Abt

Abt Associates Inc.

The Study of Thimerosal and Autism

Documentation and Codebook for the Child Vaccination Histories File:

Cambridge, MA
Lexington, MA
Hadley, MA
Bethesda, MD
Chicago, IL

September 3, 2010

Prepared for
Frank DeStefano
National Immunization
Program
Centers for Disease Control
and Prevention
Atlanta, GA 30329

Prepared by
Cristofer Price
Yeqin He

Abt Associates Inc.
Suite 800 North
4550 Montgomery Avenue
Bethesda, MD 20814-3343

This documentation was prepared by Cristofer Price and Yeqin He of Abt Associates Inc. for the Immunization Safety Office (ISO) of the Centers for Disease Control and Prevention (CDC) Atlanta, GA 30333. Questions about the documentation, substantive questions, or technical issues regarding the data file should be directed to the CDC ISO, MS D26, 1600 Clifton Road, Atlanta, Georgia 30333 (404-639-8256).

Suggested Citation for Study of Thimerosal and Autism Public Use Data Set:

Price, C.S., He, Y. (2010) The Study of Thimerosal and Autism: Data Set Documentation. Bethesda MD: Abt Associates, Inc; Prepared for the National Immunization Program Centers for Disease Control and Prevention Atlanta, GA

Documentation and Codebook
for the
Child Vaccination Histories File

1. Introduction to the Child Vaccination Histories File

The Child Vaccination Histories File contains the data that were used to construct the measures of early childhood (postnatal) exposure to ethylmercury from thimerosal-containing vaccines and immune globulin preparations received by the study children during the age range spanning birth to 20 months. The Child Vaccination Histories File is provided with the public use data in order to make the calculation of exposure amounts transparent, and so that analysts have the potential to calculate alternative measures of exposure[1]. For a full description of the creation of the measures prenatal exposure to ethylmercury from thimerosal-containing vaccines and immune globulin preparations received children's mothers during their pregnancies with the study children, see Section 7.3 of the Technical Report (Volume 1).

The Child Vaccination Histories File contains the vaccination histories of the n=1,095 children in the analysis data set[2]. Each row of the Child Vaccination Histories File represents a record of a vaccine received on a particular day. Thus, the file has many records per child. The file includes each child's "resolved vaccine history", and also includes the raw, original, non-cleaned vaccine data from each of two data sources. The resolved vaccine histories were obtained from cleaning the raw, original data and resolving any discrepancies among the two data sources and any discrepancies between the records and recommended childhood vaccination schedules. Data cleaning procedures are described in Section 7.3 of the Technical Report (Volume 1).

All previous analyses were, and any future analyses should be based on resolved vaccine histories. The raw, non-cleaned records of vaccination receipts are known to contain errors. The primary rationale for providing the raw data along with the resolved histories is for documentation purposes and transparency. Data users can compare the raw data to the resolved vaccine histories to gain a better understanding of the data cleaning process

---

[1] For example, a future analyst might want to create a measure that cumulate exposure over some time period other than the ones used for the current study.
[2] See Documentation and Codebook for Main Analysis File

that was used, and to make their own judgments as to the appropriateness of the techniques.

The resolved vaccine histories include only vaccine receipts during the age range spanning from birth to two years. Although the raw, non-cleaned records include records of receipts at older ages, the data cleaning process focused exclusively on receipts that occurred between birth and age two-years. Therefore, if any analyses in the future were to be based on vaccines received by children older than two years, then the analysts would have to clean the raw data using procedures similar to those described Section 7.3 of the Technical Report (Volume 1).

2. File Formats and Variable Descriptions

The Childhood Vaccination Histories File is provided in three formats: 1) ASCII text format, 2) SAS transport file, and 3) Excel spreadsheet. For analysis purposes, the first two formats are recommended. The excel spreadsheet is provided because it is in a convenient format for visual inspection of the vaccine histories. There are multiple records per ID on this data set. There are 1,095 unique ChildIDs in the data set.

Exhibit 1.
Variables Included in the Child Vaccination Histories File

Variable	Type	Label	Source[a] Med. Chrt.	Source[a] Comp. Aut.	Notes[b]
childID	Num	ChildID for Merging Data Sets			
Res_Vacdays1	Num	Resolved: Age in days at vac receipt	X	X	See Vol 1 Section 7.3.1, Exhibit 7.3.1.1
Res_VacType	Char	Resolved: Type of vaccine received	X	X	See Vol 1 Section 7.3.1, Exhibit 7.3.1.1
Res_MFR	Char	Resolved: Vaccine manufacturer	X	X	See Vol 1 Section 7.3.1, Exhibit 7.3.1.1
res_year	Num	Resolved: Year Vac Recieved	X	X	See Vol 1 Section 7.3.4, Exhibit 7.3.4.1
MercAmt	Num	Mercury amount (in vac receipt)			See Vol 1 Section 7.3.4, Exhibit 7.3.4.1
RecptWtKG1	Num	Chld weight in KGs at time of vac receipt	X		
Amt_wt1	Num	Merc amount / Weight in KGs at vac receipt	X	X	
HepBPolio_R1	Num	Cleaning rule: HepB/Polio Rule 1			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
HepBPolio_R2	Num	Cleaning rule: HepB/Polio Rule 2			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
DTPHib_R1	Num	Cleaning rule: DTP - HIB Rule 1			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
DTPHib_R2	Num	Cleaning rule: DTP - HIB Rule 2			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
BadDate	Num	Cleaning rule: Bad date			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
Lookup	Num	Cleaning rule: Special case lookup			See Vol 1 Section 7.3.3, Exhibit 7.3.3.1
Ch_VacDays1	Num	Chart: Age in days at vac receipt	X		
Ch_Vactype	Char	Chart: Type of vaccine received	X		
Ch_VacText	Char	Chart: Original text on vaccine type	X		
Ch_Mfr	Char	Chart: Manufacturer	X		
Ch_Lot	Char	Chart: Lot number	X		
Cmptr_VacDays1	Num	Computer-automated: Age in days at vac receipt		X	
Cmptr_VacType	Char	Computer-automated: Type of vaccine received		X	
Cmptr_VacCode	Char	Computer-automated: Vaccine code		X	Vaccine Safety Datalink (VSD) codes for vaccines received.

Exhibit 1.
Variables Included in the Child Vaccination Histories File

Variable	Type	Label	Source[a] Med. Chrt.	Source[a] Comp. Aut.	Notes[b]
					See Exhibit 2, below.
Cmptr_mfr	Char	Computer-automated: Manufacturer		X	
Cmptr_lot	Char	Computer-automated: lot number		X	
SortDays1	Num	Age in days (used to sort file)	X	X	Sort data set by ChildID then Sortdays1 to have most useful layout for viewing vaccine history

[a] Data Source: Medical chart abstraction, or computer automated data set (VSD data).
[b] Vol I and section numbers correspond to Technical Report Volumes I (Price et al, 2009), which is included with this data set.

Exhibit 2
Codes Used in Computer-automated Data Set to Indicate Vaccine Type

Code	Vaccine Type
1	01=diphtheria tetanus toxoids and pertussis
2	02=poliovirus , live, oral
3	03=measles mumps and rubella virus
4	04=measles and rubella virus
5	05=measles virus
6	06=rubella virus
7	07=mumps virus
8	08=hepatitis B , pediatric or pediatric/adolescent dosage
9	09=tetanus and diphtheria toxoids, adsorbedfor adult use
10	10=poliovirus , inactivated
11	11=pertussis
13	13=tetanus immune globulin
14	14=immune globulin, NOS
15	15=influenza virus , split virus (incl. purified surface antigen)
16	16=influenza virus , whole virus
17	17=Hib, conjugate NOS
18	18=rabies , for intramuscular injection
20	20=DTaP (diphth., tet. andacel. pert.)
21	21=varicella virus
22	22 = DTP-Hib
23	23=plague
25	25=typhoid , live, oral
28	28 = DT (peds) (diphtheria and tetanus)
30	30=hepatitis B immune globulin
31	31=hepatitis A , pediatric dosage, NOS
32	32=meningococcal polysaccharide (MPSV4)
33	33=pneumococcal polysaccharide
34	34=rabies immune globulin
35	35=tetanus toxoid, adsorbed
36	36=varicella zoster immune globulin
37	37=yellow fever
38	38=rubella and mumps virus
39	39=Japanese encephalitis
40	40=rabies , for intradermal injection
41	41=typhoid , parenteral, other than acetone-killed, dried
42	42=hepatitis B , adolescent/high riskinfant dosage
43	43=hepatitis B , adult dosage
44	44=hepatitis B , dialysis patient dosage
45	45=hepatitis B , NOS
46	46=Hib, PRP-D conjugate
47	47=Hib, HbOC conjugate
48	48=Hib, PRP-T conjugate
49	49=Hib, PRP-OMP conjugate
51	51=Hib, conjugate and Hepatitis B

Exhibit 2
Codes Used in Computer-automated Data Set to Indicate Vaccine Type

Code	Vaccine Type
52	52=hepatitis A , adult dosage
53	53=typhoid , parenteral, acetone-killed,dried (U.S. military)
57	57=hantavirus
58	58=hepatitis C
59	59=hepatitis E
60	60=herpes simplex virus, type 2
61	61=human immunodeficiency virus
62	62=human papilloma virus , quadrivalent
63	63=Junin virus
64	64=leishmaniasis
65	65=leprosy
66	66=Lyme disease
67	67=malaria
68	68=melanoma
69	69=parainfluenza-3 virus
70	70=Q fever
71	71=respiratory syncytial virus immune globulin, intravenous
72	72=rheumatic fever
73	73=Rift Valley fever
74	74=rotavirus, live, tetravalent
75	75=vaccinia (smallpox)
76	76=Staphylococcus bacteriophage lysate
77	77=tick-borne encephalitis
78	78=tularemia
79	79=vaccinia immune globulin
80	80=Venezuelan equine encephalitis,live, attenuated
81	81=Venezuelan equine encephalitis, inactivated
83	83=hepatitis A , pediatric/adolescent dosage,2 dose schedule
83R	83R = HepA
84	84=hepatitis A , pediatric/adolescent dosage,3 dose schedule
85	85=hepatitis A , NOS
86	86=immune globulin, intramuscular
87	87=immune globulin, intravenous
88	88=influenza virus , NOS
89	89=poliovirus , NOS
90	90=rabies , NOS
91	91=typhoid , NOS
92	92=Venezuelan equine encephalitis , NOS
93	93=respiratory syncytial virus monoclonal antibody (palivizumab), intramuscular
94	94=measles, mumps, rubella, andvaricella virus
99	99=RESERVED - do not use
100	100=pneumococcal conjugate , polyvalent
101	101=typhoid Vi capsular polysaccharide
102	102=DTP- Haemophilus influenzaetype b conjugate and hepatitis b
103	103=meningococcal C conjugate
104	104=hepatitis A and hepatitis B

Exhibit 2
Codes Used in Computer-automated Data Set to Indicate Vaccine Type

Code	Vaccine Type
105	105=vaccinia (smallpox) , diluted
108	108=meningococcal , NOS
109	109=pneumococcal , NOS
110	110 = DTaP-Hep B-IPV (DTaP hep poliov)
111	111=influenza virus , live, attenuated,for intranasal use
112	112=tetanus toxoid, NOS
113	113=tetanus and diphtheria toxoids,adsorbed, preservative free, for adult use
114	114=meningococcal polysaccharide (groupsA, C, Y and W-135) diphtheria toxoidconjugate (MCV4)
115	115=tetanus toxoid, reduced diphtheria toxoid, and acellular pertussis , adsorbed
116	116=rotavirus, live, pentavalent
117	117=varicella zoster immune globulin (Investigational New Drug)
118	118=human papilloma virus , bivalent
119	119=rotavirus, live, monovalent
121	121=zoster vaccine, live
122	122=rotavirus, NOS
998	998=no administered
999	999=unknown or immune globulin
HBP	HBP = Hib polysaccharide
MM	MM = Measles/Mumps
DHB	DHB = DTaP-HepB
X02	X02 = Experimental DTaP (Acelimune) (ug=25)
X03	X03 = Experimental EXPHDTP (hemoB,DTP) (ug=25)
X10	X10 = Experimental Meningicoccal (ug=0)

Codes obtained from: VSD Dynamic Data File (DDF) Dictionary Vaccine Safety Data Link Project (Revision date: 8-3-2006)