

Use of Laboratory Test Data to Describe Histoplasmosis Epidemiology in the United States

Joann L. Cloud^{1,2}, Andrew Wilson², Marc Roger Couturier³, Talmage M. Holmes¹

¹Department of Public Health, Walden University, 155 5th Ave S. Minneapolis, MN 55401; ²ARUP Institute For Clinical and Experimental Pathology, 500 Chipeta Way, Salt Lake City, UT 84108; ³Department of Pathology, University of Utah, ARUP Laboratories, 500 Chipeta Way, Salt Lake City, UT 84108

Revised Abstract

Histoplasmosis can be a devastating invasive fungal disease caused by *Histoplasma capsulatum*. Twenty-six states are considered endemic for *H. capsulatum*, as defined using skin testing data collected from US military recruits between 1958-1969. Today, histoplasmosis surveillance is a notifiable disease in only 11 states, limiting our knowledge of histoplasmosis epidemiology. We explored the use of *Histoplasma* antigen laboratory test (HALT) data to describe the epidemiology of histoplasmosis. HALT data were extracted from the laboratory information system of ARUP Laboratories for the years 2007, 2008, and 2009. Data were cleaned and analyzed using version 9.2 of the SAS System (SAS Institute Inc, Cary, NC, USA). Positivity proportions and binomial distribution estimates (BDE) for HALT were calculated using exact binomial methods. We assessed the precision of histoplasmosis proportion estimates for each state using 99% upper confidence limits. HALT data were available for 45 states (n=31201) with TX, CA, OH, AR, and KY, constituting >15,000 tests. For the years 2007-2009, 14 states provided high resolution estimates (BDE<0.0573) of histoplasmosis prevalence. The states were ranked in order of prevalence. Most cases of histoplasmosis consistently occurred in the age group 40-49 years. This study was delimited to HALT data from a single reference laboratory. With complete laboratory data from all labs in the US, HALT data may be a useful source for histoplasmosis surveillance.

Introduction

Histoplasmosis is an invasive fungal disease caused by *Histoplasma capsulatum* associated with certain geographic regions, including the Ohio and Mississippi River valleys of the United States. The fungus thrives in soil with high nitrogen content, such as in regions of bird roosts and bat habitats. Upon ground disruption, fungal spores become airborne, subjecting humans and animals to infection through inhalation. After inhalation into the lungs, the fungus changes from a mycelial (i.e. mold) phase to a yeast form that can readily enter the blood stream and disseminate to any part of the body. Histoplasmosis is an opportunistic disease and immunocompromised individuals are at greater risk for becoming ill after exposure.

No universal histoplasmosis surveillance system exists for the United States. Some individual states receive reported cases and report annual summaries, but many of the endemic states do not track histoplasmosis cases. Nonendemic states would benefit from a histoplasmosis surveillance system as well, because of the high rate of travel and the high number of immunocompromised persons.

This study aims to describe the epidemiology of histoplasmosis using *Histoplasma* antigen laboratory test (HALT) data. The NIAID Mycoses Study Group states histoplasmosis cases are defined by laboratory confirmation (De Pauw et al., 2008). The HALT is the most common laboratory test used in conjunction with the clinical description to diagnose histoplasmosis.

This study was approved by the University of Utah IRB and the Walden University IRB.

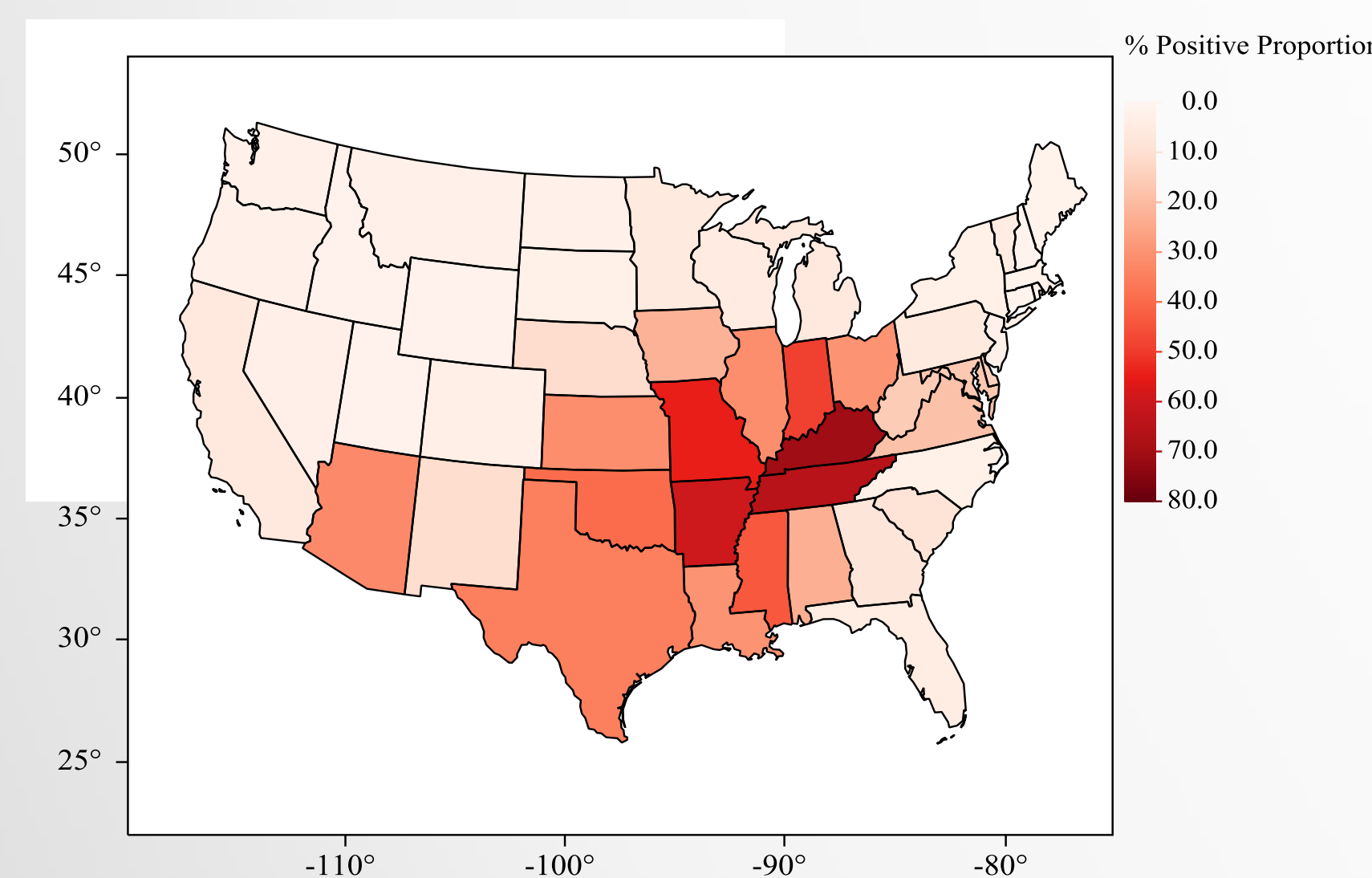


Figure 1. Histoplasmosis exposure study using skin tests performed on US Navy recruits, ages 17-21, during the years 1958-1969. This map was generated using the data published by Edwards et al. (1973) and shows the degree of positive proportions (% positive) of skin testing by state.

Methods

Study Period

- Study period includes years 2007, 2008, and 2009

Data Sources

Histoplasma Antigen Laboratory Test (HALT) Data

Data Collection

- HALT data were extracted from the database of ARUP Laboratories, a clinical reference laboratory
 - Patient identification
 - Histoplasma Antigen Test Result
 - Age
 - Gender
 - Geographic location
- Data were cleaned and deduplicated using SAS, version 9.2 (SAS Institute Inc., Cary, NC, USA)
 - Multiple patients – one per year
 - Mixture of positive & negative results- 1 positive counted
 - Quality control eliminated
 - Human data only (animals eliminated)

Data Analysis

- Data were compiled into a spreadsheet (Microsoft Excel, 2010)
- Data were uploaded into the JMP Pro 9 software (SAS, Inc.)
- Positive proportions (per state) were calculated
- Binomial distribution estimates (per state) were calculated
- States with stable HALT data were determined
 - Binomial distribution estimates weighted by total tests
 - Cutoff established at upper 99% confidence limit
- Individual cases were evaluated for positive proportions of HALT by
 - Age
 - Gender
 - Geographic distribution

Notifiable Disease Data

Data Collection

- States with notifiable disease data were identified using the Council of State and Territorial Epidemiologists (CSTE) website (<http://www.cste.org/>)
- Of these states, states in common with those with stable HALT data were noted
- Individual state department of health websites were visited for data collection
- Histoplasmosis counts and rates were recorded

Data Analysis

- Data were compiled into a spreadsheet (Microsoft Excel, 2010)
 - Total numbers of cases per year
 - Rate per 100,000 population per year
 - Rate by gender
 - Rate by age

Histoplasma Antigen Laboratory Test (HALT) Data

Table 1. HALT Data by State

State	Number of Positive HALT	Total Number of HALT	HALT Positive Proportion	Binomial Distribution Estimate
AL	4	278	0.0144	0.6288
AR	59	2173	0.0272	0.5346
AZ	14	657	0.0213	0.5704
CA	58	4210	0.0138	0.5348
CO	6	373	0.0161	0.6063
CT	1	39	0.0256	0.7358
FL	18	1230	0.0146	0.5622
GA	8	792	0.0101	0.5925
HI	0	83	0.0000	1.0000
IA	5	282	0.0177	0.6160
ID	1	25	0.0400	0.7358
IL	34	1609	0.0211	0.5454
IN	9	593	0.0152	0.5874
KS	2	40	0.0500	0.6767
KY	25	1787	0.0140	0.5529
LA	6	314	0.0191	0.6063
MA	14	1033	0.0136	0.5704
MD	0	38	0.0000	1.0000
ME	0	68	0.0000	1.0000
MI	8	489	0.0164	0.5926
MN	0	13	0.0000	1.0000
MO	11	187	0.0588	0.5793
MS	0	6	0.0000	1.0000
MT	0	4	0.0000	1.0000
NC	5	714	0.0070	0.6160
ND	0	17	0.0000	1.0000
NE	17	755	0.0225	0.5640
NJ	28	350	0.0800	0.5501
NM	4	390	0.0103	0.6288
NV	0	50	0.0000	1.0000
NY	17	976	0.0177	0.5348
OH	42	2214	0.0190	0.5409
OK	1	84	0.0119	0.7358
OR	5	386	0.0130	0.6160
PA	3	186	0.0161	0.6472
SC	8	519	0.0154	0.5926
SD	5	238	0.0210	0.6160
TN	21	1189	0.0177	0.5577
TX	170	5496	0.0309	0.5204
UT	8	273	0.0293	0.5926
VA	1	210	0.0048	0.7358
WA	2	37	0.0541	0.6767
WI	14	746	0.0188	0.5704
WV	2	43	0.0465	0.6767
WY	0	5	0.0000	1.0000

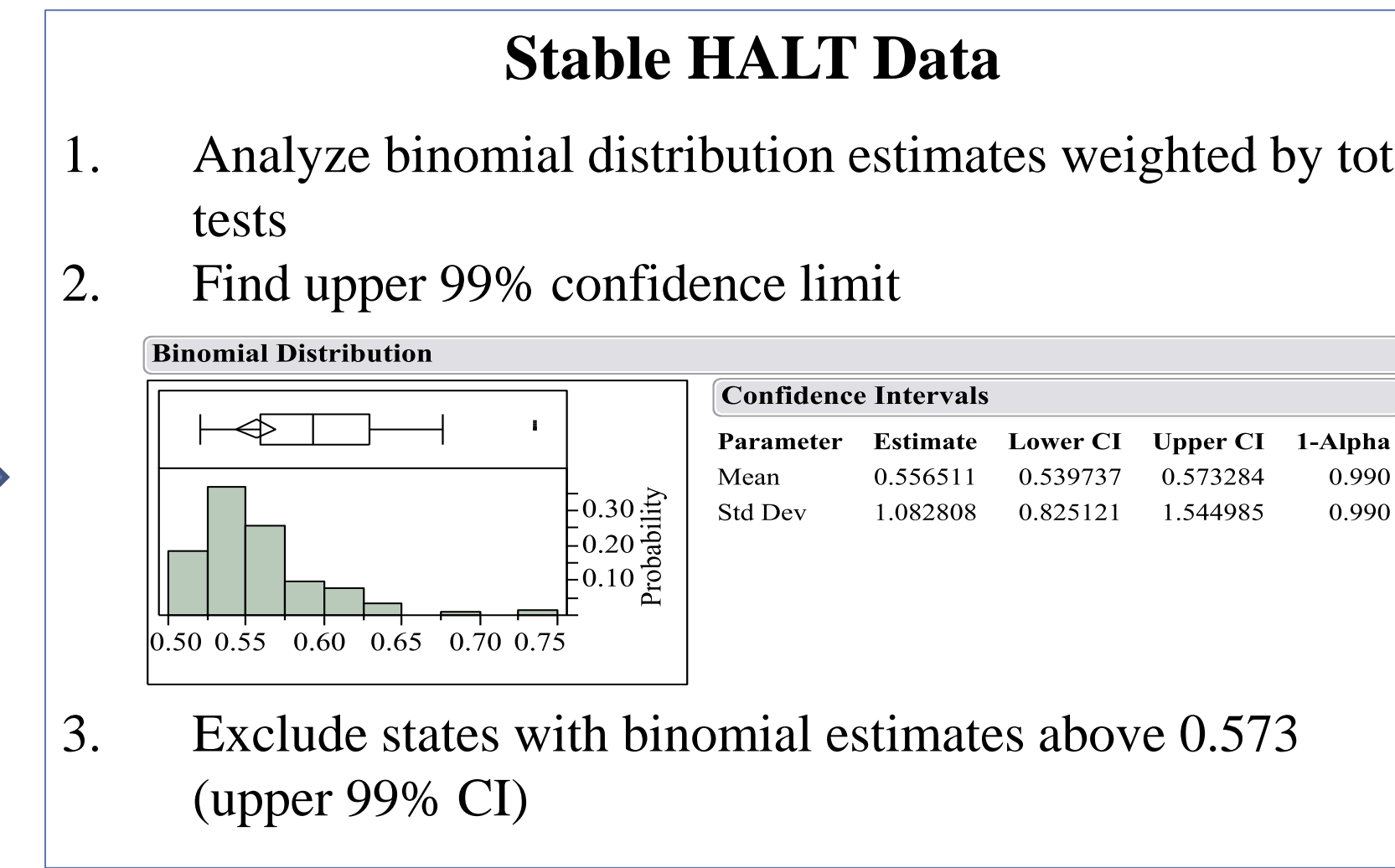


Figure 2. Generating Stable HALT Data

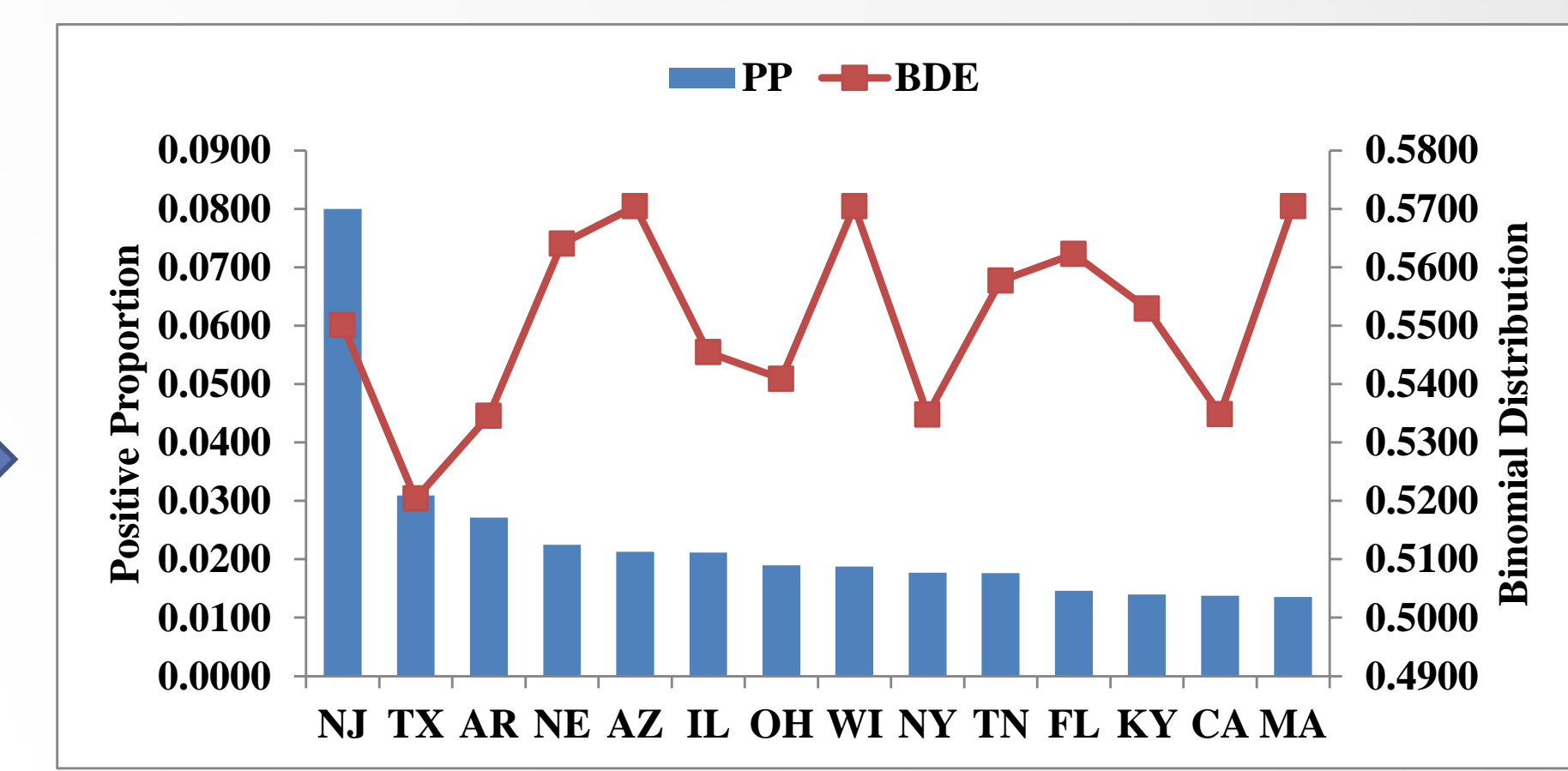


Figure 3. States with Stable HALT Data
Note: PP, Positive Proportion; BDE, binomial distribution estimate

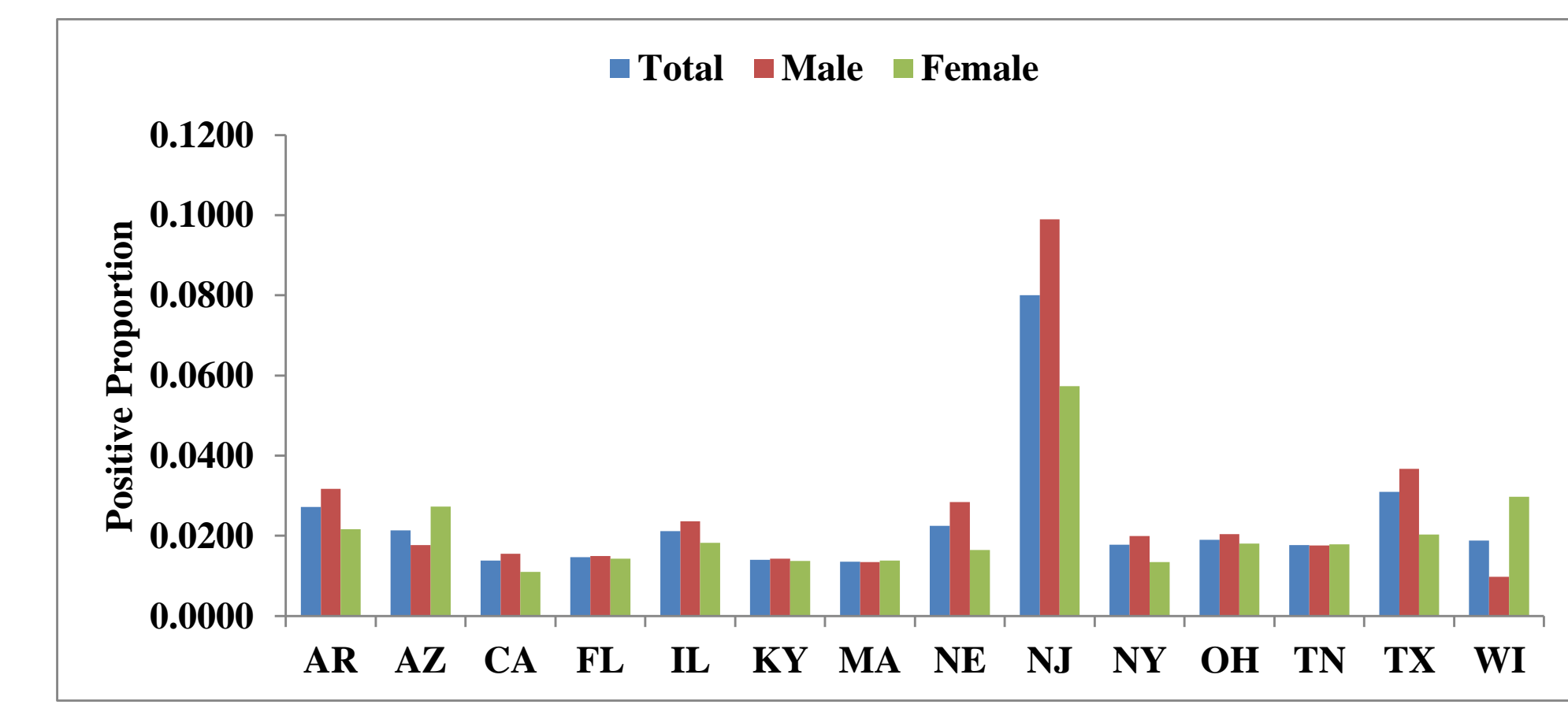


Figure 4. Positive Proportion of HALT Data by Gender per State (2007-09)

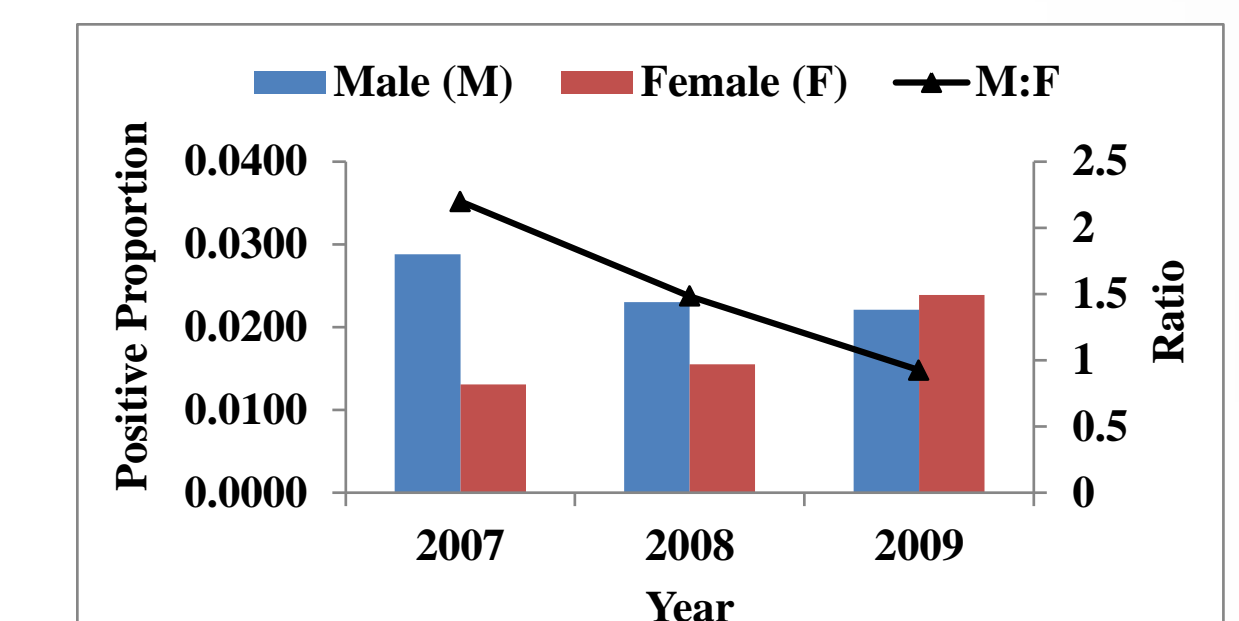


Figure 5. Positive Proportion of HALT Data by Gender per Year
Note: M:F, male-to-female

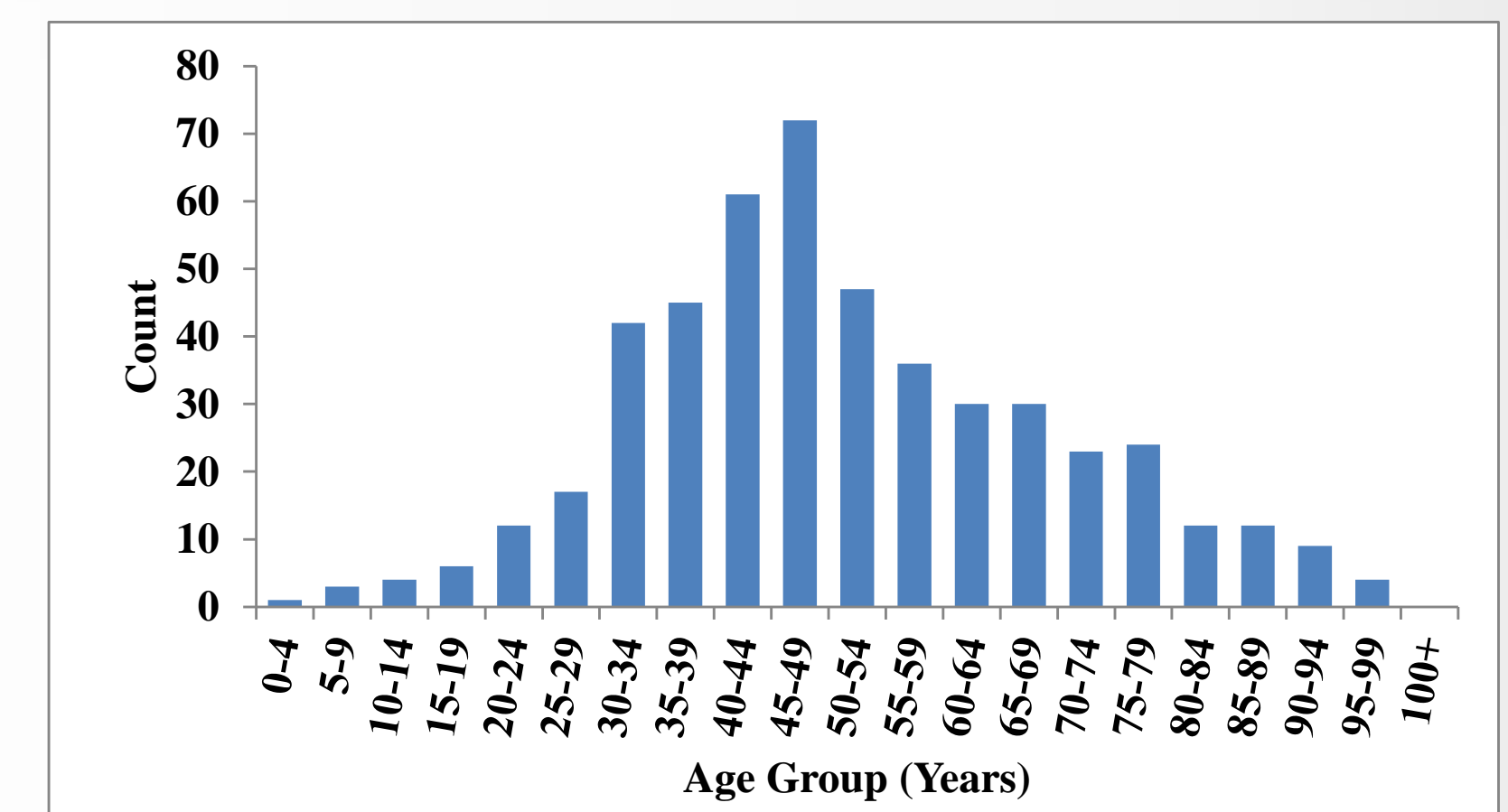


Figure 6. Frequency of Positive HALT Data by Age Category (2007-09)

Notifiable Disease Data

Table 2. State Notifiable Disease Data for Histoplasmosis Reported Cases

State	Number of Cases				Average	Rate*
	2007	2008	2009	2008		
IL	123	150	102	125	2.94	
IN	116	89	136	114	5.31	
MI	126	93	58	92	2.78	
AR	80	46	46	57	5.98	
MS	31	55	41	42	4.31	
KY	46	43	32	40	2.82	
AL	35	8	7	17	1.06	
DE	1	1	1	1	0.34	
PA	2	0	1	1	0.02	
WI	NA	NA	NA	NA	NA	
MN	NA	NA	NA	NA	NA	

*per 100,000 population; 2008 population estimates (US Census Bureau)

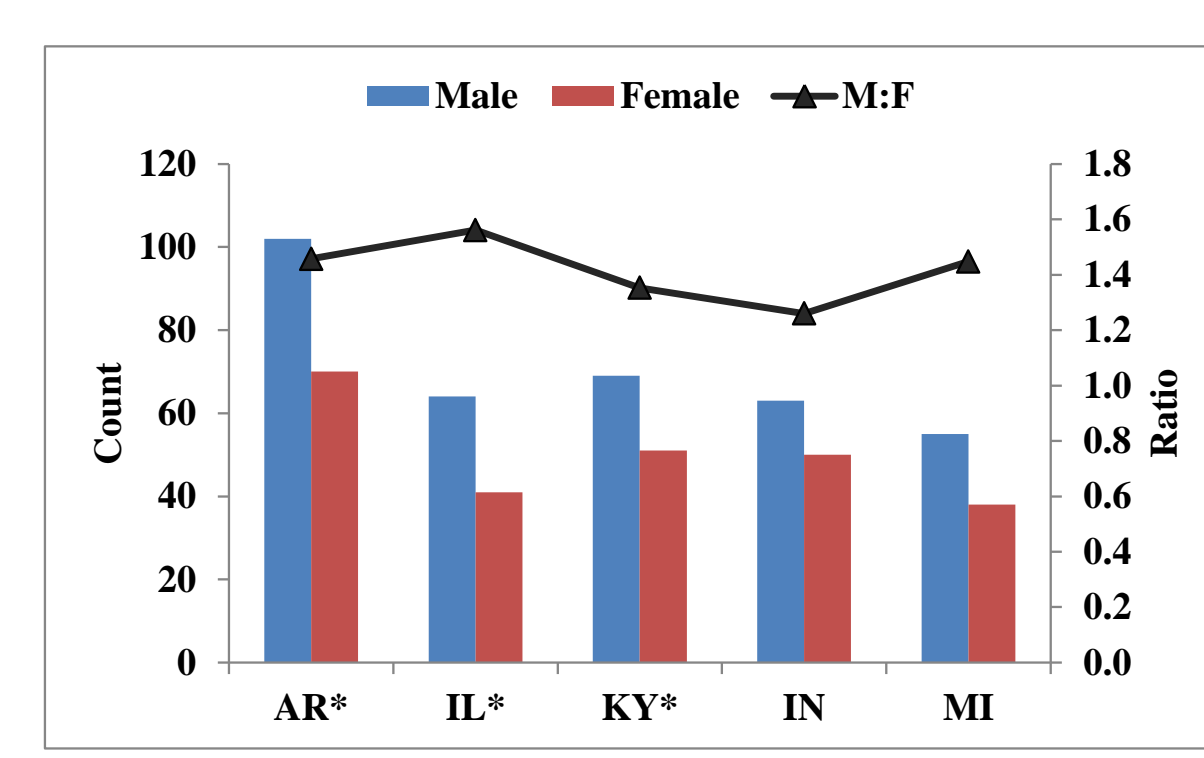


Figure 8. Reported Histoplasmosis Cases by Gender per State (2007-09)
Note: *states with HALT data

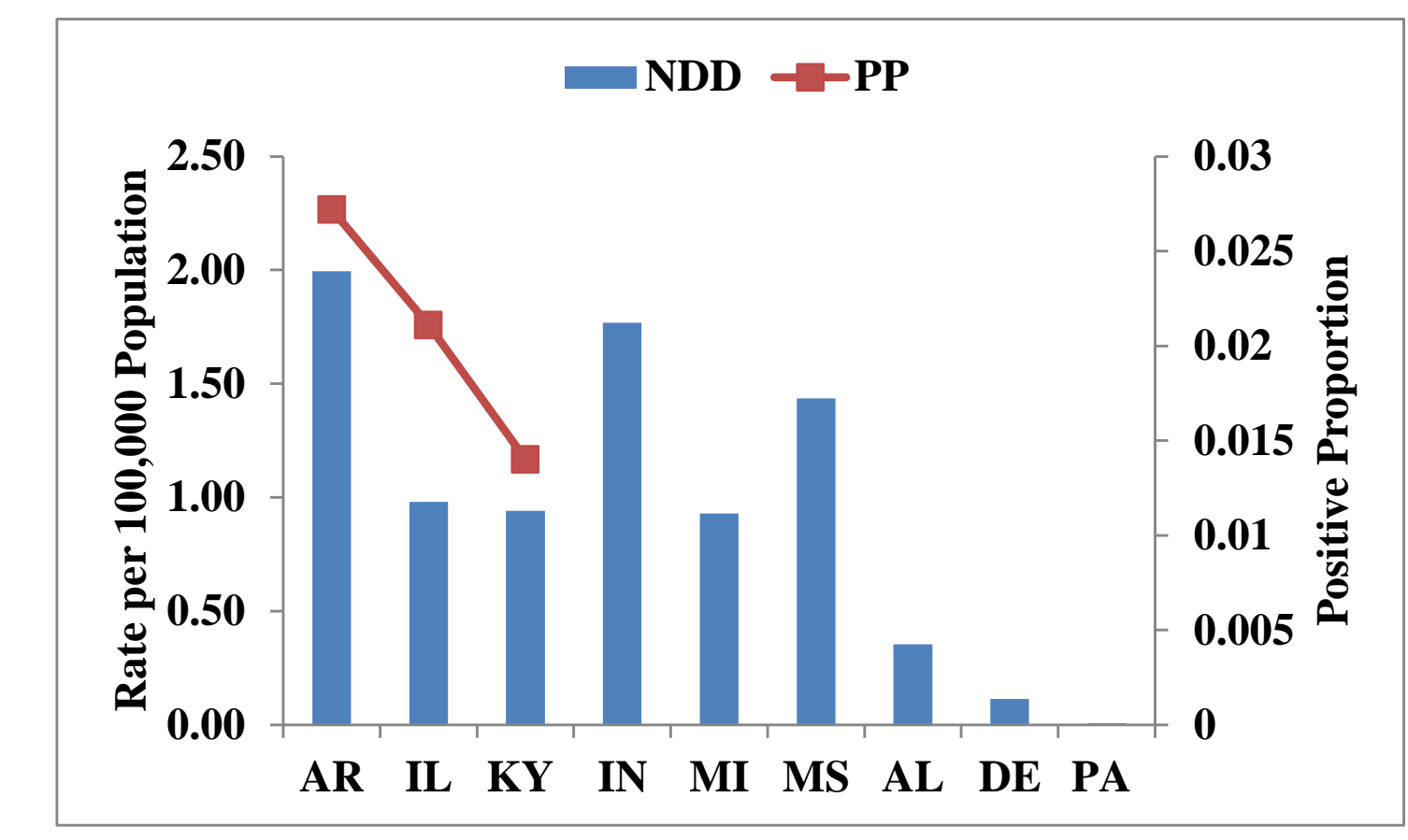


Figure 7. Average Rate of Reported Cases of Histoplasmosis by State and HALT Positive Proportions for AR, IL, and KY (2007-09)
Note: NDD, notifiable disease data; 2008 population estimates (US Census Bureau)

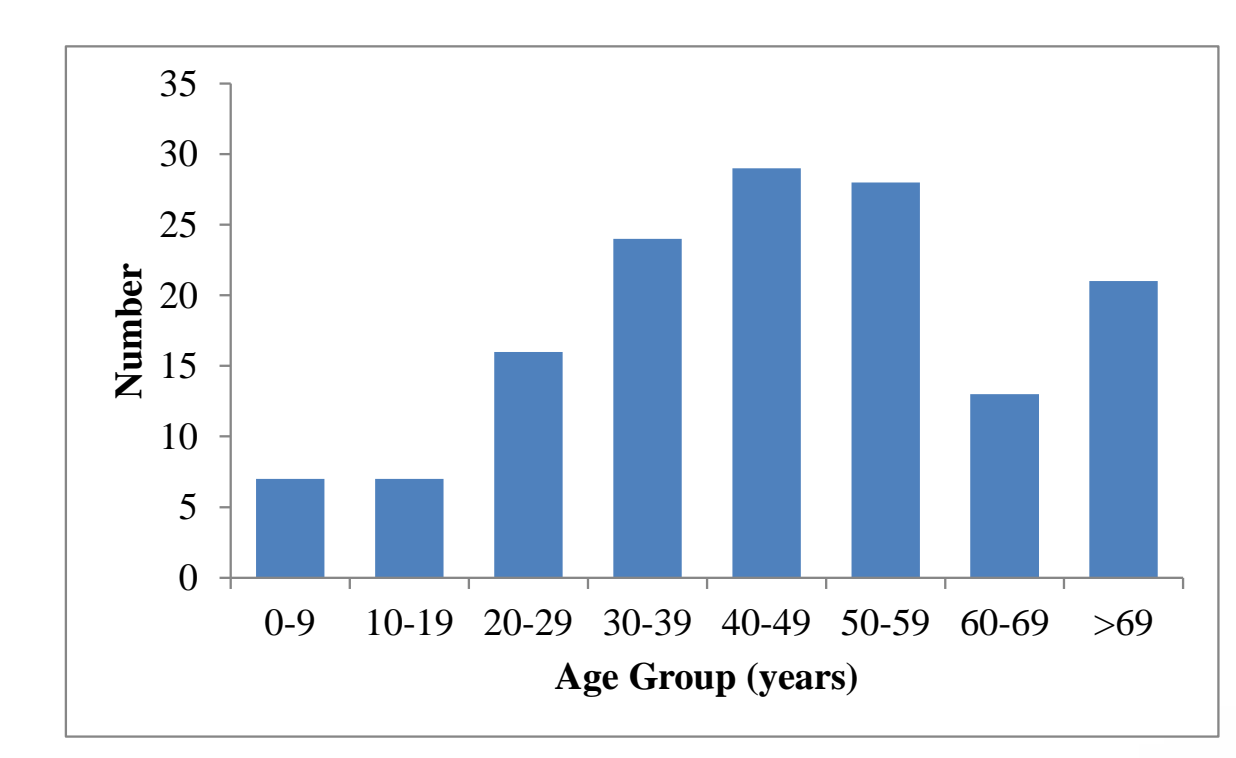


Figure 9. Frequency of Reported Histoplasmosis Cases by Age Group (AR, IL, KY; 2007-09)

Summary and Conclusions

Histoplasma antigen testing is widely used to diagnose the disease. HALT positive proportions were used to define histoplasmosis epidemiology. In this study, states with stable HALT data were determined using weighted binomial distribution estimates. NJ appears to be an outlier with extremely high positive proportion, possibly due to laboratory quality control testing data mixed with patient data.

The following descriptive epidemiology was obtained using HALT data from AR, AZ, CA, FL, IL, KY, MA, NE, NJ, NY, OH, TN, TX, WI for the years 2007-2009:

- NJ, TX, and AR have the highest positive proportions of HALT results
- More males than females have positive HALT results (average male-to-female ratio 1.2)
- The average age for positive HALT results is 47, with most occurring in ages 30-50 years

Histoplasmosis is a reportable disease for 11 states (9 states summarize annual data):

- AR has the highest rate (5.98 cases per 100,000 population); IN and MS have the next highest rate per 100,000 population (5.31 and 4.31, respectively)
- There are more reported cases of histoplasmosis in males than females (average male-to-female ratio 1.4)
- The most frequent age group for reported cases of histoplasmosis is 40-49 years.

Limitations and Suggestions:

- The major limitation is incomplete data. The study is delimited to HALT data from one reference laboratory.
- The use of positive proportions relies on the assumption that all laboratories have the same positive proportion of HALT results.
- Notifiable disease data are limited to 9 states, preventing a reliable comparison with HALT data. Only 3 states (AR, IL, and KY) have both HALT and notifiable disease data.
- Two-source capture-recapture analysis (Sandifer & Goldoft, 2010) to assess the completeness of the HALT data source is suggested.

For rare infectious diseases, the use of laboratory test data is a logical data source for surveillance in this era of electronic disease reporting. Notifiable disease data are limited by state department of health resources. For monitoring of the occurrence of histoplasmosis, HALT data could be useful if laboratory test data from all laboratories are compiled for a more complete dataset.