper copies of the medical chart) to OptumInsight
- transfer the abstraction database via secure FTP to OptumInsight per established practices

2. Abstraction Training

Teleconference training sessions were conducted with the abstraction firm. All four abstractors working on the study attended the training. The initial training included a review of the study background, objectives, timeline and deliverables, including procedures for inter-rater reliability assessments. Dr. Kaiser presented an overall description of ASD as well as discussed ASD triggers that may be noted in the charts. The specific information from the chart that was necessary to record and where to record it was also presented. Confidentiality procedures and the need for removal of PHI were reviewed. In addition, the chart abstraction form and manual (discussed below) were reviewed with detailed explanation about how abstractors should complete each study question.

A formal testing of the abstractors' findings occurred during a second training session. A sample of mock charts was supplied by Dr. Kaiser. These charts had been used to train abstractors in the ADDM Network. Each abstractor abstracted each of the mock charts and the results were reviewed by Dr. Kaiser. This enabled the abstractors to gain experience with the chart abstraction form and allowed an opportunity to bring up and discuss issues for each item on the chart abstraction form. This also allowed Dr. Kaiser to determine the quality of each abstractor. Based on this formal testing, it was determined that the quality for each abstractor met our stringent standards of rigor and reliability for this project.

3. Provider Identification and Participation

Based on the study sample selected as described in Section III.C above, the name and address of the providers identified were obtained from the OptumInsight claims database. Following the health plan medical director notification period (discussed above in Section III.B), OptumInsight

provided the names of the providers and the respective plan members selected as study subjects to the abstraction firm. Selected providers were sent a participation request cover letter approved by the health plan affiliated with OptumInsight and signed by the health plan national medical director (NMD). The letter requested the provider's support in arranging medical chart review of identified subjects, explained the study purpose and method for subject identification, indicated the study approvals received, and provided the contact information of the abstraction firm contracted to conduct the study.

The abstraction firm then contacted the providers' office directly to request access to the medical charts. The abstraction firm worked with the providers' offices to obtain copies of medical charts for the chart abstraction period specified in Section III.C.4 above (minimum of six months to a maximum of 36 months).

4. Chart Screening Criteria

Once a chart was procured from the provider, the abstractors applied screening criteria to the chart as an initial step to determine if the chart could be deemed 'valid' for our study. Although a provider may have met the above selection criteria and responded to our request to share the medical chart, it was still important to ensure that charts selected for the study met some minimal criteria. Our screening criteria are basically three-fold: 1) apply chart study eligibility criteria that was used in identifying the subject in the claims database (discussed in III.C.1 above); 2) ensure confidence that the chart is representative of the subject identified in the claims and 3) the chart contains key elements so that the chart can be potentially informative for the purposes of this study. If the chart met these criteria, the chart was abstracted.

a. Chart study eligibility & claims-chart match criteria

The chart study eligibility criteria were applied to the chart to ensure the chart also met the chart study inclusion criteria. In addition, we applied additional criteria to ensure that the chart corresponded to the subject identified in the claims.

- Chart study eligibility criteria:
 - Between 2 and 20 years of age; and
 - No evidence of Rett's or CDD.
- Quality-check criteria: Date of birth and gender on the chart match the claims-based information for the subject.

b. Informative chart criteria

The motivation behind these criteria was based on our experience that some charts are much more skeletal than others and may not contain informative notes for abstraction. Secondly, even if a chart included detailed notes from encounters, it is possible that encounters are exclusively for acute care matters unrelated to developmental issues. Therefore, the chart must have had one of the following in order to be deemed 'valid'. Once the abstractor saw evidence of any one of these items, the chart was flagged for complete abstraction:

- ASD diagnosis or potential ASD. This includes any mention of an ASD condition in the chart, which could vary from a specific diagnosis of ASD (with or without the ICD-9-CM code) to just mentioning potential evidence of ASD.

- Documentation of a social, developmental, or behavioral trigger associated with ASD.
- Testing, assessment or a comprehensive evaluation. A comprehensive evaluation was defined as developmental assessments that included psychiatric/psychological evaluations, neurological evaluations, developmental assessments, and/or speech pathology evaluations. The evaluations were required to include a description or summary of the child's developmental status. A description of each necessary item was included in the chart abstraction manual.

It is recognized that these criteria had the potential to bias predictive value (increase positive predictive value and decrease negative predictive value) because this could favor selection of charts 'confirming' ASD into our sample. However, charts that were excluded for reasons of being uninformative were tracked. While they were not abstracted in full, it can be concluded that they would not have confirmed ASD and consequently can be added back into tables and statistical analyses below as charts not confirming ASD (unconfirmed).

5. Chart Data Abstraction

If the chart passed the screening criteria described above, the abstractors proceeded with abstracting the clinical elements of interest from the medical chart and entering the data into an Access database (clinical elements are listed and defined in Section III.E.1). Two primary documents, a medical chart abstraction form and an abstraction manual, were developed to guide the abstractors in this process. The medical chart abstraction form is included in Appendix A. The form specified the time period in which data should be abstracted and allowed for multiple data points to be abstracted for relevant elements. An electronic data collection database (Access), modeled after the form, was developed for electronic entry of abstracted data. The abstraction manual described the source, location, and type of data to examine for each data element in the chart abstraction form. In addition, the manual described aspects of data collection important to understanding how to populate each field within the abstraction database.

As chart abstractions were completed, the abstraction firm entered the data directly into the Access database and was responsible for validating the data entry (e.g., checking for errors). The method of recording data (direct entry) was decided by the abstraction firm. Once data collection was complete, the chart abstraction database was sent to OptumInsight. Verification of medical chart abstraction completeness was conducted by OptumInsight upon receipt of the abstracted data. Electronic copies of the medical charts were redacted as to black out subject name and sent to OptumInsight in PDF format. Once the files were downloaded from a secure FTP site, they were saved on a restricted access network drive compliant with laws and standards related to the protection of private health information.

6. Assignment of Final Case Categorization of ASD ('gold standard' for this study)

The data abstracted from the charts and entered into the Access database was then reviewed by Dr. Marygrace Kaiser to assign the final case categorization of ASD to each of the 418 subjects. ****

**** Dr. Kaiser is a former ADDM principal investigator with seven years experience as an expert reviewer in the CDC ADDM network. She used a coding guide developed on the basis of Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (Text Revision) (DSM-IV-TR) criteria to determine whether each identified subject met the

Based upon criteria adapted from those utilized by the CDC ADDM network criteria, Dr. Kaiser assigned each subject to one of four categories: Level 1 confirmed, Level 2 confirmed, ASD ruled-out, and Unconfirmed. See Section III.E.2 for the definitions of the confirmation levels. This final case categorization of ASD was then used as the 'gold standard' to which we compared our claims-based ASD case algorithms.

In addition, five variables were recorded by Dr. Kaiser to provide greater context regarding the clinical assessment. These included: 1) the number of evaluations mentioned in the abstracted chart data; 2) the number of evaluations that mention an ASD diagnosis; 3) the degree of certainty that the child is an ASD case from clinical judgment of the chart overall; 4) the degree of impairment associated with ASD; and 5) an ASD DSM IV checklist. See Section III.E.3 for the definitions of these variables.

7. Quality Assurance

The chart abstraction process is complicated and involves multiple parties reviewing detailed information. Therefore, to ensure a high-quality result from the chart study, our study design included a Pilot Study and an Inter-rater Reliability (IRR) Assessment.

a. Pilot Study

The Team conducted a pilot study of 15 chart abstractions in the early stages of the chart abstraction process to confirm the information desired from the charts was available, assess whether any revisions to the chart abstraction form and database were needed, and ultimately determine the feasibility of conducting the chart abstraction for the approximately 400 remaining charts. Based on review of the pilot study charts, it was determined that the data collection instrument worked adequately in capturing relevant information from the charts and the chart study, as designed. Minor modifications to the clinical review portion of the form, such as including additional information on the confidence of the ASD diagnosis, were made. Dr. Kaiser then provided additional training of the abstractors as needed to ensure consistency between abstractors and study objectives. Chart procurement continued for the larger study while the chart pilot data was reviewed but abstraction was placed on hold during this time. See Appendix B for the results of the pilot study.

b. Inter-rater Reliability Assessment

We conducted inter-rater reliability (IRR) at various points during the chart abstraction process on two different elements: 1) the abstraction of clinical elements of interest from the chart; and 2) the final case categorization of ASD.

► Abstraction of Clinical Elements from the Chart

In addition to assessing reliability during training, a total of 40 completely abstracted charts were submitted for IRR assessment throughout the course of the study. This amounts to 9.6% (40/418) of all abstracted charts. A total of four abstractors evaluated and abstracted chart data for this project.

ASD case definition. Behaviors in the social, communication and behavior domains associated with ASD were identified in addition to evidence of delays before age three years, previous ASD diagnoses and autism-specific behaviors of sufficient quality or intensity to be highly indicative of an ASD.

For the purpose of IRR, following an extensive training session, the lead abstractor abstracted all 40 charts and the remaining three abstractors abstracted an equal portion of the 40 charts.

Reliability testing with all abstractors was assessed to assure that the reliability rate of all abstractors remains at 85% or higher. The double abstracted charts were reviewed by OptumInsight and Dr. Kaiser to ensure that the required IRR achieved a minimum of 85% agreement throughout the course of the project.

The first 15 charts underwent IRR during the pilot study. To ensure all abstractors were included in the first IRR assessment, the vendor was required to have all four abstractors assigned to the study abstract all 15 medical charts used in the pilot study. The vendor returned the completed database which included the pilot study abstractions for each abstractor to the project Team who performed the IRR calculation.

To determine the reliability of abstracted open-text fields, Dr. Kaiser and her team examined the text fields to determine if the information abstracted from the medical charts (such as clinical notes and behavioral descriptions) were consistent across abstractors. IRR was recorded for each open-ended question as a yes (the abstracted comments from each abstractor agreed sufficiently) or no (the abstracted comments from each abstractor did not sufficiently agree). Dr. Kaiser provided the documentation/criteria that were used to arrive at a yes or no assessment. Following review of open-ended text fields, Dr. Kaiser sent the reliability information database to OptumInsight.

For the 15 pilot study charts, the IRR met the minimum 85% agreement for all questions, with the exception of Q15 (behavioral descriptions), as the IRR for the open-ended text field related to this item was 73.3% (See Appendix B for IRR by question). Initially, abstractors were not consistently abstracting the same behavioral information into this field. It was the case in all four abstracted charts that the abstractors did not thoroughly abstract a portion of the provider notation related to behaviors. Additional training was provided on Q15 as well as Q14 (diagnosis of a non-ASD condition associated with ASD) and Q18 (evidence or notation of a developmental plateau or regression).

A second, third and fourth IRR were conducted on 10, 10 and 5 charts, respectively, during various times throughout the abstraction process in order to determine if the abstractors remained consistent. The resulting IRR at each of the subsequent reviews met the minimum 85% agreement for all questions.

► Final Case Categorization of ASD

A subsample of the charts reviewed by Dr. Kaiser was selected for an IRR review of final case categorization of ASD. For this IRR review, a colleague of Dr. Kaiser, Dr. Vanessa Gonzalez, also an experienced ADDM expert reviewer and a consultant to this project, reviewed approximately 10% (42) of all medical charts. Dr. Gonzalez completed her own final case categorization of ASD for these 42 medical charts for comparison to Dr. Kaiser's case determinations. Table 4 below cross tabulates Dr. Kaiser's and Dr. Gonzalez's final ASD case categorization. The percent agreement for the final case definitions was 90.5% (38 of 42 cases) and the Kappa statistic, commonly interpreted as quantifying agreement above that is expected by chance, was 85.4% (95% CI 63.9%, 100%). Of the four instances where there was not agreement, two differed only on level of confirmation while the other two were cases that met Level 2 criteria for one reviewer and

were viewed as unconfirmed by another. If Kappa is recalculated based on collapsing the table to simply “confirmed vs. unconfirmed” the estimate becomes 88.9%. Typically, a Kappa at or above 80% is considered to represent good inter-rater agreement.

Table 4: Comparison of Final Case Categorization of ASD Inter-rater Agreement

Reviewer 2	Reviewer 1				Total
	Confirmed L1	Confirmed L2	Unconfirmed	ASD Ruled-out	
Confirmed L1	17	1	0	0	18
Confirmed L2	1	9	0	0	10
Unconfirmed	0	2	12	0	14
ASD Ruled-out	0	0	0	0	0
Total	18	12	12	0	42

E. Description of Claims and Chart Variables

1. Chart Data Variables

As described above, the information from a subject’s medical chart was assessed up to 6 months prior to and up to 30 months after the index date to identify the following chart data variables.

- **Medical chart screen failure.** Whether a child passed the screening for medical chart abstraction. The reason for failing the screening (discrepancies between claim record and chart in date of birth, age as of index date, gender, evidence of Rett’s or CDD, or non-informative medical chart) was determined. The screens needed for an informative medical chart included evidence of an ASD diagnosis (see next variable definition), evidence of ASD testing or assessment, a comprehensive evaluation or evidence of social, developmental or behavior triggers associated with ASD as described in further detail above.
- **Diagnosis of ASD.** Whether a child had evidence of diagnosis of ASD (autistic disorder, specified PDD, PDD-NOS) in the chart. This could vary from a specific diagnosis (with or without an ICD-9-CM code) to just mentioning potential evidence of ASD. The date of the diagnosis and age at diagnosis was also determined.
- **ASD-associated condition.** Whether a child had a diagnosis of an ASD-associated condition (Table 1) in the chart.
- **Behavioral descriptions of ASD.** The presence of behavioral descriptions indicative of ASD or consistent with ASD.
- **History of developmental delay.** Whether a child had evidence of developmental delay before the age of three years.
- **Referral for ASD assessment.** Whether a child was referred for testing or an assessment related to autism or ASD.
- **Developmental plateau.** Whether a child had evidence of a developmental plateau or developmental regression.

- **Select other mental and behavioral health comorbidities.** The presence of comorbidities often related to ASD. These included attention deficit with or without hyperactivity disorder (ADHD/ADD), depression, bipolar/manic disorder, obsessive compulsive disorder (OCD), schizophrenia, epilepsy, intellectual disability and Tourette's syndrome.

2. *Variables Determined by Clinical Review of Abstracted Chart Data*

- **Final Case Categorization of ASD ('gold standard' of ASD case for this study).** Final determination of a diagnosis of ASD (autism, Asperger's or PDD-NOS) based on the chart data. In order to capture the context surrounding clinical outcome determination and to allow for greater sensitivity in overall clinical assessment, the clinical outcome variable includes two Levels of ASD confirmation. The date of the initial diagnosis and age at diagnosis were also determined. Subjects were defined as:
 1. Level 1 Confirmed – met our adaptation of the CDC-ADDM Project case definition criteria for ASD. The CDC-ADDM project classifies a child as having an ASD if there is evidence from historical record of either 1) the DSM-IV-TR criteria in the social, communication, and behavior domains and evidence of delays before age 3 years or 2) the social and either communication or behavior criteria for PDD-NOS OR Asperger disorder and at least one of the autism-specific behaviors of a sufficient quality or intensity to be highly indicative of an ASD.
 2. Level 2 Confirmed – A) presence of an ASD diagnosis in the chart OR B) some evidence of social and either communication or behavior criteria for ASD but not enough description to qualify as Level 1.
 3. Unconfirmed – not enough evidence to confirm ASD nor evidence that ASD has been ruled out for the subject. This includes any chart that does not include any statement regarding ASD.
 4. ASD ruled-out – clear evidence or a definitive statement that the child did not meet diagnostic criteria for ASD
- **Number of evaluations in medical chart.** The number of evaluations identified in the medical chart.
- **Number of evaluations that have a mention of an ASD diagnosis.** The number of evaluations in the medical chart where there is mention of a diagnosis of ASD.
- **Degree of certainty of an ASD diagnosis.** The degree of certainty based on chart review that the child has ASD on a 5-point scale ranging from 'not sure' to 'very sure.'

3. *Claims Variables*

As discussed in Section II. Objectives, claims-based data elements are incorporated into this chart study for two purposes: 1) to compare the true and false positive ASD cases to determine if any claims data could be used to further refine the claims-based ASD case algorithms; and 2) to assess the generalizability of the findings from the chart study sample to the broader study population. The clinical characteristics were assessed during the 6 month pre-index period while the health care and medication utilization variables were assessed during the 6 month post-index period. The full definitions for select variables are included in Appendix C.

- **Age at index year.** Using subjects' year of birth, subjects' age in years as of the year of the index date (first medical claim for ASD (Likely and Possible ASD) or ASD-associated condition (Enriched controls).
- **Age group at index year.** Subjects' age group as of the index year (year of first medical claim for ASD (Likely and Possible ASD) or ASD-associated condition (controls). Subjects with and without ASD were categorized as 2-8; 9-17; and 18-20 years of age.
- **Gender.** Gender from enrollment data.
- **Geographic location.** The United States region in which the study subject is enrolled in a health plan as of the index date. States were categorized into geographic regions in accordance with the U.S. Census Bureau's region designations. The regions are presented in Appendix C, Table C-1.
- **Behavioral health comorbidities.** Whether subjects had medical claims for selected behavioral health conditions. A binary variable (0/1) for each condition was created based on the presence of 2 or more medical claims with the relevant diagnosis code(s) (in any position). Conditions assessed were anxiety, depression, intellectual disability, obsessive compulsive disorder, schizophrenia, epilepsy and other seizure disorders, bipolar disorder, and Tourette syndrome. See Appendix C, Table C-2 for a list of ICD-9-CM diagnosis codes used.
- **Most frequent medications.** The top 30 medications classes based on a frequency of filled prescriptions according to a proprietary classification system. If a subject had multiple counts of the same medication class, it was counted only once.
- **Select psychotropic medications.** Whether a subject had at least 1 pharmacy or medical claim for anxiolytic, antidepressant, mood stabilizing, traditional antipsychotic, atypical antipsychotic, or anticonvulsant medications. Appendix C, Table C-3 lists the select medications examined.⁺⁺⁺
- **Attention deficit hyperactivity disorder medications.** Whether a subject had at least 1 pharmacy or medical claim for an attention deficit hyperactivity disorder medication. See Appendix C, Table C-3 for medications included.⁺⁺⁺
- **Risperidone.** Whether a subject had at least 1 pharmacy or medical claim for risperidone. ⁺⁺⁺ See Appendix C, Table C-3.
- **Oxytocin.** Whether a subject had at least 1 pharmacy or medical claim for oxytocin.⁺⁺⁺ See Appendix C, Table C-3.
- **Unique medications.** A total count of unique medication prescriptions filled.
- **Total medication dispensings.** A total count of all prescription fills.
- **All-cause health care office visits.** A count of a subject's office visits (e.g., provider offices, health clinics) was calculated. Office visits were calculated as at most 1 per provider per day.

⁺⁺⁺ Pharmacy claims were identified using a proprietary pharmacy coding system that rolls up multiple National Drug Codes (NDCs) to the ingredient (i.e., generic drug) level. Additionally, relevant HCPCS codes on medical claims were used as indicated in Table 5.

F. Analytic Strategy

1. Descriptive Analyses

Descriptive comparisons were conducted on both chart and claims variables to support many of the objectives outlined in Section II. Numbers and percentages are provided for dichotomous and polychotomous variables. Means, medians and standard deviations are provided for continuous variables. Key chart variables were stratified by cohort (Likely ASD, Possible ASD and Enriched controls). Claims data variables were stratified by true and false positive ASD cases to determine if any claims data could be used to further refine the claims-based ASD case algorithms.

Descriptive analyses of claims data variables also include comparisons of samples and the populations from which they were drawn along various dimensions, to assess the generalizability of our results. Whenever applicable, chi-square tests were conducted to calculate the P-values to indicate whether differences between comparison groups of interest were statistically significant.

2. Estimating Predictive Values

As stated above, the main objective of the chart study was to evaluate the claims-based ASD case algorithms used to identify children with ASD in our claims-based analyses for Tasks B, C, and D. To assess and represent the performance of claims-based ASD case algorithms in large samples, positive and negative predictive values (and 95% confidence intervals) are the most informative statistics. The positive predictive value (PPV) is the probability that a subject identified by a given claims-based algorithm actually has ASD as confirmed by the clinical review of the medical chart ('gold standard'), and the negative predictive value (NPV) is the probability that a subject who does not have ASD based on the claims-based algorithm actually does not have ASD, again according to the 'gold standard.' Specifically, PPV was calculated by dividing the number of true positives by the sum of true positives and false positives. Similarly, NPV was calculated by dividing the number of true negatives by the sum of true negatives and false negatives.

For claims-based research about children with ASD, positive predictive value is of greater importance compared to negative predictive value. If the positive predictive value is low, claims-based cases may not accurately represent all children with ASD in the population. Although ASD is more common than previously believed, it is still relatively rare in the population, meaning that we can expect a priori that the negative predictive value of claims-based ASD case algorithms will be high (i.e., the vast majority of children without a claims-based indicator of ASD will truly not have ASD). Thus, large samples of children who do not meet claims-based criteria for ASD will be reasonably representative of children without ASD in the population. In order to confirm this expectation, however, we also included in our validation study a smaller sample of subjects, the Enriched controls. If the negative predictive value in this sample is high, we can be further assured that negative predictive value in the general population will be high.

As discussed above, the 'gold standard' for ASD case confirmation for this study was based on the clinical review of the medical charts. The results of the clinical review fell into one of the following four categories: Level 1 confirmed, Level 2 confirmed, ASD ruled-out, or Unconfirmed, which are further defined below. For the purposes of calculating positive predictive value we considered two categories, ASD ruled-out and ASD Unconfirmed, to be non-cases. We also included children whose charts were examined but were uninformative (n=14) to be non-cases for the purposes of calculating PPV. Although the charts that were unconfirmed could prove to be

true ASD if full clinical information was available, we are considering them to be non-cases in order to derive the most conservative estimate of PPV.

The positive predictive value of our claims-based ASD case algorithms was calculated four ways as summarized in Table 5 below. We calculated the positive predictive value for the Likely ASD algorithm (>1 claim with ASD diagnostic code) and for Likely or Possible ASD algorithms combined (1 or more claim with ASD diagnostic code). A confirmed case for the 'gold standard' was defined in two ways: Level 1 confirmed (more stringent) and Level 1 or 2 confirmed.

Table 5: Positive Predictive Values calculated based on two claims-based ASD algorithms and two 'gold standard' confirmed case definitions

Claims-based ASD case algorithms	'Gold Standard'	
	Level 1 confirmed	Level 1 or 2 confirmed
Likely ASD	PPV	PPV
Likely or Possible ASD	PPV	PPV

Each of the four positive predictive values in the table above was also calculated on a weighted basis to account for our sampling stratification described in section III.C. In order to calculate the weighted version of these statistics, PPV and NPV were calculated for each cell in Table 2. Those values were then multiplied by the appropriate sampling fraction for that cell and averaged to determine an overall PPV and NPV. Finally, we calculated the positive predictive value of subgroups in the chart study according to sampling characteristics (age group, length of enrollment) and by provider tier. Examination of charts and identification of cases among the Enriched control sample will be conducted to determine negative predictive value.

Note that sensitivity and specificity were not calculated in this study. We believe predictive value has the most meaning in a validation study context, although sensitivity and specificity are, of course, commonly referred to in clinical applications. We do not present sensitivity and specificity, first, because our conditional sampling on claims-based ASD case algorithms means that we would need to reweight results. Moreover, the fact that only those with indicators of ASD-associated conditions were included in our sample of individuals not meeting claims-based ASD criteria (a prudent decision in the exploration of predictive value) means that we have no data on the number of false negatives in the large population of individuals not meeting claims-based ASD criteria without ASD-associated diagnoses. While we can comfortably assume that the *proportion* with ASD in this group is very small (and consequently we are also comfortable with the assertion that overall negative predictive value is high), given that this group comprises a large number of individuals (n=138,876), even small fluctuations in this very small proportion will affect the count of false negatives to a degree where sensitivity estimates are substantially influenced. So, without data here, sensitivity estimation is impossible.

IV. Results

This section presents the results tables. The main sections include: a) an overview of the sampling results; b) informative chart findings; c) final case categorization of ASD; d) positive and negative predictive values; e) descriptive comparisons between of true and false positives; and f) descriptive tables comparing the generalizability of the chart sample to the overall sample of the three cohorts in the claims.

A. Sampling

Detailed information about the sample selection according to each inclusion/exclusion criteria for each cohort is presented in Appendix D. For the Likely ASD cohort (>1 claim with ASD diagnostic code) there were 15,400 enrollees with two or more claims for ASD, Asperger's or PDD-NOS and who met age and minimum continuous enrollment period inclusion (and did not have claims for Rett or CDD). Of these, 5,781 (37.5%) subjects were covered under a fully-insured plan and approved for studies involving personal health information, thus making charts potentially accessible. Similarly, for the Possible ASD cohort (1 claim with ASD diagnostic code) and the Enriched control (1 or more claim with an ASD-associated condition diagnostic code) cohorts respectively there were 6,205 and 36,007 meeting basic criteria and, of these, 2,444 (39.9%) and 14,779(41.0%) from plans where charts were potentially accessible. The total sample of children who met study criteria are presented by cohort and sampling strata in Table 6. It should be noted that fewer than 150 children were identified as belonging to the longer-enrolled-rich claims categories for the Possible ASD cohort for both age groups.

Table 6. Eligible Study Subjects by Cohort and Sampling Strata.

	Total (N=23,004)		Younger (< age 8)						Older (≥ age 8)					
			Shorter enrolled (<18 months) (N=3,645)		Longer enrolled (≥ 18 months) (N=6,952)		Longer enrolled - rich claims (N=2,839)		Shorter enrolled (<18 months) (N=2,538)		Longer enrolled (≥ 18 months) (N=5,655)		Longer enrolled - rich claims (N=1,375)	
			n	%	n	%	n	%	n	%	n	%	n	%
Likely ASD	5,781	100.00	631	10.92	1,327	22.95	861	14.89	677	11.71	1,835	31.74	450	7.78
Possible ASD	2,444	100.00	393	16.08	522	21.36	142	5.81	521	21.32	742	30.36	124	5.07
Enriched Controls	14,779	100.00	2,621	17.73	5,103	34.53	1,836	12.42	1,340	9.07	3,078	20.83	801	5.42

The resulting sample selected for chart procurement from the eligible study subjects (based on the sampling strategy presented in Table 2) is presented in Table 7. For each cell, a goal of 6 times the targeted study sample (Table 2) was established. To make up for shortages in the longer-enrolled-rich claims categories, longer-enrolled non-rich claims categories were oversampled.

Table 7. Distribution of a 6 to 1 Random Sample of Eligible Children Available for Chart Selection.

	Total (N=2,400)		Younger (< age 8)						Older (≥ age 8)					
			Shorter enrolled (<18 months) (N=480)		Longer enrolled (≥ 18 months) (N=758)		Longer enrolled - rich claims (N=682)		Shorter enrolled (<18 months) (N=114)		Longer enrolled (≥ 18 months) (N=151)		Longer enrolled - rich claims (N=215)	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Likely ASD	1,056	100.00	210	19.89	210	19.89	420	39.77	54	5.11	54	5.11	108	10.23
Possible ASD	1,044	100.00	210	20.11	488	46.74	142	13.60	48	4.60	79	7.57	77	7.38
Enriched Controls	300	100.00	60	20.00	60	20.00	120	40.00	12	4.00	18	6.00	30	10.00

To begin chart acquisition, a random sample of subjects in each category equal to the number targeted in each cell (Table 2) was selected. Attempts were made to acquire and abstract the charts for these subjects. As unavailable and ineligible charts were identified, additional subjects within the incomplete cells were then selected for chart procurement and abstraction. When the targeted cell count was met, procurement of charts ceased for that cell. However, in the case of the longer-enrolled-rich claims cells among the Possible ASD cohort, an insufficient number of subjects with charts were available to meet targets and children from the longer-enrolled categories within the Possible ASD cohort were selected as replacements.

Of the 2,400 subjects randomly sampled for chart abstraction, 113 were removed from the sample because of provider involvement in HEDIS, leaving 2,287 eligible subjects in the sample (Figure 2). The selected provider for a total of 1,537 subjects (selected at random) was contacted. These included 678 children in the Likely cohort, 576 children in the Possible ASD cohort, and 283 in the Enriched control cohort. Among these, charts for a total of 544 subjects were procured. Charts for a total of 462 subjects were screened prior to complete review and charts for 44 subjects were excluded prior to abstraction for several reasons. One subject was excluded because the gender/DOB information in the chart did not match that in claims, another was excluded because there was an indication of Rett Syndrome in the chart, 28 were excluded because the chart suggested the subject was under age 2 at the time of their first diagnosis of ASD or ASD-associated condition, and 14 were excluded because their charts were considered uninformative and were not fully abstracted and who were thus presumed to be non-cases. They will be included in analyses where appropriate. The PPV and NPV calculations include these subjects and are based on a final sample size of 432 subjects, while the remaining analyses are based on fully abstracted charts and have a final sample size of 418 subjects.

Summaries of the chart procurement and medical chart screening results are provided in Appendix D.

Figure 2. Sample Selection

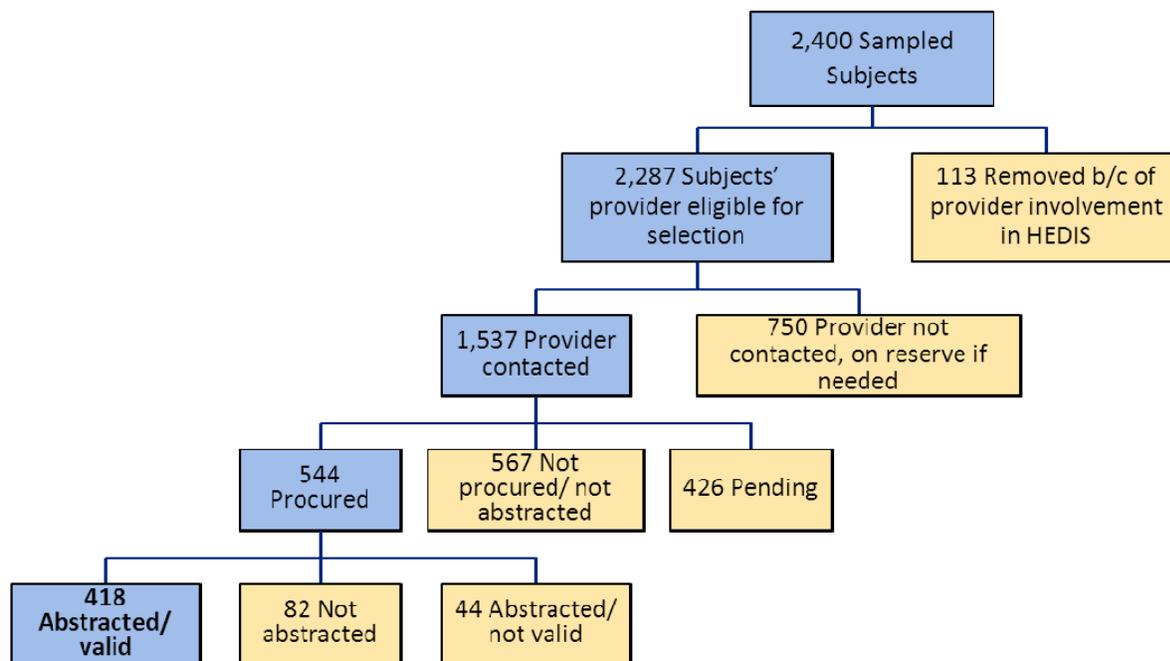


Table 8 presents the number of subjects for whom charts were abstracted and the proportion of the targeted sampling strata (Table 2) that was achieved. The targeted sample size was achieved in all cells except three. Targets were not achieved in older children with longer-enrolled-rich claims in the Possible and Likely ASD strata as well as in younger children with longer-enrolled-rich claims with Possible ASD. Note that as shown above in Table 2, these strata were those for which choosing a six-fold greater than the target sample size of 400 was inadequate to identify the targeted number of charts. Also note that we compared the generalizability of this final sample to the eligible sample. These results are discussed in greater detail in Section IV.F below.

Table 8. Proportion of Targeted Sample Size Achieved by Cohort and Strata

	Total		Younger (< age 8)						Older (≥ age 8)					
			Shorter enrolled (<18 months)		Longer enrolled (≥ 18 months)		Longer enrolled - rich claims		Shorter enrolled (<18 months)		Longer enrolled (≥ 18 months)		Longer enrolled - rich claims	
	n	%	n	%	n	%	n	%	N	%	n	%	n	%
Likely ASD	180	102.27	35	100.00	36	102.86	73	104.29	9	100.00	12	133.33	15	83.33
Possible ASD	180	103.45	35	100.00	84	240.00	24	34.29	11	137.50	14	155.56	12	70.59
Enriched Controls	58	116.00	15	150.00	13	130.00	19	95.00	2	100.00	3	100.00	6	120.00
Total	418	104.50	85	106.25	133	166.25	116	72.50	22	115.79	29	138.09	33	82.50

B. Informative Chart Findings

As described earlier in this report, chart abstraction focused on several variables key to assessing ASD status, including evidence of an ASD diagnosis and other diagnoses mentioned in the chart, behavioral assessments, evidence of developmental delay, ASD-related referrals, evidence of developmental plateau, tests for ASD, etc. These chart findings are presented by cohort in Table 9.

Table 9. Chart Variables by Cohort (Chart Data)

	Total (N=418)		Likely ASD (N=180)		Possible ASD (N=180)		Enriched Controls (N=58)	
	n	%	n	%	n	%	n	%
Diagnosis of ASD*	256	61.24	153	85.00	103	57.22	0	0.00
Autism	151	36.12	91	50.56	60	33.33	0	0.00
Asperger's Syndrome	40	9.57	24	13.33	16	8.89	0	0.00
PDD-NOS	119	28.47	74	41.11	45	25.00	0	0.00
ASD	76	18.18	56	31.11	20	11.11	0	0.00
Other	0	0.00	0	0.00	0	0.00	0	0.00
ASD-Associated Condition								
Intellectual Disability	12	2.87	7	3.89	5	2.78	0	0.00
Specific Condition Associated with ASD	2	0.48	2	1.11	0	0.00	0	0.00
General Codes Associated with ASD	50	11.96	27	15.00	11	6.11	12	20.69
Specific Developmental Delays Associated with ASD	248	59.33	122	67.78	97	53.89	29	50.00
Behavioral Descriptions of ASD	367	87.80	172	95.56	145	80.56	50	86.21
History of Developmental Delay	160	38.28	83	46.11	54	30.00	23	39.66
Referral for ASD Assessment	133	31.82	65	36.11	58	32.22	10	17.24
Developmental Plateau or Regression	43	10.29	29	16.11	12	6.67	2	3.45
Test or Assessments for ASD	148	35.41	71	39.44	55	30.56	22	37.93
Other Conditions/Health Issues	183	43.78	81	45.00	68	37.78	34	58.62

* A subject could have multiple diagnoses noted in the chart for ASD (e.g., one for autism and one for ASD). Therefore, the values on this row are not the sum of the rows beneath it.

In line with the claims sampling requirements for the Enriched controls cohort (for whom no ASD diagnostic code was found within the claims data), none of the children in this cohort had a diagnosis of ASD mentioned or otherwise indicated within their medical chart, thus confirming a robust negative predictive value. Compared to children in the Possible ASD cohort (who had only one claim with an ASD diagnostic code), children in the Likely ASD cohort (who had two or more claims with an ASD diagnostic code) were more likely to have a diagnosis of ASD indicated or mentioned in their medical chart (85.0% vs. 57.2%). Children in the Enriched controls cohort were more likely to have general codes associated with ASD noted in their charts compared to the ASD cohorts.

The majority of subjects in each of the cohorts had behavioral descriptions within their charts, especially the Likely ASD cohort (95.6%, compared to 86.2% for the Enriched controls and 80.6% for the Possible ASD cohort). Evidence of developmental delays, tests/assessments for ASD, and

referrals for ASD assessment were also somewhat common overall (38.3%, 35.4% and 31.8%, respectively). Children in the ASD cohorts were more likely to have evidence of referrals for ASD assessments and have evidence of developmental plateau or regression. Still a notable proportion of the children in the Enriched control group had evidence of developmental delays (39.7%) and tests/assessments for ASD (37.9%).

Appendix E, Tables E-1 through E-4 present these chart data by cohort stratified by provider tier and age group. The additional questions providing the degree of certainty of the diagnosis of ASD as well as other additional information informing the diagnosis are also presented in Appendix E.

In addition to chart variables described above, chart abstractors also identified whether there was evidence in the subjects' medical charts of other conditions and health issues frequently associated with ASD, including ADHD, depression, bipolar/manic depression, OCD, schizophrenia, epilepsy, intellectual disability, and Tourette's Syndrome. These results are presented by cohort in Table 10 below.

Overall, with the exceptions of ADHD and epilepsy – which were more prevalent among the Enriched controls – the occurrence of the conditions was similar across cohorts. Overall, just under a third of the subjects had evidence of ADHD, 12.0% had evidence of epilepsy, and 6.0% had evidence of OCD within their chart. Three percent or less of subjects had evidence of the other conditions. The proportion with ADHD and epilepsy was approximately 40.0% and 21.0%, respectively, among Enriched controls.

Table 10. Other Conditions/Health Issues by Cohort (Chart Data).

	Total (N=418)		Likely ASD (N=180)		Possible ASD (N=180)		Enriched Controls (N=58)	
	n	%	n	%	n	%	n	%
Other Conditions/ Health Issues	183	43.78	81	45.00	68	37.78	34	58.62
ADHD	125	29.90	52	28.89	50	27.78	23	39.66
Depression	13	3.11	5	2.78	6	3.33	2	3.45
Bipolar/Manic Depression	4	0.96	3	1.67	0	0.00	1	1.72
OCD	23	5.50	10	5.56	11	6.11	2	3.45
Schizophrenia	0	0.00	0	0.00	0	0.00	0	0.00
Epilepsy	51	12.20	26	14.44	13	7.22	12	20.69
Intellectual disability	13	3.11	8	4.44	5	2.78	0	0.00
Tourette's Syndrome	3	0.72	2	1.11	0	0.00	1	1.72

C. Final Assessment of ASD ('gold standard')

Although we did not have complete medical charts for the children in this study, the key objective of this chart study was to conduct a clinical review of the available abstracted medical chart data to assign a final case categorization of ASD. Based on a clinical review of the chart data, subjects were categorized into one of four ASD groups: Level 1 confirmed ASD, Level 2 confirmed ASD, Unconfirmed, and ASD ruled-out. As described earlier, Level 1 confirmed included subjects who

had evidence from historical record of either 1) the DSM-IV-TR criteria in the social, communication, and behavior domains and evidence of delays before age 3 years or 2) the social and either communication or behavior criteria for PDD-NOS OR Asperger disorder and at least one of the autism-specific behaviors of a sufficient quality or intensity to be highly indicative of an ASD. Level 2 confirmed included subjects with documentation of an ASD diagnosis in the chart OR B) some evidence of social and either communication or behavior criteria for ASD but not a complete enough description to qualify as Level 1. The unconfirmed group did not have enough evidence to confirm ASD (Level 1 or 2) nor evidence that ASD had been ruled out for the subject. This included any chart that does not include any statement or information regarding ASD. Finally, ASD ruled-out included subjects with clear evidence or a definitive statement that the child did NOT meet diagnostic criteria for ASD (indicating some form of subjective or objective evaluation).

Table 11 presents the final case categorization of ASD of study subjects by cohort. Note, this table includes the 14 charts for subjects whose charts were determined 'uninformative' and therefore were not reviewed but are assumed to be a non-case. Overall, 37.3% of our study subjects were classified as Level 1 confirmed, and approximately a quarter of all subjects met the requirements for Level 2 confirmation. Just over a third (34.5%) were placed in the unconfirmed group. Only 1.4% of the total sample was determined to be ASD ruled-out cases.

As expected, ASD confirmation status varied significantly by cohort. Among the children in the Likely ASD cohort, a large proportion received a Level 1 ASD confirmation (61.0%), and another 26.4% were classified as Level 2 confirmed, for a total of 87.4% with evidence of ASD in their chart. Just over 10% of the Likely ASD cohort were left unconfirmed, and only one subject in this cohort was actually considered an ASD ruled-out case.

Based on the clinical review of the chart data, the majority of Possible ASD cohort members (61.6%) were also confirmed as having ASD; however, fewer subjects in this group met Level 1 ASD confirmation requirements (26.3%) and more were classified as unconfirmed cases (35.8%).

Only one of the Enriched control subjects (1.7%) were classified as having ASD, and all the rest (98.3%) were classified as ASD unconfirmed. The only subject that was classified as having ASD met only the Level 2 confirmation requirements.

The final case categorization of ASD by provider tier status, age group and enrollment categories are presented in Appendix E, Tables E-6 through E-11.

Table 11. Final Case Categorization of ASD by Cohort (Based on Chart Data)

	Total (N=432)*		Likely ASD (N=182)*		Possible ASD (N=190)*		Enriched Controls (N=60)*	
	n	%	n	%	n	%	n	%
Final Case Categorization of ASD								
Level 1 ASD Confirmation	161	37.27	111	60.99	50	26.32	0	0.00
Level 2 ASD Confirmation	116	26.85	48	26.37	67	35.26	1	1.67
ASD Ruled-out	6	1.39	1	0.55	5	2.63	0	0.00
Unconfirmed ASD*	149	34.49	22	12.09	68	35.79	59	98.33

*Including 14 subjects whose charts were not reviewed but are assumed to be a non-case are included in the Unconfirmed ASD row (2 Likely ASD; 10 Possible ASD; and 2 Enriched controls).

D. Estimated Predictive Values of Alternate Claims-Based ASD Case Algorithms

As stated above, the main objective of the chart study is to evaluate our claims-based ASD case algorithms used to identify children with ASD for our claims-based analyses in Tasks B, C, and D. In this section, we report and discuss the positive and negative predictive values. The positive predictive value is the probability that a subject identified as having ASD by the claims-based ASD case algorithm has ASD as confirmed by the clinical review of the medical chart ('gold standard' results presented in prior section), and the negative predictive value is the probability that a subject identified as not having ASD by the claims-based ASD case algorithm actually does not have ASD according to the 'gold standard.' Again, PPV was calculated by dividing the number of true positives by the sum of true positives and false positives. Similarly, NPV was calculated by dividing the number of true negatives by the sum of true negatives and false negatives.

1. Positive Predictive Value

The results of the clinical review assigned each study subject to one of the following four categories: Level 1 confirmed, Level 2 confirmed, ASD ruled-out, or Unconfirmed. Principal analyses of positive predictive values are presented for Likely ASD and Likely or Possible ASD children using Level 1 or Level 2 criteria as confirming ASD case status in Tables 12-15. We also present results for Level 1 confirmation only to estimate PPV using the most stringent criteria. For the purposes of calculating PPV and NPV, we consider both unconfirmed and ASD ruled-out as 'gold standard' non-cases.

The unweighted and weighted PPVs are presented in Table 12. Given the formula for calculating PPV, the unweighted PPV for the Likely ASD cohort based on Level 1 is the same as the percentage presented in Table 11 above (61%). In comparison, based on Level 1 confirmed, the PPV for the Likely or Possible cohort is 43.3%. Using the Level 1 or Level 2 criteria, the PPV was 87.4 in the Likely ASD cohort and 74.2 in the Likely or Possible ASD cohorts. Weighting the PPVs to adjust for our sampling strategy, as described in Section III.F., had little effect on the PPVs. For example, the weighted PPV based on Level 1 or Level 2 confirmed for the Likely or Possible cohort was 76.49 compared to the unweighted PPV of 74.19 (the 95% CI overlapped). Thus, going forward, only unweighted PPVs will be presented.

Table 12. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions in full sample*.

		Unweighted	Weighted
Level 1 ASD confirmation			
Likely ASD	Valid N	182	
	Confirmed N	111	
	PPV(%)	60.99	60.93
	Lower 95% CI	53.50	53.09
	Upper 95% CI	68.12	68.77
Likely or Possible ASD	Valid N	372	
	Confirmed N	161	
	PPV(%)	43.28	44.95
	Lower 95% CI	38.18	38.65
	Upper 95% CI	48.49	51.25
Level 1 or Level 2 ASD confirmation			
Likely ASD	Valid N	182	
	Confirmed N	159	
	PPV(%)	87.36	87.34
	Lower 95% CI	81.64	78.22
	Upper 95% CI	91.82	96.46
Likely or Possible ASD	Valid N	372	
	Confirmed N	276	
	PPV(%)	74.19	76.49
	Lower 95% CI	69.43	68.35
	Upper 95% CI	78.57	84.63

*CI calculated as Exact Binomial

The remaining tables in this section present the unweighted PPV stratified by the two sampling strata criteria (age and length of enrollment) and provider tier.

Table 13 presents the PPV calculations by age group. When using the Level 1 ASD criteria, the PPV was generally higher in younger children (64.38 for the Likely ASD cohort and 45.15 for the Likely and Possible ASD cohort) compared to older children (47.22 and 35.62, respectively). However, when using the Level 1 or Level 2 criteria, the PPV was somewhat higher in the older children. This is due to the higher proportion of younger children with Level 1 ASD confirmation while older children were more likely to have a Level 2 ASD confirmation (Appendix E).

Table 13. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions, by age group*.

		Younger (< age 8)	Older (≥ age 8)
Level 1 ASD confirmation			
Likely ASD	Valid N	146	36
	Confirmed N	94	17
	PPV(%)	64.38	47.22
	Lower 95% CI	56.04	30.41
	Upper 95% CI	72.13	64.51
Likely or Possible ASD	Valid N	299	73
	Confirmed N	135	26
	PPV(%)	45.15	35.62
	Lower 95% CI	39.42	24.75
	Upper 95% CI	50.98	47.69
Level 1 or Level 2 ASD confirmation			
Likely ASD	Valid N	146	36
	Confirmed N	126	33
	PPV(%)	86.30	91.67
	Lower 95% CI	79.64	77.53
	Upper 95% CI	91.43	98.25
Likely or Possible ASD	Valid N	299	73
	Confirmed N	216	60
	PPV(%)	72.24	82.19
	Lower 95% CI	66.79	71.47
	Upper 95% CI	77.24	90.16

*CI calculated as Exact Binomial

Table 14 presents the PPV by length of enrollment period. In children with Likely ASD, the PPVs were generally lower in the longer enrolled children with rich claims compared to the other two groups. The higher estimate in the shorter-enrolled group was unexpected. However, 95% CIs on the PPVs for the longer and shorter-enrolled overlap substantially and the differences between the estimates are most likely related to chance variation.

Table 14. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions stratified by enrollment categories*.

		Shorter enrolled (<18 months)	Longer enrolled (≥ 18 months)	Longer enrolled - rich claims
Level 1 ASD confirmation				
Likely ASD	Valid N	44	48	90
	Confirmed N	28	31	52
	PPV(%)	63.64	64.58	57.78
	Lower 95% CI	47.77	49.46	46.91
	Upper 95% CI	77.59	77.84	68.12
Likely or Possible ASD	Valid N	93	153	126
	Confirmed N	45	56	60
	PPV(%)	48.39	36.60	47.62
	Lower 95% CI	37.89	28.97	38.65
	Upper 95% CI	58.99	44.76	56.70
Level 1 or Level 2 ASD confirmation				
Likely ASD	Valid N	44	48	90
	Confirmed N	41	46	72
	PPV(%)	93.18	95.83	80.00
	Lower 95% CI	81.34	85.75	70.25
	Upper 95% CI	98.57	99.49	87.69
Likely or Possible ASD	Valid N	93	153	126
	Confirmed N	74	109	93
	PPV(%)	79.57	71.24	73.81
	Lower 95% CI	69.95	63.38	65.23
	Upper 95% CI	87.23	78.26	81.24

*CI calculated as Exact Binomial

The PPVs are presented by provider tier status in Table 15. When using the Level 1 ASD criteria, the PPVs for Likely or Possible were considerably higher in children with a provider in the 1st tier compared to children with a provider in the 2nd or 3rd tier (64.02 vs. 26.92). This was as expected, given we assumed that specialists were more likely to have made and documented a diagnosis for ASD. When using the Level 1 or 2 confirmation criteria, the PPVs for Likely or Possible remained higher for 1st tier than for 2nd and 3rd tier providers (83.54 vs. 66.83), but the difference was smaller.

Table 15. Positive Predictive Values (%) and 95% CIs for alternate claims-based case definitions using two 'gold standard' definitions stratified by provider tier*.

		Total	1st Tier	2nd & 3rd Tier
Level 1 ASD confirmation				
Likely ASD	Valid N	182	103	79
	Confirmed N	111	73	38
	PPV(%)	60.99	70.87	48.10
	Lower 95% CI	53.50	61.10	36.71
	Upper 95% CI	68.12	79.41	59.64
Likely or Possible ASD	Valid N	372	164	208
	Confirmed N	161	105	56
	PPV(%)	43.28	64.02	26.92
	Lower 95% CI	38.18	56.17	21.02
	Upper 95% CI	48.49	71.36	33.49
Level 1 or Level 2 ASD confirmation				
Likely ASD	Valid N	182	103	79
	Confirmed N	159	90	69
	PPV(%)	87.36	87.38	87.34
	Lower 95% CI	81.64	79.38	77.95
	Upper 95% CI	91.82	93.11	93.76
Likely or Possible ASD	Valid N	372	164	208
	Confirmed N	276	137	139
	PPV(%)	74.19	83.54	66.83
	Lower 95% CI	69.43	76.96	59.98
	Upper 95% CI	78.57	88.86	73.18

*CI calculated as Exact Binomial

2. Negative Predictive Value in Subjects with ASD-associated Conditions

As mentioned, negative predictive value for a claims-based ASD case algorithm is expected to be high. Should there be false negatives, it is anticipated they would occur disproportionately among individuals with some claims-based evidence of other conditions associated with ASD. As presented in Table 10 above, we collected data on a small sample (n=60) of such subjects (Enriched controls). Of the 60, there was only 1 subject who met gold-standard criteria (Level 2 confirmation). Consequently, this approach suggests that negative predictive value for claims based claims-based ASD case algorithm will be high.

E. Comparison of True and False Positives

As described in the study objectives above, we compared the presence of certain claims-based data elements between true and false positive ASD cases to determine if other elements of claims data might be useful to modify or refine the claims-based algorithms. To develop such claims-based indicators in our validation study sample, we would want to identify ASD-associated diagnostic or other codes (medication, procedure, etc.) that could be added to claims-based ASD

case algorithm to improve the PPV and NPV by either differentially including as claims-based cases subjects who now appear as false negatives or differentially remove as cases those who are currently false positives. Since our study sample included only one false negative (the one subject in the Enriched controls meeting Level 2 confirmation criteria), our ability to empirically explore ways to capture such individuals is limited. Therefore, this exploration focused on seeing if there were additional claims-based variables commonly seen in false positives that were infrequently seen in true positives. Ostensibly, codes for these indicators could then be incorporated into improved claims-based ASD case algorithm as exclusion codes in Tasks B, C, and D. Although we undertook such exploration, our limited sample size prevented a comprehensive exploration of potential variables and thus these results have to be considered with some caution.

The categorization of true and false positives in this section is based on Level 1 or Level 2 confirmed criteria. Analyses are presented separately for the Likely ASD cohort and the Likely and Possible ASD cohorts combined. Specifically, members of the Likely ASD cohort were categorized as true positives if they met the criteria for either Level 1 or Level 2 confirmation. Likely ASD cohort members who did not meet either set of criteria (i.e., were unconfirmed or ASD ruled-out cases) were categorized as false positives.

Table 16 presents the behavioral health comorbidities in children in the Likely ASD cohort by case classification status. There were only 21 false positives among the Likely ASD cohort. Therefore, these results should be interpreted with caution as percentages appear large based on small cell sizes. There were no significant differences between true positives and false positives in the proportion of children with behavioral health comorbidities.

Table 16. Behavioral Health Comorbidities by Case Classification Status (Claims Data) - Likely ASD.

Behavioral Health Comorbidities	True Positives (N=159)		False Positives (N=21)		p-value
	n	%	n	%	
Anxiety	5	3.14	0	0.00	0.410
Attention Deficit (with or without hyperactivity)	10	6.29	2	9.52	0.577
Bipolar Disorder	4	2.52	2	9.52	0.093
Depression	2	1.26	0	0.00	0.605
Epilepsy and other Seizure Disorders	8	5.03	3	14.29	0.096
Intellectual Disability	0	0.00	0	0.00	–
Obsessive-Compulsive Disorder	1	0.63	0	0.00	0.716
Schizophrenia	0	0.00	0	0.00	–
Tourette Syndrome	0	0.00	0	0.00	–

The behavioral health comorbidities in the Likely or Possible ASD cohorts (1 or more claim with ASD diagnostic code) are presented by case classification status in Table 17. As in Table 16, there were no significant differences between true and false positives in the proportion of children with behavioral health comorbidities in most cases, with the exception of attention disorders and bipolar disorder, which were more prevalent among false positives compared to true positives.

Table 17. Behavioral Health Comorbidities by Case Classification Status (Claims Data) - Likely or Possible ASD.

	True Positives (N=276)		False Positives (N=84)		p-value
	n	%	n	%	
Behavioral Health Comorbidities					
Anxiety	8	2.90	2	2.38	0.800
Attention Deficit (with or without hyperactivity)	20	7.25	13	15.48	0.022
Bipolar Disorder	5	1.81	5	5.95	0.043
Depression	4	1.45	2	2.38	0.559
Epilepsy and other Seizure Disorders	11	3.99	7	8.33	0.109
Intellectual Disability	2	0.72	1	1.19	0.681
Obsessive-Compulsive Disorder	1	0.36	0	0.00	0.581
Schizophrenia	1	0.36	0	0.00	0.581
Tourette Syndrome	0	0.00	0	0.00	–

Table 18 presents a comparison of post-index psychotropic medication use in children in the Likely ASD cohort by case classification status. The proportions of children with claims for psychotropic medication were not statistically significantly different between true and false positives for most medications. However, while not statistically significant, a higher proportion of false positives (14%) had a prescription for risperidone than the true positives (4.4%), suggesting that risperidone may have been prescribed for conditions other than ASD itself.

Table 18. Post-Index Psychotropic Medications by Case Classification Status (Claims Data) - Likely ASD

	True Positives (N=159)		False Positives (N=21)		p-value
	n	%	n	%	
Anti Depressants	16	10.06	1	4.76	0.435
Anticonvulsant/Antiepileptics	21	13.21	5	23.81	0.194
Antipsychotics (Traditional, 1st Generation)	0	0.00	0	0.00	–
Antipsychotics (Atypical)	15	9.43	3	14.29	0.486
risperidone	7	4.40	3	14.29	0.063
Anxiolytics	9	5.66	1	4.76	0.866
Attention Deficit Medications	40	25.16	6	28.57	0.736
Hormones					
Oxytocin	0	0.00	0	0.00	–
Mood Stabilizers	0	0.00	1	4.76	0.006
	mean	SD	mean	SD	–
Total Number of Unique Medications	2.59	2.57	2.57	2.40	0.973
Total Number of Medication Dispensings	6.09	7.57	6.52	6.79	0.802

A comparison of post-index psychotropic medications in children in the Likely and Possible ASD cohorts is presented by case classification status in Table 19. The proportions of children with psychotropic medications were similar across case classification status. In contrast to the likely ASD cohort in the prior table, the proportion of children with a prescription for risperidone does not differ greatly between the true and false positives when likely and possible ASD cohorts are combined.

Table 19. Post-Index Psychotropic Medications by Case Classification Status (Claims Data) - Likely and Possible ASD.

	True Positives (N=276)		False Positives (N=84)		p-value
	n	%	n	%	
Anti Depressants	23	8.33	7	8.33	1.000
Anticonvulsant/Antiepileptics	26	9.42	12	14.29	0.204
Antipsychotics (Traditional, 1st Generation)	1	0.36	0	0.00	0.581
Antipsychotics (Atypical)	18	6.52	7	8.33	0.567
risperidone	9	3.26	5	5.95	0.264
Anxiolytics	13	4.71	5	5.95	0.647
Attention Deficit Medications	63	22.83	23	27.38	0.391
Hormones					
Oxytocin	0	0.00	0	0.00	–
Mood Stabilizers	0	0.00	1	1.19	0.069
	mean	SD	mean	SD	–
Total Number of Unique Medications	2.30	2.33	2.71	2.96	0.243
Total Number of Medication Dispensings	5.16	6.63	6.38	7.06	0.145

Table 20 presents the most frequent medications received in children in the Likely ASD cohort by case classification status. True positives were less likely to receive anticonvulsants, combination narcotics/analgesics and bowel evacuants compared to false positives.

Table 20. Top 30 Most Frequent Medications^{###} by Case Classification Status (Claims Data) - Likely ASD

	True Positives (N=159)		False Positives (N=21)		p-value
	n	%	n	%	
Penicillins	49	30.82	9	42.86	0.267
Miscellaneous psychotherapeutic agents	19	11.95	1	4.76	0.325
Erythromycins & other macrolides	18	11.32	5	23.81	0.107
Antihistamines	23	14.47	4	19.05	0.580
Beta agonists inhalers	14	8.81	3	14.29	0.420
Third generation cephalosporins	14	8.81	4	19.05	0.141
Anticonvulsants	15	9.43	5	23.81	0.049
Selective serotonin reuptake inhibitors	10	6.29	0	0.00	0.237
Adrenal hormones	12	7.55	2	9.52	0.751
Antitussive combinations	15	9.43	2	9.52	0.989
Miscellaneous antipsychotics	9	5.66	2	9.52	0.487
Antibiotics	16	10.06	2	9.52	0.938
Miscellaneous pulmonary agents	12	7.55	1	4.76	0.643
Intranasal steroids	6	3.77	1	4.76	0.826
Inhaled corticosteroids	6	3.77	1	4.76	0.826
Second generation cephalosporins	5	3.14	1	4.76	0.698
Combination narcotic /analgesics	5	3.14	4	19.05	0.002
Decongestant / antihistamines	8	5.03	1	4.76	0.958
Adrenergic antagonists & related drugs	10	6.29	1	4.76	0.784
Miscellaneous otic preparations	6	3.77	2	9.52	0.229
Vitamins & hematinics	4	2.52	1	4.76	0.556
First generation cephalosporins	4	2.52	1	4.76	0.556
Otic steroid / antibiotic	5	3.14	2	9.52	0.155
Topical corticosteroids medium potency	7	4.40	0	0.00	0.327
Topical antifungals	7	4.40	1	4.76	0.940
Sulfa's & related agents	4	2.52	2	9.52	0.093
Bowel evacuants	7	4.40	4	19.05	0.008
Topical antibacterials	10	6.29	1	4.76	0.784
Miscellaneous antidepressants	0	0.00	0	0.00	–
Therapy for acne	3	1.89	0	0.00	0.526

^{###} Most frequent medications were defined based on the Overall Likely cohort.

The most frequent medications received in children in the Likely and Possible ASD cohorts are presented by case classification status in Table 21. True positives were less likely to receive combination narcotics/analgesics and miscellaneous antidepressants compared to false positives.

**Table 21. Top 30 Most Frequent Medications
by Case Classification Status (Claims Data) - Likely and Possible ASD**

	True Positives (N=276)		False Positives (N=84)		p-value
	n	%	n	%	
Penicillins	83	30.07	33	39.29	0.114
Miscellaneous psychotherapeutic agents	32	11.59	9	10.71	0.824
Erythromycins & other macrolides	33	11.96	16	19.05	0.097
Antihistamines	29	10.51	10	11.90	0.718
Beta agonists inhalers	25	9.06	10	11.90	0.441
Third generation cephalosporins	29	10.51	12	14.29	0.340
Anticonvulsants	22	7.97	11	13.10	0.154
Selective serotonin reuptake inhibitors	13	4.71	2	2.38	0.350
Adrenal hormones	17	6.16	4	4.76	0.632
Antitussive combinations	23	8.33	9	10.71	0.502
Miscellaneous antipsychotics	13	4.71	5	5.95	0.647
Antibiotics	29	10.51	6	7.14	0.362
Miscellaneous pulmonary agents	20	7.25	6	7.14	0.974
Intranasal steroids	9	3.26	4	4.76	0.519
Inhaled corticosteroids	9	3.26	3	3.57	0.890
Second generation cephalosporins	6	2.17	5	5.95	0.078
Combination narcotic /analgesics	8	2.90	8	9.52	0.010
Decongestant / antihistamines	14	5.07	6	7.14	0.468
Adrenergic antagonists & related drugs	15	5.43	4	4.76	0.809
Miscellaneous otic preparations	10	3.62	4	4.76	0.636
Vitamins & hematinics	5	1.81	4	4.76	0.129
First generation cephalosporins	10	3.62	5	5.95	0.350
Otic steroid / antibiotic	8	2.90	5	5.95	0.189
Topical corticosteroids medium potency	8	2.90	2	2.38	0.800
Topical antifungals	9	3.26	1	1.19	0.312
Sulfa's & related agents	8	2.90	5	5.95	0.189
Bowel evacuants	8	2.90	5	5.95	0.189
Topical antibacterials	13	4.71	7	8.33	0.204
Miscellaneous antidepressants	0	0.00	2	2.38	0.010
Therapy for acne	5	1.81	2	2.38	0.741

F. Generalizability

The three tables in this section compare demographic and clinical characteristics between the overall eligible sample, the sample selected for possible chart selection (6 times the target sample), the charts procured, and the final chart sample by cohort. The objective of these tables is to determine if the final study sample is representative of the eligible population. Overall, the final sample is younger, with a higher proportion of females and children from the south, and with a

lower proportion of children with behavioral health comorbidities. While some of these differences seem to be related to provider non-response or charts failing screen criteria (reasons for subject not being included in final sample), most were related to our sampling strategy and thus intentional and expected. Appendix F includes the procurement rate by provider tier and subject demographic characteristics.

Table 22 presents the demographic characteristics in the Likely ASD cohort. The final sample had lower mean age, a greater proportion of children from the South, a smaller proportion of children with ADD and depression, and a higher proportion of children with epilepsy compared to children in the overall eligible sample.

Table 22. Subject Demographic Characteristics by Abstraction Status (Claims data) - Likely ASD.

	Likely ASD							
	Eligible sample (N=5,781)		Over-sample (N=1,056)		Provider contacted (N=678)		Final sample (N=180)	
	mean	SD	mean	SD	mean	SD	mean	SD
Age	8.42	4.72	5.89	3.79	5.80	3.77	5.73	3.44
	n	%	n	%	n	%	n	%
Age Categories								
2-8	3,201	55.37	871	82.48	562	82.89	151	83.89
9-17	2,349	40.63	166	15.72	102	15.04	27	15.00
18-20	231	4.00	19	1.80	14	2.06	2	1.11
Gender								
Male	4,709	81.46	871	82.48	561	82.74	148	82.22
Female	1,072	18.54	185	17.52	117	17.26	32	17.78
Geographic Region								
Northeast	825	14.27	161	15.25	90	13.27	20	11.11
Midwest	1,792	31.00	262	24.81	156	23.01	39	21.67
South	2,505	43.33	514	48.67	354	52.21	103	57.22
West	659	11.40	119	11.27	78	11.50	18	10.00
ASD Related Co-morbid Conditions								
Anxiety	290	5.02	40	3.79	28	4.13	5	2.78
Attention Deficit (w/or wo/ hyperactivity)	750	12.97	109	10.32	56	8.26	12	6.67
Bipolar Disorder	223	3.86	29	2.75	17	2.51	6	3.33
Depression	254	4.39	29	2.75	15	2.21	2	1.11
Epilepsy and other Seizure Disorders	92	1.59	36	3.41	28	4.13	11	6.11
Intellectual Disability	32	0.55	5	0.47	4	0.59	0	0.00
Obsessive-Compulsive Disorder	61	1.06	9	0.85	5	0.74	1	0.56
Schizophrenia	10	0.17	1	0.09	0	0.00	0	0.00
Tourette Syndrome	18	0.31	3	0.28	3	0.44	0	0.00

A comparison of demographic and clinical characteristics between the overall eligible sample, the sample selected for possible chart selection (6 times the target sample), the charts procured and the final chart sample is presented for children in the Possible ASD cohort in Table 23. The final sample had a higher proportion of males and a greater proportion of children from the South compared to children in the overall eligible sample.

Table 23. Subject Demographic Characteristics by Abstraction Status (Claims data) - Possible ASD.

	Possible ASD							
	Eligible sample (N=2,444)		Over- sample (N=1,044)		Provider contacted (N=576)		Final sample (N=180)	
	mean	SD	mean	SD	mean	SD	mean	SD
Age	9.09	4.75	6.14	3.70	6.12	3.64	6.01	3.44
	n	%	n	%	n	%	n	%
Age Categories								
2-8	1,212	49.59	867	83.05	472	81.94	149	82.78
9-17	1,106	45.25	160	15.33	97	16.84	29	16.11
18-20	126	5.16	17	1.63	7	1.22	2	1.11
Gender								
Male	1,866	76.35	775	74.23	433	75.17	123	68.33
Female	578	23.65	269	25.77	143	24.83	57	31.67
Geographic Region								
Northeast	354	14.48	152	14.56	76	13.19	19	10.56
Midwest	726	29.71	304	29.12	164	28.47	48	26.67
South	1,072	43.86	474	45.40	280	48.61	91	50.56
West	292	11.95	114	10.92	56	9.72	22	12.22
Behavioral Health Comorbidities								
Anxiety	101	4.13	31	2.97	13	2.26	5	2.78
Attention Deficit (w/ or wo/ hyperactivity)	266	10.88	98	9.39	61	10.59	21	11.67
Bipolar Disorder	93	3.81	25	2.39	15	2.60	4	2.22
Depression	104	4.26	27	2.59	13	2.26	4	2.22
Epilepsy and other Seizure Disorders	42	1.72	28	2.68	18	3.13	7	3.89
Intellectual Disability	20	0.82	11	1.05	8	1.39	3	1.67
Obsessive-Compulsive Disorder	13	0.53	3	0.29	1	0.17	0	0.00
Schizophrenia	5	0.20	1	0.10	1	0.17	1	0.56
Tourette Syndrome	8	0.33	2	0.19	1	0.17	0	0.00

Table 24 presents the demographic characteristics between the overall eligible sample, the sample selected for possible chart selection (6 times the target sample), the charts procured and the final chart sample in the Enriched control cohort. The final sample had a lower proportion of males, a greater proportion of children from the South, and a higher proportion of children with depression compared to children in the overall eligible sample.

Table 24. Subject Demographic Characteristics by Abstraction Status (Claims data) - Enriched Controls.

	Enriched Control							
	Eligible sample (N=14,779)		Over- sample (N=300)		Provider contacted (N=283)		Final sample (N=58)	
	mean	SD	mean	SD	mean	SD	mean	SD
Age	6.65	4.85	5.63	4.26	5.70	4.28	6.12	4.27
	n	%	n	%	n	%	n	%
Age Categories								
2-8	10,295	69.66	248	82.67	235	83.04	48	82.76
9-17	4,005	27.10	46	15.33	42	14.84	9	15.52
18-20	479	3.24	6	2.00	6	2.12	1	1.72
Gender								
Male	9,904	67.01	191	63.67	181	63.96	34	58.62
Female	4,875	32.99	109	36.33	102	36.04	24	41.38
Geographic Region								
Northeast	1,902	12.87	41	13.67	40	14.13	8	13.79
Midwest	4,262	28.84	73	24.33	68	24.03	7	12.07
South	7,025	47.53	153	51.00	145	51.24	38	65.52
West	1,590	10.76	33	11.00	30	10.60	5	8.62
ASD Related Co-morbid Conditions								
Anxiety	296	2.00	5	1.67	5	1.77	1	1.72
Attention Deficit (w/ or wo/ hyperactivity)	802	5.43	10	3.33	10	3.53	2	3.45
Bipolar Disorder	112	0.76	2	0.67	1	0.35	0	0.00
Depression	300	2.03	5	1.67	5	1.77	2	3.45
Epilepsy and other Seizure Disorders	316	2.14	9	3.00	9	3.18	3	5.17
Intellectual Disability	57	0.39	0	0.00	0	0.00	0	0.00
Obsessive-Compulsive Disorder	34	0.23	0	0.00	0	0.00	0	0.00
Schizophrenia	7	0.05	0	0.00	0	0.00	0	0.00
Tourette Syndrome	14	0.09	0	0.00	0	0.00	0	0.00

V. Discussion and Conclusion

This report summarizes the ability of our claims-based ASD case algorithms to accurately identify children with ASD within the research claims databases. We do this by estimating the probability that a subject identified by our claims-based ASD case algorithm as having ASD is in fact a case based on our ‘gold standard’. In general, our claims-based ASD case algorithms identify ASD cases and non-cases fairly accurately. The positive predictive value of our Likely ASD algorithm, when compared to a ‘gold standard’ (Level 1 or 2), is 87.4%. And, as expected, the negative predictive value is high; only one of the 60 Enriched controls was a false negative.

In the rest of the discussion, we seek to provide context for the results presented above, elaborate on these results, and consider the implications of the chart study for Tasks B, C and D. We also address limitations of our study.

A. Comparison to the Literature

We undertook what is, to our knowledge, the first validation study of claims-based ASD case identification in a large sample representative of the insured US population. There has been only one previously published study that attempted to empirically assess the validity of administrative health databases for autism diagnoses.¹⁰ This Canadian study examined the performance of administrative data for ASD case identification in a sample of children referred to an autism specialty clinic. ‘Gold standard’ diagnoses made by a team of trained clinicians were compared to ASD determinations based on seven algorithms derived from combinations of single or multiple claims with an ASD diagnostic code from three administrative databases (i.e., hospital data, physician billing data, and outpatient mental health data). The study calculated the sensitivity and specificity for each of the seven algorithms and concluded that defining “cases” as a single appearance or mention of an ASD diagnosis in any of the three databases resulted in the best combination of sensitivity and specificity: 69.3% and 77.3% respectively.¹⁰

There are a number of potential concerns or issues regarding the applicability of the findings from the Canadian study to the US, privately insured population. First, the administrative data used in Canada is not derived in a private insurance environment so the coding conventions, though based on the same international classification of codes, may be driven by substantively different forces. Second, and probably most important, is that the study population in the Canadian study was drawn from individuals seen at a single ASD specialty clinic. Thus, subjects confirmed as cases receiving specialty care are unlikely to represent all cases in the population. As a result, it could be that sensitivity is high among a referred population but may change if used to identify children with ASD in a general population, many of whom are not unequivocally diagnosed or referred for specialty care. Lastly, in assessing the performance of claims-based ASD case algorithms in large samples for research purposes, arguably the most relevant statistics are positive and negative predictive values, as discussed in detail in Section III.F Analytic Strategy above.

B. Summary of Results

1. Positive Predictive Value

Although we present data for both Level 1 and Level 2 confirmation separately, and Level 1 confirmation is closest to the chart-based criteria used in the CDC ADDM autism surveillance projects, the most useful gold-standard in this study is Level 1 *or* Level 2 confirmed. Given we have

access to medical charts alone (without educational records as included in the ADDM surveillance study) and charts from a single provider covering a limited time period, the likelihood that these charts would be robust enough to provide opportunity for Level 1 confirmation is probably somewhat inherently limited. Therefore, considering we have incomplete information on every subject, requiring only Level 1 confirmation as the ‘gold standard’ may miss many children who have ASD. Level 1 confirmation also appears to be correlated to some extent to the type of provider, as 68.6% of Likely ASD and Possible ASD subjects with Tier 1 provider charts reviewed were Level 1 confirmed while just 28.4% of those with Tier 2 or 3 provider charts reviewed were Level 1 confirmed (based in data from Table 14). We can’t fully know the direction of the causal effect, if any, behind this correlation because subjects with more apparent ASD symptoms may be more likely to find their way to Tier 1 providers, but considering all factors, a focus on Level 1 or Level 2 confirmation seems most reasonable.

The unweighted PPV for the claims-based Likely ASD case algorithm using Level 1 or Level 2 confirmation was 87.4%. The overall weighted PPV, which accounts for different sampling strata by age and claims richness, was very close to the unweighted estimate at 87.3 (estimates were similarly close for all combinations of claims-based ASD case algorithms and confirmation level criteria). Consequently, we are comfortable focusing on interpretation of unweighted PPVs and all PPVs discussed from this point forward are unweighted.

When the claims-based ASD case algorithm is narrowed to include only the Likely ASD cohort, PPV increased from 74.2% to 87.4%. This suggests that caution should be used when relying on a claims-based ASD case algorithms based on the appearance of only one claim with an ASD diagnostic code. Under such a definition, over one quarter of those meeting criteria were not confirmed (Level 1 or Level 2) through the chart review. Consequently, in the remaining analytic Tasks for this project we will evaluate outcomes only for the Likely ASD cohort. The Possible ASD cohort will be held in reserve, and may be used to supplement some particular analyses that have small sample sizes.

PPV was examined in younger versus older subjects because of concerns that charts for older subjects are more distal to the period of intense diagnostic assessment and might contain inadequate information to provide an opportunity to confirm case status. Consistent with this, lower PPVs based on Level 1 confirmation were observed in older subjects (Table 13). Level 1 or Level 2 PPVs were slightly higher for older children, likely reflecting the fact that older children have had more opportunity to acquire an ASD diagnosis, mention of which is sufficient to meet confirmation criteria under Level 2. Level 1 or Level 2 PPV for Likely ASD in younger subjects was 86.3% and 91.7% in older subjects.

Because claims-based ASD case algorithms may also be associated with length of enrollment (with longer-enrolled individuals essentially under surveillance longer with greater opportunity for claims with ASD diagnostic codes to be observed), we explored PPV in subjects with different enrollment period lengths. In general, PPV was similar across subjects enrolled <18 months and ≥18 months suggesting that “surveillance bias” around claims-based ASD case identification is not strong. Level 1 or Level 2 PPV for the Likely or Possible claims-based ASD case algorithm was 79.6% in the shorter-enrolled and 71.2% and 73.4% in the two longer enrolled groups, with substantial overlap between 95% CIs on the PPVs.

2. Negative Predictive Value

As mentioned, validation studies face challenges in terms of precisely measuring NPV due to the low expected frequency of confirmed ASD cases in a random sample of subjects without a claim with an ASD diagnostic code. At the same time this does imply that the operating expectation is that NPV for a claims-based algorithm is most likely quite high and, consequently, samples of subjects without a claim with an ASD diagnostic code probably do represent non-ASD populations very well. Our check on this assumption of high NPV was to review charts in a small sample of subjects without a claim with an ASD diagnostic code who had a claim with an ASD-associated condition diagnostic code. As reported above, in this Enriched control sample of 60 subjects only 1 met criteria at Level 2, which supports the assumption that NPV of claims-based ASD case algorithms will likely be high. It is, however, possible that our selection criteria for the Enriched cohort did not accurately identify those most likely to be false negatives via claims-based ASD case algorithms. In general, we saw among Enriched controls a plausible range of neurodevelopmentally-related codes including ADHD, depression and epilepsy. Few had intellectual impairment – but this is known to be under-reported in administrative data – especially data based on health care data only.¹¹

3. Potential to Augment Claims-based ASD Case Algorithm

As described above, we also undertook analyses to explore the potential to improve PPV by augmenting a claims-based ASD case algorithm approach with claims information beyond ASD ICD-9-CM codes. Based on the conditions and medication use variables descriptively presented in our analysis, we have no clear recommendation on the augmentation of claims-based ASD case algorithms beyond ASD ICD-9-CM codes. However, there are some points of note. First, risperidone was more often seen among false positives than true positives (approaching statistical significance in the Likely ASD group). This association between false positive status and risperidone use may be related to the use of atypical antipsychotics, and risperidone specifically, for ASD-associated conditions or indications as well as ASD itself. Although we did not incorporate risperidone use into our claims-based case definition for the chart review study, we did incorporate it into the definition for Task A (a subject with one claim with an ASD diagnostic code and one prescription for risperidone was included in the Likely ASD group). This observation in the validation sample suggests that, perhaps, incorporation of risperidone use as a rule-in code for the Likely ASD group may have been ill-advised. Only 3.4% (1,189 of 34,754) of the Likely ASD cases in the Task A sample were flagged as Likely ASD based on only one claim with an ASD diagnostic code and a prescription for risperidone use. Nonetheless, we have decided, for analyses in subsequent Tasks, to reclassify these children as Possible ASD. Further, we would suggest caution in the incorporation of risperidone use in other claims-based ASD case algorithms.

4. Generalizability

Overall, the final samples do differ with respect to some demographic and clinical characteristics than the eligible population. Specifically, the final samples over-represented the South. This was especially true for the Likely ASD group (43% for eligible sample vs. 57% for the final chart sample) and Enriched control groups (48% for eligible sample vs. 66% for the final chart sample). Our final Likely ASD and Enriched control sample also included a greater proportion of female subjects than the eligible population. In addition, the proportion of the samples with specific behavioral health comorbidities also differed between the eligible samples and the final study

samples. The final samples for all three cohorts had a higher proportion of children with epilepsy than the eligible samples. The Likely and Possible ASD final samples also had a lower proportion of anxiety and depression. The Likely ASD final sample had a lower proportion of children with ADHD than the eligible sample (6.7% vs. 13.0%).

Regional differences in the presence of behavioral health comorbidities were introduced when we drew our initial sample pool of 2400 eligible subjects for potential chart review based on the two sampling strata presented in Table 2. Differences in gender were generally introduced at the chart procurement stage. As discussed above, when drawing our initial sample pool, we planned to over-sample children younger than age 8 and subjects with longer enrollment and richer claims. So, as expected, our final sample is younger but this oversampling would not be expected to drive differences in gender. Even so, overall, the chart study cohort seems to reasonably represent the eligible population in our claims database.

The eligible population, however, is also a select group of source population given our initial data source (e.g., children enrolled in a particular health plan). A comparison of our study's eligible population to the US population, as well as to a national sample of children with ASD from the National Survey of Children's Health (NSCH) 2007 was conducted as part of the Task A Claims Analyses Report. We found that while the eligible population for this study is similar to the population of privately insured individuals, the data are not representative of the entire US population of children with ASD. The results of the complete comparison are included in Appendix F.

C. Strengths and Limitations

As with any research, strengths and limitations apply. In this section, we summarize the main considerations.

1. Strengths

a. Sample size

The analyses from medical charts re-emphasize the critical importance of the scale of the study populations we are able to bring to bear on the research hypotheses of Tasks A, B, C and D. As discussed in our companion report on the Task A claims analysis, our data provide a very large, national database of children with ASD, their siblings and parents, with counts for children with ASD exceeding 80,000 (across the OptumInsight and Impact data sets) and over 57,000 associated siblings and 80,000 associated parents of children with ASD (OptumInsight only). We are unaware of any other studies, with the exception of certain Medicaid claims-based studies, which have drawn on such a large and diverse population of children with ASD,^{§§§§} while studies of family members of children with ASD have relied on far smaller cohorts, typically fewer than 300 individuals.¹²

This large sample size is accompanied by a breadth of age distribution with about 13% of the children with ASD aged 0-1 and 60% aged 2-10. Further, over one-third have continuous enrollment of 48 months or greater.

^{§§§§} Other studies have had fewer than 60,000 subjects (Mandel, Shimabukuru, Croen, Liptak, Birenbaum, Flanders, Leslie)

Below we discuss the limitations of the data, but to a significant degree, those limitations may be addressable within the context of the large starting population. For example, to the extent that geographic mix inhibits generalizability, our starting population is sufficiently large that judicious re-weighting by geographic cell may overcome this limitation (to be confirmed by further investigation).

Further, the large starting population, together with the level of resources made available by NIMH, supported the selection of a large and diverse group of subjects for the chart review. The size and diversity of these subjects in turn enabled us to look at a variety of dimensions in the abstraction process, such as provider characteristics, length of observation period, volume of claims, and of course, the multiple claims-based ASD case algorithms. While not a part of the scope of the current study, the detail available from the large chart sample may be useful in the development of further refinements to the claims-based ASD case algorithm (such as a statistical predictive model for claims-based ASD identification). This aspect may be favorably compared to the limited scope of other studies to test claims-based ASD case algorithms against medical charts.^{13 14 15 16 17 18}

In addition, the large chart sample supported the various analyses of the robustness of the PPV estimates reported above, giving us confidence in our approach, as well as revealing potential ways for further refining our methods for Tasks B, C and D.

b. PPV and NPV

The high and robust PPVs give us considerable confidence to proceed with Tasks B, C and D, as well as buttressing our conclusions in the Task A claims analysis. While based on the results of this study we may elect to proceed differently in our approaches for Tasks B, C and D, especially in considering some further refinements to cohort selection criteria, we are starting from a good foundation of a large population of children with ASD. And it is unlikely we have left behind, in our selection criteria, significant segments of the population of children with ASD. These benefits are in a sense, double-barreled, since parents and siblings of these children are also the subject of the research – this increases our confidence in our algorithms for the children and increases our confidence that these family members are indeed family members to a child with ASD.

2. Limitations

a. Imperfect 'gold standard' or its application

The clinical review of medical charts that served as the 'gold standard' for this study is an imperfect 'gold standard' due to a few factors.

First, due to the logistics of accessing medical charts for subjects in multiple locations, it was not feasible to access a child's entire medical chart. This study was limited to acquiring a single medical chart from a single provider confined to the period of 6 months prior to and 30 months after the first claim with an ASD diagnostic code in the claims database. Because only one medical chart could be abstracted from a single provider for each child, the length, quantity of information and quality of the medical chart varied greatly. This was the main impetus for our approach to the implementation of a minimal threshold for an informative chart (Section III.D.4).

Nevertheless, it remains possible that had a chart for another provider treating an unconfirmed case been accessed, it could have supported either confirmation or rule-out.

Second, this study did not account for the length of medical chart available for review. As previously stated, up to 36 months of chart information for each subject was reviewed for this study. However, some charts had less than 36 months of information available. Since the length of medical chart may be related to the amount of information available it may also be related to the likelihood of an ASD diagnosis and the degree of certainty of that diagnosis.

b. Other potential bias in PPV calculation

Certain factors that are difficult to measure could result in bias of the calculated PPV. Some observers contend that diagnoses presented on claims, and perhaps those recorded in charts are impacted by the payment process. Establishing the eligibility of a child for special programs, or avoiding stigmatizing a child, could cause a provider to submit claims that are in line with information recorded in charts, but not truly representative of the child's clinical status. Consequently, the direction of such potential bias is not clear. And such bias is not measurable by tools available to us.

c. Generalizability

Overall, the final samples do differ on some demographic and clinical characteristics from the study population in the Task A claims analysis as well as the starting sample meeting the selection criteria for this chart study. This may affect the generalizability of our results. As previously stated, most of these difference are clearly due to our sampling strategy (outlined in Table 2). Of course, this should not obscure that the population size and granularity of our data enables this detailed level of comparison, which might not be apparent in other studies. Given the multiple factors at play and that the weighted and unweighted PPVs are similar, we believe that the differences in the characteristics of children observed at various points in the sampling process are not of large concern for Tasks B, C and D. Nevertheless, we must keep in mind these differences as we develop the detailed methods for these Tasks and report out findings.

D. Implications and Recommendations

Overall, the chart study results endorse the ability to use claims data for research about ASD in children and associated health outcomes and utilization. Claims data is able to identify children who have actually been validly diagnosed with ASD, which may also prove useful for research about the etiology of ASD and the role of claims-based risk factors. Furthermore, these findings support using large, existing claims databases in general, particularly important for poorly understood and heterogeneous conditions such as ASD, thus paving the way to greater knowledge of this oft-disabling condition in an efficient, and timely way without additional burden to children and families.

In addition, the results of this chart study have implications for project Tasks B, C, and D:

- We will revise the Likely ASD criteria to only include children with two or more claims with an ASD diagnostic code. The presence of a risperidone prescription and one claim with an ASD diagnostic code will no longer be considered in the Likely criteria. Possible ASD criterion would remain one claim with an ASD diagnostic code. As a result of this change, those children who were categorized as Likely ASD using the criterion of one ASD diagnostic code and a risperidone prescription (3.4% of Likely ASD cohort; 1,189 of

34,754 study subjects) would be re-categorized as Possible ASD. The total number of children identified as having ASD would be unchanged.

- We would discourage future claims-based ASD case algorithms from incorporating risperidone unless further work is done on the ability of risperidone use to identify an ASD diagnosis.
- The PPV increases from 74.2% to 87.4% when the Possible ASD cohort is not included in the case definition. Consequently, in the remaining analytic Tasks for this project we will primarily use the Likely ASD cohort for analyses. The Possible ASD cohort will be held in reserve and may be used to supplement particular subgroup analyses that have small sample sizes.

In summary, we will primarily use the two-claim (Likely ASD) claims-based case algorithm based on the presence of ICD-9 codes to identify children with ASD and their family members in Tasks B, C, and D. This is because of the differences in the PPV of the two algorithms as well as the differences in demographic and clinical conditions reported in Task A Claims Study Report delivered to NIMH on October 17, 2011.

VI. References

- ¹ Leslie DL, Martin A. Health care expenditures associated with autism spectrum disorders. *Arch Pediatr Adolesc Med* 2007;161(4):350-5.
- ² Croen LA, Najjar DV, Ray GT, Lotspeich L, Bernal P. A comparison of health care utilization and costs of children with and without autism spectrum disorders in a large group-model health plan. *Pediatrics* 2006;118(4):e1203-e1211.
- ³ Mandell DS, Morales KH, Marcus SC, Stahmer AC, Doshi J, Polsky DE. Psychotropic medication use among Medicaid-enrolled children with autism spectrum disorders. *Pediatrics* 2008;121(3):e441-e448.
- ⁴ Mandell DS, Cao J, Ittenbach R, Pinto-Martin J. Medicaid expenditures for children with autistic spectrum disorders: 1994 to 1999. *J Autism Dev Disord* 2006;36(4):475-85.
- ⁵ Wang, Li, Leslie, Douglas. Health Care Expenditures for Children With Autism Spectrum Disorders in Medicaid. *Journal of the American Academy of Child and Adolescent Psychiatry* 2010; 49(11):1165-1171.
- ⁶ Centers for Medicare & Medicaid Services. "Medicare Claims Processing Manual, Chapter 26: Completing and Processing Form CMS-1500 Data Set." Available at: <https://www.cms.gov/manuals/downloads/clm104c26.pdf>. Accessed September 7, 2011.
- ⁷ Cecil G. Sheps Center for Health Services Research. "Implementation of the UB-04." Available at: http://www.shepscenter.unc.edu/research_programs/hosp_discharge/links/ub04_fact_sheet.pdf. Accessed September 7, 2011.
- ⁸ National Uniform Billing Committee. "History of the NUBC." Available at: <http://www.nubc.org/history.html>. Accessed September 7, 2011.
- ⁹ Gearing RE, Mian IA, Barber J, Ickowicz A. A methodology for conducting retrospective chart review research in child and adolescent psychiatry. *J Can Acad Child Adolesc Psychiatry* 2006;15(3):126-34.
- ¹⁰ Dodds L, Spencer A, Shea S, et al. Validity of autism diagnoses using administrative health data. *Chronic Dis Can* 2009;29(3):102-7.
- ¹¹ Ouellette-Kuntz H, Shooshtari S, Temple B, et al. Estimating Administrative Prevalence Of Intellectual Disabilities in Manitoba. *Journal on Developmental Disabilities* 2009;15(3):69-80.
- ¹² Gray, j *Intell Dis* 2006:970-6; Lecavalier, J *Intell Dis* 2006: 172-83; Metzler, J *Ped Psych* 2008:380-6; Autism Speaks is at the time of this writing collecting self-report data from parents on their website with allegedly about 4,000 participants: <http://www.autismspeaks.org/news/news-item/ian-research-report-family-stress-%E2%80%94-part-1>) accessed August 17, 2011.
- ¹³ Andrade SE, Graham DJ, Staffa JA, et al. Health plan administrative databases can efficiently identify serious myopathy and rhabdomyolysis. *J Clin Epidemiol* 2005;58(2):171-4.
- ¹⁴ Wahl PM, Rodgers K, Schneeweiss S, et al. Validation of claims-based diagnostic and procedure codes for cardiovascular and gastrointestinal serious adverse events in a commercially-insured population. *Pharmacoepidemiol Drug Saf* 2010;19(6):596-603.
- ¹⁵ Spettell CM, Wall TC, Allison J, et al. Identifying physician-recognized depression from administrative data: consequences for quality measurement. *Health Serv Res* 2003;38(4):1081-102.
- ¹⁶ Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. *J Clin Epidemiol* 2004;57(2):131-41.
- ¹⁷ Mapel DW, Frost FJ, Hurley JS, et al. An algorithm for the identification of undiagnosed COPD cases using administrative claims data. *J Manag Care Pharm* 2006;12(6):457-65.
- ¹⁸ Winkelmayr WC, Schneeweiss S, Mogun H, Patrick AR, Avorn J, Solomon DH. Identification of individuals with CKD from Medicare claims data: a validation study. *Am J Kidney Dis* 2005;46(2):225-32.