AGENT-BASED MODELING OF COCCIDIOIDOMYCOSIS

by

Carrie Marie Anderson

B.S., University of Minnesota, 1997

M.S., University of Minnesota, 1999

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This dissertation was presented

by

Carrie Marie Anderson

It was defended on

April 3, 2013

and approved by

Dissertation Advisor:
Stephen R. Wisniewski, Ph.D., Professor, Department of Epidemiology
Graduate School of Public Health, University of Pittsburgh

Committee Member:
Evelyn O. Talbott, Dr.P.H., Professor, Department of Epidemiology
Graduate School of Public Health, University of Pittsburgh

Committee Member:
Eunha Shim, Ph.D., Assistant Professor, Department of Epidemiology
Graduate School of Public Health, University of Pittsburgh

Committee Member:
Douglas F. Lake, Ph.D., Associate Professor, School of Life Sciences
College of Liberal Arts and Sciences, Arizona State University

Committee Member:
John J. Grefenstette, Ph.D., Professor, Department of Biostatistics
Graduate School of Public Health, University of Pittsburgh

Copyright © by Carrie Marie Anderson

2013

AGENT-BASED MODELING OF COCCIDIOIDOMYCOSIS

Carrie Marie Anderson, Ph.D.

University of Pittsburgh, 2013

Coccidioidomycosis is a fungal infection with an estimated yearly incidence of 150,000 cases in the United States. Up to 50 percent of those cases are estimated to occur in Maricopa County, Arizona, the geographical focus of this dissertation. Maricopa County is a hotspot for coccidioidomycosis due to its unique environmental and climactic conditions including its soil type, geology, dust storms, and temperature. Although there has been a large amount of research on the epidemiology of coccidioidomycosis in Maricopa County and elsewhere, forecast modeling of disease incidence has not been well established. Current analyses focus on demographic and environmental factors and their effect sizes but have limited use for modeling the public health impact of events such as dust storms or vaccination strategies.

However by incorporating results from previous studies, and including historical data from the Centers for Disease Control and Prevention and Maricopa County Department of Health, stochastic epidemiological agent-based modeling of coccidioidomycosis can be successfully performed. The development and validation of such a model and its public health significance in forecasting coccidioidomycosis incidence are described in this dissertation. Among the findings, a moderately sized dust storm in Maricopa County would be expected to increase coccidioidomycosis morbidity by 4,676 cases and mortality by 42 cases. The development of a vaccine against coccidioidomycosis could decrease annual morbidity by 5,979 cases if individuals get vaccinated at rates comparable to influenza. Even a vaccination campaign that is one-fourth as effective as an influenza campaign would still have a significant impact on public health with a reduction in annual morbidity by 2,361 cases.

Further with the development of the web-based tool described in this dissertation, public health researchers and epidemiologists can use the model to forecast disease morbidity and mortality for other endemic regions. The tool can also be used to forecast disease burden for

hypothetical scenarios such as the development of a vaccine against coccidioidomycosis. In summary this dissertation uses stochastic epidemiological agent-based modeling to forecast incidence and assess the public health impact of vaccination and natural events such as dust storms, and provides a valuable tool for epidemiologists and researchers.

TABLE OF CONTENTS

1.0 INTRODUCTION 1

1.1 SPECIFIC AIMS 2

2.0 BACKGROUND 7

2.1 PATHOGEN 7

2.2 HISTORY 8

2.3 ENVIRONMENT 9

2.4 OCCUPATION 12

2.4.1 Prison 14

2.5 DEMOGRAPHICS 15

2.5.1 Pregnancy 17

2.5.2 Race / Ethnicity 18

2.5.3 Genetics 19

2.6 CO-MORBIDITY 19

2.6.1 Transplantation 21

2.7 BEHAVIORS 22

2.8 CASE STUDIES 23

2.8.1 Misdiagnosis 23

2.9 COST ANALYSIS 24

2.10 TESTS FOR COCCIDIOIDOMYCOSIS 25

2.11 SURVEILLANCE IN ARIZONA 26

2.12 VACCINATION 27

3.0 METHODOLOGY 29

3.1 STOCHASTIC MODEL OVERVIEW 29

3.2 MATHEMATICAL AND PROBABILISTIC NOTATION 30

3.2.1 Mathematical Notation 30

3.2.2 Probability Distributions 30

3.2.3 Kronecker's Delta Distribution 31

3.2.4 Uniform Distribution 31

3.3 POPULATION 31

3.3.1 Demographic and Risk Factors 32

3.3.2 State Variable 33

3.3.3 Clinical Variables 33

3.3.4 Temporal Variables 34

3.3.5 Strata 34

3.4 ENVIRONMENTAL FACTORS 35

3.5 DISEASE MODEL 35

3.6 POPULATION DYNAMICS 37

3.6.1 Births 38

3.6.2 Deaths 39

3.6.3 Updates 40

3.7 POPULATION INITIALIZATION 48

3.7.1 Demographic and Risk Factors (X) 48

3.7.2 State Variable (*S*) 49
3.7.3 Clinical Variables (*C*) 49
3.7.4 Temporal Variables (*T*) 50
3.8 MODEL INSTANTIATION 50
3.8.1 Age Strata 57
3.8.2 Initial Population Sizes 57
3.8.3 Natality Rates 57
3.8.4 Mortality Rates 58
3.8.5 Infection Rates 58
3.8.6 Distributions for Risk Factors and Prior Infection 59
3.9 MODEL VALIDATION 61
3.10 SENSITIVITY ANALYSIS 61
3.11 PUBLIC HEALTH IMPLICATIONS 62
3.11.1 Vaccination Strategy 62
3.11.2 Vaccination Coverage 63
3.11.3 Weather Events 64
3.12 ONLINE TOOL 66
4.0 RESULTS 69
4.1 MARICOPA COUNTY, ARIZONA 69
4.2 MODEL INSTANTIATION 72
4.3 MODEL VALIDATION 77
4.3.1 Face Validity 78
4.3.2 Independent Validation 89

4.4 SENSITIVITY ANALYSIS 91

4.5 PUBLIC HEALTH IMPLICATIONS 93

4.5.1 Vaccination Strategy 93

4.5.2 Vaccination Coverage 95

4.5.3 Weather Events 95

4.6 ONLINE TOOL 101

5.0 DISCUSSION 105

5.1 MODEL VALIDATION 105

5.2 SENSITIVITY ANALYSIS 107

5.3 PUBLIC HEALTH IMPLICATIONS 108

5.4 ONLINE TOOL 109

5.5 FUTURE POTENTIAL USES 110

6.0 CONCLUSIONS 112

6.1 PUBLIC HEALTH SIGNIFICANCE 113

BIBLIOGRAPHY 114

LIST OF TABLES

Table 1. Data sources and resolution. 51

Table 2. Summary statistics for PM10 across 13 stations by month for 2009 and 2010. 52

Table 3. Parameters examined in one-way sensitivity analyses. 62

Table 4. Age group by gender and coccidioidomycosis for Maricopa County, 2009. 70

Table 5. Race/ethnicity by gender and coccidioidomycosis for Maricopa County, 2009. 71

Table 6. Age group by race/ethnicity for Maricopa County, 2009. 72

Table 7. Model inputs compared to simulation estimates for Maricopa County, 2010. 79

Table 8. Actual and forecasted number of reported cases for Maricopa County, 2010. 91

Table 9. One-way sensitivity analyses: mean number of infections, symptomatic infections, and case reports..... 92

Table 10. Directory listing of web-based tool. 102

LIST OF FIGURES

Figure 1. PM10 data for 13 weather stations in Maricopa County, Arizona. Top: 2009. Bottom: 2010..... 53

Figure 2. Daily PM10, temperature, and precipitation for Maricopa County, 2009..... 54

Figure 3. Daily PM10, temperature, and precipitation for Maricopa County, 2010..... 55

Figure 4. Temperature and rainfall time series for Maricopa County, 2010. 56

Figure 5. PM10 data in Maricopa County. Top: PM10 during a 2009 weather event. Middle: PM10 in 2010. Bottom: PM10 for 2010 with injected weather event. 65

Figure 6. Population count for Maricopa County, 2009. 74

Figure 7. Diabetes probabilities for Maricopa County, 2009. 75

Figure 8. Reported infection probabilities for Maricopa County, 2009. 76

Figure 9. Vaccination probabilities for Maricopa County, 2009. 77

Figure 10. Gender distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009..... 81

Figure 11. Age distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009..... 82

Figure 12. Race/ethnicity distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009..... 83

Figure 13. Simulated number of new infections by week. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 84

Figure 14. Cross-correlation function between date of infection and PM10. 85
Figure 15. Simulated number of new symptomatic infections by week of symptom onset. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 86
Figure 16. Cross-correlation function between date of infection and date of symptom onset. 87
Figure 17. Simulated number of new coccidioidomycosis case reports by week of report. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 88
Figure 18. Cross-correlation function between date of symptom onset and date of case report. . 89
Figure 19. Symptomatic infections (top), infections resulting in death (middle), and reported cases (bottom) by vaccination strategy and coverage. 94
Figure 20. New infections by week with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 96
Figure 21. New infections with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event. 97
Figure 22. Symptomatic infections by date of symptom onset with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 98
Figure 23. New symptomatic infections with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event. 99

Figure 24. New reported infections by date of report with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals. 100
Figure 25. New reported infections by week of report with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event. 101

1.0 INTRODUCTION

Coccidioidomycosis, commonly known as Valley Fever, San Joaquin Fever, or Desert Rheumatism, is an infection from the aerosolized spores of fungi: *Coccidioides immitis* or *Coccidioides posadasii* (Parish and Blair 2008). Every year approximately 150,000 new coccidioidal infections of humans are estimated to occur in the United States, and a large portion of these infections occur in the southwestern United States where the fungi reside in soil (Galgiani, Ampel et al. 2005). Cases of coccidioidomycosis in Arizona account for around 60 percent of all reported coccidioidal infections in the United States (Centers for Disease Control and Prevention 2011), and about 72 to 82 percent of those reported cases occur in Maricopa County, Arizona (Park, Sigel et al. 2005; Sunenshine qtd. in Foote 2011).

Roughly 40 percent of coccidioidal infections are symptomatic (Smith, Beard et al. 1946; Pappagianis, Sun et al. 1993; Louie, Ng et al. 1999; Chiller, Galgiani et al. 2003). Approximately 10 to 15 percent of symptomatic cases have significant manifestations such as chronic pulmonary infections (Einstein and Johnson 1993; Pappagianis, Sun et al. 1993). There are some cases of coccidioidal infection spread through the body (e.g., skin lesions). These disseminated infections occur in approximately one half to one percent of all cases (Einstein and Johnson 1993; Pappagianis, Sun et al. 1993), but dissemination rates may actually be as high as five percent (Crum, Lederman et al. 2004). Mildly symptomatic coccidioidal patients present with flu-like symptoms (Galgiani, Ampel et al. 2005) but infection can spread to the meninges and brain (Mendel, Milefchik et al. 1994; Wang, Jerng et al. 2005). Coccidioidal infections can also spread into bones, joints, and nearly all major organs (Einstein and Johnson 1993; Blair 2007). Meningitis is the main cause of death (Einstein and Johnson 1993).

Patients with localized, acute pulmonary infections typically only require the monitoring of symptoms (e.g., sweating, fever, cough, headache, muscle aches, joint pain, or rash) but not treatment (Galgiani, Ampel et al. 2005). However, patients with extensive spread of infection, or patients at risk for further complications, may require treatment with antifungal medications,

surgical resection, or medication and resection combined (Galgiani, Ampel et al. 2005; Jaroszewski, Halabi et al. 2009). For such individuals, therapy may be required for several months to several years, and in some patients, especially those with coccidioidal meningitis, lifelong treatment with antifungal medication may be required (Galgiani, Ampel et al. 2005).

In nearly all cases, infection occurs by inhaling spores aerosolized from soil (Blair 2007). In 1977, a dust storm in California carried dust from the southern San Joaquin Valley to northern and coastal regions of the state, causing cases of coccidioidomycosis to quadruple in California and increase 20-fold in the San Joaquin Valley (Pappagianis and Einstein 1978; Flynn, Hoeprich et al. 1979). A 6.7 magnitude earthquake in Northridge, California in 1994 caused an outbreak of coccidioidomycosis with 203 outbreak-associated cases including three fatalities (Schneider, Hajjeh et al. 1997). Additionally, traveling to an endemic area poses risk. Fifty-nine members of a church group agreed to skin testing for coccidioidomycosis after traveling to Mexico where soil excavation occurred, and 27 of the 59 were found to have a positive coccidioidal result (Cairns, Blythe et al. 2000). Kamei et al. found 31 cases (87.1% male and 12.9% female) of imported coccidioidomycosis in Japan through 2001 (Kamei, Sano et al. 2003).

Outbreaks have also been documented following archeological excavations (Werner, Pappagianis et al. 1972), including Utah where reflective heat, arid weather, and the sifting of sandy, alkaline soil are thought to be contributing factors in ten of 18 persons contracting coccidioidomycosis (Petersen, Marshall et al. 2004). Similarly at risk are construction and agricultural workers, as well as military personnel, especially those around aerosolized dust (Schmelzer and Tabershaw 1968; Crum, Lamb et al. 2002). Further, an analysis of cases by Park et al. in Arizona between 1998 and 2001 found that environmental variables were associated with cases including temperature and ambient levels of dust (Park, Sigel et al. 2005). Outbreaks have also been negatively associated with precipitation (Einstein and Johnson 1993; Tamerius and Comrie 2011), supporting the "grow and blow" hypothesis (Comrie and Glueck 2007).

1.1 SPECIFIC AIMS

Up to 50 percent of the yearly coccidioidal cases in the United States are estimated to occur in Maricopa County, Arizona with roughly 40 percent of all cases presenting clinical symptoms.

Because its environmental and climactic conditions are well suited for *Coccidiodes* growth and reproduction, Maricopa County is a hotbed for coccidioidomycosis. As will be discussed in depth in Chapter 2, *Coccidiodes* prefer to grow in alkaline, non-clay, and temperate soil with high salinity. Further, Maricopa County is fairly unique with its flat lands, open areas, and moderate vegetation that make the region prone to dust storms that help to aerosolize infective *Coccidiodes* spores. Because of the high incidence rate, we will examine the coccidioidomycosis incidence in Maricopa County through the development of a computational agent-based model. Most of the previous studies of coccidioidomycosis in Maricopa County and elsewhere analyze incidence data using summary statistic or regression-based approaches without forecasting the public health impact of events such as dust storms or vaccination strategies. While such analytic approaches can be very effective at identifying key contributors to incidence rates and identifying effect sizes, they are not so useful for predicting future coccidioidomycosis incidence. The ability to create such forecasts would be a considerable asset to public health, as epidemiologists and other researchers seek answers to questions such as, "How many additional cases of coccidioidomycosis can we expect from the most recent dust storm? When and where will these coccidioidal cases occur, and which demographic groups will be most affected?" To assist in answering these types of questions, we will construct an epidemiological agent-based model of coccidioidomycosis. This approach has two advantages:

a) We will better understand coccidioidomycosis by using not only available data but also expert domain knowledge and coccidioidal results from previous studies. By synthesizing all of this information into a single framework, we can better study the contributions of various factors and generate better forecasts of disease.
b) We will be able to generate forecasts of disease incidence under a variety of scenarios that will include factors such as weather events and vaccination strategies. Further, forecasted disease incidence can include error bounds to help assess the public health impact that various scenarios have on the range of disease incidence.

In order to forecast incidence using various scenarios, we first need to develop and program an agent-based model of coccidioidomycosis. Our aims, objectives, and hypotheses are as follows.

Aim 1: Model Development: We will develop and implement a stochastic epidemiological agent-based model of coccidioidomycosis that forecasts disease incidence and variability on public health.

Objective 1: A stochastic epidemiological agent-based model can be constructed and parameterized using available data and expert domain knowledge from coccidioidal studies.
Approach 1: We will develop an epidemiological model that uses historical coccidioidomycosis incidence data and synthesizes knowledge from previous studies on coccidioidomycosis. A subset of the incidence data (Maricopa County, 2010) will be held out from model development and used for validation. As our model will be quantitative and stochastic, it will be capable of not only providing a single forecasted epidemic curve, but also mean epidemic curves and forecast intervals, for various parameterized scenarios including the public health impact of weather events and vaccination strategies. Due to the complexity of the model, it would be difficult or impossible to implement the model in existing stochastic modeling software packages. Thus, the model will be implemented in the statistical programming language R (R Core Team 2012).

Aim 2: Model Validity: We will access face validity of the model, validate the model using incidence data (Maricopa County, 2010) not previously incorporated into the model, and perform sensitivity analysis to examine the impact of parameter value choices.

Objective 2a: The model will demonstrate previously reported demographic, occupational, co-morbidity and behavioral risk factors for coccidioidomycosis. In particular, occupations with high dust exposure and high dust levels as measured by environmental sensors will be associated with coccidioidomycosis infections, and the demographic distribution of infections will be consistent with previously reported data for Maricopa County. The model will show that African-American race, male, diabetic, immunocompromised status, smoking and age are risk factors for developing severe pulmonary symptoms. The model will also demonstrate that African-American race, diabetes, and smoking are risk factors for disseminated disease.
Approach 2a: We will assess face validity by examining whether model simulations reveal risk factors in the same direction and approximate magnitude as those reported in previous studies.

As the model's development is based upon incidence data and previous studies, this is in essence a check that the model properly incorporates existing knowledge as would logically be expected.

Objective 2b: In an independent test of the model using Maricopa County 2010 data, forecast intervals calculated by the model will contain true reported incidence counts, in total and by gender, age, and race/ethnicity. Forecast intervals for incidence rates will also contain true reported incidence rates.

Approach 2b: We will employ data (not previously used to develop the model) to provide an independent validation of the model. Specifically, incidence counts and rates from Maricopa County, Arizona in 2010, in total and by gender, age and race/ethnicity, will be used as an independent test set for model validation.

Objective 2c: The model will respond predictably to changes in key parameter values.

Approach 2c: We will run multiple simulations using a base case, a low value, and a high value for each of several key parameters. We will assess the impact of value choices on model output.

Aim 3: Public Health Impact: We will use the model to forecast incidence and investigate the public health impact of vaccination strategies, vaccination coverage, and a weather event.

Objective 3: The model can be utilized to forecast incidence and disease burden and to assess the public health impact of vaccination strategies, vaccination coverage, and a weather event.

Approach 3: We will use the model to forecast disease incidence and burden for Maricopa County, Arizona in 2010 based on data available prior to 2010. We will investigate the public health impact of possible vaccination of all individuals, children only, and adults only at vaccination rates comparable to influenza and at lesser vaccination rates. We will investigate the public health impact of a moderate-sized dust storm by using PM10 data from a real dust storm to create a synthetic dust storm during 2010. We will then assess the increase in incidence and disease burden attributable to the dust storm.

Aim 4: Online Tool: We will develop an interactive web-based tool that researchers and policy makers can use to guide coccidioidal decisions regarding outbreak prevention and management.

Objective 4: An interactive web-based tool can be developed that will incorporate our model and forecast coccidioidomycosis incidence under various conditions chosen by the end user.
Approach 4: We will develop a web-based tool that will accept inputs from users, call R to simulate from the model, and return forecasts on coccidioidomycosis incidence.

2.0 BACKGROUND

2.1 PATHOGEN

There are an estimated 1.5 million species of fungi worldwide (Hawksworth 1991; Hawksworth 2001). In the United States alone, there are an estimates 13,000 species of fungi that are phytopathogenic, causing plant pathology (Madden and Wheelis 2003), but only about 200 species of fungi are able to cause human infection (McGinnis 1991). Two of those species are *Coccidioides immitis* (Californian) and *Coccidioides posadasii* (non-Californian) in the genus *Coccidioides* of the family Onygenaceae in the order Onygenales (*Trichophyton rubrum*, the fungus that causes athlete's foot, is in this same order) of the kingdom Fungi. Coccidioidomycosis is a fungal disease that is caused by these two coccidioidal species of fungus. This disease was first discovered in 1892 in Argentina and became recognized as a hazard during the severe drought of the 1930s (Rutherford and Barrett 1996).

Coccidioides immitis and *Coccidioides posadasii* are commonly thought of as thermally dimorphic species, capable of taking two forms depending on temperature (Saubolle, McKellar et al. 2007). At 77°F (25°C) the fungi form branching filaments called hyphae that amass into a network of hyphae called mycelium (Larone 1995). Multiple networks of mycelium, called mycelia, can be found in soil in the southwestern United States, Mexico, and Central and South America. The fungal development of mycelia in soil is called the saprophytic phase. It is during this phase that the hyphae separate into pieces called arthroconidia (spores) that are the primary source of dissemination for these fungi (Barrera 1986; Kirkland and Fierer 1996). Most cases of coccidioidomycosis, the fungal infection commonly known as Valley Fever, are caused by the inhalation of spores of arthroconidia arising from soil disturbances (Blair 2007).

At 98.6°F (37°C) the fungi form round cellular forms called spherules that fill with 800 to 1,000 small spores (endospores) that are released at the end of the spherules lifecycle (Ampel

2000; Saubolle 2000). Each of the 800 to 1,000 released endospores can produce its own spherule containing another 800 to 1,000 endospores (Saubolle, McKellar et al. 2007). This process is temperature dependent as well as carbon dioxide dependent (Klotz, Drutz et al. 1984) and can occur in most mammals including dogs, cats, llamas, dolphins, zoo animals (Valley Fever Center for Excellence 2010) and some reptiles (Shubitz 2007). Without the presence of carbon dioxide, endospores can instead form mycelia (Klotz, Drutz et al. 1984), and it is possible to have mycelium growth in the human body (Osaki, Morishita et al. 2005).

When spores of arthroconidia are inhaled, the host lung tissue provides a mix of temperature and carbon dioxide that is conducive for spherule formation (Klotz, Drutz et al. 1984). Rarely, fungi can be also introduced to host tissue via a subcutaneous wound or organ transplantation (Miller, Hendren et al. 2004; Blair 2007; Saubolle, McKellar et al. 2007; Gaidici and Saubolle 2009). The fungal development in a host is called the parasitic phase, and released endospores are capable of disseminating throughout the body via the lymphatic system and bloodstream (Saubolle, McKellar et al. 2007). The parasitic phase continues until death occurs or control is achieved by drug therapy or immune response (Saubolle 1996).

2.2 HISTORY

The southwestern United States and northern Mexico were dubbed the Lower Sonoran Life Zone (Merriam 1898). *Coccidioides* species are naturally occurring fungi in the arid and semi-arid regions of the Unites States, Mexico, and Central and South America, including the Lower Sonoran Life Zone. The first case of coccidioidomycosis was reported in 1892 and was thought initially to be tuberculosis, but telltale spherules were present in tissue samples (Espinel-Ingroff 2003). In 1896, two cases of disseminated coccidioidomycosis were reported in California, the first cases of coccidioidomycosis in the United States (Espinel-Ingroff 2003). Between 1900 and 1907, dimorphic fungi (arthroconidia versus spherule formation) were first understood including *Coccidioides immitis* (Espinel-Ingroff 2003). In 1918, the first case of coccidioidomycosis in animals (cattle) was reported (Morrow 2006), in 1919 the first case of coccidioidomycosis in California state prisons was reported (Pappagianis 2007), but it wasn't until 1937 that "Valley

Fever" was recognized as coccidioidomycosis (Dickson 1937). In 1941, Farness reported the first case of coccidioidomycosis in a pregnant woman (Farness 1941). Reported in 1960, Maddy et al. described the geographical distribution of the *Coccidioides* species by testing 11,643 cattle in Arizona (2,859 were positive for disease) and found the Arizonan distribution was analogous to the Lower Sonoran Life Zone described by Merriam (Maddy, Crecelius et al. 1960).

Until the 1990s *Coccidioides immitis* was thought to be the only species in the genus *Coccidioides* (Saubolle, McKellar et al. 2007) but genetic differences in species were found (Fisher, Koenig et al. 2002). The newly discovered *Coccidioides posadasii* grows slower than (average colony size is approximately 5 mm smaller at growth week two) *Coccidioides immitis* in higher salt concentrations (0.034M NaCl versus 0.136M NaCl; M = 1 mole of NaCl ≈ 58g), only *Coccidioides immitis* has thus been found in the San Joaquin Valley, California region, though *Coccidioides posadasii* is found in across more geographical locations (Fisher, Koenig et al. 2002) including overlapping regions of *Coccidioides immitis* (Thompson 2011). Both species are now thought to have been genetic isolates for the last 11 million years (Fisher, Koenig et al. 2002). However, both species are morphologically identical (Fisher, Koenig et al. 2002) and manifest into indistinguishable disease (Galgiani, Ampel et al. 2005; Thompson 2011). Recently, two bison mandibles (around 8,500 years old) were tested and found to have evidence of coccidioidal infection (Morrow 2006). Morrow claims that contamination of the mandibles is unlikely given they were found in Nebraska, a non-endemic coccidioidal area.

2.3 ENVIRONMENT

Coccidioides immitis and *Coccidioides posadasii* live in areas of varied vegetation and do not appear to depend on the types of vegetation present on land (Lacy and Swatek 1974; Fisher, Bultman et al. 2007). Instead, Fisher et al. argue that water, soil type, and temperature at a depth 0.79 inches (2 cm) to 7.90 inches (20 cm) into the earth (where *Coccidioides* species reside) are influential factors that control the growth of the *Coccidioides* species in soil (Fisher, Bultman et al. 2007). In order for fungal growth, Fisher et al. state that there needs to be oxygen, organic material (carbon and nitrogen), and water (precipitation or water body). Water should filter

through the depth of the fungi, as well as pass through the soil without saturation, and the temperature should range between 68°F (20°C) and 104°F (40°C) for optimal growth (Fisher, Bultman et al. 2007). Soils with a high clay concentration are not common in the *Coccidioides* habitat (Lacy and Swatek 1974; Fisher, Bultman et al. 2007), and *Coccidioides* fungi prefer an alkaline soil (pH range: 6.1-8.6) (Lacy and Swatek 1974) with a salinity three times higher than non *Coccidioides* inhabited soil (Elconin, Egeberg et al. 1964).

Coccidioides species are difficult however to isolate in air and soil (Ajello, Maddy et al. 1965; Kolivras, Johnson et al. 2001) and, even recently, the "reports of positive isolations of the causal agent, *Coccidioides* spp. from environmental samples have been scarce" (Baptista-rosas, Catalán-dibene et al. 2012). One reason positive samples are hard to find is that fungal specimens are small. A grain of sand is about 0.15 mm in diameter whereas one arthroconidia segment (one spore) is about 1.3 percent of that size at approximately 0.002 mm wide (Fisher, Bultman et al. 2007). Another reason positive samples are hard to find is that, while large geographic regions are home it to the *Coccidioides* species, the species live in patches rather than in continuous coverage (Lacy and Swatek 1974; College of Veterinary Medicine 2010; Irfan 2012). Tabor states, "You can move over just a few feet and you can't find it" (qtd. in Irfan 2012). Wind, however, is known to spread the spores of arthroconidia (Campbell 1980). Over 300 new cases of coccidioidomycosis, including ones in non-endemic areas, were the result of a dust storm that occurred 12 weeks prior (Pappagianis and Einstein 1978).

Because there is a lack of positive environmental samples of the *Coccidioides* species, researchers use other available data such as health reports and environmental data to follow disease. Tamerius states, "The only things we have to go by are the actual health data" in reference to the disease tracking (qtd. in Irfan 2012). Comrie offers that "disease incidence data offer the best (and only) available multiyear time series for comparison with climatic conditions" (Comrie 2005). Comrie uses disease incidence, precipitation levels, and PM10 data (used as a surrogate for spore distribution) from Arizona to show a bimodal seasonality (June-July and October-November) of coccidioidomycosis, that PM10 and precipitation are inversely related (less PM10 during wet periods), and that concomitant coccidioidomycosis rates and PM10 are positively associated during winter and prior to monsoon season (Comrie 2005). PM10 is a measure of particular matter (up to 10 micrometers in diameter) often expressed in ug/m3 units.

Further environmental aspects of coccidioidomycosis were studied. Summarizing data from the Medical Electronic Disease Surveillance Intelligence System (MEDSIS), an Arizona disease surveillance system, Sunenshine et al. reported a seasonal affect in Arizona, where more cases typically appear between October and January and May to August (Sunenshine, Anderson et al. 2007). Tamerius and Comrie used autocorrelation to examine how coccidioidomycosis relates to weather (Tamerius and Comrie 2011). They obtained data from the Arizona Department of Health Services (ADHS), the Arizona Meteorological Network (AZMET), the National Weather Service, the Maricopa County Air Quality Department, and the Pima County Department of Environmental Quality. They aggregated case data by month, and monthly averages were calculated for PM10 and climatic variables (Tamerius and Comrie 2011). They calculated autocorrelations and found a primary exposure season running from August to March, a secondary exposure season from April to June, and a strong association between exposure rates in the primary season to precipitation from 36 months past (Tamerius and Comrie 2011), but their statistical analysis failed to account for the large number of associations tested. Tamerius and Comrie also created a statistical model for exposure rates during the primary exposure season with seasonal precipitation totals as independent variables; however, only the intercept was statistically significant (α=0.05) in their model (Tamerius and Comrie 2011).

Park et al. used Poisson regression to determine whether environmental factors and an increase in monthly cases of coccidioidomycosis were associated (Park, Sigel et al. 2005). They obtained data from the National Electronic Telecommunications System for Surveillance (NETSS), National Climatic Data Center, and the Maricopa County Department of Environmental Quality. Park et al. calculated risk ratios for coccidioidomycosis for various parameters and found that when there is more ambient dust, the risk of coccidioidomycosis is higher ($p<0.001$), when the average temperature during the previous three months is higher, the risk of coccidioidomycosis is higher ($p=0.009$), and when cumulative rainfall during the previous seven months is less, the risk of coccidioidomycosis is higher ($p<0.001$) (Park, Sigel et al. 2005).

Rodents were also thought to be the source of coccidioidomycosis, but this was assuredly refuted decades ago (Smith, Beard et al. 1946; Wilson, Smith et al. 1953; Ajello 1967). There is evidence however of the *Coccidioides* species living in decaying animals "by growing through the carcass[es] as mycelium" (Sharpton, Stajich et al. 2009) and that the fungi, protected by bodily juices, could grow again in soil (Sorensen 1964). Maddy and Crecelius showed that

previously unaffected soil became contaminated with fungi after the burial of tissue from animals that were infected with coccidioidomycosis (Maddy and Crecelius 1967). Barker et al. found samples taken near rodent tunnels to have a noticeable rate of positivity for *Coccidioides* species (Barker, Tabor et al. 2012), and Lacy and Swatek observed increased animal activity on their sample locations (compared to adjacent locations) during a soil study (Lacy and Swatek 1974). Barker et al. found genetic variability among the *Coccidioides* species and report that it could be due to fungal reproduction (currently not fully understood) or may due to arthroconidia dispersion caused by weather, humans, or animals (Barker, Tabor et al. 2012).

2.4 OCCUPATION

Schmelzer and Tabershaw report several occupations that pose as exposure factors for coccidioidomycosis; these occupations include agriculture, construction, pipeline, highway, and utility operations (Schmelzer and Tabershaw 1968). Converse and Reed performed a study of animals as "biological air samplers" and found dog infectivity synonymous with agricultural and construction workers because of the proximity to soil and found monkey infectivity synonymous with "average human infections" due to distance from soil (Converse and Reed 1966). Gehlbach et al. explain that inflected dust in cotton can be an occupational exposure and source of infection (Gehlbach, Hamilton et al. 1973). Campbell mentions dust from running airplane propellers on military airfields (Campbell 1980) and dust from archeology also poses a risk. In 1972, seventeen of 34 students began having symptoms of coccidioidal infection within 15 to 18 days of the archeological excavation (Werner and Pappagianis 1973).

Military personnel and workers on military bases similarly face risk. As of 1999, a population of over 350,000 military personnel was stationed in *Coccidioides* endemic areas of the United States (Olivere, Meier et al. 1999). In 2001, ten of 22 Navy SEALs who were training in an endemic area of California had evidence of coccidioidal infection and all ten had "prominent symptoms" of coccidioidomycosis (90%: fever; 80%: cough; 70%: sweats; 60%: chills, weight loss; 50%: muscle aches; 40%: headache, labored breathing) (Crum, Lamb et al. 2002). Eight of the ten men (nine Caucasian, one Hispanic) first reported symptoms within two

to three weeks of arriving at the training base, five of the ten men had abnormal chest x-rays, five of the ten men were treated with antifungal medication for up to six months, and none of the men developed disseminated disease, though one man had his diving duties restricted because of coccidioidomycosis related lung involvement (Crum, Lamb et al. 2002).

In the mid-1940s Smith et al. performed monthly coccidioidomycosis testing of military personnel across four military airfields (Minter, Gardner, Lemoore, Merced) in California until the fields closed (Minter, Gardner, Lemoore) or coccidioidomycosis was shown to not be a problem (Merced) (Smith, Beard et al. 1946). Two groups were formed: uninfected (as determined by blood/serology) and infected (both symptomatic and un-symptomatic) (Smith, Beard et al. 1946). Monthly tallies were calculated using the number newly infected divided by the number tested, and summing the tallies, 25.10% of military personnel at Minter Field were found to have been infected from July 1941 through June 1942. Also at Minter Field, 1200 personnel initially tested negative, but due to the "rapid turnover of personnel" only 221 of the original 1200 were available 12 months later for retesting; of those 221 retests, 52 were found to be positive – a number similar to the sum of the monthly tallies (Smith, Beard et al. 1946).

Using summations of monthly tallies from the all four fields (not all fields available all years), Smith et al. found infection rates in 1941-1942 to be 25.10% (Minter as mentioned) and 20.99% (Gardner) with a precipitation during previous rainy season (October through March) of 1.96 inches, infection rates in 1942-1943 to be 13.51% (Minter), 11.57% (Gardner), 12.43% (Lemoore), and 0.63% (Merced) with 5.04 inches of precipitation during the previous rainy season (Smith, Beard et al. 1946), and infection rates in 1943-1944 to be 15.61% (Minter), 13.58% (Gardner), 6.21% (Lemoore), and 0.36% (Merced) with 5964 inches of precipitation during the previous rainy season (Smith, Beard et al. 1946). Smith et al. note that in 1941 the Minter and Gardner military airfields were being constructed and that associated dust could account for the higher rates during the 1941-1942 test period (Smith, Beard et al. 1946).

In 2007, ten of twelve construction workers (eleven men, one woman, eleven Caucasian, one Hispanic) evacuating soil at a California military base had evidence of coccidioidal infection, and seven of these ten symptomatic workers sought treatment from 21 or more medical personnel (Cummings, McDowell et al. 2010). The excavation work was typically performed in four 10-hour shifts per week from October 8-17 and during October 19-28 symptoms were first reported (100%: cough, chest pain; 90%: fever; 80%: sweats; 70%: muscle pain and/or joint

pain; 50%: headache; 40%: rash) (Cummings, McDowell et al. 2010). Seven of the ten affected workers had abnormal chest x-rays, and one worker (Caucasian) developed disseminated coccidioidomycosis of the skin (Cummings, McDowell et al. 2010).

2.4.1 Prison

Burwell et al. reviewed inmate medical records from January 2003 through October 2004 to find cases of coccidioidomycosis (n=79), cases were factored into groups (no treatment (n=4), ≤ 4 weeks ("early") since beginning antifungal treatment (n=32), > 4 weeks ("late") since beginning antifungal treatment (n=43)), and Wilcoxon, log rank, and χ^2 tests were used to test for differences in outcome between demographics of those who received treatment (n=75) (Burwell, Park et al. 2009). Symptoms most frequently reported were sweating (25 or 32%), fever (40 or 51%) and cough (57 or 72%), 51 of 79 inmate patients (75%) had an abnormal x-ray, and 55 of 79 inmate patients (70%) had first been given antibiotic treatment before being diagnosed with coccidioidomycosis (Burwell, Park et al. 2009). All the inmate patients were male, 34 (43%) were African American, 22 (28%) were Hispanic, 12 (15%) were Caucasian, 7 (9%) had diabetes, 2 (3%) were immunosuppressed, but Burwell et al. did not find any significant differences between those treated "early" versus "late" (Burwell, Park et al. 2009).

However, inmates at California state prisons located in endemic areas have higher rates of infection than the surrounding non-incarcerated population. Taft Federal Correctional Institution is located roughly an hour southwest of Bakersfield, California, in Kern County, a coccidioidomycosis endemic area. The six-year average rate (years 2005 through 2010) of coccidioidal infection in Taft Federal Correctional Institution (1,179 per 100,000) is 7.46 times higher than in Kern County (158 per 100,000), which includes Bakersfield, California (Fresno County Community Health Division 2011). For years 2004 to 2010, one percent of the Fresno County, California population, including the city of Fresno in California, resided as inmates at the Pleasant Valley State Prison, but these inmates made up 43 percent of coccidioidomycosis cases in Fresno County, and between 2007 to 2010, fourteen percent of the Kings County, California population resided as state prisoners (in Avenal State Prison, located in Avenal California, and California State Prison at Corcoran, located in Corcoran California), but these

prisoners made up 58 Percent of the coccidioidomycosis cases in Kings County (Fresno County Community Health Division 2011). Pleasant Valley State Prison's incidence rate in 2005 was estimated to be 3,000 per 100,000 (Pappagianis 2007).

2.5 DEMOGRAPHICS

Dodge et al. performed a randomized, stratified, sample of the Tucson, Arizona area from 1977 to 1979, and selected individuals who were willing to take two skin tests, a different reagent per arm (skin tests are currently not performed in the United States due to lack of available reagent), to determine reactivity in an endemic population (Dodge, Lebowitz et al. 1985). Dodge et al. found one reagent to be more sensitive and report that of 1,639 persons tested (33.4% tested positive), 739 were male (34.8% tested positive) and 900 were female (32.3% tested positive); broken down by age groups, there were 92 males (39.1 tested positive) and 91 females (35.2 tested positive) of 183 persons <15 years old (37.2 tested positive), there were 360 males (43.6 tested positive) and 400 females (43.8 tested positive) of 760 persons 15-54 years old (43.7 tested positive), and there were 287 males (22.3 tested positive) and 409 females (20.5 tested positive) of 696 persons >54 years old (21.3 tested positive) (Dodge, Lebowitz et al. 1985).

Summarizing available data from the Arizonan Medical Electronic Disease Surveillance Intelligence System (MEDSIS), Sunenshine et al. reported that males have a slightly higher risk of coccidioidal infection (55% male cases in 2005), and the risk for the elderly has increased over time, with 123 new coccidioidal cases over six years in the ≥65 year old age group (Sunenshine, Anderson et al. 2007). Flaherman et al. calculated relative risks (RR) using the Inpatient Hospital Discharge Data set from the California Office of Statewide Health Planning and Development (OSHPD) (Flaherman, Hector et al. 2007). Flaherman et al. found that pregnancy (RR=2.5; CI: 2.03-3.08), the elderly (age 56-69: RR=2.12; CI: 2.01-2.26; age ≥70: RR=2.74; CI: 2.54-2.97), African Americans (RR=2.68; CI: 2.48-2.91), and males (RR=2.14; CI: 2.03-2.27) were associated with an increased risk for coccidioidal related hospitalization (Flaherman, Hector et al. 2007). By summarizing the California Department of Public Health (CDPH) surveillance data from 2001-2009, Hector et al. found 65 percent of the 20,931

coccidioidal cases were male, 48 percent were Hispanic (versus the Californian population being only 35 percent Hispanic), and 12 percent were African American (versus the Californian population being only six percent African American) (Hector, Rutherford et al. 2011).

Rosenstein et al. performed a case-control study (Kern County, California) to determine risk factors for sever pulmonary and disseminated disease in patients with coccidioidomycosis; all patients enrolled in the study (n=380) had coccidioidomycosis (Rosenstein, Emery et al. 2001). Controls (n=270) had mild illness (flu-like), and there were two sets of cases: those with pulmonary disease (n=77; hospitalized with pneumonia) and those with disseminated disease (n=33; extrapulmonary or lesions). Rosenstein et al. found that, when compared univariately to controls, the pulmonary cases were older ($p \leq 0.05$) and the disseminated cases were more likely to be African American (22% (n=7) vs 6% (n=16)) or pregnant (27% (n=3) vs 2% (n=3)) (Rosenstein, Emery et al. 2001). Rosenstein et al. also found that the disseminated cases, in a multivariate model with controls, were more likely to be African American (OR: 4.6; CI: 1.4-15) or have income less than $15K (OR: 2.4; CI: 1.1-5.7) (Rosenstein, Emery et al. 2001).

Tabor and O'Rourke performed logistic regression to evaluate risks for coccidioidomycosis (Tabor and O'Rourke 2010). They surveyed 5,460 households (n=14,105) and divided the greater Tucson, Arizona area into three elevation groups (basin, riparian, and foothills) and two demographic strata (low Hispanic, high Hispanic). Their bivariate analysis showed that individuals age 45-54 were 4.72 times more likely (CI: 2.80-7.95), and individuals age >64 were 5.64 more likely (CI: 3.21-9.92), to have coccidioidomycosis than individuals less than 25 years of age (Tabor and O'Rourke 2010). They also showed that a residence in a low Hispanic, foothills area was 2.73 times more likely (CI: 1.36-5.49), and a residence in a low Hispanic, riparian area was 2.48 time more likely (CI: 1.26-4.87), to have resident with coccidioidomycosis than a resident of a low Hispanic basin area or a high Hispanic area respectively (Tabor and O'Rourke 2010). Tabor and O'Rourke also performed multivariate analyses and show that age again was a risk factor, as was being African American, non-Hispanic (OR: 2.70; CI: 1.55-4.72) (Tabor and O'Rourke 2010). Spinello et al. state that a risk factor for disseminated coccidioidal disease is ethnicity but that cases of pulmonary coccidioidal disease are "largely reflective of the population" (Spinello, Johnson et al. 2007).

2.5.1 Pregnancy

Pregnancy is a known coccidioidal risk factor (Bercovitch, Catanzaro et al. 2011). Crum and Ballon-Landa reviewed literature for cases of coccidioidomycosis in pregnancy women; of the cases they wound, two of five women (40%) were found to have disseminated coccidioidomycosis before becoming pregnant, four of eight women (50%) were found to have disseminated coccidioidomycosis during the first trimester of pregnancy, eight of 13 women (62%) were found to have disseminated coccidioidomycosis during the second trimester, 24 of 25 women (96%) were found to have disseminated coccidioidomycosis during the third trimester, and five of seven women (71%) were found to have postpartum disseminated coccidioidomycosis (Crum and Ballon-Landa 2006). Of the women with disseminated coccidioidomycosis, 17 of 18 were African Americans (94%), 7 of 9 were Hispanic (78%) and 13 of 23 Caucasian (56%) (Crum and Ballon-Landa 2006). Thirty-six percent of the women died, and lack of antifungal medication played a role in the number of deaths; only one in 18 deaths (6%) occurred in women receiving antifungal therapy (Crum and Ballon-Landa 2006).

Caldwell et al. claim there is reporting bias of coccidioidal disease in pregnant women (Caldwell, Arsura et al. 2000). Of 32 women (20 were Hispanic (63%), six were Caucasian (19%), five were African American (16%), and one was marked as other (3%)), Caldwell et al. found 12 women (37.5%) were diagnosed with coccidioidomycosis in the third trimester, three women (9%) had disseminated coccidioidomycosis (skin, skin and bone, meninges), 25 women (78%) had normal deliveries, 28 women had live births, and 23 of the women (72%) recovered from coccidioidomycosis without receiving any treatment, and none of the women died (Caldwell, Arsura et al. 2000). Crum and Ballon-Landa state that a risk factor for disseminated coccidioidal disease is pregnancy but that "pregnancy may not be a risk factor for the acquisition of coccidioidomycosis" (Crum and Ballon-Landa 2006). Immunological and hormonal changes in pregnant women are thought to be the cause of the increased coccidioidal dissemination seen in pregnancy (Powell, Drutz et al. 1983; Spinello, Johnson et al. 2007)

Neonatal coccidioidal infection is rare; as of 2007, only 15 neonatal cases were found in a literature review (Hooper, Lu et al. 2007). The main mode of transmission is through vaginal secretions containing coccidioidal infection (Crum and Ballon-Landa 2006). At approximately 25 weeks gestation, a 22-year old woman gave birth vaginally to a baby girl, who at two days

old, was hospitalized with coccidioidal infection, treated numerous time, but died at 34 days of age; the mother had disseminated coccidioidomycosis (Linsangan and Ross 1999). Transmission however is not always vaginal. At approximately 32 weeks gestation, a 20-year old women gave birth by cesarean to a baby girl, who at two weeks of age, was hospitalized with coccidioidal infection, treated numerous times, and discharged at 1.5 months of age; the mother had coccidioidomycosis including of the placenta (Charlton, Ramsdell et al. 1999).

2.5.2 Race / Ethnicity

Disseminated coccidioidomycosis is more likely to occur in persons of African American, Filipino/Asian, Hispanic race or ethnicity (Cox and Magee 2004) and potentially Native Americans (Sievers 1974; Huang, Bristow et al. 2012). However, the exact mechanism by which African Americans acquire disseminated coccidioidal disease is not known (Ruddy, Mayer et al. 2011). There are concerns about drawing inferences due to race or ethnicity due to the small sample sizes of severe cases in racial or ethnic groups, differences in access to medical care among races/ethnicities, and lack of accounting for racial heterogeneity in studies (Ruddy, Mayer et al. 2011). However, Ruddy et al. agree that African Americans are at greater risk for disseminated coccidioidomycosis, but they do not subscribe that same risk to the acquisition of the disease in African Americans (Ruddy, Mayer et al. 2011).

Nonetheless, population increases in endemic areas lend themselves to more construction and increased exposure from dust (Ruddy, Mayer et al. 2011). From 2000 to 2010, the population of African Americans in Arizona increased from 158,873 (3.1%; one race) to 259,008 (4.1%; one race) and from 185,599 (3.6%; race alone or in combination with other races) to 318,665 (5.0%; race alone or in combination with other races); the population of Arizona was 5,130,632 in 2000 and 6,392,017 in 2010 (United States Census Bureau 2010). In 2007, of the 36 percent of cases (total of 4,832 cases) reported to the Arizona Department of Health Services (ADHS), the incidence of coccidioidomycosis was nearly twice the rate for African Americans (53 per 100,000) than for Caucasians (37 per 100,00) and higher than American Indian/Alaska Natives (28 per 100,000), Asian/Hawaiian/Pacific Islanders (31 per 100,000), and Hispanics (15 per 100,000) (Arizona Department of Health Services 2007).

2.5.3 Genetics

Louie et al. propose that genetics influence the severity of coccidioidomycosis in humans (Louie, Ng et al. 1999). Deresinski et al. show an association between disseminated coccidioidal disease and blood group B (type B in Asians: 25.4%; type B in African Americans: 19% (American Red Cross 2012)) and human leukocyte antigen (HLA) genes (Deresinski, Pappagianis et al. 1979). As HLA genes are partly responsible for the inflammatory and immune responses in humans, and blood group B has been shown to be associated with coccidioidal infection, Louie et al. perform a case-control study to access genetic factors in relation to severity of coccidioidal disease (Louie, Ng et al. 1999). Cases (mild: n=83 and severe: n=109) were enrolled in California from the Kern County Health Department and Medical Center, numerous controls were matched via a literature review, and ABO blood type and four loci (DRB1, DQA1, DQB1, DPB1) of the HLA genes were considered (Louie, Ng et al. 1999).

Using the Hardy-Weinberg test for frequency with significance at $p<0.05$, Louie et al. found that loci DQB1 (in Hispanics with mild coccidioidomycosis) and DPB1 (in Hispanics and Caucasian with mild coccidioidomycosis and in Caucasians with severe coccidioidomycosis) "were not in equilibrium" (not as would be expected) (Louie, Ng et al. 1999). Louie et al. also found, using χ^2 tests, that Hispanics had differences ($p<0.001$) in disease severity based on blood type (Louie, Ng et al. 1999). While blood type B was less frequent in Hispanics with severe coccidioidomycosis than controls (OR: 0.36; $p=0.021$), and blood type A was more frequent in Hispanics with mild coccidioidomycosis than controls (OR: 5.53; $p<0.001$) (when comparing severity of disease in Hispanics cases) blood type B was associated with severe coccidioidal disease than blood type A (OR_A: 0.31; $p=0.024$; OR_B: 5.2; $p=0.039$); no such differences were found in Caucasians ($p=0.80$) or African Americans ($p=0.21$) (Louie, Ng et al. 1999).

2.6 CO-MORBIDITY

MEDSIS was used by Tsang et al. to find 493 participants for an enhanced surveillance study; of these participants, only 33 percent (n=164) had no co-morbid condition (Tsang, Anderson et al.

2010). Thirteen percent (n=62) had heart disease, 18 percent (n=90) had lung disease, 14 percent (n=70) had a malignancy, 2 percent (n=11) had a transplant, 2 percent (n=9) had HIV, and 15 percent (n=72) had diabetes (Tsang, Anderson et al. 2010). Flaherman et al., utilizing OSHPD, found that immune status (HIV: RR=13.9; CI: 12.5-15.5; AIDS: RR=34.5; CI: 31.0-38.4) was associated with an increased risk for coccidioidomycosis related hospitalization (Flaherman, Hector et al. 2007). Fish et al. found in their retrospective study that 32 of 77 HIV patients (41.6%) with concomitant coccidioidomycosis died with a range of 32 months (Fish, Ampel et al. 1990), and Rempe et al. found in their retrospective study that of 33 patients having coccidioidal fungemia (in blood), 29 of which had HIV, 24 were dead within one month (Rempe, Sachdev et al. 2007). In the Kern Country case-control study done by Rosenstein et al., patients with severe pulmonary disease and coccidioidomycosis were more likely to have diabetes (OR: 3.3; CI: 1.3-8.1) than patients with mild coccidioidomycosis (Rosenstein, Emery et al. 2001). However, Park et al. used NETSS to survey case patients based on when the onset their disease occurred (during a high or low coccidioidomycosis incidence period) and found that having diabetes at the onset of illness was not more prevalent during high incidence periods than low incidence periods (18% vs 17%; p=0.8) (Park, Sigel et al. 2005).

Leake et al. performed univariate and multivariate regression and evaluated 89 coccidioidal case patients with two sets of controls: 91 geographic controls (same county of residence as cases) and 58 lab negative controls (same county of residence as cases and with a negative coccidioidal lab report) (Leake, Mosley et al. 2000). They found that case patients lived in Arizona a median of 6.5 years, geographic controls lived in Arizona 19.5 years ($p<0.001$), lab negative controls lived in Arizona a median of 11 years ($p<0.01$), and individuals who lived in Arizona the shortest were at highest risk for acquiring coccidioidomycosis: <4 years (OR: 7.6; CI: 2.8-20.8), 4-12 years (OR: 4.7; CI: 1.9-11.5), 13-25 years (OR: 2.7; CI: 1.1-7.0) (Leake, Mosley et al. 2000). Using univariate analysis, Leake et al. found that case patients were more likely than geographic controls to have congestive heart failure (OR: 3.2; CI: 1.1-9.4), use corticosteroids (OR: 3.1; CI: 1.1-9.2), have cancer (OR: 2.4; CI: 1.2-4.8), and have an immunologic condition (OR: 10.5; CI: 1.3-84.9); no significant differences were found when using the lab-negative controls (Leake, Mosley et al. 2000). In the multivariate analysis, Leake et al. omitted immunologic conditions (only 5% of the study population) and found congestive heart failure (OR: 8.3; CI: 1.3-54.7), corticosteroid use (OR: 6.8; CI: 1.2-39.7), and non-skin

cancers (OR: 3.2; CI: 1.1-9.0) to be more likely in case patients than geographic controls; as before, no differences were found by using the lab-negative controls (Leake, Mosley et al. 2000).

2.6.1 Transplantation

Since 1970, organ transplantation in Arizona has been an active arena (Blair 2007). Between 1970 and 1979, eighteen patients of 260 patients identified as having kidney transplant (6.92%) were infected with coccidioidomycosis, and twelve of those eighteen had disseminated disease (Cohen, Galgiani et al. 1982). In 1987 the Mayo Clinic opened an Arizonan branch, and in 1999 surgeons at this clinic performed the first liver transplant in the Phoenix metro area (Mayo Clinic 2006). As of mid-2006, over 1100 transplants (n=1188) have been performed at the Mayo Clinic in Arizona, and only 18 of the 1188 patients (1.52%) were infected with coccidioidomycosis (Blair 2007). The decrease in coccidioidomycosis may be due to improved antifungal medication (introduction of azoles), better antirejection therapy (cellular level instead of global targeting), and routine screening tests that are performed "at the time of pre-transplant evaluation, at the time of transplantation surgery, four months after transplantation, and annually thereafter" (Blair 2007). In coccidioidomycosis endemic areas, the incidence of disease in transplant patients is now only one to three percent (Vikram and Blair 2009) and mirrors the rate of endemic area disease in non-transplant individuals (Blair 2007).

However, differences between transplant patients and non-transplant individuals in endemic areas exist. Rates of disseminated coccidioidomycosis can be as high at 75 percent in transplant patients with coccidioidomycosis (Cohen, Galgiani et al. 1982) whereas disseminated disease occurs in one half to five percent of all cases (Einstein and Johnson 1993; Pappagianis, Sun et al. 1993; Crum, Lederman et al. 2004). Locations of dissemination also differ; in non-transplant individuals, "skin, joints, bones, and meninges" are typical bodily locations for disseminated coccidioidomycosis (Galgiani 1993), whereas in transplant patients those locations, as well as the transplanted organ, spleen, liver, thyroid and pancreas, can be affected (Blair and Logan 2001). Of the transplant patients afflicted with coccidioidomycosis, approximately 50 percent are cases of reactivated coccidioidal disease occurring within one year of surgery (Logan, Blair et al. 2001). Transplant patients with coccidioidomycosis have a greater risk of

death; Braddy et al. reported two deaths (33%) in six cases of coccidioidomycosis from 205 transplant patients (Braddy, Heilman et al. 2006) and Cohen et al. reported 13 deaths (72%) in 18 cases of coccidioidomycosis from 260 transplant patients (Cohen, Galgiani et al. 1982).

Most cases of coccidioidomycosis in transplant patients are likely caused by immunosuppression therapy (new disease from exposure or reactivation or latent disease), but cases of coccidioidomycosis from donor transfer have been reported, though such transfer is not heavily documented (Tripathy, Yung et al. 2002; Wright, Pappagianis et al. 2003). White et al. discuss three transplant patients: a liver transplant in a 46-year old man who died day 17 post-transplant, a kidney transplant in a 28-year old man who died day 19 post-transplant, and a kidney transplant in a 58-year old woman, who after the previous deaths, was started on antifungal medication and remained asymptomatic (Wright, Pappagianis et al. 2003). All three patients received organs from the same donor, a 36-year old man with a coccidioidal history (Wright, Pappagianis et al. 2003). Tripathy et al. discuss a lung transplant in a 21-year old man who received organs from an Arizonan donor, developed coccidioidomycosis, but was placed on antifungal medication and survived (Tripathy, Yung et al. 2002).

2.7 BEHAVIORS

MEDSIS was used by Tsang et al. to find 493 participants for an enhanced surveillance study (Tsang, Anderson et al. 2010). Of these participants, 15 percent (n=76) were active smokers, 41 percent (n=203) were past smokers, and 41 percent (n=202) were never smokers; n=12 unknown smoking status (Tsang, Anderson et al. 2010). In the Kern Country case-control study done by Rosenstein et al., patients with severe pulmonary disease and coccidioidomycosis were more likely to smoke (OR: 2.4; CI: 1.1-5.4) than patients with mild coccidioidomycosis (Rosenstein, Emery et al. 2001). However, Park et al. used NETSS to survey case patients based on when the onset their disease occurred (during a high or low coccidioidomycosis incidence period) and found that smoking at the onset of illness was not more prevalent during high incidence periods than low incidence periods (21% vs 22%; p=0.8) (Park, Sigel et al. 2005). In both the univariate analysis and multivariate analysis by Leake, Mosley et al., coccidioidal cases were more likely to

smoke (univariate OR: 2.4; CI: 1.3-4.5 – multivariate OR: 3.7; CI: 1.4-9.6) than geographic controls from the same county of residence as the coccidioidal cases (Leake, Mosley et al. 2000).

2.8 CASE STUDIES

There have been a number of case studies. Osaki et al. reported the partial lung resection (to remove a fungal ball) in a 33-year old Japanese man who had lived in Bakersfield, California at the onset of his coccidioidomycosis illness (Osaki, Morishita et al. 2005). Dykes et al. report on a 78-year old man with coccidioidomycosis of the testicle that resulted in an orchiectomy (Dykes, Stone et al. 2005). Ruggles also reports of a case of testicular coccidioidomycosis in a 46-year old male resulting in an orchiectomy (Ruggles 2008). Gaidici and Saubolle discuss the case of a 37-year old veterinary assistant who was diagnosed with coccidioidomycosis of the arm after being bitten by a cat that was infected with coccidioidomycosis (Gaidici and Saubolle 2009). Kumar et al. discuss the case of a 68-year old man with coccidioidomycosis of the elbow who underwent a synovectomy (Kumar, Narasimhan et al. 2011). Miller et al. present the death of a 61-year old patient who received a bilateral lung transplant from a donor who had traveled to Mexico; the death of this patient was almost certainly due to disseminated coccidioidomycosis "secondary to immunosuppression treatment" (Miller, Hendren et al. 2004).

2.8.1 Misdiagnosis

Lung cancer was thought to be the probable diagnosis of a 71-year old man who received a lung lobectomy of tissue infected with coccidioidomycosis; the man had lived in Tucson, Arizona only a few months before experiencing symptoms of disease (Wang, Wen et al. 2011). Malignancy of the lung was also thought to be the diagnosis of a 34-year old male who underwent "broad excision" surgery of the lung after repeatedly traveling to Arizona for business in the year prior to surgery (Goegebuer, Nackaerts et al. 2009). Malignant bone lesions were initially considered in a patient with disseminated coccidioidomycosis of the skeleton; the patient was traveling in Arizona a couple of months before diagnosis (Arora, Taneja et al. 2012).

Further, a 65-year old patient who previously lived in Arizona and tested false-negative for *Coccidioides* species (after mentioning Valley Fever to her physicians) was misdiagnosed with lung cancer and underwent surgical resection (Petrini, Skold et al. 2003). Ugurlu et al. discuss a 77-year old woman, who was originally treated with antibiotics and incision/drainage treatment, was found to have coccidioidomycosis of the eyelid (Ugurlu, de Alba Campomanes et al. 2005). Lee et al. present a case of misdiagnosis in a 72-year old man with coccidioidomycosis who was thought to initially have carcinoma of the lung (Lee, Wilcox et al. 2008). Ugurlu et al. state that differential diagnoses should include coccidioidomycosis, as it "can mimic bacterial infection or malignant neoplasm" (Ugurlu, de Alba Campomanes et al. 2005).

2.9 COST ANALYSIS

Caldwell et al. reviewed 536 laboratory confirmed cases of coccidioidomycosis obtained from the Kern County Health Department, California, and estimated an overall average cost of $8,096 (hospitalization: 63%; clinic visits: 18%; lost wages: 12%; drug treatment: 7%) per patient, with an average cost of $5,400 per non-dissemination versus an average cost of $48,000 per dissemination (Caldwell, Welch et al. 1996). Caldwell et al. also found that four percent of the disseminated coccidioidal infections where responsible for 23 percent of the total cost, and for the 7,130 cases occurring from 1991 to 1993, they estimated a $56 million cost (approximately $82 million in 2012 dollars) not including costs beyond one year of illness (Caldwell, Welch et al. 1996). MEDSIS was used by Tsang et al. to find 493 participants from January 2007 through February 2008 for an enhanced surveillance study, and of these participants, 23 percent (n=111) used the emergency department as their first treatment center, 26 percent (n=128) sought treatment from a medical facility more than 10 times, 41 percent (n=200) were hospitalized with a median length of 6 days (IQR: 4-10 days; max: 306 days), and of those employed (n=225), 74 percent (n=167) missed worked on average 31 days (IRQ: 5-30 days; max: 365 days) (Tsang, Anderson et al. 2010). Using the Arizona Hospital Discharge Database for 2007, Tsang, Anderson et al. found that a primary diagnosis of coccidioidomycosis was responsible for 1,093 hospital visitations (total cost: $59 million; median cost per visit: $33,000), and that a primary

diagnosis or secondary diagnosis of coccidioidomycosis was responsible for 1,735 hospital visitations (total cost: $86 million; median cost per visit: $30,000) (Tsang, Anderson et al. 2010).

Using Decision Maker software, Barnato et al. create a Markov model to simulate the potential consequences of a yet to be developed vaccination; they consider two age cohorts (children: ≤17 year of age; adults: 18-65 years old) across 10 epidemic counties (8 in Arizona and 2 in California) (Barnato, Sanders et al. 2001). As they are interested in average costs (e.g., cost associated with coccidioidal meningitis) and not individual effects, they do not consider risk factors such as gender, immunosuppression conditions, co-morbidity, pregnancy, or smoking; race is considered in calculating the expected value of dissemination across all peoples, and age in considering in determining immunity (children: 14.5%; adults: 47.5%) (Barnato, Sanders et al. 2001). They run their Markov model under three scenarios (not vaccinating, screening and then vaccinating, vaccinating everyone) and find that vaccinating all children is the dominant strategy (most effective at least cost), increasing life expectancy by 0.47 days (and decreasing average cost of illness by $33) over not vaccinating children; in adults, life expectancy days increase by 0.16 days and average cost of illness increases by $123 (Barnato, Sanders et al. 2001).

2.10 TESTS FOR COCCIDIOIDOMYCOSIS

There are a number of serological tests that can be used to detect the coccidioidal antibodies IgM and IgG. Tests for IgM include immunodiffusion, enzyme immunoassay, latex agglutination, or tube precipitin; tests for IgG include immunodiffusion, enzyme immunoassay, or complement fixation (Arizona Department of Health Services 2009). For determining coccidioidal infection within the first month of infection, IgM is more detectable, but as length of disease progresses, IgM diminishes and IgG becomes the dominate antibody (Pappagianis and Zimmer 1990). Nevertheless, false positive and negative results are possible. Kuberski et al. reviewed 32 patient records, and of 17 of 32 patients found to have positive IgM and negative IgG when tested with enzyme immunoassay, only three of these patients (17.6%) were probable for coccidioidal disease (82.4% false-positive rate) (Kuberski, Herrig et al. 2010). On the other hand, Blair and Currier conclude that enzyme immunoassay testing results in low false positives after reviewing

405 patients records from 19999 to 2003; they found 28 patients having positive IgM when tested with enzyme immunoassay, and not one of the 28 test results was falsely positive for coccidioidomycosis (Blair and Currier 2008). However, sensitivity is "modest at best" because many patients have delayed immune responses (Blair, Coakley et al. 2006; Lake 2012).

2.11 SURVEILLANCE IN ARIZONA

Physician reports to Arizona Department of Health Services (ADHS) of coccidioidomycosis started in the 1930s (Hector, Rutherford et al. 2011). Prior to 1994, cases of coccidioidomycosis reported to the ADHS were not required to be confirmed via laboratory testing, but beginning in 1994, ADHS adopted the Council of State and Territorial Epidemiologists (CSTE) coccidioidal case definition (compatible symptomatology and laboratory confirmation) (Centers for Disease Control and Prevention 1996). In 1995, reported cases of coccidioidomycosis became nationally notifiable (Centers for Disease Control and Prevention 2011) and in 1997, cases of laboratory confirmed coccidioidomycosis became notifiable at the state level (Tsang, Anderson et al. 2010). Currently ADHS uses CSTE laboratory confirmation only for case surveillance with the understating that coccidioidal counts are underestimates and biased toward individuals with medical coverage and disease severity (Tsang, Anderson et al. 2010). Laboratory tests for reportable disease include "cultural, histopathologic, or molecular evidence" or "immunologic evidence" (detection of either immunoglobulin M (IgM) or immunoglobulin G (IgG) antibody responses using varied testing procedures including enzyme immunoassay) or currently unavailable skin tests (due to lack of reagents in the United States) (Arizona Department of Health Services 2009). Beginning in 2009, enzyme immunoassay test results were reported not in combination with other tests as had previously been reported, and this change increased the number of reported coccidioidal cases in Arizona (73 per 100,000 in 2008 and 186 per 100,000 in 2010) (Hector, Rutherford et al. 2011; Benedum and Tsang 2012). However, severe underreporting is evident. Chang, Anderson et al. found that serological testing for coccidioidomycosis was infrequent in two healthcare facilities (2% with 95% CI: 0.04%-8% and 13% with 95% CI: 6%-22%) serving community-acquired pneumonia (CAP) patients in

Phoenix, Maricopa County, Arizona, and of 60 selected randomly CAP patients, coccidioidal disease was present in only 15% of patients (95% CI: 8%-26%) (Chang, Anderson et al. 2008).

2.12 VACCINATION

The current treatment options include azoles (five-sided antifungal agent with at least one nitrogen and one or more oxygen, nitrogen or sulfur) and amphotericin B (large ringed antifungal agent containing hydroxyl groups and double bonded carbons) (Thompson 2011) but there is currently no vaccine to protect individuals against the acquisition of coccidioidomycosis (Cole, Xue et al. 2004; Valley Fever Center for Excellence 2010). Levine et al. described the outcome of an experimental inoculation of mice with "spheruleendospore [sic] vaccine" (a type of vaccine thought to be "superior in mice" than arthroconidia or mycelium based vaccines); after vaccination more than half of the mice that were exposed to a lethal dose of free arthroconidia survived for six months (Levine, Cobb et al. 1961). Pappagianis et al. discuss human (volunteer) trials (Levine and Smith 1967; Pappagianis, Levine et al. 1967) with a "whole spherule vaccine" where six of 78 inoculated persons developed antibodies, but that humans could only tolerate 1/400 (0.25%) of the dose needed to create an immunological response in mice (Pappagianis, Hector et al. 1979), likely due to the formaldehyde used in the preparation (Lake 2012). They further describe their work to create a "subcellular extract" (instead of using the entire spherule, they use "phosphate-buffered saline" to extract antigen from walls of spherules "killed with formaldehyde") in hopes of creating a tolerable vaccine for humans that would generate a stronger immunological response in the population (Pappagianis, Hector et al. 1979).

Additionally, Lecara et al. evaluated a vaccine in mice using mycelium cell wall antigen and found that mice were protected against coccidioidomycosis when challenged with arthroconidia (Lecara, Cox et al. 1983). With a 0.5 mg injection, 20 of 25 mice (80%) survived 35 days ($p<0.0001$), with a 1.5 mg injection, 12 of 25 (48%) mice survived 35 days ($p<0.02$), and with a 3.0 mg injection, 11 of 25 (44%) mice survived 35 days ($p<0.01$) (Lecara, Cox et al. 1983). However, all mice (surviving and non-surviving) were infected with coccidioidomycosis (Lecara, Cox et al. 1983). Via intranasal challenge, Lecara et al. found partial protection in mice

(0.2 mg dose: 16 (68%) of 25 mice survived 40 days ($p>0.05$); 1.0 mg dose: 20 (83%) of 24 mice survived 40 days ($p<0.05$)) (Lecara, Cox et al. 1983). However, all mice surviving the 1.0 mg intranasal challenge were infected with coccidioidomycosis (Lecara, Cox et al. 1983). Pappagianis, in continuation of his previous work with formaldehyde and spherule cell walls, conducted a Phase III double blind clinical trial from 1980 to 1985 with 2,867 study participants who were all skin test negative and absent any coccidioidal history (Pappagianis 1993). Participants were randomized into two groups: vaccine (n=1,436) and placebo (n=1431). The vaccine group received three injections of the spherule derived vaccine, and the placebo group received three injections of a sodium chloride solution (Pappagianis 1993). Cases of coccidioidomycosis (symptomatic or serological) were then compared between groups (Pappagianis 1993). Of the vaccine participants, nine individuals were serologically positive, and nine individuals were symptomatically positive, for coccidioidomycosis, and of the placebo participants, twelve individuals were serologically positive, and thirteen individuals were symptomatically positive, for coccidioidomycosis (Pappagianis 1993). However, no statistical significance in cases between the vaccine and placebo groups was found, and disease severity was indiscernible between the groups as well (Pappagianis 1993).

There is renewed interest in nikkomycin Z (complex antifungal agent with three five-sided rings in combination with a mix of carbon, hydrogen, nitrogen, and oxygen bonds) as a potential vaccination for coccidioidomycosis (Thompson 2011). Nikkomycin Z was discovered by Bayer in the 1970s (Hector 2011), has been used in mice experiments were it was "well tolerated" (Hector, Zimmer et al. 1990), but development was stopped until 2005 when rights to nikkomycin Z were acquired by the University of Arizona (Valley Fever Center for Excellence 2008). Nikkomycin Z works by inhibiting the enzyme growth in chitin, the fibrous substance making up spherule cell walls, of fungal species (Thompson 2011). Nikkomycin Z was also tolerated well by dogs with seven of nine dogs showing improvement of their coccidioidal symptoms after two months of treatment (Valley Fever Center for Excellence 2012). Nikkomycin Z clinical trials are planned, though future studies are dependent on the manufacturing of additional nikkomycin Z (Valley Fever Center for Excellence 2012).

3.0 METHODOLOGY

3.1 STOCHASTIC MODEL OVERVIEW

We develop a stochastic, agent-based epidemiological model for coccidioidomycosis (Aim 1). In this model each person in a population, such as a zip code or county, is modeled as an individual agent with his own demographic and clinical characteristics. At any moment in time, each living person in the population is categorized as either *susceptible to infection* or *infected*, where the latter term indicates that the individual was previously infected regardless of whether the individual ever was or ever will be symptomatic. Thus, an individual belongs to one of three states at any moment in time: susceptible to infection, infected, or dead.

Further, we run our model over a given time interval. First we initialize the population by setting the state and all individual-level characteristics for each individual. As the model is simulated over time, individuals may be born, susceptible individuals may become infected, and live individuals may die. The probabilities of these occurrences depend on individual demographic and clinical characteristics as well as on climate, weather, and environmental conditions. Additionally in our model the individual-level characteristics impact both the severity and clinical course of the infection in infected individuals.

3.2 MATHEMATICAL AND PROBABILISTIC NOTATION

In this section we define mathematical and probabilistic notation we use to describe our model.

3.2.1 Mathematical Notation

In this work we use the bold font to denote vectors, which may either be constant vectors or random vectors. We define the zero vector, denoted $\mathbf{0}$, as a vector of zeros of the appropriate length. In most circumstances the length of such a zero vector will be clear from the context; otherwise it will be explicitly indicated using a subscript. For example, a zero vector of length n would be denoted $\mathbf{0}_n$. We use the letter I to denote an indicator function for a subset of some space $\mathcal{X}$. Specifically, for a subset $A \subset \mathcal{X}$, the function I_A is defined for all $x \in \mathcal{X}$,

$$I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}.$$

Similarly, we use the letter I to denote a real-valued function of a Boolean variable whose value is one if the Boolean argument is true and zero if the argument is false, i.e.,

$$I(B) = \begin{cases} 1 & \text{if } B \text{ is } \textbf{true} \\ 0 & \text{if } B \text{ is } \textbf{false} \end{cases}.$$

3.2.2 Probability Distributions

We denote probability distributions in several different ways. For common, named distributions with standard parameterizations we indicate the distribution for a random variable using the twiddle notation (e.g., $X \sim N(2,0.5)$ would indicate that X has a normal distribution with mean parameter 2 and variance parameter 0.5). The same holds for conditional distributions (e.g., $X|(Y = y) \sim N(y, 1)$ indicates that the conditional distribution for X given Y is normal with mean Y and variance 1). For discrete distributions, we may write out explicitly the probability mass function, that is, writing the probability that the random variable equals particular values. If a random variable can only take two values, we may write only the probability that the variable equals one of the values and leave it as understood that the probability of the other possible value is one minus that first probability. For example, if $X$is a discrete random variable

that can only take on the values of 0 and 1, we may simply write $P(X = 1) = 0.4$ and leave it implied that $P(X = 0) = 0.6$. For random variables with continuous distributions we may write the probability density function, the cumulative distribution function, or the survival function.

3.2.3 Kronecker's Delta Distribution

Kronecker's delta distribution is a point mass distribution. If X is a random variable such that $P(X = c) = 1$ for some constant c then X is said to have a point mass distribution at c and we may write, $X \sim \delta_c$. Similarly, if $\boldsymbol{X}$ is a random vector such that $P(\boldsymbol{X} = \boldsymbol{c}) = 1$ for some constant vector $\boldsymbol{c}$ then $\boldsymbol{X}$ is also said to have a point mass distribution at $\boldsymbol{c}$ and we may write, $\boldsymbol{X} \sim \delta_{\boldsymbol{c}}$.

3.2.4 Uniform Distribution

The uniform distribution over a continuous interval (a, b) is denote by $U(a, b)$. As it is a continuous distribution, the uniform distribution over half-closed and fully-closed intervals, $(a,b]$, $[a,b)$ and $[a, b]$ are the same distribution, $U(a, b)$.

3.3 POPULATION

The population in the model refers to a set of individuals that reside in a particular geographic area during a fixed time interval. For example, the population may consist of all individuals that were residents of Maricopa County, Arizona at any time between January 1, 2011 and December 31, 2011. Additionally, we make assumptions about the population. First, we assume that the population only consists of residents of the geographic region. This assumption is in place because most of the epidemiological and demographic statistics focus solely on residents and we therefore also focus only on residents. It is important to remain aware of this assumption, particularly when using the model to estimate the economic impact of policy decisions. A second assumption is that the population is closed, that is there is no migration in or out of the region during the time interval. Individuals may only enter the population via birth and

individuals may only leave the population via death. This assumption of a closed population is a reasonable simplifying assumption for short time intervals such as a single year. At a fixed moment in time, all of the characteristics of an individual are captured in a vector $\boldsymbol{Y}$. To simplify the notation, we subdivide the vector $\boldsymbol{Y}$ into four components, $\boldsymbol{Y} = (\boldsymbol{X}, S, \boldsymbol{C}, \boldsymbol{T})$ where $\boldsymbol{X}$ is a vector of demographic and risk factors for either an initial infection or for complications from infection, S is a state variable that indicates whether an individual is susceptible to infection, infected, or dead, $\boldsymbol{C}$ is a vector of clinical variables that describe, for an infected individual, the symptoms and severity of the disease, and $\boldsymbol{T}$ is a vector of temporal variables that describe the time course of the disease.

3.3.1 Demographic and Risk Factors

We express the vector of demographic and risk factors as follows: $\boldsymbol{X} = (X_A, X_G, X_R, X_D, X_C, X_S, X_V, X_L)$. We next describe the individual components of $\boldsymbol{X}$.

Age. The variable X_A denotes the current age of an individual, in years. This is a continuous variable; hence, an individual may have an age of 26.43.

Gender. The variable X_G denotes the gender, where $X_G = 1$ for female and $X_G = 2$ for male.

Race/Ethnicity. The variable X_R denotes the race/ethnicity group of an individual. We use the following race/ethnicity groups in accordance with the 2009 Maricopa County Health Status Report (MCHSR) (Maricopa County Department of Public Health 2011): White ($X_R = 1$), Hispanic ($X_R = 2$), African-American ($X_R = 3$), American Indian ($X_R = 4$), Asian ($X_R = 5$), and Other ($X_R = 6$). Individuals in two or more racial/ethnicity groups are classified as "Other."

Dust Exposure. The variable X_D denotes the dust exposure group of an individual. In the present model, we include a "Normal" group ($X_D = 0$) and an "Elevated Risk" group ($X_D = 1$) that would include workers in construction, agriculture, and military personnel.

Co-morbidity. The variable X_C denotes the co-morbidity group of an individual. The groups are as follows: $X_C = 0$ for no co-morbid condition, $X_C = 1$ for diabetic, immunocompetent, $X_C = 2$ for non-diabetic, immunocompromised, and $X_C = 3$ for diabetic and immunocompromised.

Smoking. The variable X_S denotes smoking status. Currently, $X_S = 1$ for individuals that smoked in the previous 6 months and $X_S = 0$ otherwise.

Vaccination. The variable X_V is an indicator variable for whether the individual was vaccinated against coccidioidomycosis. Thus $X_V = 1$ for vaccinated individuals, and $X_V = 0$ otherwise.

Location. The variable X_L is a categorical variable that indicates the weather station that best represents the individual's primary location, which might refer to a home or work location.

3.3.2 State Variable

The variable S denotes the current state of an individual. Individuals that are susceptible to infection belong to state $S = 1$; individuals that have previously been infected belong to state $S = 2$; and individuals that have died due to any cause belong to state $S = 0$.

3.3.3 Clinical Variables

We express the vector of clinical variables as follows: $\boldsymbol{C} = (C_S, C_P, C_D, C_M)$. These variables indicate the symptoms and outcome for cases of coccidioidomycosis. We next describe the individual components of $\boldsymbol{C}$.

Symptomatic. The variable C_S is an indicator variable for whether an infected individual has had or will have symptoms. Thus $C_S = 1$ for symptomatic cases, and $C_S = 0$ for asymptomatic, sub-clinical cases and for non-infected individuals.

Severe Pulmonary Coccidioidomycosis. The variable C_P is an indicator variable for whether an infected individual has severe pulmonary coccidioidomycosis, which may include cavities, nodules, and chronic pulmonary coccidioidomycosis.

Disseminated Coccidioidomycosis. The variable C_D is an indicator variable for whether an infected individual has extrapulmonary disseminated coccidioidomycosis.

Mortality. The variable C_M is an indicator variable for whether an infected individual died as a result of coccidioidomycosis.

3.3.4 Temporal Variables

We express the vector of temporal variables as follows: $\boldsymbol{T} = (T_I, T_{S0}, T_{S1}, T_D, T_M)$. We next describe the individual components of $\boldsymbol{T}$.

Time of Infection. The variable T_I denotes the time of infection. The time of infection is defined to be zero for non-infected individuals.

Time of Symptoms. The variables T_{S0} and T_{S1} denote the times of symptom onset and cessation of symptoms, respectively. These variables are defined to be zero for non-infected and asymptomatic infected individuals.

Time of Dissemination. The variable T_D denotes the time of diagnosis for disseminated infection. This variable is defined to be zero for individuals without disseminated disease.

Time of Mortality. The variable T_M denotes the time of death due to coccidioidomycosis. This variable is defined to be zero for individuals that do not die from coccidioidomycosis.

3.3.5 Strata

As coccidioidomycosis incidence statistics are often reported by age group rather than precise age, it is convenient to define age groups for risk stratification. Let *m* denote the number of age strata, which are defined by the constants $a_0, \ldots, a_m$ such that age stratum j contains ages in the interval $A_j = (a_{j-1}, a_j]$ where all ages are continuous variables and expressed in years. Now with the stratification by age, the population is further stratified by age, gender and

race/ethnicity. We denote the number of living individuals in age stratum j, gender g and race/ethnicity r at time τ_i by $N_{i,j,g,r}$. Formally,

$$N_{i,j,g,r} = \#\{\boldsymbol{Y} \in \Pi_i | X_A \in A_j, X_G = g, X_R = r, S > 0\},$$

where # denotes the cardinality of the set.

3.4 ENVIRONMENTAL FACTORS

A number of environmental factors have been shown to be related to incidence of coccidioidomycosis. We adopt the model of Park et al. that uses four environmental variables to forecast coccidioidomycosis infections. These variables, which we organize into a vector $\boldsymbol{E} = (E_{R2}, E_{R7}, E_P, E_T)$, are as follows.

2-Month Rainfall. The variable E_{R2} denotes the cumulative rainfall during the previous 2 months.

7-Month Rainfall. The variable E_{R7} denotes the cumulative rainfall during the previous 7 months.

PM10. The variable E_P denotes the current daily average PM10.

Temperature. The variable E_T denotes the average temperature during the previous 3 months in degrees Fahrenheit.

3.5 DISEASE MODEL

Conceptually, we think of the disease and its consequences as happening in the following order. First, an individual becomes infected. Second, each individual that is infected has some probability of being symptomatic. Third, a symptomatic individual may develop certain symptoms and outcomes. Each symptomatic individual has some probability of having severe pulmonary symptoms. This infected individual also has some probability of developing disseminated coccidioidomycosis as well as having some probability of dying from the disease.

These latter events are clearly not independent, as individuals with disseminated disease have a much greater likelihood of death than those with mild flu-like symptoms. In this section we describe the basic structure and define the notation of the probabilistic disease model. Detailed formulae for specific probabilities are provided in subsequent sections.

The course of the disease begins with the initial infection. The probability of annual infection for a susceptible individual is denoted by ρ. This probability of infection accounts for individual-level characteristics including stratum. The daily probability of infection is based on ρ but adjusted for the current and historical environmental conditions. The precise equation used in our model is given in Section 3.6.3.2. The probability ρ is computed as follows:

$$\rho = \frac{p_{ISR}/p_{Susc}}{p_{S|I} \times \sum_l p_{Gl|IS} p_{R|ISGl}}.$$

The numerator is the probability of infection, symptoms and report given that the individual is susceptible. p_{ISR} is estimated based upon historical rates of reported infections. p_{Susc} must be estimated, possibly using skin test results data. The denominator above is simply the probability that the individual is symptomatic and reported given infection. The factors in the denominator are described below. For an infected individual, we next determine whether the individual is symptomatic. For a non-vaccinated individual, this probability is denoted by $p_{S|I}$. For a vaccinated individual, the probability is adjusted according to the vaccine effectiveness. The precise equation used in our model is given in Section 3.6.3.4. For a symptomatic individual, the probabilities of various symptoms and outcomes depend on individual-level characteristics.

The probabilities for a given individual are denoted as $p_{Pulm|IS}$, the probability of severe pulmonary symptoms (i.e., $C_P = 1$), given that the individual is infected and symptomatic, $p_{Diss|IS}$, the probability of developing disseminated coccidioidomycosis (i.e., $C_D = 1$), given that the individual is infected and symptomatic, and $p_{Death|IS}$, the probability of mortality from coccidioidomycosis (i.e., $C_M = 1$), given that the individual is infected and symptomatic. Operationally, these probabilities are calculated population-wide, i.e., for a "typical" individual in the population but are then adjusted to account for individual characteristics. The population-wide probabilities are denoted as $p^0_{Pulm|IS}$, $p^0_{Diss|IS}$, and $p^0_{Death|IS}$ respectively, and we estimate these probabilities to be 20.3%, 8.7% and 1.1% respectively, based upon Rosenstein et al. (Rosenstein, Emery et al. 2001). We group symptomatic individuals into one of four mutually exclusive and exhaustive symptom/outcome groups. Group 1 consists of individuals with mild

symptoms only ($C_P = C_D = C_M = 0$). Group 2 contains individuals with non-lethal severe pulmonary symptoms ($C_P = 1$ and $C_D = C_M = 0$). Individual in Group 3 develop non-lethal dissemination with or without severe pulmonary symptoms $\big((C_P = 0 \text{ or } C_P = 1) \text{ and } C_D = 1 \text{ and } C_M = 0\big)$. Finally, Group 4 contains individuals that die from the disease $\big((C_P = 0 \text{ or } C_P = 1) \text{ and } (C_D = 0 \text{ or } C_D = 1) \text{ and } C_M = 1\big)$.

We compute the probabilities of these four symptom/outcome groups as follows. The probability that the individual has mild symptoms only (Group 1) is computed as $p_{G1|IS} = \big(1 - p_{Pulm|IS}\big) \times \big(1 - p_{Diss|IS}\big) \times \big(1 - p_{Death|IS}\big)$, the probability of non-lethal severe pulmonary symptoms (Group 2) as $p_{G2|IS} = p_{Pulm|IS} \times \big(1 - p_{Diss|IS}\big) \times \big(1 - p_{Death|IS}\big)$, the probability of non-lethal disseminated disease (Group 3) as $p_{G3|IS} = p_{Diss|IS} \times \big(1 - p_{Death|IS}\big)$, and the probability of death from coccidioidomycosis (Group 4) as $p_{G4|IS} = p_{Death|IS}$. Note that these computations do *not* assume full mutual independence of severe pulmonary symptoms, dissemination, and death. The probability of death, for example, will be set to be higher for an individual with disseminated disease than an individual without disseminated disease, and these values will remain consistent with the above symptom/outcome group probabilities. The likelihood that a case is reported to the health department is then set to depend on the symptom/outcome group. We let $p_{R|ISGl}$ denote the probability that an infected, symptomatic individual in Group l is reported to the health department. In our base case, we set $p_{R|ISG1}$=0.25, $p_{R|ISG2}$=0.80, $p_{R|ISG3}$=0.95, and $p_{R|ISG4}$=0.99. There was little literature upon which to base these values, so these values are examined in a sensitivity analysis.

3.6 POPULATION DYNAMICS

Over time the population will change in many ways. Individuals will be born and individuals will die. The characteristics of individuals in the population will also change. Individuals will age, some individuals will become infected, and some of those will suffer various symptoms and consequences from the infection including death. In this section we describe how the population changes over time in our model. We consider that the model will be run over a fixed time span.

We discretize the time span into k intervals of equal length L, which is expressed in years. We denote the start of the time span by τ_0. It follows that interval i is $\Delta_i = (\tau_{i-1}, \tau_i]$, where $\tau_i = \tau_0 + i \times L$ for $i = 1, \ldots, k$. The full time span over which the model will run is denoted by $\Delta^* = (\tau_0, \tau_k]$. We denote the population at time τ_i by Π_i. Formally, Π_i is a set of $\boldsymbol{Y}$ vectors which describe the characteristics of each individual in the population at time τ_i. Note that some of these characteristics are dependent on the specific time τ_i, such as age (X_A) and state (S), while other characteristics may refer to past or future events such as having symptomatic coccidioidomycosis (C_S).

Because we discretize time, we express the population dynamics as the generation of the population Π_i from Π_{i-1}. There are three major mechanisms of change:

- Births – new individuals are introduced into Π_i that were not represented in Π_{i-1}.
- Deaths – some individuals in Π_{i-1} may not live past time τ_i. Rather than exclude those individuals from Π_i, we simply update their state variable to indicate their death $(S = 0)$.
- Updates – most individuals in Π_{i-1} will also be represented in Π_i with updated characteristics.

In order to describe population dynamics we need to follow individuals from Π_{i-1} to Π_i. Notationally we will use the prime character to denote updated variables. Hence, an individual $\boldsymbol{Y} \in \Pi_{i-1}$ becomes $\boldsymbol{Y}' \in \Pi_i$. This notational convention also applies to the sub-vectors of $\boldsymbol{Y}$ and their constituent variables.

3.6.1 Births

We begin by defining $\beta_{j,r}$ to be the birth rate per woman per year in age stratum A_j and of race/ethnicity r. We also define α_i to be the probability that a birth is in time interval Δ_i given that the birth is in the full time span Δ^*. The α_i and $\beta_{j,r}$ values could be either estimated from historical data or explicitly set. In Section 3.8.3 we describe how we can estimate α_i and $\beta_{j,r}$ from historical data. Hence the probability of a birth to a mother in age stratum A_j and of race/ethnicity r during time interval Δ_i is $\beta_{j,r} \times \alpha_i$ assuming that the date of the birth is independent of the age and race/ethnicity of the birth mother. We denote the number of births to

mothers in age stratum A_j and of race/ethnicity r during time interval Δ_i by $B_{i,j,r}$. Assuming that women give birth independently, $B_{i,j,r}$ follows a binomial distribution with parameters $N_{i,j,1,r}$ and $\beta_{j,r} \times \alpha_i$. Here we use the Poisson approximation to the binomial distribution and simulate,

$$B_{i,j,r} \sim Pois\left(N_{i,j,1,r} \times \beta_{j,r} \times \alpha_i\right).$$

Then for $l = 1, \ldots, B_{i,j,r}$, we define a new individual $\boldsymbol{Y}_{i,j,r,l} = \left(\boldsymbol{X}_{i,j,r,l}, S_{i,j,r,l}, \boldsymbol{C}_{i,j,r,l}, \boldsymbol{T}_{i,j,r,l}\right)$ as follows. We set the vector of demographic and risk factors according to $\boldsymbol{X}_{i,j,r,l} = \left(X_{A,l}, X_{G,l}, X_{R,l} = r, X_{D,l} = 0, X_{C,l} = 0, X_{S,l} = 0, X_{V,l} = 0\right)$ where $X_{A,l} \sim U(0, L)$ and $P\left(X_{G,l} = 1\right) = P\left(X_{G,l} = 2\right) = 0.5$. Note here that we assume that the child's race/ethnicity is the same as the mother. We also assume that the gender of the child is independent of the age and race/ethnicity of the mother and independent of the date of the birth. We set the initial state to be susceptible, i.e., $S_{i,j,r,l} = 1$. Finally, we set $\boldsymbol{C}_{i,j,r,l} = \boldsymbol{0}$ and $\boldsymbol{T}_{i,j,r,l} = \boldsymbol{0}$.

3.6.2 Deaths

Death may occur due to coccidioidomycosis or due to other causes. In this section we consider the modeling of deaths due to other causes. Informally we define ϕ_{ijgr} as the probability of death due to other causes during time interval Δ_i for an individual in age stratum A_j, of gender g and of race/ethnicity r. This cannot be the formal definition due to the following complication. When individuals are simulated to become infected, the existence and times of future disease-related events such as dissemination and death are also simulated. The rates of such events as well as the distribution of times for these events are based on observed data. Such observed data obviously presume that an individual did not die prior to such an event. Therefore, as we simulate the existence and times of future events, it becomes implied that an individual will not die due to other causes prior to their events occurring. If, for example, an individual with coccidioidomycosis is simulated to have disseminated disease six months after infection but is allowed to die only one month after infection, the simulated rates of dissemination would, on average, be less than observed rates. We therefore formally define ϕ_{ijgr} as follows,

$$\phi_{ijgr} = P\left(S' = 0 \middle| \boldsymbol{Y} \in \Pi_{i-1}, S > 0, X_A \in A_j, X_G = g, X_R = r, T_{MAX} \leq \tau_{i-1}\right),$$

where T_{MAX} is defined to be the maximum value of the vector $\boldsymbol{T}$ excluding the time of report, T_R. That is, $T_{MAX} = \max\{T_I, T_{S0}, T_{S1}, T_D, T_M\} = \max\{T_I, T_{S1}, T_D, T_M\}$. Observe that the definition of the probability ϕ_{ijgr} is for an individual that is eligible for death for the entire interval Δ_i. In our simulation model we will also need the probability of death for when an individual is only eligible for death during a portion of the interval Δ_i, which will be defined below to be proportional to ϕ_{ijgr}.

3.6.3 Updates

In this section we describe the update from Π_{i-1} to Π_i. Specifically, for an individual $\boldsymbol{Y} = (\boldsymbol{X}, S, \boldsymbol{C}, \boldsymbol{T}) \in \Pi_{i-1}$ with an updated vector $\boldsymbol{Y}' = (\boldsymbol{X}', S', \boldsymbol{C}', \boldsymbol{T}') \in \Pi_i$, we define the conditional distribution of $\boldsymbol{Y}'$ given $\boldsymbol{Y}$. To define this conditional distribution, we consider the value of S, the prior state of the individual.

3.6.3.1 Transition from State 0 (Death)

Individuals that are already dead remain dead, and all of their associated variables remain unchanged. Thus, $P(\boldsymbol{Y}' = \boldsymbol{Y} | \boldsymbol{Y} \in \Pi_{i-1}, S = 0) = 1$.

3.6.3.2 Transition from State 1 (Susceptible)

Individuals that are susceptible to infection may transition to any of the other two states, infected and death, or remain susceptible. It is also possible for an individual to become infected and die within the same time interval. Hence we introduce a temporary state variable $\overline{S}$ that represents the initial destination state and a temporary time variable $\overline{T}$ that represents the time of the transition. For an individual that becomes infected and dies within the same interval, $S = 1, \overline{S} = 2$ and $S' = 0$. The conditional distribution of $\overline{S}$ is as follows,

$$P\left(\overline{S} = \bar{s} \middle| \boldsymbol{Y} \in \Pi_{i-1}, S = 1, X_A \in A_j, X_G = g, X_R = r, \boldsymbol{E}\right)$$
$$= \begin{cases} \min\{\phi_{ijgr}, (1 - p_i^*(\boldsymbol{Y}, \boldsymbol{E}))\}, & \bar{s} = 0 \\ 1 - p_i^*(\boldsymbol{Y}, \boldsymbol{E}) - \min\{\phi_{ijgr}, (1 - p_i^*(\boldsymbol{Y}, \boldsymbol{E}))\}, & \bar{s} = 1 \\ p_i^*(\boldsymbol{Y}, \boldsymbol{E}), & \bar{s} = 2 \end{cases}.$$

In the above, $p_i^*(\boldsymbol{Y}, \boldsymbol{E})$ denotes the probability that an individual $\boldsymbol{Y}$ becomes infected in the time interval Δ_i given the environmental conditions $\boldsymbol{E}$. As environmental conditions may vary greatly within the interval Δ_i, we define $p_{i,d}^*(\boldsymbol{Y}, \boldsymbol{E})$ to be the probability that an individual $\boldsymbol{Y}$ becomes infected on the d^{th} day in the time interval Δ_i given the environmental conditions $\boldsymbol{E}$. Therefore, $p_i^*(\boldsymbol{Y}, \boldsymbol{E}) = \sum_d p_{i,d}^*(\boldsymbol{Y}, \boldsymbol{E})$. The daily probability is defined as,

$$\log\left(p_{i,d}^*(\boldsymbol{Y}, \boldsymbol{E})\right)$$
$$= \min\Big\{0, \log\left(\rho_{j,g,r} \div 365\right) + \log(1.015) \times (E_P - \mu_{EP}) + \log(1.012)$$
$$\times (E_T - \mu_{ET}) + \log(0.554) \times \left(\frac{E_{R2}}{E_{R7}} - \mu_{ER2/R7}\right) + \log(0.860) \times (E_{R7} - \mu_{ER7})$$
$$+ \log(RR_D) \times (X_D - \mu_{XD})\Big\},$$

where $\rho_{j,g,r}$ is the population-wide annual probability of infection for a susceptible individual in age stratum A_j, gender g and of race/ethnicity r. In the above, all of the environmental variables are for the d^{th} day in the time interval Δ_i, and the values of $\mu_{EP}, \mu_{ET}, \mu_{ER2/R7}$ and μ_{ER7} denote the daily average values of the corresponding variables based upon historical data. Also, RR_D denotes the relative risk of infection for the elevated risk dust exposure group versus the normal risk group, and μ_{XD} is the estimated proportion of the population in the elevated risk dust exposure group. The environmental coefficients in the probability formula for E_P, E_T, $\frac{E_{R2}}{E_{R7}}$, and E_{R7} are all from Park et al. (Park, Sigel et al. 2005). Schneider et al. found that the relative risk of infection for individuals that reported being directly in a dust cloud following an earthquake was 2.0 compared to individuals that did not report being in a dust cloud (Schneider, Hajjeh et al. 1997). Therefore, in our base case we set the value of $RR_D = 2^{5/7} = 1.64$ to reflect high exposure during a typical 5-day work week. We examine this value in a sensitivity analysis using relative risks of $RR_D = 1.5^{5/7} = 1.34$ and $RR_D = 2.5^{5/7} = 1.92$. Note that seasonal variations in infection rates are modeled through the environmental variables.

If $\overline{S} = 0$, then the individual has died, $S' = 0$ and $\overline{T}$ represents the time of death. We simulate $\overline{T}$ from a uniform distribution on Δ_i and update the remaining variables accordingly. Hence the following updates occur:

$$S'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 0\big) \sim \delta_0\,,$$

$$\overline{T}\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 0, S' = 0\big) \sim U(\tau_{i-1}, \tau_i)\,,$$

$$X'_A\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, X_A = x_A, S = 1, \overline{S} = 0, S' = 0, \overline{T} = \overline{t}\big) \sim \delta_{x_A + \overline{t} - \tau_{i-1}}\,,$$

$$X'_Z\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, X_Z = x_Z, S = 1, \overline{S} = 0, S' = 0, \overline{T} = \overline{t}\big) \sim \delta_{x_Z} \text{ for } Z \in \{G, R, D, C, S, V, L\}\,,$$

$$\boldsymbol{C}'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 0, S' = 0, \overline{T} = \overline{t}\big) \sim \delta_{\boldsymbol{0}}\,, \text{ and}$$

$$\boldsymbol{T}'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 0, S' = 0, \overline{T} = \overline{t}\big) \sim \delta_{\boldsymbol{0}}\,.$$

If $\overline{S} = 1$, then the individual remains susceptible, $S' = 1$ and $\overline{T}$ will simply not be defined. The following updates occur:

$$S'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 1\big) \sim \delta_1\,,$$

$$X'_A\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, X_A = x_A, S = 1, \overline{S} = 1, S' = 1\big) \sim \delta_{x_A + L}\,,$$

$$X'_Z\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, X_Z = x_Z, S = 1, \overline{S} = 1, S' = 1\big) \sim \delta_{x_Z} \text{ for } Z \in \{G, R, D, C, S, V, L\}\,,$$

$$\boldsymbol{C}'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 1, S' = 1\big) \sim \delta_{\boldsymbol{0}}\,, \text{ and}$$

$$\boldsymbol{T}'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 1, S' = 1\big) \sim \delta_{\boldsymbol{0}}\,.$$

If $\overline{S} = 2$, then the individual has become infected and $\overline{T}$ represents the time of infection. Here we assume that the probability of infection is approximately constant throughout the short interval Δ_i and simulate $\overline{T}$ uniformly on Δ_i. That is,

$$\overline{T}\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2\big) \sim\ U(\tau_{i-1}, \tau_i)\,.$$

We describe in Section 3.6.3.4 the simulation of the vectors $\boldsymbol{C}'$ and $\boldsymbol{T}'$ of clinical and temporal variables associated with coccidioidomycosis. Given $\boldsymbol{T}'$, the final state S' can be determined from T'_M because S' is either 2 or 0, and we already account for the likelihood of death by non-coccidioidal causes in the probability that $\overline{S} = 0$. Hence, the only way that $S = 1$, $\overline{S} = 2$ and $S' = 0$ is if the infected individual dies from coccidioidomycosis between the time of infection $\overline{T}$ and τ_i. That is,

$$S'\big|\big(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \boldsymbol{C}' = \boldsymbol{c}', \boldsymbol{T}' = \boldsymbol{t}', \overline{T} = \overline{t}\big) \sim\ \delta_{s'}\,.$$

where $s' = 2 \times \left(1 - I_{(\overline{t}, \tau_i]}(t'_M)\right)$. We then simulate the vector $\boldsymbol{X}'$ as follows:

$$X_A' \big| \left(\boldsymbol{Y} \in \Pi_{i-1}, X_A = x_A, S = 1, \overline{S} = 2, S' = 2, \overline{T} = \overline{t} \right) \sim \delta_{x_A + L} \, ,$$

$$X_A' \big| \left(\boldsymbol{Y} \in \Pi_{i-1}, X_A = x_A, S = 1, \overline{S} = 2, S' = 0, \overline{T} = \overline{t} \right) \sim \delta_{x_A + \overline{t} - \tau_{i-1}} \, ,$$

$$X_Z' \big| \left(\boldsymbol{Y} \in \Pi_{i-1}, X_Z = x_Z, S = 1, \overline{S} = 2, S' = 0, \overline{T} = \overline{t} \right) \sim \delta_{x_Z} \text{ for } Z \in \{G, R, D, C, S, V, L\} \text{, and}$$

$$X_Z' \big| \left(\boldsymbol{Y} \in \Pi_{i-1}, X_Z = x_Z, S = 1, \overline{S} = 2, S' = 2, \overline{T} = \overline{t} \right) \sim \delta_{x_Z} \text{ for } Z \in \{G, R, D, C, S, V, L\} \, .$$

Finally, we note that the variables X_D, X_C, X_S X_V, and X_L remain unchanged for all transitions from the susceptible state. As the model is intended to be run over a relatively short time span, for example 1 year, the occupation, co-morbidity and smoking variables are expected to be relatively stable. The model could be enhanced to reflect the dynamics of these variables. The variable X_V also remains unchanged as in the current version, no individuals are vaccinated.

3.6.3.3 Transition from State 2 (Infected)

From the infected state, an individual can either remain in the infected state or die. That death could be either related to coccidioidomycosis or other causes. Recall from Section 3.6.2 that T_{MAX} is the maximum of the vector $\boldsymbol{T}$ excluding the time of report T_R, that is, $T_{MAX} = \max\{T_I, T_{S0}, T_{S1}, T_D, T_M\} = \max\{T_I, T_{S1}, T_D, T_M\}$ and note that $T_M > 0$ implies that $T_M = T_{MAX}$. It follows that for an individual in the infected state that either (a) $T_M \in \Delta_i$, (b) $T_{MAX} > \tau_i$ or (c) $T_{MAX} \leq \tau_i$ and $T_M = 0$. If $T_M \in \Delta_i$ then the individual dies due to coccidioidomycosis and transitions to the death state. That is,

$$P(S' = 0 | \boldsymbol{Y} \in \Pi_{i-1}, S = 2, T_M \in \Delta_i) = 1 \, .$$

If $T_{MAX} > \tau_i$ then the individual is not eligible for death and therefore must remain in the infected state. That is,

$$P(S' = 2 | \boldsymbol{Y} \in \Pi_{i-1}, S = 2, T_{MAX} > \tau_i) = 1 \, .$$

If $T_{MAX} \leq \tau_i$ and $T_M = 0$ then the individual is eligible for death due to other causes for either all or part of Δ_i. If $T_{MAX} \leq \tau_{i-1}$ then the individual is eligible for death due to other causes for the entire interval and the probability of death is,

$$P\left(S' = 0 \middle| \boldsymbol{Y} \in \Pi_{i-1}, S = 2, X_A \in A_j, X_G = g, X_R = r, T_{MAX} \leq \tau_{i-1}, T_M = 0\right) = \phi_{ijgr} \, .$$

If the individual is only eligible for a portion of the interval, we set the probability of death to be proportional to ϕ_{ijgr}. Hence, if $T_{MAX} \leq \tau_i$ and $T_M = 0$ then the probability of death is,

$$P\left(S' = 0 \middle| \boldsymbol{Y} \in \Pi_{i-1}, S = 2, X_A \in A_j, X_G = g, X_R = r, T_{MAX} \leq \tau_i, T_M = 0\right)$$
$$= \phi_{ijgr} \times \frac{\tau_i - \max\{T_{MAX}, \tau_{i-1}\}}{L},$$

and the probability of remaining in the infected state is,

$$P\left(S' = 2 \middle| \boldsymbol{Y} \in \Pi_{i-1}, S = 2, X_A \in A_j, X_G = g, X_R = r, T_{MAX} \leq \tau_i, T_M = 0\right)$$
$$= 1 - \left(\phi_{ijgr} \times \frac{\tau_i - \max\{T_{MAX}, \tau_{i-1}\}}{L}\right).$$

As above, the variables X_D, X_C, X_S, X_V and X_L remain unchanged for all transitions from the infected state to any other state.

3.6.3.4 Coccidioidomycosis Clinical and Temporal Variables

In this section we describe the simulation of the vectors $\boldsymbol{C}'$ and $\boldsymbol{T}'$ of clinical and temporal variables associated with coccidioidomycosis. Specifically, we describe the joint distribution:

$$(\boldsymbol{C}', \boldsymbol{T}') \left| \left(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}\right)\right. .$$

Approximately 40% of infected individuals are symptomatic (Pappagianis, Sun et al. 1993; Louie, Ng et al. 1999; Chiller, Galgiani et al. 2003). This percentage would be reduced if an effective, distributable vaccine against coccidioidomycosis were developed. We therefore set,

$$P\left(C_S' = 1 | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}\right) = p_{S|I} \times (1 - \theta_{Vac} \times X_V) .$$

In the above, our base case value of $p_{S|I} = 0.40$ to reflect the 40% symptomatic rate. We explore the effect of this parameter in a sensitivity analysis using values of $p_{S|I} = 0.30$ and $p_{S|I} = 050$. Our base case value of vaccine effectiveness was set to $\theta_{Vac} = 0.75$ to reflect a vaccine that reduces the rate of symptoms by 75% in vaccinated individuals. In sensitivity analyses we explore changing the vaccine effectiveness to $\theta_{Vac} = 0.60$ and $\theta_{Vac} = 0.90$. Note that in this model, individuals are either vaccinated and immediately protected (pending effectiveness of vaccine) or they are not vaccinated.

In this work, in regards to vaccination, we are focused on the long term public health impact of vaccine availability rather than the short term impact during vaccine introduction. The probability that a symptomatic case has severe pulmonary symptoms (cavities, nodules, chronic pulmonary coccidioidomycosis) is a function of the individual's risk factors. We set,

$$\log\left(p_{Pulm|IS}\right) = \log\left(P\left(C'_P = 1|\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1\right)\right)$$
$$= min\{0, \log\left(p^0_{Pulm|IS}\right) + \log(1.03) \times (X_A - 37.4) + \log(1.19)$$
$$\times (X_G - 1 - 0.545) + \log(1.53) \times (I(X_R = 3) - 0.079) + \log(2.81)$$
$$\times (I(X_C \in \{1,3\}) - 0.111) + \log(5.26) \times (I(X_C \in \{2,3\}) - 0.016) + \log(2.34)$$
$$\times (X_S - 0.147)\}.$$

The above equation uses relative risks derived from Rosenstein et al. that reflects increased risk for severe pulmonary symptoms associated with increased age, males, African Americans, diabetes, immunocompromised, and smoking (Rosenstein, Emery et al. 2001). The minimum function ensures that the resulting probabilities never exceed one.

The probability of extrapulmonary dissemination is a function of the individual's risk factors. One risk factor for dissemination is pregnancy. We intentionally omit pregnancy from the vector $\boldsymbol{X}$ of demographic and risk factors because of the dynamic nature of the variable. Instead, we determine the pregnancy state for an individual, denoted by P^*, upon infection. One implication of this approach is that there is no direct relationship between births and pregnancies aside from the rates at which they occur. We simulate P^* as follows:

$$P\left(P^* = 1 \middle| \boldsymbol{Y} \in \Pi_{i-1}, X_A \in A_j, X_G = g, X_R = r, S = 1, \overline{S} = 2\right) = \begin{cases} 0 & \text{if } g = 2 \\ 0.75 \times \beta_{j,r} & \text{if } g = 1 \end{cases}.$$

Note that in the above, we could replace 0.75 with a summation over α; for simplicity we approximate the summation with 0.75.

Given P^* we simulate the probability of dissemination according to the following formula, which utilizes relative risks and demographic characteristics from Rosenstein et al. (Rosenstein, Emery et al. 2001):

$$\log\left(p_{Diss|IS}\right) = \log\left(P\left(C'_D = 1|\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1\right)\right)$$
$$= min\{0, \log\left(p^0_{Diss|IS}\right) + \log(3.58) \times (I(X_R = 3) - 0.079) + \log(2.34)$$
$$\times (I(X_C \in \{1,3\}) - 0.111) + \log(1.64) \times (X_S - 0.147) + \log(8.18)$$
$$\times (P^* - 0.021)\}.$$

Again, the minimum function ensures that the probabilities do not exceed one. The probability of mortality is a function of the individual's risk factors. We set,

$$P\left(C'_M = 1|\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1, C'_D = c'_D\right) = \frac{p_{Death|IS}}{p_{Diss|IS}} \times c'_D.$$

The above reflects the fact that virtually all deaths are from disseminated coccidioidomycosis. Finally we consider the probability that a case is reported. We presume that the probability of report depends on the severity of symptoms and disease outcome. These probabilities are set as follows for the base case:

$$p_{R|ISG1} = \left(C'_R = 1 | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1, C'_P = 0, C'_D = 0, C'_M = 0\right) = 0.25 ,$$

$$p_{R|ISG2} = \left(C'_R = 1 | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1, C'_P = 1, C'_D = 0, C'_M = 0\right) = 0.80 ,$$

$$p_{R|ISG3} = \left(C'_R = 1 | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1, C'_D = 1, C'_M = 0\right) = 0.95 , \text{ and}$$

$$p_{R|ISG4} = \left(C'_R = 1 | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}, C'_S = 1, C'_M = 1\right) = 0.99 .$$

We next consider the distribution of $\boldsymbol{T}'$. We begin with the time of infection:

$$T'_I | \left(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, \overline{T} = \overline{t}\right) \sim \delta_{\overline{t}} .$$

Then, the time from infection to symptom onset is typically 1-3 weeks (Galgiani, Ampel et al. 2005). We set the time from infection to symptom onset to be uniform from 0.5 to 3.5 weeks. That is,

$$T'_{S0} | \left(\boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, C'_S = 1, T'_I = t'_I\right) \sim U(t'_I + 3.5/365, t'_I + 24.5/365) .$$

Next the duration of pulmonary symptoms depends on whether the pulmonary infection is mild or severe. For mild infections we base the duration of symptoms on Crum et al. who investigated an outbreak among Navy SEALS(Crum, Lamb et al. 2002). In this relatively healthy population, the median duration of symptoms was 19 days, with a range of 1 to 63 days. In accordance with these data, we set the conditional density function for $T'_{S1}$to be,

$$f_{T'_{S1}}\left(t'_{S1} | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, T'_{S0} = t'_{S0}, C'_S = 1, C'_P = 0\right)$$

$$= \begin{cases} 365/36 & t'_{S0} + \frac{1}{365} < t'_{S1} \leq t'_{S0} + \frac{19}{365} \\ 365/88 & t'_{S0} + \frac{19}{365} < t'_{S1} \leq t'_{S0} + \frac{63}{365} \\ 0 & t'_{S1} \leq t'_{S0} + \frac{1}{365} \text{ or } t'_{S1} > \frac{63}{365} \end{cases} .$$

For severe infections we base the duration of symptoms on Tsang et al. who surveyed reported cases in Arizona (Tsang, Anderson et al. 2010). They reported a median duration of symptoms of 120 days and an interquartile range of 49 to 198 days. Based on these data we set the conditional density function for symptom end as follows:

$$f_{T'_{S_1}}\left(t'_{S1}|\boldsymbol{Y} \in \Pi_{i-1}, S=1, \overline{S}=2, T'_{S0}=t'_{S0}, C'_S=1, C'_P=1\right)$$

$$= \begin{cases} 365/120 & t'_{S0}+\frac{19}{365} < t'_{S1} \le t'_{S0}+\frac{49}{365} \\ 365/284 & t'_{S0}+\frac{49}{365} < t'_{S1} \le t'_{S0}+\frac{120}{365} \\ 365/312 & t'_{S0}+\frac{120}{365} < t'_{S1} \le t'_{S0}+\frac{198}{365} \\ 365/896 & t'_{S0}+\frac{198}{365} < t'_{S1} \le t'_{S0}+\frac{422}{365} \\ 0 & t'_{S1} \le t'_{S0}+\frac{19}{365} \text{ or } t'_{S1} > \frac{422}{365} \end{cases}.$$

In the above, we use 19 days as the minimum duration of symptoms based upon the median duration for mild symptoms. The maximum duration of 422 days was determined using the common definition of outlier, the third quartile plus 1.5 times the interquartile range.

Adams et al. reported statistics on the duration of time from initial onset of symptoms to dissemination (Adam, Elliott et al. 2009). They reported that 61% of disseminated cases occur within 2 months, 21% between 3 and 9 months after onset, 4% between 1 and 2 years, and 13% occur more than 3 years after onset. In addition, they reported that the longest duration between symptom onset and dissemination was 10 years. Based on their statistics, we simulate the time of diagnosis of disseminated coccidioidomycosis as follows:

$$f_{T'_D}\left(t'_D|\boldsymbol{Y} \in \Pi_{i-1}, S=1, \overline{S}=2, T'_{S0}=t'_{S0}, C'_S=1, C'_D=1\right)$$

$$= \begin{cases} 0 & t'_D \le t'_{S0} \\ 0.61 \times \frac{1}{0.21} & t'_{S0} < t'_D \le t'_{S0}+0.21 \\ 0.21 \times \frac{1}{.54} & t'_{S0}+0.21 < t'_D \le t'_{S0}+0.75 \\ 0.05 \times \frac{1}{1.25} & t'_{S0}+0.75 < t'_D \le t'_{S0}+2 \\ 0.13 \times \left(0.375 \times e^{-0.375\left(t'_D-t'_{S0}-2\right)}\right) & t'_D > t'_{S0}+2 \end{cases}.$$

The above distribution is a mixture distribution that reflects that the duration between onset of symptoms and dissemination is distributed with probabilities of 0.61, 0.21 and 0.05 of being uniform on the intervals of 0-2.5 months, 2.5-9 months, and 9 months to 2 years. With probability 0.13 the duration has a shifted exponential distribution with rate parameter 0.375. This rate parameter was chosen such that, given that the duration is greater than 2 years, the conditional probability that the duration is greater than 10 years is 0.05. With this value, the conditional average duration is 4 years, 8 months.

For individuals that die from the disease, the duration of time from dissemination to death was set to be exponentially distributed with rate parameter 6, which implies that the average duration is 2 months. That is,

$$f_{T'_M}\left(t'_M | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, T'_D = t'_D, C'_S = 1, C'_D = 1, C'_M = 1\right) = \begin{cases} 6e^{-6(t'_M - t'_D)} & t'_M > t'_D \\ 0 & t'_M \leq t'_D \end{cases}.$$

Finally, for cases that are reported to the health department, Tsang et al. found that the median number of days from symptom onset to diagnosis was 55 (Tsang, Anderson et al. 2010). Based on those data, we assume that the duration of time from symptom onset to case report follows an exponential distribution with a median of 55 days. This distribution has a mean of 79 days and has the following density function:

$$f_{T'_R}\left(t'_R | \boldsymbol{Y} \in \Pi_{i-1}, S = 1, \overline{S} = 2, T'_{S0} = t'_{S0}, C'_S = 1, C'_R = 1\right) = \begin{cases} \frac{365}{79} e^{-\frac{365}{79}(t'_R - t'_{S0})} & t'_R > t'_{S0} \\ 0 & t'_R \leq t'_{S0} \end{cases}.$$

3.7 POPULATION INITIALIZATION

In the previous section we describe how the population is updated over time in our model. In this section we describe how the population was initialized, that is, how we construct Π_0. We consider as given that the initial population sizes for each stratum, $N_{0,j,g,r}$, are known. We describe in Section 3.8.2 how these population sizes may be set for a particular run of the model. We iterate over the strata and for the stratum consisting of age stratum A_j, gender g, and race/ethnicity r, we create a subset $\Pi_{0,j,g,r}$ of Π_0 corresponding to this stratum, that is,

$$\Pi_{0,j,g,r} = \left\{\boldsymbol{Y} = (\boldsymbol{X}, S, \boldsymbol{C}, \boldsymbol{T}) \in \Pi_0 \middle| X_A \in A_j, X_G = g, X_R = r\right\}.$$

We populate $\Pi_{0,j,g,r}$ with $N_{0,j,g,r}$ members and set the values of $\boldsymbol{X}$, S, $\boldsymbol{C}$ and $\boldsymbol{T}$ for each individual as described below.

3.7.1 Demographic and Risk Factors ($\boldsymbol{X}$)

The age variable X_A was simulated according to a uniform distribution on A_j. The gender variable X_G and race/ethnicity variable X_R were set to g and r, respectively. The variables X_D,

X_C, X_S, X_V and X_L were simulated according to their respective distributions, P_D, P_C, P_S and P_V, which may be set using region-specific data as described in Section 3.8. Hence,

$$X_A \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim U(a_{j-1}, a_j) ,$$

$$X_G \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim \delta_g ,$$

$$X_R \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim \delta_r ,$$

$$X_D \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_D ,$$

$$X_C \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_C ,$$

$$X_S \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_S ,$$

$$X_V \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_V \text{ , and}$$

$$X_L \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_L .$$

3.7.2 State Variable ($\boldsymbol{S}$)

The state variable S was simulated according to the distribution P_I which may be set using region-specific data as described in Section 3.8. Therefore,

$$S \big| \boldsymbol{Y} \in \Pi_{0,j,g,r} \sim P_I .$$

3.7.3 Clinical Variables ($\boldsymbol{C}$)

In the present version of the model, we set the vector of clinical variables to the zero vector. Following a burn-in period, the distribution of clinical variables should reflect natural variations in the clinical characteristics of infected individuals. We recommend at least five months of burn-in in order to prevent under-forecasting of reported cases during the run period of interest.

3.7.4 Temporal Variables (*T*)

As with the clinical variables, the temporal variables are set to the zero vector and then updated during a burn-in period.

3.8 MODEL INSTANTIATION

Model instantiation refers to the process of setting all required inputs necessary to run the model. The most substantial part of instantiation is the setting of various region-specific parameters in the model, including population estimates, coccidioidomycosis incidence rates, prevalence rates for risk factors, pregnancy rates $(\beta_{j,r} \text{ and } \alpha_i)$ and mortality rates $(\phi_{i,j,g,r})$. In many circumstances these quantities can be estimated using historical data from the region under study. In other cases, data from other geographical regions may be used in its place. Finally, if no data are available, these quantities could be hand-set and then considered for sensitivity analysis.

In this section we describe how we set these inputs to run the model for Maricopa County, Arizona in 2010 using only incidence data available prior to 2010. We focus on Maricopa County because roughly half of all United States cases occur in Maricopa County. It is important to note that there is nothing critical about the specific locale and time in the model, aside from the use of population, incidence, natality, and mortality data. The model could be readily instantiated for any other location and time if appropriate inputs are provided. Retaining this flexibility in the model is critical for the proposed work as one aim of the dissertation is to construct a tool that public health can use to forecast incidence.

Table 1 provides a summary of the data sources that we use to instantiate and validate our stochastic agent-based model. Data sources include Centers for Disease Control and Prevention, United States Census Bureau, United States Environmental Protection Agency, National Climatic Data Center, and Maricopa County Department of Public Health, and Arizona Department of Health Services. For each type of data, the spatial, temporal and strata resolution of the data are indicated and the source of the data is listed. Additional details on the data and their source are provided in the corresponding sections where the data are described.

Table 1. Data sources and resolution.

Data	Resolution	Source
Population Counts	County-level counts by age, gender, and race/ethnicity	United States Census Bureau via Maricopa County Department of Public Health
Natality Rates	Annual county-level counts by age and race/ethnicity of birth mother State-level counts by day of birth	Maricopa County Department of Public Health Arizona Department of Health Services
Mortality Rates	Annual county-level counts by age and race/ethnicity, and by age and gender	Maricopa County Department of Public Health
Coccidioidomycosis Incidence	Annual county-level counts by age, by gender, and by race/ethnicity	Maricopa County Department of Public Health
PM10	Daily concentration at thirteen weather stations in county	United States Environmental Protection Agency
Temperature, Rainfall	Daily data in county	National Climatic Data Center (National Oceanic and Atmospheric Administration)
Diabetes Prevalence	National prevalence rates by age, gender, and race/ethnicity	Centers for Disease Control and Prevention
HIV Prevalence	National prevalence counts by age, by gender, and by race/ethnicity National population counts by age, gender and race	Centers for Disease Control and Prevention United States Census Bureau
Smoking Prevalence	National prevalence rates by age and gender, and by gender and race/ethnicity	Centers for Disease Control and Prevention
Occupation	Annual national rates and counts by age, gender and race/ethnicity	United States Bureau of Labor Statistics
Flu Vaccination	Annual national rates by age and gender, and by age and race/ethnicity	Centers for Disease Control and Prevention

To forecast incidence, we provide the model with environmental data for 2010. The rationale is that we want to use the most applicable data possible, as it would not make sense to forecast 2010 incidence using 2009 PM10 in a case where we had 2010 PM10 data. As such, we would like to be able to examine the implications of different environmental conditions, for example a dust storm, and need to provide the model with actual or hypothetical environmental data for the period of analysis. Here we use average PM10 data from 13 Air Quality Subsystem (AQS) stations in Maricopa County that each reported daily mean PM10 for 2010 (United States Environmental Protection Agency 2012). Those stations were AQS 04-013-0019, -2001, -3002, -3010, -4003, -4004, -4006, -4009, -4010, -4011, -4016, -9812, and -9997. Table 2 contains the mean, standard deviation, minimum and maximum across the 13 stations by month for 2009 and 2010. The daily PM10 data for the 13 stations are displayed graphically in Figure 1. Several AQS stations in Maricopa County were excluded because they only provided weekly data. We use weather data from both the weather station 722749 at the Chandler Municipal Airport and the Chandler Airport Precipitation Gage (Device ID 6585) (National Climatic Data Center 2012).

Table 2. Summary statistics for PM10 across 13 stations by month for 2009 and 2010.

	2009				2010			
Month	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
January	31.6	15.7	5.0	83.0	26.5	17.4	4.0	99.0
February	28.9	14.4	3.0	88.0	19.5	9.6	3.0	54.0
March	39.0	21.0	11.0	209.0	21.6	10.1	5.0	75.0
April	35.2	20.2	6.0	213.0	29.3	11.6	8.0	98.0
May	33.9	12.8	10.0	78.0	32.0	11.8	11.0	107.0
June	31.7	9.9	13.0	93.0	32.0	9.7	16.0	80.0
July	48.4	48.5	13.0	439.0	28.6	10.4	9.0	83.0
August	39.7	23.6	8.0	142.0	24.3	10.0	8.0	71.0
September	33.5	18.7	1.0	174.0	33.2	10.1	12.0	65.0
October	43.3	22.0	15.0	220.0	30.5	20.0	8.0	158.0
November	51.5	20.0	14.0	111.0	34.1	13.9	6.0	87.0
December	32.7	14.4	6.0	132.0	35.3	21.5	5.0	113.0

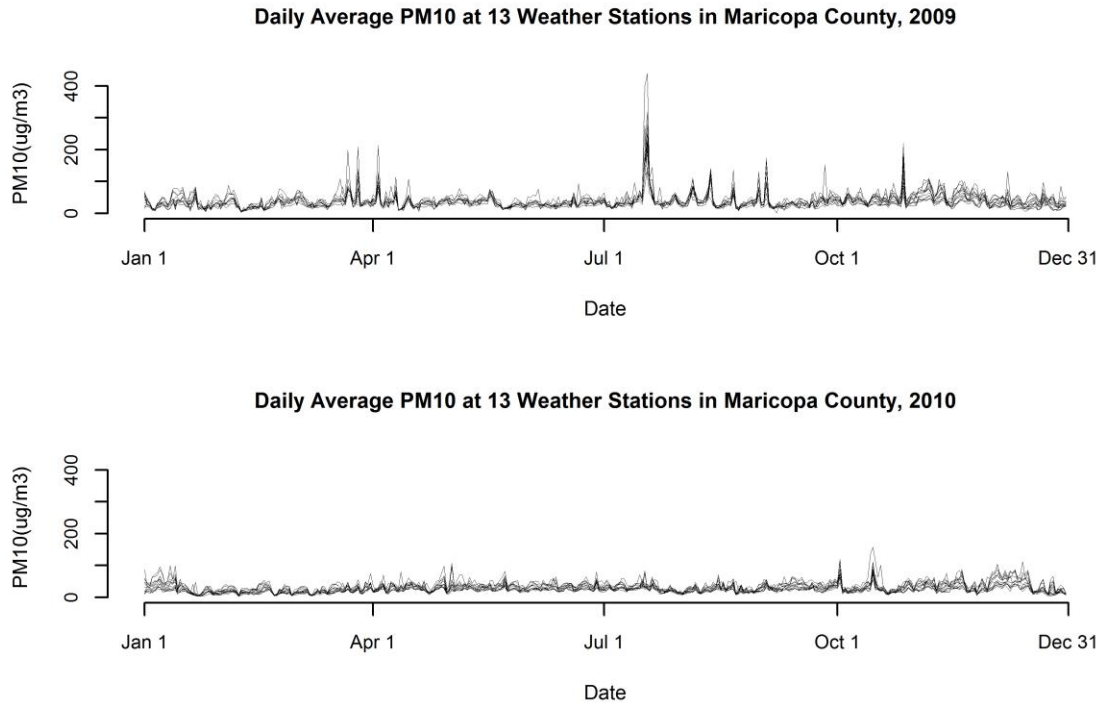

Figure 1. PM10 data for 13 weather stations in Maricopa County, Arizona. Top: 2009. Bottom: 2010.

In Figures 2 and 3 we show PM10, temperature, and precipitation data for Maricopa County in 2009 and 2010, respectively. The effect of dust storm in July of 2009 on PM10 can be readily observed in Figure 2. In contrast, PM10 was relatively low in 2010. The seasonal variation in temperature and infrequent precipitation in Maricopa County is evident in both Figures 2 and 3. The environmental time series that are utilized by the model (see Section 3.6.3.2) are PM10, average temperature during the previous three months, the ratio of the cumulative rainfall during the previous two months to the cumulative rainfall during the previous seven months, and the cumulative rainfall during the previous seven months. The latter three time series are displayed graphically in Figure 4 for 2010. The seasonal variation in temperature and rainfall in Maricopa County are exhibited in the figure.

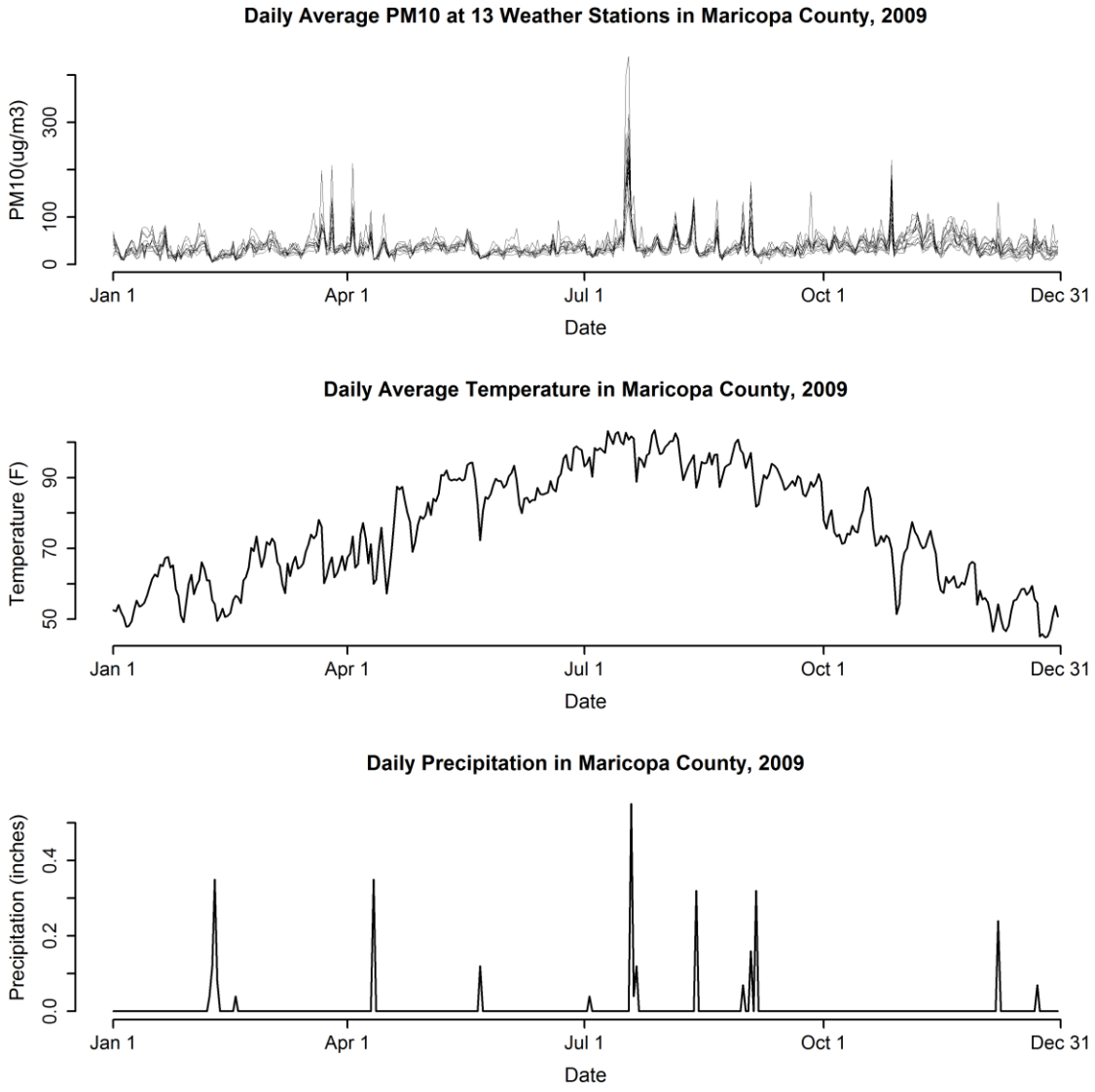

Figure 2. Daily PM10, temperature, and precipitation for Maricopa County, 2009.

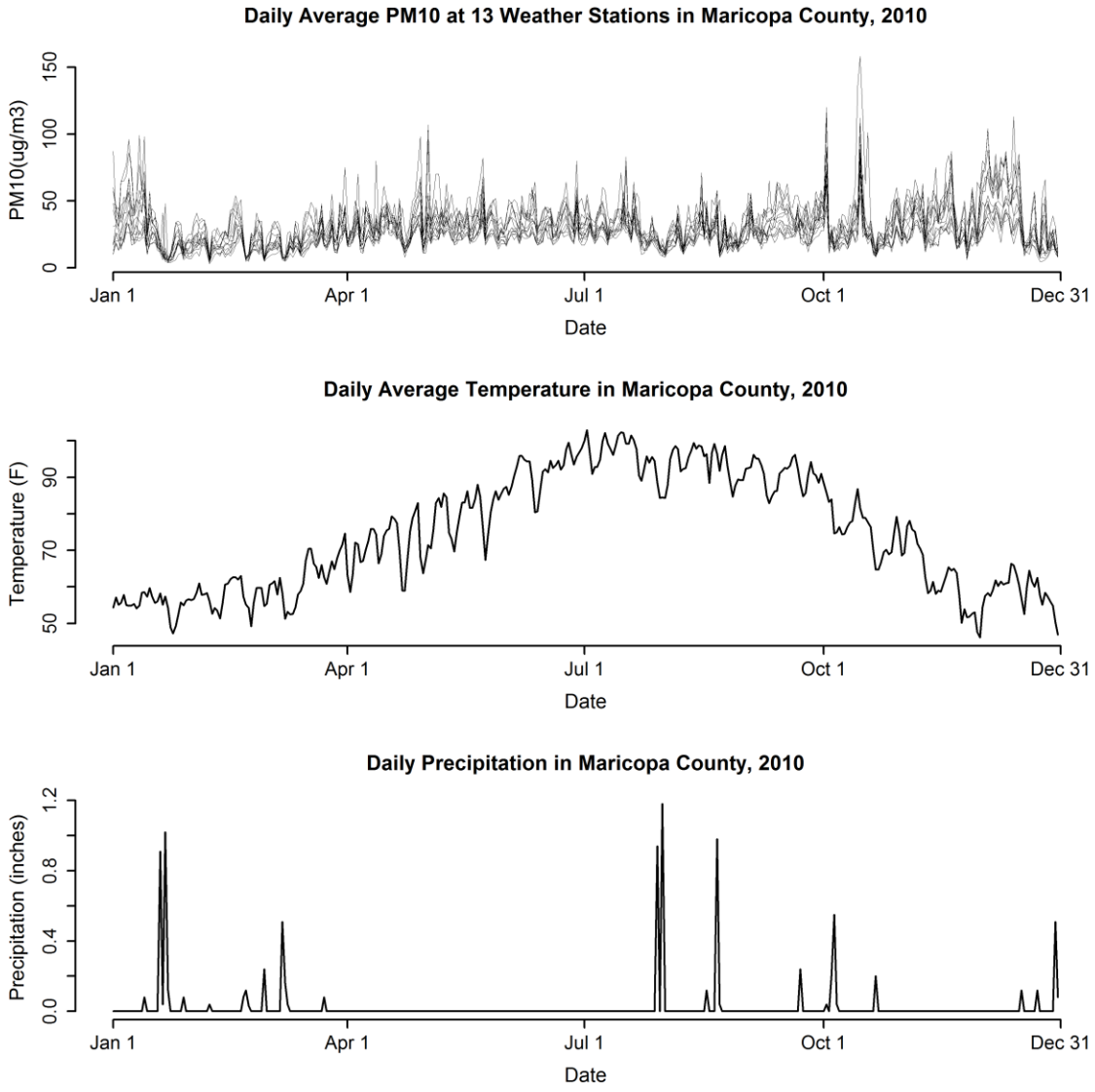

Figure 3. Daily PM10, temperature, and precipitation for Maricopa County, 2010.

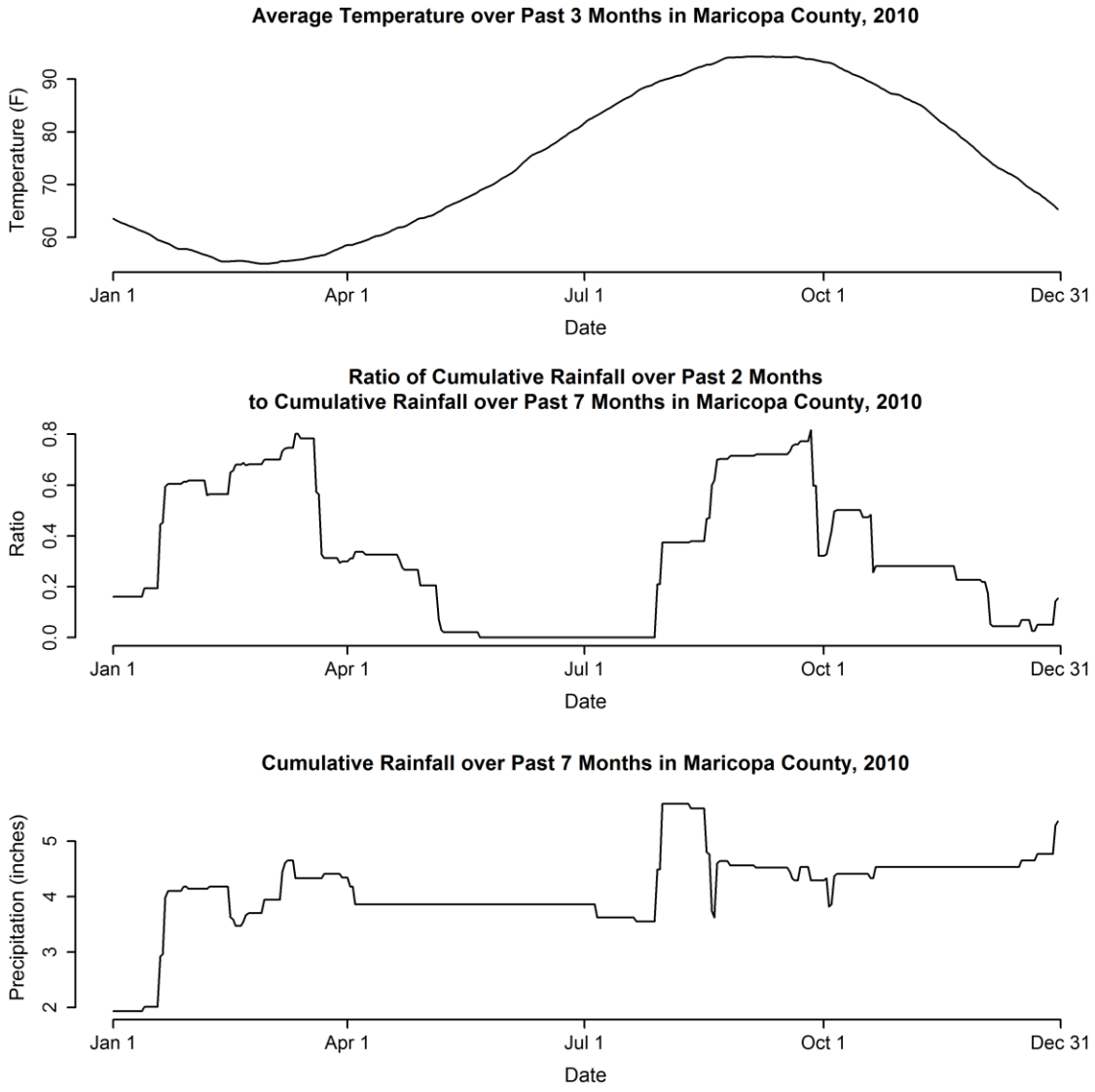

Figure 4. Temperature and rainfall time series for Maricopa County, 2010.

3.8.1 Age Strata

We adopt the age strata that were used for population statistics in the 2009 Maricopa County Health Status Report (MCHSR) (Maricopa County Department of Public Health 2011). They are: 0-4, 5-9, 10-14, 15-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, and 85+. Hence $m = 12$ with $a_0 = 0$, $a_1 = 5$, $a_2 = 10$, $a_3 = 15$, $a_4 = 20$, $a_5 = 25$, $a_6 = 35$, $a_7 = 45$, $a_8 = 55$, $a_9 = 65$, $a_{10} = 75$, $a_{11} = 85$, and $a_{12} = \infty$.

3.8.2 Initial Population Sizes

Initial population sizes by age, gender, and race/ethnicity strata ($N_{0,j,g,r}$) were obtained from the 2009 United States Census Data for Maricopa County (Maricopa County Department of Public Health 2011). See Tables 2, 3, and 4. Note that these population estimates are extrapolations from the 2000 United States Census. While we could obtain better population estimates by interpolating the 2000 and 2010 United States Census estimates, we refrain from doing so in order to run the model in a completely predictive fashion.

3.8.3 Natality Rates

The 2009 MCHSR provides natality statistics by age and race/ethnicity of birth mother (Maricopa County Department of Public Health 2011). The age strata used were <15, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, and 45+. To construct estimates of $\beta_{j,r}$ for the age strata defined in Section 3.8.1, we assume that all births to mothers under the age of 15 were to mothers in the 10-14 age stratum, aggregate the 25-29 and 30-34 age strata, aggregate the 35-39 and 40-44 age strata, and assume that all births to mothers 45 and over were to mothers in the 45-54 age stratum. We then use these derived natality statistics and the initial population sizes $N_{0,j,1,r}$ to compute natality rates as point estimates for $\beta_{j,r}$. We average daily birth data from the Arizona Health Status and Vital Statistics Table 1B-18 for 2005-2009 (Arizona Department of Health Services 2012) to estimate α_i, which reflects seasonal variations in natality rates.

3.8.4 Mortality Rates

The 2009 MCHSR (Maricopa County Department of Public Health 2011) provides mortality statistics by (1) age and gender, and (2) age and race/ethnicity. The age and race/ethnicity strata used for both breakdowns are the same as our age and race/ethnicity strata. We use iterative proportional fitting to estimate the number of deaths by age, gender and race/ethnicity (Fienberg 1970) and then use $N_{0,j,g,r}$ to compute yearly mortality rates. We assume that mortality rates are constant across time and thus multiply these rates by L to construct point estimates for $\phi_{i,j,g,r}$.

3.8.5 Infection Rates

The 2009 MCHSR (Maricopa County Department of Public Health 2011) only provides:

a) Number of reported coccidioidomycosis cases broken down by age strata, using age strata of <1, 1-4, 5-9, 10-14, 15-19, 20-24,25-34, 35-44, 45-54,55-64, 65-84, and 85+.
b) Number of reported coccidioidomycosis cases broken down by race/ethnicity.
c) Number of reported coccidioidomycosis cases broken down by gender.

We estimate the number of reported infections in each of our age, gender and race/ethnicity strata for Section 3.5 as follows. We derive the marginal number of reported infections in each age stratum by using the age strata breakdown in 2009 MCHSR, aggregating the <1 and 1-4 strata, and distributing the 65-84 cases into the 65-74 and 75-84 age strata in proportion to the 2009 United States Census Population estimates (Maricopa County Department of Public Health 2011). Reported cases of unknown age were assigned proportionally according the number of infections in each age stratum. We derive the marginal number of reported infections for each gender by using the gender breakdown in 2009 MCHSR (Maricopa County Department of Public Health 2011) and distribute reported cases of unknown gender in proportion to the number of reported infections for each gender. The large number of reported infections with unknown or other race was treated by using Bayes Theorem to impute the race stratum for individuals with unknown race. Computationally, we adjust the number of reported infections in

the other race stratum to be proportional to the number of individuals in the other race stratum. The excess reported infections were then distributed across the race strata, excluding the other stratum, in proportion to the infection rates in those strata. The joint number of reported infections by age, gender and race/ethnicity were estimated by using iterative proportional fitting (Fienberg 1970). Finally, the number of reported infections in each stratum was divided by the population size for that stratum to derive rates of reported infections. These rates provide the base estimates for p_{ISR}. However, analysis of the previous five years of reported infections counts shows that reported infection counts are over-dispersed (variance greater than the mean) by a factor of 80. Some but not all of this over-dispersion can be explained by variation in weather and environmental data. To conservatively account for the remaining over-dispersion in infection counts these base rates are multiplied by a gamma random variable with shape and scale parameters such that the reported infection counts will be over-dispersed by a factor of 80.

3.8.6 Distributions for Risk Factors and Prior Infection

In this section we describe how we set P_D, P_C, P_S, P_V, P_L and P_I, which are the distributions for the variables X_D, X_C, X_S, X_V, X_L and S in Π_0. The variable X_D is an indicator for increased risk of dust exposure, primarily through occupation. We use Bureau of Labor Statistics (BLS) for 2009 to obtain counts of workers in construction and extraction occupations by gender and race (United States Bureau of Labor Statistics 2009). We also use BLS data for 2009 to extract information on the distribution of age across the construction occupation by gender and race (United States Bureau of Labor Statistics 2009). We use these data to estimate the probability of high dust exposure for each gender g and race r as follows. Let W denote the event that an individual is of working age, and $\overline{W}$ denote the event that the individual is not of working age. Then, assuming that only workers fall into the high dust exposure group we have,

$$P(High\ Dust|g,r) = P(High\ Dust|g,r,W) \times P(W|g,r)\ .$$

The first factor, $P(High\ Dust|g,r,W)$ is from BLS 2009 Table 10. The second factor is from Maricopa County population data in 2009 MCHSR (Maricopa County Department of Public Health 2011). We can then compute the probability of high dust exposure for each age, gender and race stratum as follows:

$$P(High\ Dust|j,g,r) = \frac{P(A_j|g,r,High\ Dust) \times P(High\ Dust|g,r)}{P(A_j|g,r)}.$$

In the above, $P(A_j|g,r,High\ Dust)$ is estimated from BLS 2009 Table 14 (United States Bureau of Labor Statistics 2009) and $P(A_j|g,r)$ is estimated from Maricopa County population data in 2009 MCHSR (Maricopa County Department of Public Health 2011).

The variable X_C holds the co-morbidity group, which is defined in terms of diabetes and immunocompromised patients. Here we use national prevalence rates for diabetes and HIV from the CDC and assume that HIV and diabetes are independent. Diabetes prevalence rates were obtained the CDC (Centers for Disease Control and Prevention 2010). HIV infection rates were obtained from Table 5 in the HIV Surveillance Report by the CDC (Centers for Disease Control and Prevention 2010). The variable X_S is an indicator for smoking, specifically whether the individual has smoked over the last 6 months. National smoking rates from the CDC were used to estimate the probability of smoking within strata for adults (Centers for Disease Control and Prevention 2012) and children (Centers for Disease Control and Prevention 2012). The variable X_V is an indicator for vaccinated. For our base case we assume that no vaccine exists and hence set X_V to zero for all individuals. For scenarios with vaccination (child only, adult only, both child and adult), we utilize national influenza vaccination rates and scale them according to the presumed vaccination coverage (Centers for Disease Control and Prevention 2011).

The state variable S indicates whether an individual is dead $(S = 0)$, susceptible to infection $(S = 1)$, or infected $(S = 2)$. As we only simulate living individuals in Π_0, it suffices to indicate the probability that an individual was infected prior to τ_0, which equals $1 - p_{Susc}$. Dodge et al. provided percentages of positive skin tests across gender and three age strata based on a sample of individuals in Tucson, Arizona (Dodge, Lebowitz et al. 1985). Galgiani estimated that 75% of the population in Maricopa County is susceptible to coccidioidomycosis infection (Galgiani 2012). We therefore scale the percentages from Dodge proportionally such that 25% of the Maricopa County population had prior infection (Dodge, Lebowitz et al. 1985).

3.9 MODEL VALIDATION

For Aim 2 of this dissertation we examine the validity of the model. Specifically we assess both the face validity of the model (Objective 2a) and conduct an independent validation (Objective 2b). First, we assess the face validity by examining whether the outputs from the model matched the relative risks and other characteristics that were used to construct the model. Second, we use the model to forecast incidence data and compare the results to actual counts of reported cases. In this independent validation, the model was trained using data through 2009 and tested by forecasting for 2010. For both validation purposes we use the model to forecast incidence data for Maricopa County, Arizona in 2010 (Maricopa County Department of Public Health 2012) by simulating the complete population of approximately 4 million residents for the year. This simulation was repeated ten times, requiring over 250 hours of CPU time.

3.10 SENSITIVITY ANALYSIS

As part of Aim 2 of this dissertation, we examine the role of several parameters in one-way sensitivity analyses. The parameters examined are listed in Table 3 below. The table also provides the values of the parameters for the base case, as well as the alternative values used in the sensitivity analysis.

Table 3. Parameters examined in one-way sensitivity analyses.

Parameter	Base Case Value	Alternative Values
$p_{S\|I}$, probability of symptoms given infected without vaccination.	0.4	0.3, 0.5
$p_{R\|ISGl}$, the probabilities of report given infected, symptomatic and falling within one of four symptom/outcome groups.	Odds ratio of 1.0	Odds ratios of 0.5 and 1.5, reflecting an increase or decrease by 50% in the odds of report for all symptom/outcome groups.
RR_D, relative risk for infection for high versus normal dust exposure groups.	$2^{5/7} = 1.64$	$1.5^{5/7} = 1.34$ $2.5^{5/7} = 1.92$

3.11 PUBLIC HEALTH IMPLICATIONS

Aim 3 of this dissertation is to investigate the public health impact of vaccination strategies, vaccination coverage, and a weather event on forecasted incidence data.

3.11.1 Vaccination Strategy

To investigate the public health impact of different vaccination strategies, we examine three strategies: potential vaccination of adults only, potential vaccination of children only, and potential vaccination of all persons. For this analysis we adopt the vaccination rates for influenza reported by the Centers for Disease Control and Prevention (Centers for Disease Control and Prevention 2011). These data indicate the different rates of vaccination nationally by age, gender, and race strata. We assume that vaccination is in a steady-state, that is, that the vaccination program has already been well established. We also assume that one coccidioidal vaccination provides potential protection and once protected individuals receive lifelong protection. Note that while the influenza rates indicate the yearly vaccination rates, in our model

these vaccination rates are used to estimate the proportion of individuals in each stratum that have been vaccinated, not the percent that are vaccinated annually. Our vaccination rates were based upon annual influenza rates because it is unknown how long coccidioidomycosis vaccination would provide protection and at what rate individuals would be vaccinated. If a single inoculation is sufficient for lifelong protection, then the proportion of protected individuals in the population may be significantly higher than the annual influenza vaccination rates. In this case, our estimates for the public health impact of vaccination would be conservative, as coccidioidal vaccine uptake would continually increase until a steady state is likely reached. For the three vaccination strategies we examine, we compare the estimated number of cases and the estimated number of reported cases to no vaccination.

3.11.2 Vaccination Coverage

To investigate the public health impact of vaccination coverage, we examine the effect of different levels of vaccination coverage on incidence and reported incidence of coccidioidomycosis. We contrast high coverage, in which vaccination rates equaled those for influenza to low coverage in which the odds of vaccination in each stratum were reduced by half or more. We examine the impact of coverage under all three vaccination strategies: all individuals could receive vaccination, adults only could receive vaccination, and children only could receive vaccination.

3.11.3 Weather Events

To examine the impact of weather events on coccidioidomycosis infection and reported infection rates, we run the model to forecast incidence in 2010 but inject an artificial weather event into the PM10 data from the EPA (United States Environmental Protection Agency 2012) as shown in Figure 5. In this study, all of the weather and environmental data is identical to that used for the validation study, except for the artificial weather event. By comparing infection and reported infection rates following the artificial weather event we can assess the public health impact of such a weather event. The use of a real weather event provides a level of realism in terms of the magnitude and duration of the impact on PM10 as well as in terms of the spatial variation in PM10 readings across different weather stations.

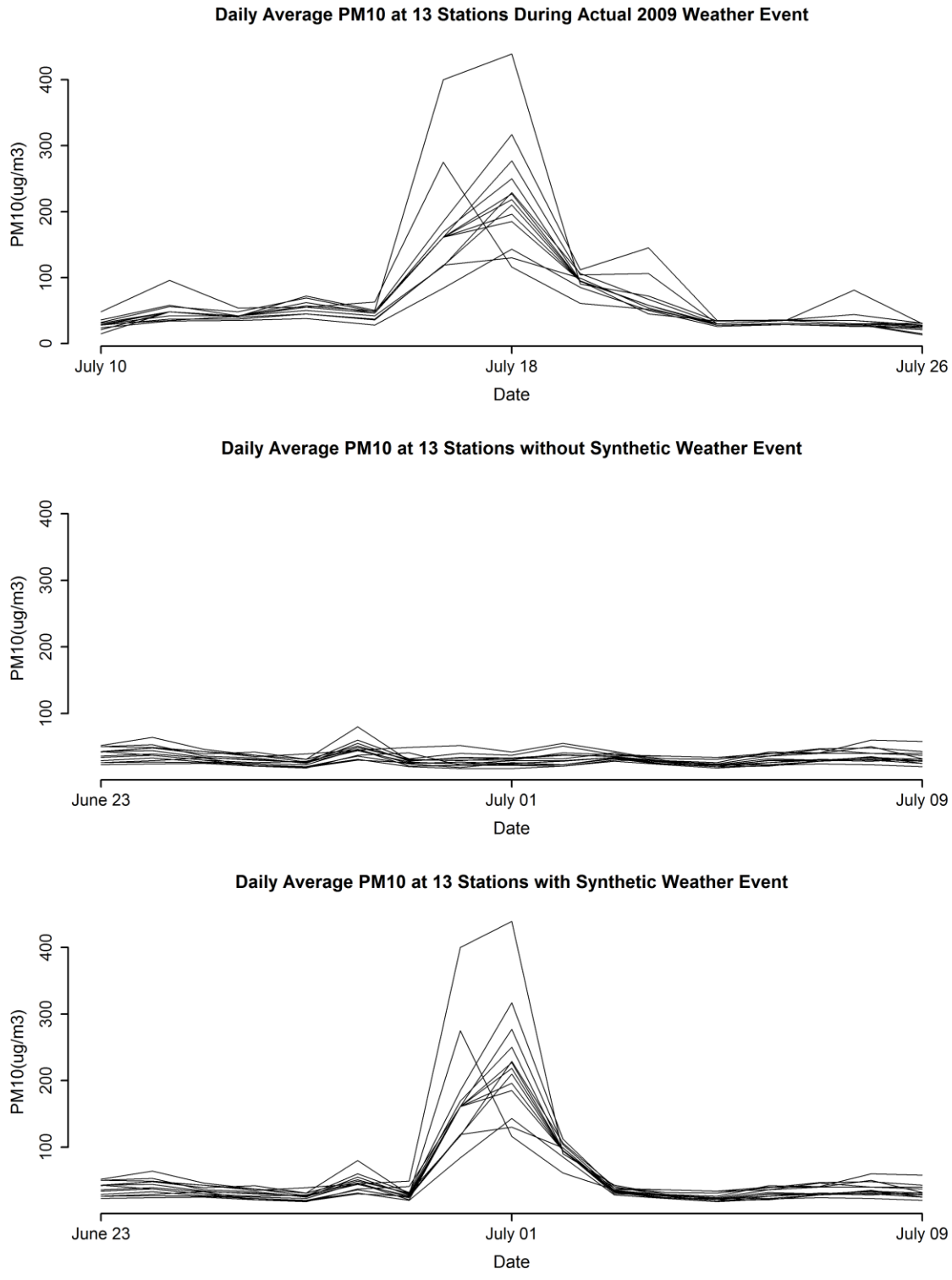

Figure 5. PM10 data in Maricopa County. Top: PM10 during a 2009 weather event. Middle: PM10 in 2010. Bottom: PM10 for 2010 with injected weather event.

3.12 ONLINE TOOL

Aim 4 of this dissertation is to develop an interactive web-based tool that researchers and policy makers could use to guide decisions regarding outbreak prevention and management. In order to run a forecasting simulation using the web-based tool, the end user may need to create and upload a number of files from historical data or estimation techniques befitting the desired simulation. The end user may need data from calendar year X (the year of interest) and calendar year $X - 1$ (the previous year) for up to 18 file uploads. The stochastic epidemiological agent-based model requires 18 files regardless: one 1 × 13 file, ten 12 × 12 files, one 52 × 1 file, four 365 × 1 files (366 × 1 for any involved leap year), and two 365 × N_ES files (366 × N_ES for any involved leap year, N_ES being the number of PM10 environmental stations). The data needed to run our stochastic epidemiological agent-based model include:

Run Parameters (one 1 × 13 file)

- Probability of symptoms given an infection
- Probability of case being reported given mild symptoms only (G1)
- Probability of case being reported given severe pulmonary symptoms without dissemination or death (G2)
- Probability of case being reported given dissemination without death (G3)
- Probability of case being reported given death (G4)
- Probability of having mild symptoms only given infected and symptomatic
- Probability of having severe pulmonary symptoms given infected and symptomatic
- Probability of having disseminated disease without death given infected and symptomatic
- Probability of mortality from coccidioidal infection given infected and symptomatic
- Vaccine effectiveness (between zero and one)
- Vaccination strategy (choose “None”, “Children”, “Adults” or “All”)
- Vaccination odds multiplier (number used to scale the odds of vaccination)
- Relative risk of infection for high dust exposure versus normal dust exposure

Age × Sex × Race/Ethnicity Values (ten 12 × 12 files)

- Population count
- Probability of mortality
- Probability of giving birth (omit males from this file)
- Probability of being in an elevated dust exposure group
- Probability of previous infection (for normal dust exposure group)
- Probability of reported infection (infected, symptomatic, reported)
- Probability of smoking at the time the model is run
- Probability of having diabetes at the time the model is run
- Probability of being immunocompromised at the time the model is run
- Probability of being vaccinated against coccidioidomycosis

Week × Conditional Birth Values (one 52 × 1 file)

- Probability of birth in each week given birth sometime in the year (α)

Day × Precipitation Values (two 365 × 1 files [366 × 1 for any involved leap year])

- Total precipitation in inches for year X
- Total precipitation in inches for year $X - 1$

Day × Temperature Values (two 365 × 1 files [366 × 1 for any involved leap year])

- Average temperature in degrees Fahrenheit for year X
- Average temperature in degrees Fahrenheit for year $X - 1$

Day × PM10 Station Values (two 365 × N_ES files [366 × N_ES for any involved leap year, N_ES being the number of PM10 environmental stations])

- Average PM10 in ug/m3 for year X
- Average PM10 in ug/m3 for year $X - 1$

The end user however does not need to create and upload every single file. We program the web-based tool to copy the 18 data files we create for our stochastic epidemiological agent-based

model, place them in a directory unique to the particular simulation along with the necessary R files, ensure that any referenced directory paths are correct, and let the end user upload and overwrite any of the 18 files with custom data files. Thus with a minimum of one required upload, the end user only needs to upload those files custom to the desired analysis. Should the end user wish to upload a file, the web-based interface will allow her to select a file for upload by clicking various browse and upload buttons. Once a file is selected for upload, the file is uploaded, cleaned for security purposes (e.g., prevent cross-site scripting attacks), and checked for dimensions. If the dimensions of the uploaded file do not match the requested data file dimensions, the end user is prompted to resolve the mismatch. The end user can continue to upload the desired files and interact with the model webpage without numerous page reloads due to our incorporation of asynchronous JavaScript. Once the end user is done uploading files, she then proceeds to click the simulate button to start the simulation process.

In order to provide the end user with the simulation results, we set up our web-based tool to create and use browser cookies to hold information relevant to the simulation. During the file upload process, the end user is prompted for an email address, which will be stored in a cookie along with a unique model identification number. The model identification number is created to be unique to a particular simulation and is stored in a cookie as reference to that particular simulation. The email cookie is used to verify an email address exists, that is then validated for legitimacy and written to a file. The latter is used only to email the end user with the simulation results, as the simulation process is not instantaneous. Once the simulation is complete, the zipped results will be emailed to the end user via a function that is called on certain page loads. Further, the end user can click a download button to retrieve simulation results current to the download request though the simulation may still be running. This allows the end user to view simulation results without having to wait for the entire process to complete.

4.0 RESULTS

4.1 MARICOPA COUNTY, ARIZONA

We choose Maricopa County, Arizona as the geographical focus of our research. Up to half of all yearly coccidioidal incidence cases in the United States occur in Maricopa County, Arizona (Park, Sigel et al. 2005; Sunenshine qtd. in Foote 2011). Age, gender, and race/ethnicity demographics and coccidioidomycosis data are listed in Tables 4, 5, and 6 (Maricopa County Department of Public Health 2011). Of the reported coccidioidal disease cases in 2009, thirty-six were without age, sixty-nine reports omitted gender, and 6,214 (77%) had unknown race. The age group with the highest infection rate is 65-84 year olds. Children under five years of age have the lowest infection rate with 8.4 cases per 100,000 children. While whites represent 72% of cases with reported race/ethnicity, African-Americans have the highest rate of reported infections with 80.9 cases per 100,000 persons. Hispanics have the lowest rate of reported infections with 21.6 cases per 100,000 persons. In Maricopa County, females represent 56% of all reported infections and have a case report rate of 223.7 per 100,000 women. In contrast, males have a case report rate of 174.9 per 100,000 men.

Table 4. Age group by gender and coccidioidomycosis for Maricopa County, 2009.

Age Group	Males		Females		Reported Number of Cocci Cases	Reported Cocci Rate per 100K
	Number	Percent	Number	Percent		
0-4	169,660	8.4	163,233	8.2	28	8.4
5-9	160,941	7.9	154,155	7.7	98	31.1
10-14	144,530	7.1	137,454	6.9	191	67.7
15-19	139,063	6.8	128,281	6.4	400	149.6
20-24	138,586	6.8	125,074	6.3	436	165.4
25-34	336,983	16.6	300,143	15.1	1,099	172.5
35-44	297,414	14.6	274,211	13.8	1,483	259.4
45-54	259,853	12.8	258,413	13.0	1,465	282.7
55-64	184,815	9.1	201,035	10.1	1,222	316.7
65-74	108,131	5.3	123,172	6.2	1,433	373.7
75-84	66,767	3.3	85,406	4.3		
85+	23,807	1.2	42,005	2.1	187	281.4
Unknown	---	---	---	---	36	---
Total	2,030,550	100.0	1,992,582	100.0	8,078	200.8

Table 5. Race/ethnicity by gender and coccidioidomycosis for Maricopa County, 2009.

Race/ Ethnicity	Males		Females		Unknown	Reported Number of Cocci Cases	Reported Cocci Rate per 100K
	Number	Percent	Number	Percent	Number		
White	1,140,454	56.2	1,173,035	58.9	---	1,348	58.3
Hispanic	678,597	33.4	602,049	30.2	---	276	21.6
African-American	91,091	4.5	88,243	4.4	---	145	80.9
American-Indian	31,615	1.6	34,066	1.7	---	44	67.0
Asian	59,992	3.0	65,885	3.3	---	51	40.5
Other	28,801	1.4	29,304	1.5	---	---	---
Unknown	---	---	---	---	---	6,214	---
Total	2,030,550	100.0	1,992,582	100	---	8,078	200.8
Reported Number of Cocci Cases	3,551	0.175	4,458	0.224	69	8,078	---
Reported Cocci Rate per 100K	174.9	---	223.7	---	---	---	200.8

Table 6. Age group by race/ethnicity for Maricopa County, 2009.

Age Group	Race/Ethnicity						Total
	White	Hispanic	Black	Indian	Asian	Other	
0-4	128,790 (5.6%)	161,497 (12.6%)	16,381 (9.1%)	6,536 (10.0%)	10,146 (8.1%)	9,543 (16.4%)	332,893 (8.3%)
5-9	130,595 (5.6%)	143,908 (11.2%)	16,640 (9.3%)	5,892 (9.0%)	9,617 (7.6%)	8,444 (14.5%)	315,096 (7.8%)
10-14	128,159 (5.5%)	118,634 (9.3%)	14,967 (8.3%)	4,975 (7.6%)	8,121 (6.5%)	7,128 (12.3%)	281,984 (7.0%)
15-19	127,100 (5.5%)	107,821 (8.4%)	14,820 (8.3%)	5,412 (8.2%)	6,662 (5.3%)	5,529 (9.5%)	267,344 (6.6%)
20-24	128,693 (5.6%)	103,771 (8.1%)	14,227 (7.9%)	5,768 (8.8%)	6,760 (5.4%)	4,441 (7.6%)	263,660 (6.6%)
25-34	315,760 (13.6%)	246,235 (19.2%)	29,509 (16.5%)	13,710 (20.9%)	23,398 (18.6%)	8,514 (14.7%)	637,126 (15.8%)
35-44	320,654 (13.9%)	185,588 (14.5%)	26,049 (14.5%)	9,723 (14.8%)	24,257 (19.3%)	5,354 (9.2%)	571,625 (14.2%)
45-54	355,177 (15.4%)	113,025 (8.8%)	22,451 (12.5%)	6,949 (10.6%)	16,466 (13.1%)	4,198 (7.2%)	518,266 (12.9%)
55-64	297,526 (12.9%)	57,753 (4.5%)	13,338 (7.4%)	3,831 (5.8%)	10,796 (8.6%)	2,606 (4.5%)	385,850 (9.6%)
65-74	189,453 (8.2%)	26,148 (2.0%)	6,575 (3.7%)	1,828 (2.8%)	5,895 (4.7%)	1,404 (2.4%)	231,303 (5.7%)
75-84	132,263 (5.7%)	12,136 (0.9%)	3,407 (1.9%)	816 (1.2%)	2,848 (2.3%)	703 (1.2%)	152,173 (3.8%)
85+	59,319 (2.6%)	4,130 (0.3%)	970 (0.5%)	241 (0.4%)	911 (0.7%)	241 (0.4%)	65,812 (1.6%)
Total	2,313,489 (57.5%)	1,280,646 (31.8%)	179,334 (4.5%)	65,681 (1.6%)	125,877 (3.1%)	58,105 (1.4%)	4,023,132 (100.0%)

4.2 MODEL INSTANTIATION

Aims 2 and 3 require instantiation of the model for a specific time and location in order to assess model validity, conduct sensitivity analyses and evaluate the public health implications of

vaccination and weather events. As discussed in the previous section, we choose Maricopa County because of the relative abundance of cases in the County. We use incidence data through 2009 to instantiate the model (Maricopa County Department of Public Health 2011). The data available for our use were limited to publically available data. Some data are extrapolated from previous years, such as with 2009 United States Census data. Other data are provided in marginal format (e.g., age by gender) so in order to find joint numbers (e.g., age, gender, and race/ethnicity) we use iterative proportional fitting (Fienberg 1970). In this section, we provide representations of data used for model instantiation. Figures 6, 7, and 8 display various counts and probabilities for the 144 strata defined by the twelve age groups, two genders, and six race/ethnicity groups. In Figure 6 we see that the 2009 Maricopa County population is predominately white and Hispanic (Maricopa County Department of Public Health 2011). The largest counts are between 25 and 65 years of age across all gender and race/ethnicity. The visible break in the graph at 25 years of age is primarily due to the shorter age groups for individuals under 25 years of age. Age groups were chosen based on the age groups used by the Maricopa County Department of Health (Maricopa County Department of Public Health 2011).

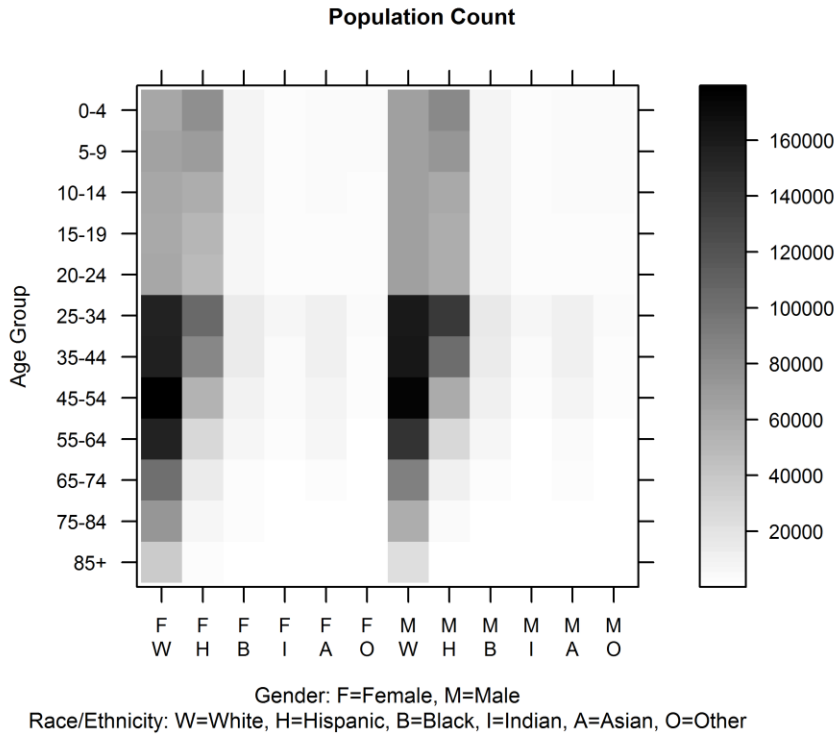

Figure 6. Population count for Maricopa County, 2009.

In Figure 7, we see that diabetes is generally an elderly disease. While type 1 diabetes is included in the graph, the overall prevalence of diabetes is much lower in children than older adults. The highest rate of diabetes is in male Native Americans. Native American and African American youth and young adults have a higher prevalence rate of diabetes than whites, Hispanics, and Asians. Further diabetes prevalence rates increase across nearly all race/ethnicity groups after 65 years of age.

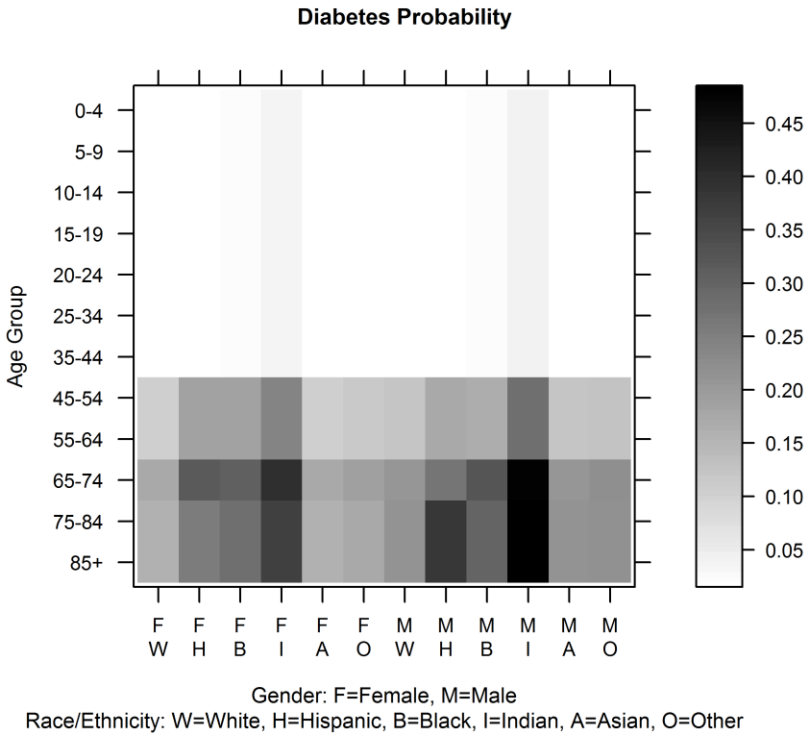

Figure 7. Diabetes probabilities for Maricopa County, 2009.

Figure 8 shows reported infection probabilities. In individuals under the age of 65, whites have the highest rates of reported coccidioidomycosis infection. Elderly blacks, Native Americans, and Hispanics have the highest rates of reported infection but these individuals represent a very small proportion of the population. See Figure 6.

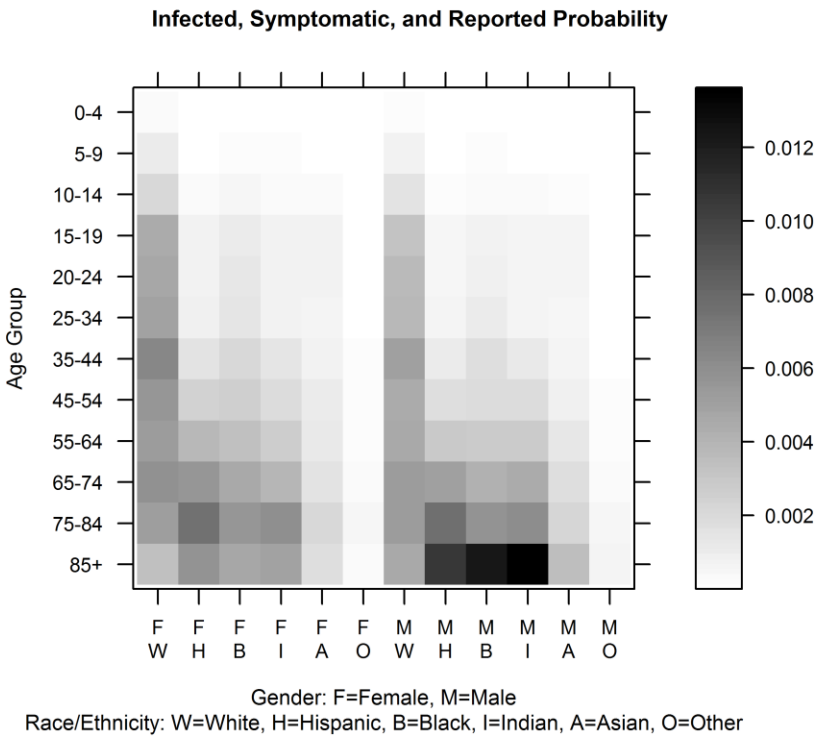

Figure 8. Reported infection probabilities for Maricopa County, 2009.

Figure 9 shows national flu vaccination rates by age, gender and race/ethnicity. Youth and elderly receive flu vaccination at nearly double the rates of young and middle-aged adults. Hispanic and African-American males have the lowest vaccination rates.

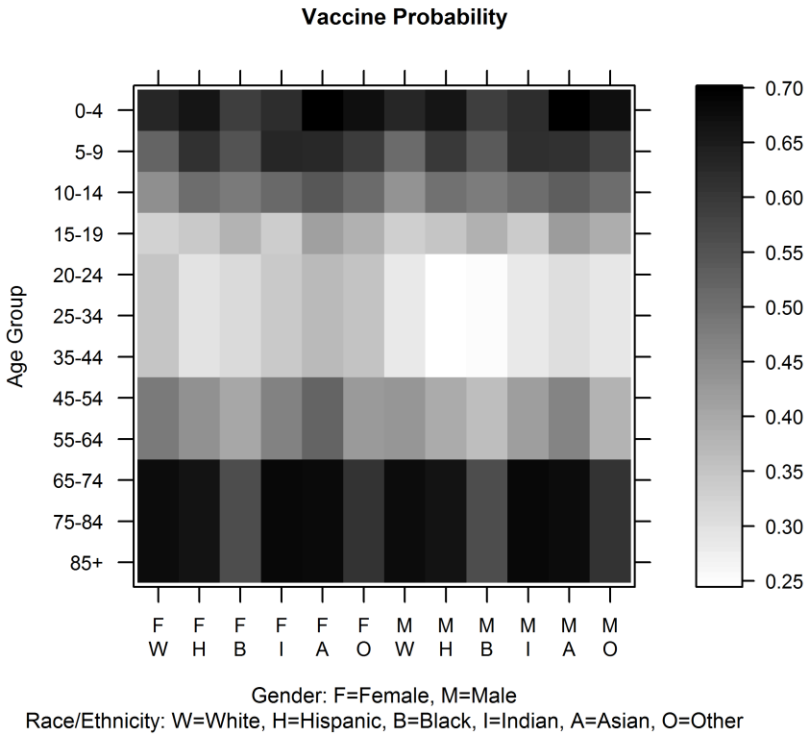

Figure 9. Vaccination probabilities for Maricopa County, 2009.

4.3 MODEL VALIDATION

Pursuant to Aim 2, we assess the validity of the model in two ways. First, we consider Objective 2a, which asserts that the model will demonstrate previously reported demographic, occupational, co-morbidity and behavioral risk factors for coccidioidomycosis. We evaluate this objective by examining whether the outputs from the model match with the relative risks and other characteristics that were used to construct the model. Second, we consider Objective 2b, which states that in an independent test of the model, forecast intervals calculated by the model will contain true reported incidence counts. We evaluate this objective by using the model to forecast incidence data for Maricopa County in 2010 (Maricopa County Department of Public Health 2012) and compare the results to actual counts of reported cases for that year.

4.3.1 Face Validity

We compare the estimated percentages and relative risks from the ten simulation runs to the corresponding percentages and relative risks used to create the model. Note that we do not expect perfect agreement because we are only comparing the marginal relative risks but the model accounts for multiple risk factors. In addition, most of the relative risks represented in the model are in reference to the population of Kern County, CA, in 1995-1996 that was studied in Rosenstein et al. (Rosenstein, Emery et al. 2001). For example, while 44% of reported infections in Maricopa County in 2009 were male (Maricopa County Department of Public Health 2011), 56% of the cases in Kern County were male (Rosenstein, Emery et al. 2001). Differences between these populations may account for some differences between the model inputs and outputs from the simulations. Regardless, we still expect relatively close agreement. For example, if these relative risks disagree by an order of magnitude, the disagreement would suggest that the model is not implemented correctly.

The results of these comparisons are presented in Table 7. Out of the fourteen parameters comparisons, ten (71%) of the estimated values were within 10% of their respective model inputs. In some cases the estimated values nearly equaled the model inputs. For example, 39.8% of the infected individuals had symptoms, compared to 40.0% as specified in the model. Also, the relative risk of severe pulmonary symptoms for males versus females was estimated exactly at 1.19. The four parameters that were not estimated to within 10% were the percent of symptomatic infections with severe pulmonary symptoms and the relative risks of severe pulmonary symptoms for diabetics versus non-diabetics, immunocompromised versus immunocompetent, and smokers versus non-smokers. These parameters are discussed in further detail in the discussion section.

Table 7. Model inputs compared to simulation estimates for Maricopa County, 2010.

Parameter	Model Input	Estimate from Simulations
Relative risk of infection for high versus low dust exposure group	1.64	1.73
Percent of infections that are symptomatic	40.0	39.8
Percent of symptomatic infections with severe pulmonary symptoms	20.3	27.2
Percent of symptomatic infections with disseminated disease	8.7	8.0
Percent of symptomatic infections resulting in death	1.1	1.0
Relative risk of severe pulmonary symptoms for African-American versus white	1.53	1.66
Odds ratio of severe pulmonary symptoms for age	1.03	1.04
Relative risk of severe pulmonary symptoms for male versus female	1.19	1.19
Relative risk of severe pulmonary symptoms for diabetic versus non-diabetic	2.81	3.43
Relative risk of severe pulmonary symptoms for immunocompromised versus immunocompetent	5.26	3.76
Relative risk of severe pulmonary symptoms for smoker versus non-smoker	2.34	1.91
Relative risk of dissemination for African-American versus white	3.58	3.85
Relative risk of dissemination for diabetic versus non-diabetic	2.34	2.45
Relative risk of dissemination for smoker versus non-smoker	1.64	1.77

We also compare the demographic distributions of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009 (Maricopa County Department of Public Health 2011). These comparisons are made for gender, age, and race/ethnicity groups. For these comparisons, we impute the demographic characteristics of any reported infections with unknown values. Figure 10 shows the proportion of actual reported infections in 2009 that are female and male, the proportion of forecasted infections that are female and male, and the proportion of forecasted case reports that are female and male. The

proportions for each gender are nearly equal, for example the proportions of female cases are between 54% and 56%. Figure 11 shows the proportion of actual case reports in 2009, the proportion of forecasted infections, and the proportion of forecasted case reports by age group. The proportions of actual and forecasted case reports are within 0.6% across the twelve age groups. The proportions of all infections differ from those reported because older individuals are more likely to have complications from infection and therefore more likely to be reported to the Department of Public Health. Figure 12 shows the proportion of actual reported infections in 2009 and the proportions of forecasted infections and case reports by race/ethnicity group. The proportions are very consistent. For example, the proportions for whites are all between 84% and 85%, for Hispanics between 11% and 13%, and all other proportions less than 3%. As observed above for age, the proportion of forecasted infections that are African-American is less than the proportions of reported infections because African-American is a risk factor for complications from infection. Thus African-American cases are more likely to generate reports.

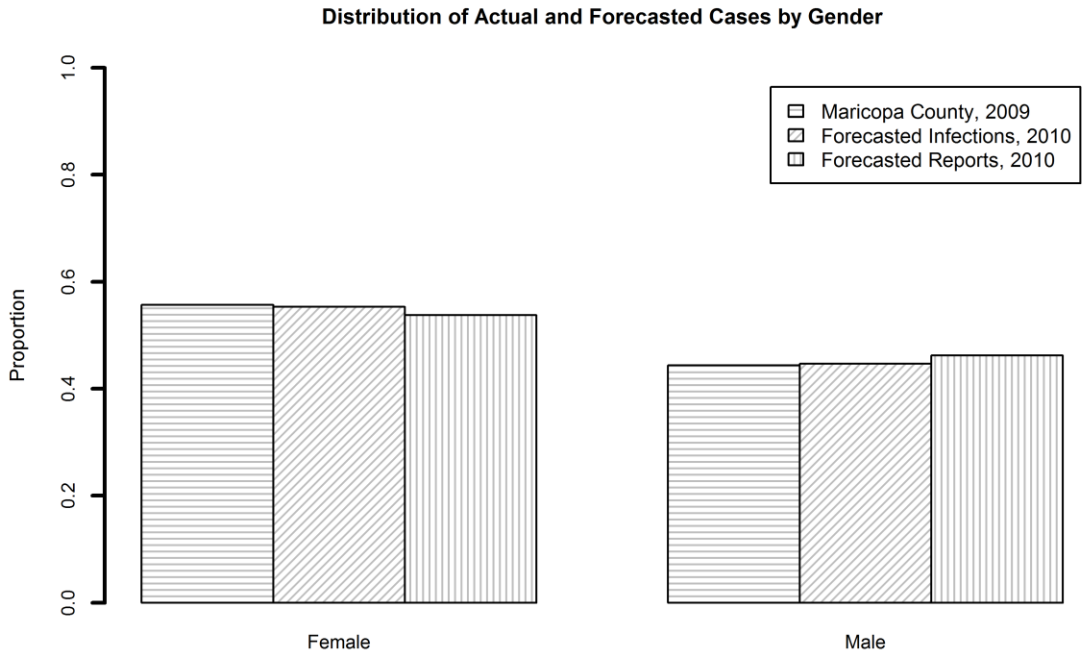

Figure 10. Gender distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009.

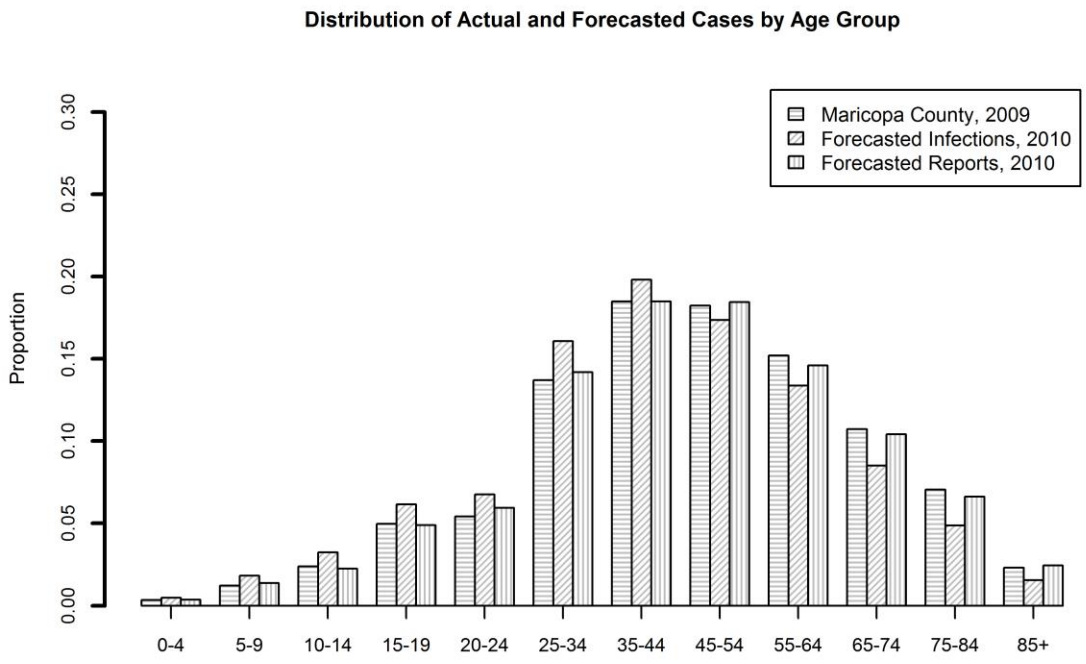

Figure 11. Age distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009.

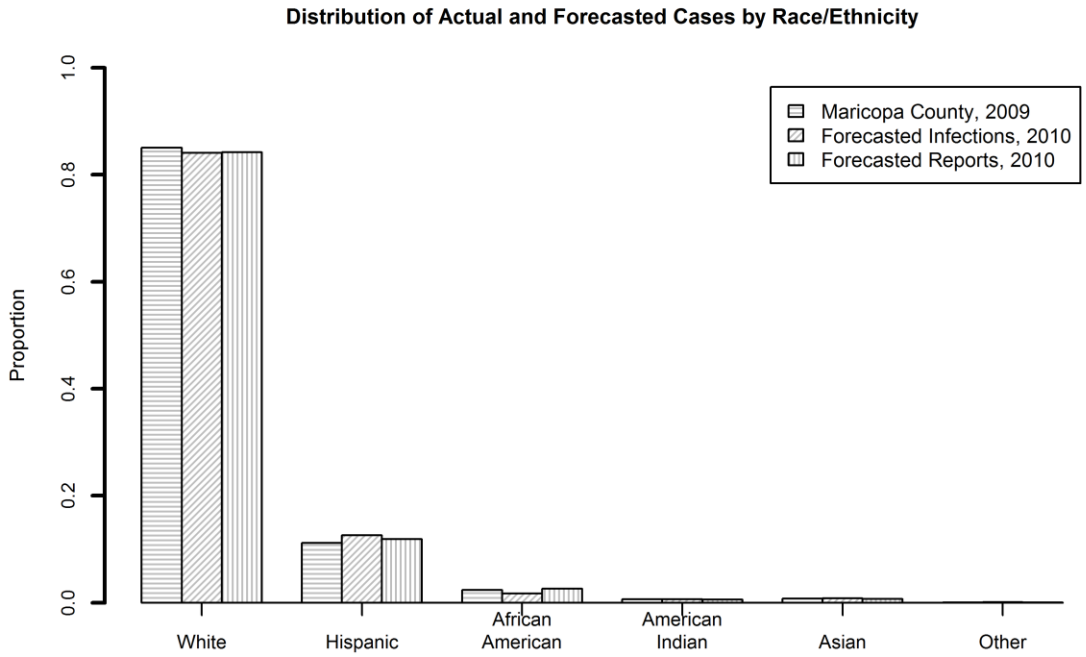

Figure 12. Race/ethnicity distribution of forecasted total infections and forecasted reported infections to actual reported infections in Maricopa County in 2009.

We further examine the temporal trends in total infections (Figure 13). The number of infections is relatively high for the first couple of weeks of 2010 followed by a significant drop in the number of new infections. Starting at week 12 (March 19-25, 2010) there is a gradual increase in the weekly number of new infections, culminating in a peak at week 29 (July 16-22) with an estimated 1390 infections. From August through December the number of weekly infections is highly variable but relatively flat, ranging from a low of 608 cases during the first week of August (July 30-August 5) to a high of 1317 cases in week 49 (December 3-9).

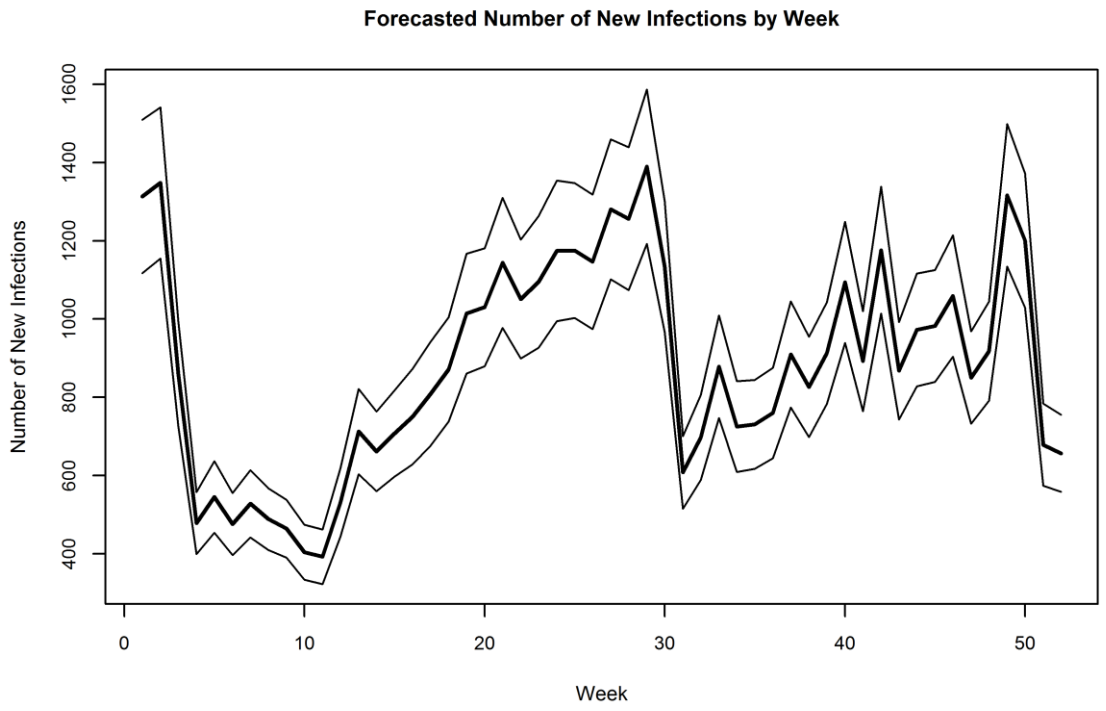

Figure 13. Simulated number of new infections by week. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

The cross-correlation function between the number of new infections and daily average PM10 is shown in Figure 14. The peak occurs at lag 0 weeks, at the which the correlation between the number of infections and the daily average PM10 is moderately strong at r=0.77. The cross-correlation exceeds 0.4 at lags of zero to five weeks, likely due to broad periods of high dust activity in January and October of 2010.

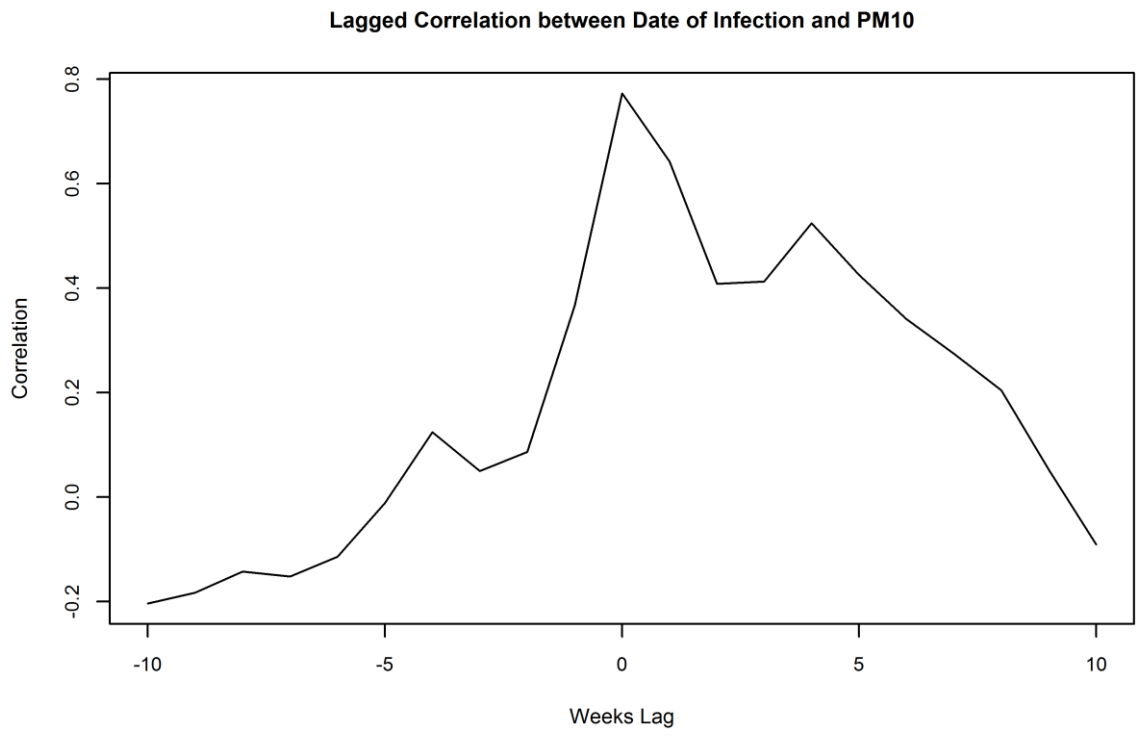

Figure 14. Cross-correlation function between date of infection and PM10.

Figure 15 displays the number of symptomatic infections by week of symptom onset. The temporal trends in symptomatic infection closely resemble that of total infections (Figure 13) but the numbers of weekly symptomatic infections is much lower than the weekly total infections as only approximately 40% of all infections are symptomatic. The week with the greatest number of new symptomatic cases is week 1 (Jan 1-7) with an estimated 558 individuals experiencing symptoms from infection. There is also a peak at week 31 (July 30-August 5) with 544 new symptomatic cases. The symptomatic infections graph is also shifted by a couple of weeks due to the incubation period of coccidioidomycosis. Figure 16 displays the cross-correlation between the number of cases by date of infection and the number of symptomatic cases by date of symptom onset. The peak correlation is 0.94, which occurs at a lag of 2 weeks and corresponds precisely to the mean incubation period in the model.

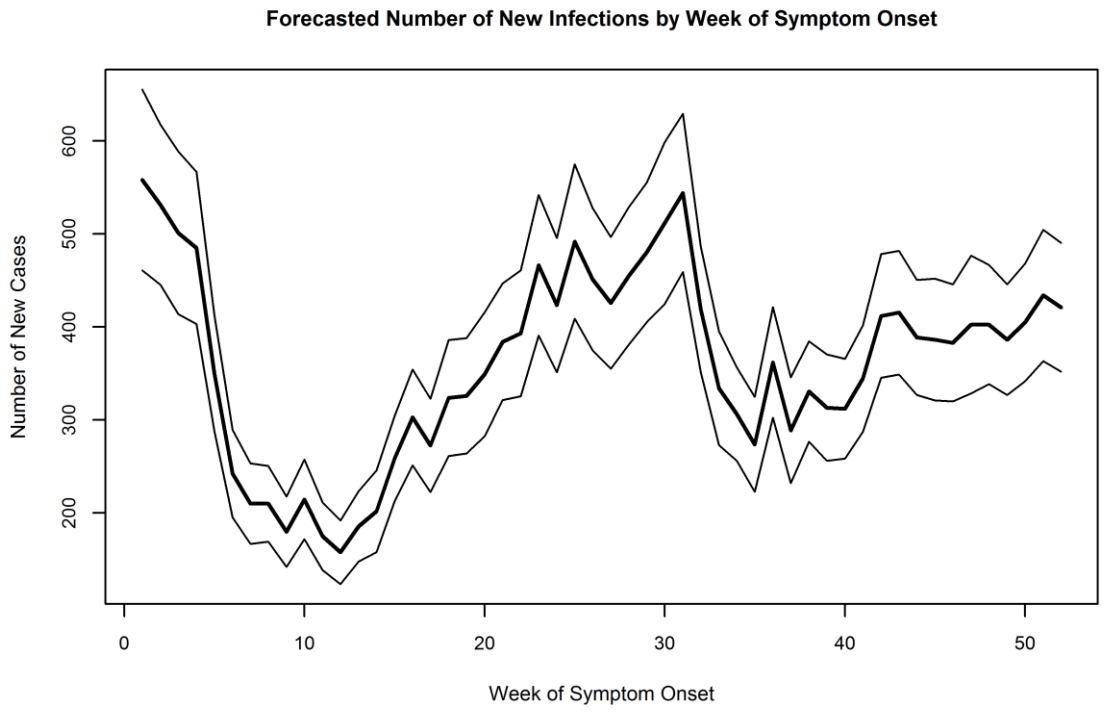

Figure 15. Simulated number of new symptomatic infections by week of symptom onset. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

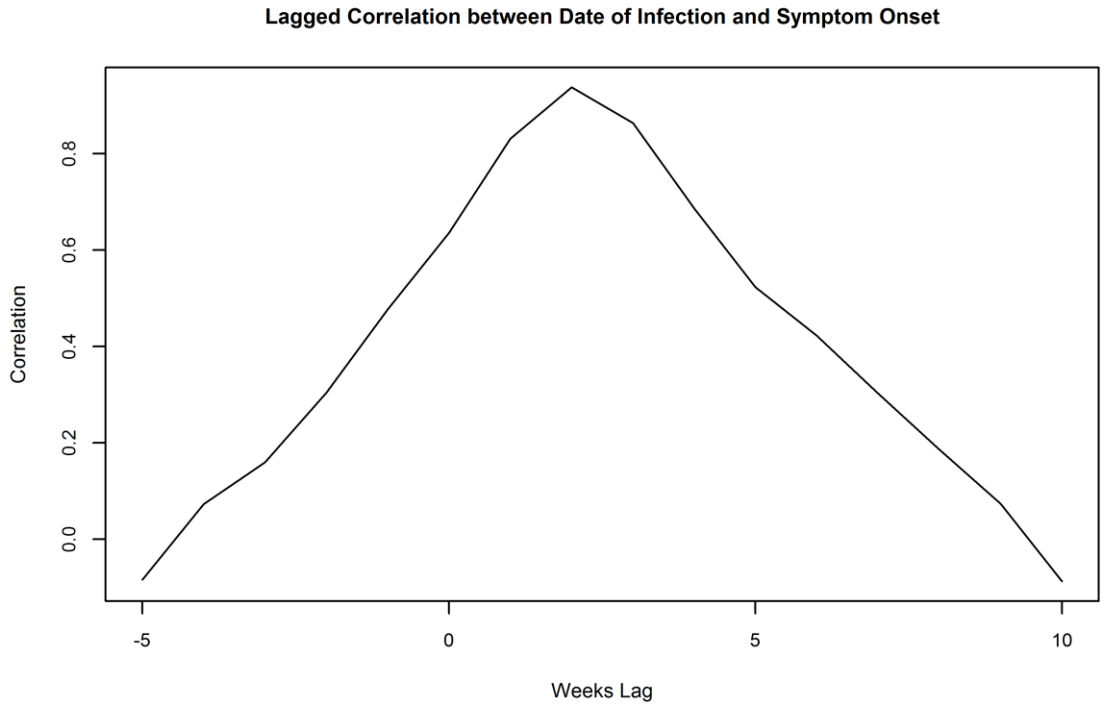

Figure 16. Cross-correlation function between date of infection and date of symptom onset.

Figure 17 displays the number of reported infections by week of report to the health department. The weekly number of reported infections ranges from 124 at week 23 (June 4-10) to 298 in week 1 (Jan 1-7). The numbers of reports is greatest in January and February of 2010. From March through December, the weekly numbers of reports is relatively flat. The peak in early 2010 can be attributed to the greater dust levels in the second half of 2009 and the long lag between infection and report. Figure 18 displays the cross-correlation function between date of symptom onset and date of case report. The greatest correlation occurs at a lag of 6 weeks, with a correlation of 0.57. The correlation exceeds 0.4 between 3 and 9 weeks lag, which is consistent with the median reporting delay of 55 days that was used in the model.

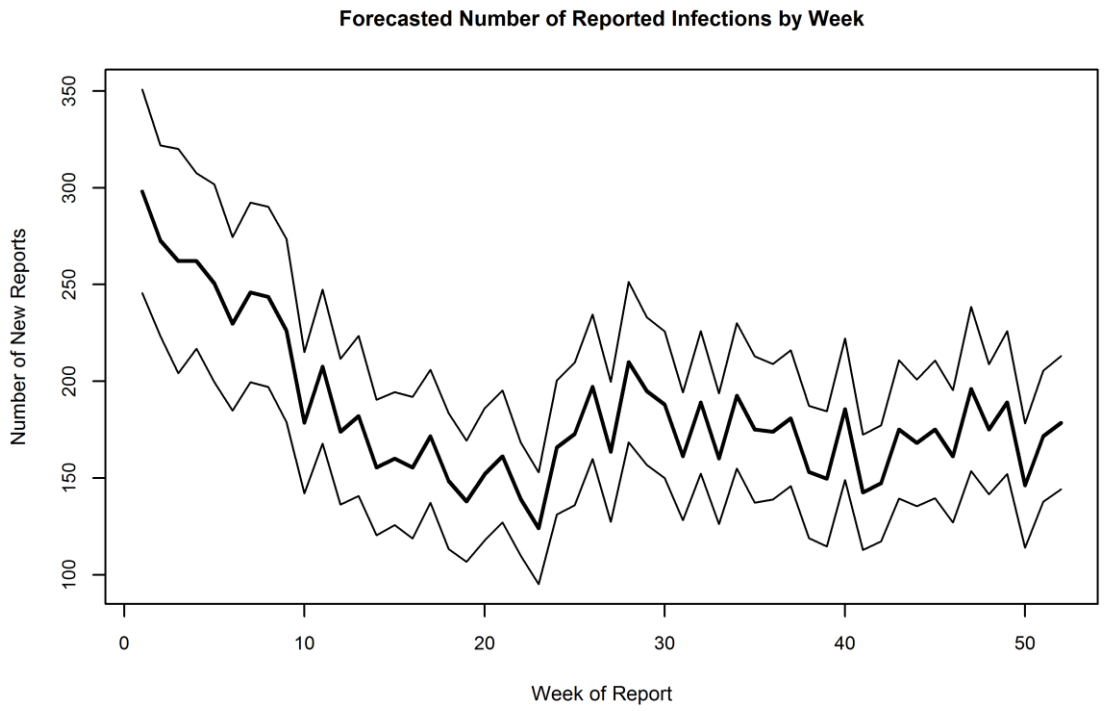

Figure 17. Simulated number of new coccidioidomycosis case reports by week of report. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

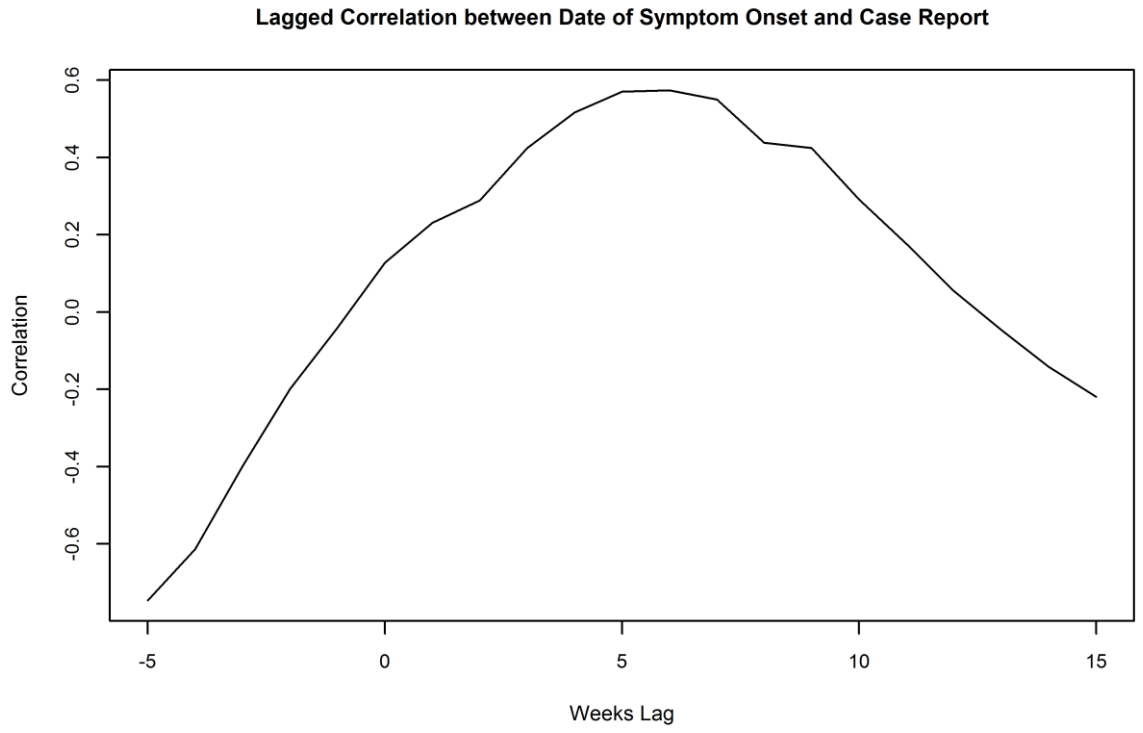

Figure 18. Cross-correlation function between date of symptom onset and date of case report.

4.3.2 Independent Validation

Five runs of the model were used to construct forecast intervals for the numbers of reported cases in 2010, in aggregate and broken down by age, gender, and race. We compare these intervals to actual counts reported to the Maricopa County Health Department (Maricopa County Department of Public Health 2012). We also compare the incidence rates per 100,000 persons to account for differing population estimates, as the forecasted incidence rates are based on extrapolated 2000 census figures while the actual incidence rates provided by the Maricopa County Health Department for 2010 (Maricopa County Department of Public Health 2012) are based on the 2010 census. The results of these comparisons are listed in Table 8. The total number of reported cases was forecasted to be 9,580, with a forecast interval of 8,354 to 10,806 cases. The actual number of reported cases in 2010 was 9,456, which is well within the 95%

forecast interval. The actual number of female and male cases, 5,464 and 3,878 respectively, also fell within the forecast intervals. The Maricopa County Health Department reported cases by age in 11 age groups. Forecast intervals for 9 of these 11 age groups contained the actual reported number of cases. The two exceptions were the 20-24 and 65-84 age groups, for which the model forecasted fewer cases than actually reported. The results were similar in terms of cases per 100,000 residents. The total incidence rate and the incidence rates by gender fell within the forecast intervals from the model. In addition, the incidence rates for 9 of the 11 age groups fell within the forecast intervals. The exceptions were the 0-4 and 20-24 age groups, for which the model forecasted fewer cases per 100,000 persons than observed. Comparisons by race/ethnicity group were not possible as 79% of the reported cases in 2010 had unknown race (Maricopa County Department of Public Health 2012).

Table 8. Actual and forecasted number of reported cases for Maricopa County, 2010.

	Actual Cases	Forecast Cases		Actual Cases per 100K	Forecast Cases per 100K	
	Number	Mean	Interval	Number	Mean	Interval
Total	9,456	9,580	8,354-10,806	247.7	238.1	207.7-268.6
Females	5,464	5,150	4,481-5,819	283.3	258.4	224.9-292.0
Males	3,878	4,430	3,860-5,001	205.4	218.2	1,90.1-246.3
0-4	45	35	22-47	15.9	10.5	6.6-14.3
5-9	106	131	103-159	37.5	41.6	32.7-50.5
10-14	209	216	176-255	75.4	76.5	62.6-90.5
15-19	512	469	391-546	184.9	175.3	146.2-204.4
20-24	449	570	481-658	168.2	216.0	182.6-249.5
25-34	1,295	1,358	1,166-1,551	239.3	213.2	183.0-243.4
35-44	1,672	1,771	1,537-2,006	318.7	309.9	268.8-351.0
45-54	1,654	1,767	1,523-2,010	328.2	340.9	293.9-387.9
55-64	1,466	1,398	1,221-1,575	368.1	362.3	316.3-408.3
65-84	1,808	1,632	1,461-1,803	448.0	425.6	381.0-470.2
85+	201	223	194-272	340.4	354.3	295.2-413.5
White	†	8,070	7,048-9,091	†	348.8	304.7-392.9
Hispanic	†	1,136	968-1,303	†	88.7	75.6-101.7
Black	†	248	201-296	†	138.4	112.0-164.8
Indian	†	52	36-68	†	79.5	54.9-104.1
Asian	†	68	49-88	†	54.4	38.7-70.0
Other	†	6	1-10	†	10.0	1.9-18.1

† Comparison is not possible because 79% of reported cases had unknown race in 2010.

4.4 SENSITIVITY ANALYSIS

Objective 2c asserts that the model will respond predictably to changes in key parameter values. To evaluate this objective, we conduct one-way sensitivity analyses. For each sensitivity analysis we report the mean number of infections, mean number of symptomatic infections, and the mean number of case reports in Table 9. While the number of total infections was very sensitive to changes in $p_{S|I}$, the numbers of symptomatic and reported infections were insensitive

to changes in $p_{S|I}$. The numbers of total and symptomatic infections were sensitive to changes in the probability of a case report given the symptom/outcome group ($p_{R|ISGl}$), the number of reported cases is insensitive. These seemingly counter-intuitive results are discussed further in the discussion section. Finally, changes of 17-18% in the relative risk of infection for individuals with high dust exposure had smaller effects of less than 3% on total infections, symptomatic infections and reported infections.

Table 9. One-way sensitivity analyses: mean number of infections, symptomatic infections, and case reports.

Parameter Value	Total Infections	Symptomatic Infections	Reported Infections
$p_{S\|I} = 0.4$ (base)	45,941	18,783	9,580
$p_{S\|I} = 0.3$ (low)	61,190	18,404	9,382
$p_{S\|I} = 0.5$ (high)	36,797	18,389	9,373
$p_{R\|ISGl}$ OR=1 (base)	45,941	18,783	9,580
$p_{R\|ISGl}$ OR=0.5 (low)	67,200	27,408	9,419
$p_{R\|ISGl}$ OR=1.5 (high)	36,810	15,122	9,192
$RR_D = 2^{5/7} = 1.64$ (base)	45,941	18,783	9,580
$RR_D = 1.5^{5/7} = 1.34$ (low)	44,939	18,375	9,360
$RR_D = 2.5^{5/7} = 1.92$ (low)	47,159	19,248	9,715

4.5 PUBLIC HEALTH IMPLICATIONS

For Aim 3 of this dissertation, we use the model to forecast incidence and investigate the public health impact of vaccination strategies, vaccination coverage, and a weather event.

4.5.1 Vaccination Strategy

The availability of an effective vaccine could significantly decrease coccidioidomycosis-related morbidity and mortality. We examine three vaccination strategies: all individuals are able to receive the vaccine, only children are eligible to receive the vaccine, and only adults are eligible to receive the vaccine. If a vaccine with 75% effectiveness became available to all individuals and was received at rates comparable to influenza, the number of annual symptomatic infections would decrease by an estimated 5,979 cases. Further the number of deaths from coccidioidomycosis would decrease by 70 and the number of reported infections would reduce by 3,180, representing approximately one-third of all reported infections. If the vaccine was only made available to adults, we would still witness nearly the same benefit. However, if the vaccine was only available to children, the number of symptomatic infections would reduce by 374, mortality would reduce by 5 deaths, and reported infections would decline by 142 cases. See Figure 19. If the vaccine was only 60% effective and available to all individuals, we would still expect to see 4,725 fewer symptomatic infections, 50 fewer deaths, and 2,510 fewer reported infections. Alternatively, if the vaccine was 90% effective and available to all, we would anticipate 7,125 fewer symptomatic infections, 82 fewer deaths and 3,776 fewer case reports. As discussed in Section 3.11.1, if a single inoculation is sufficient for lifelong protection, these estimates for the reduction in morbidity and mortality would likely be conservative.

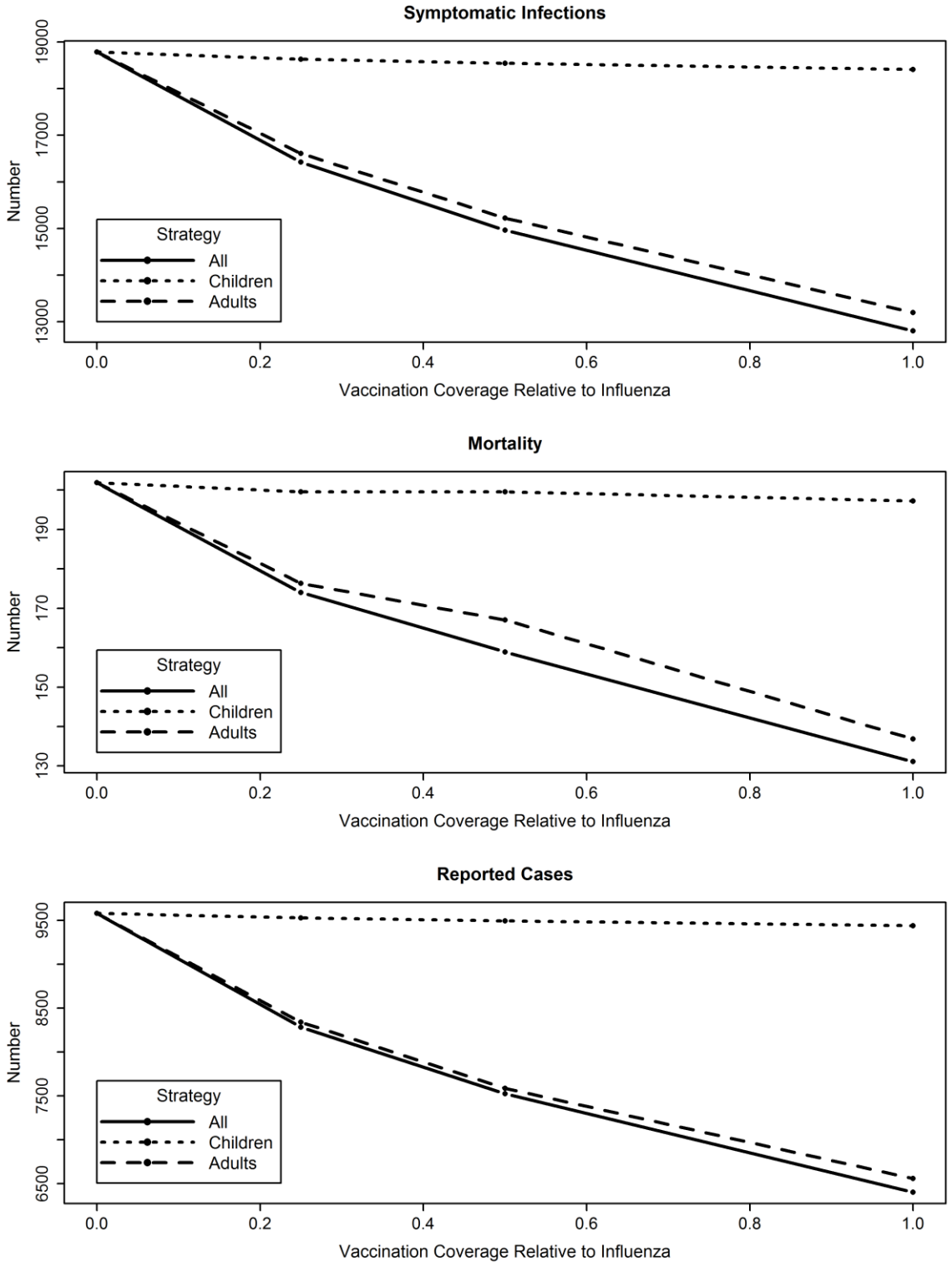

Figure 19. Symptomatic infections (top), infections resulting in death (middle), and reported cases (bottom) by vaccination strategy and coverage.

4.5.2 Vaccination Coverage

In the above, we assume that vaccination coverage was equal to influenza. With availability of a vaccine, a public health policy question could arise regarding how actively to promote vaccination. In the case of influenza, a significant effort is put forth to encourage students, workers, and elderly to be vaccinated. In this section we quantify the impact of different levels of vaccination. If the odds individuals receiving vaccination were half those of influenza, we would anticipate a reduction of 3,819 symptomatic infections in Maricopa County, 43 fewer deaths, and 2,055 fewer case reports of infection to public health. If the odds were one-quarter those of influenza, we would expect 2,361 fewer symptomatic infections, 28 fewer deaths, and 1,297 fewer case reports of infections to public health.

4.5.3 Weather Events

The synthetic weather event had a significant impact on public health in terms on number of additional infections, number of symptomatic infections and the number of reported cases. Figure 20 displays the number of new infections for 2010 with the synthetic weather event and Figure 21 compares the mean number to that found without the synthetic weather event. The synthetic weather event increased the average yearly number of new infections by 7,995 cases, from 45,941 without the synthetic weather event to 53,936 with the event. However, on the primary week of the synthetic weather event (week 26, June 26-July 1), the forecasted number of infections was 12,808, with a forecast interval of 10,965 to 14,652. This represents a one-week increase of 11,662 cases, or an 11-fold increase over the forecasted number without the event. The synthetic weather event ends on July 2, which is the first day of week 27. For this week, there was an increase of only 67 cases, from 1,281 to 1,348. Therefore, during the weeks of the weather event, we observe an additional 11,729 infections.

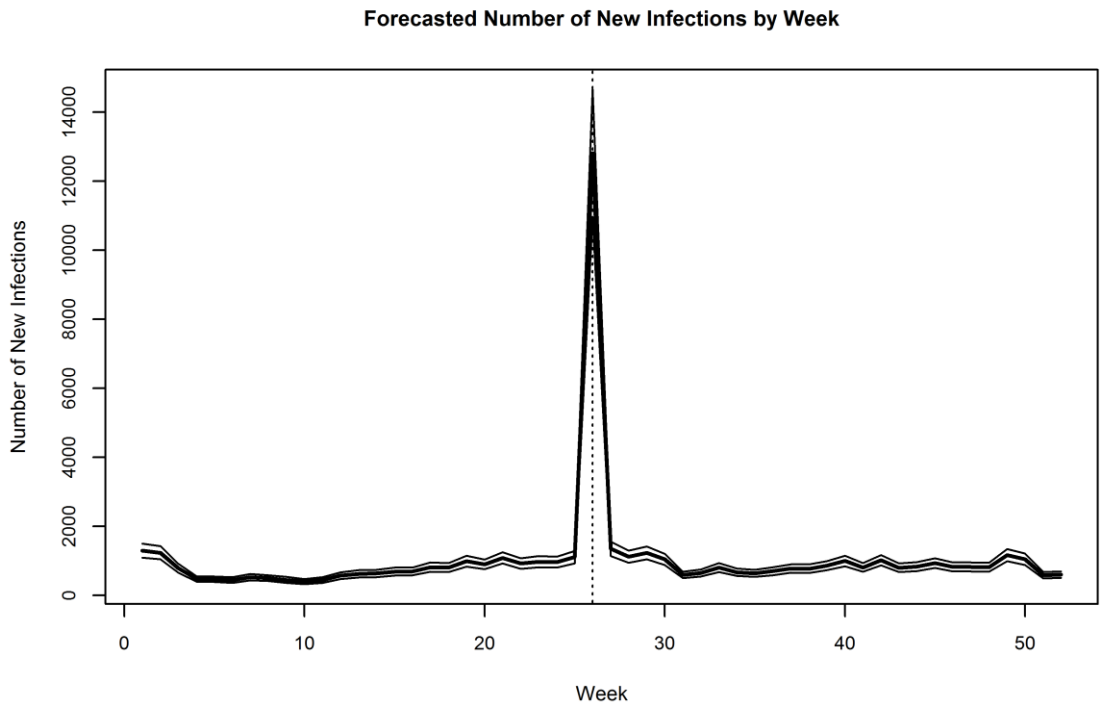

Figure 20. New infections by week with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

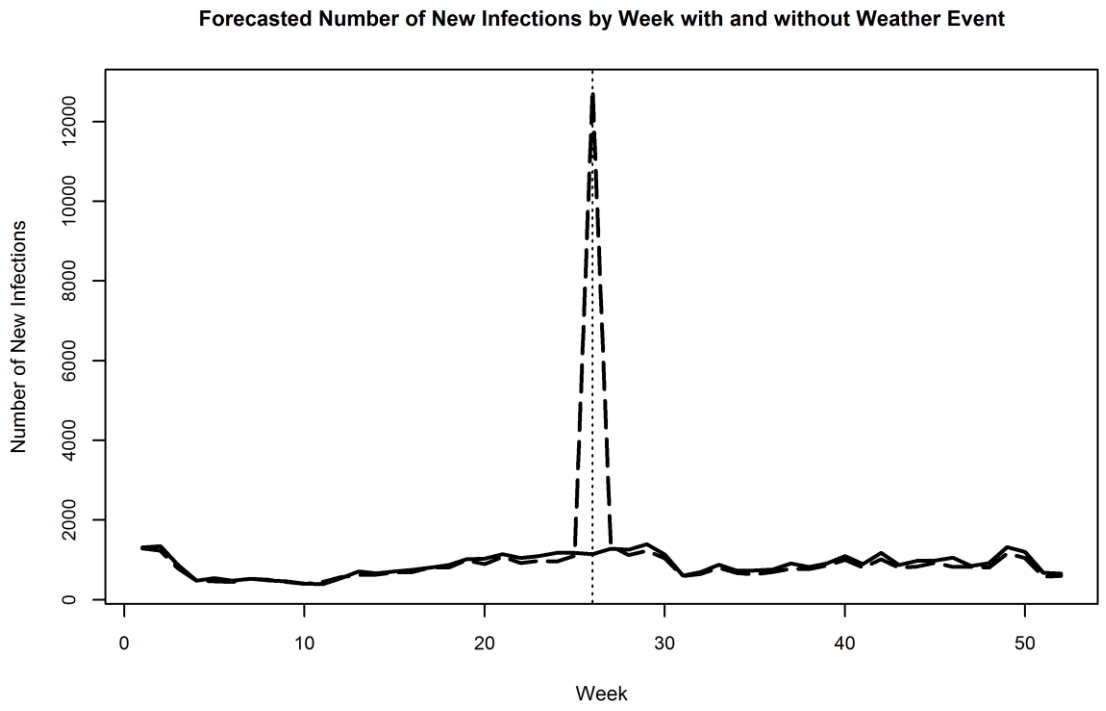

Figure 21. New infections with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event.

The event increased the yearly number of symptomatic cases by 3,363, from 18,783 to 22,146. See Figures 22 and 23. The forecasted number of symptomatic cases increases over a five-week span following the synthetic weather event. During this limited time period of five weeks, the number of symptomatic cases increased by a total of 4,673 cases. Alternatively, if we compare the number of symptomatic cases arising from infection during the two-week period containing the weather event, we find an additional 4,676 additional symptomatic cases. The forecasted number of symptomatic cases two weeks following the event is 2,048, with a forecast interval of 1,734 to 2,361 cases. This represents a 4.5-fold increase over the forecasted number without the event. The spike in cases is broader compared to that seen for new infections because there exists variability in the incubation period. The spike in symptomatic cases is also smaller in magnitude because only approximately 40% of cases are symptomatic.

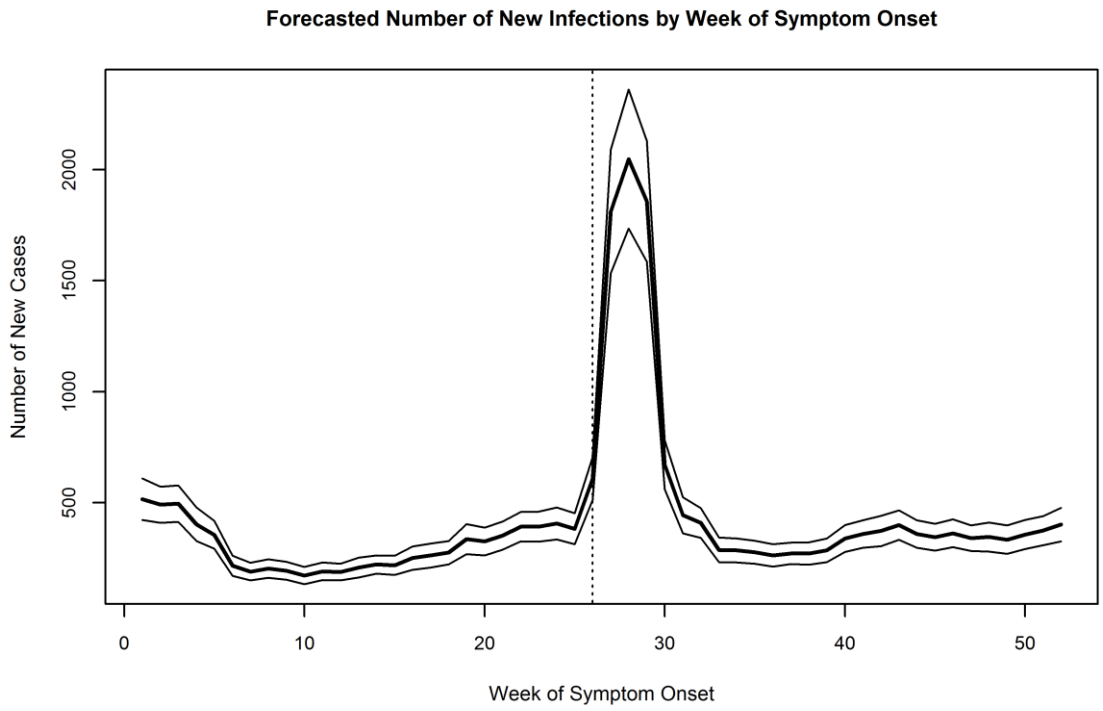

Figure 22. Symptomatic infections by date of symptom onset with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

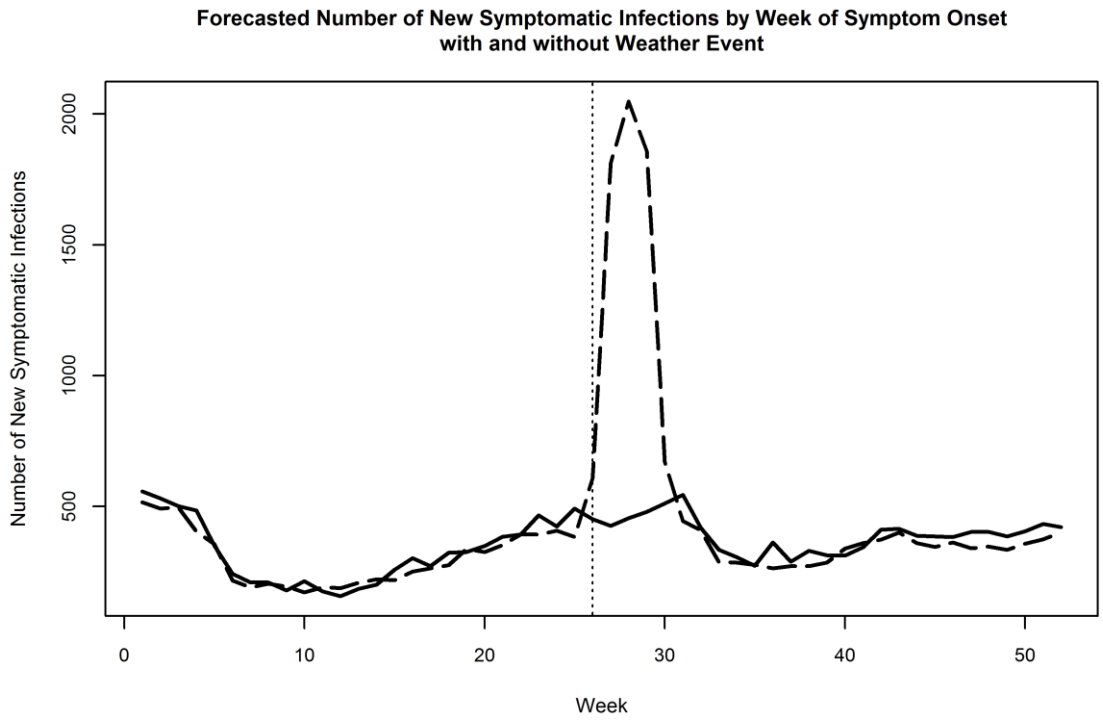

Figure 23. New symptomatic infections with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event.

The synthetic weather event also increased the yearly number of reported cases by 1,253, from 9,580 without the event to 10,833 with the event. See Figures 24 and 25 for the number of reported cases over time with the event and a comparison to the same year without the weather event. Six weeks following the event, the forecasted number of reported cases reached its maximum at 332 with a forecast interval of 277 to 387 cases, an exceedance of 76% over the forecasted number without the weather event. The number of reported cases increases at least 30% in each of third through thirteenth weeks following the event, with a maximum 1.9-fold increase in the number of reported cases 5 weeks following the event. Considering only infections that occurred during the two-week period of the weather event, we find that the weather event caused an additional 2,064 case reports.

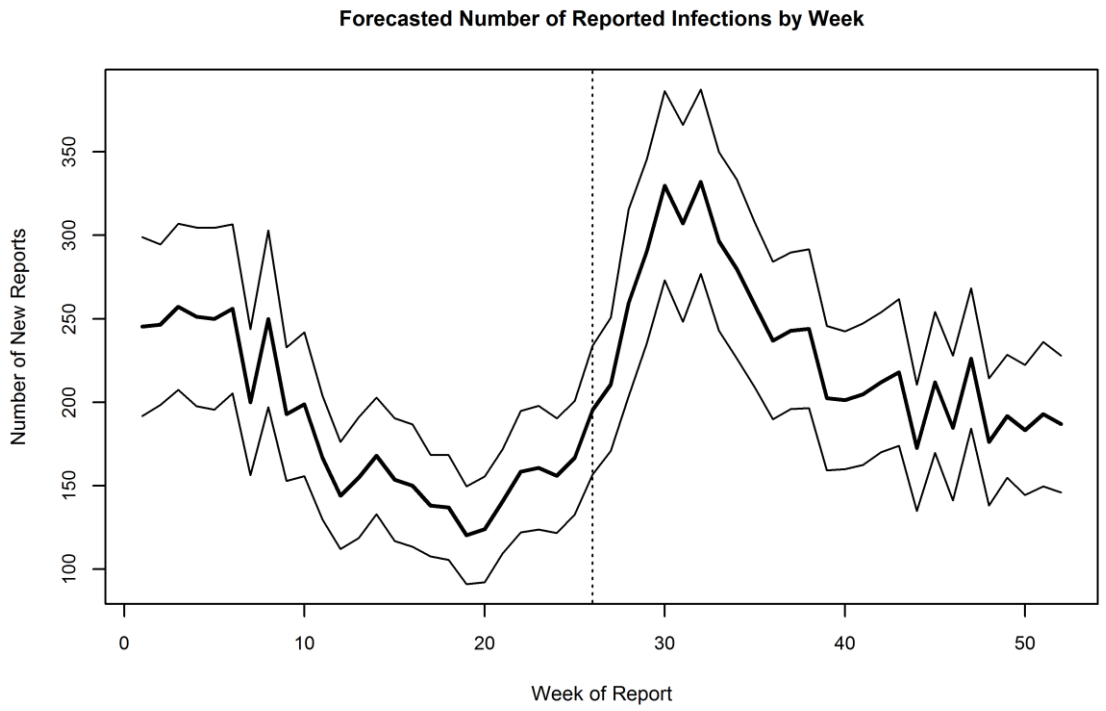

Figure 24. New reported infections by date of report with a synthetic weather event. The dashed line indicates the date of the event. Thick line represents the mean number from ten runs of the model. Thin lines represent 95% forecast intervals.

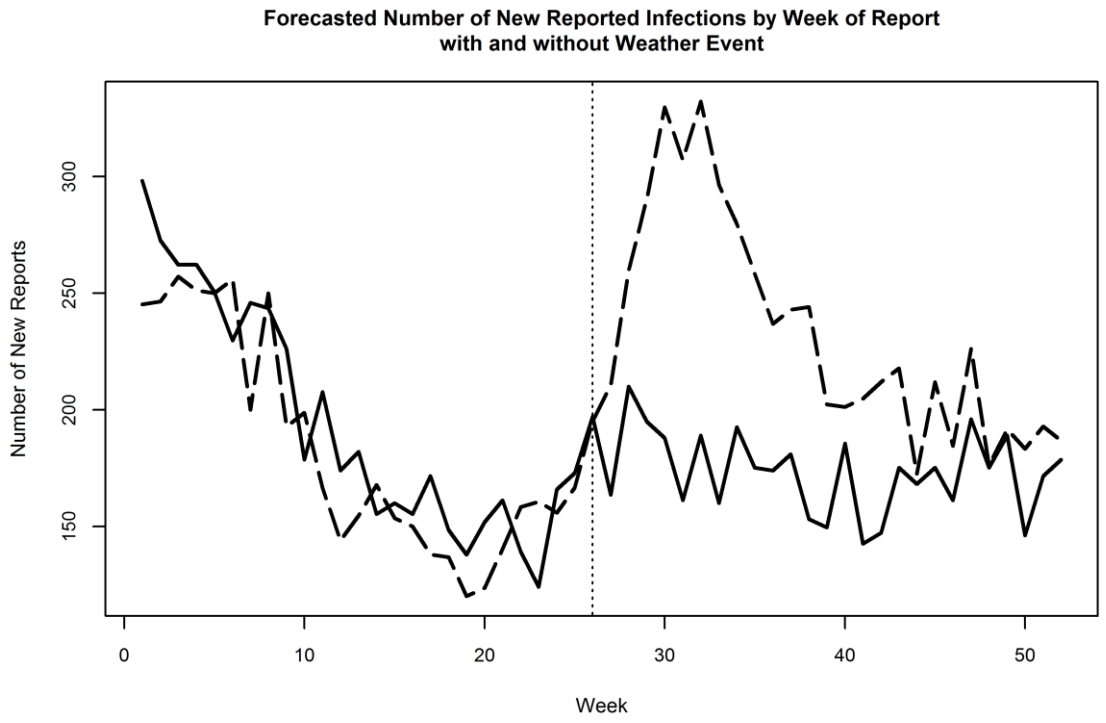

Figure 25. New reported infections by week of report with (dashed line) and without (solid line) a synthetic weather event. The dashed line indicates the date of the event.

4.6 ONLINE TOOL

For this dissertation, we create an epidemiological agent-based model for coccidioidomycosis incidence forecasting. Due to the complexity of the model, our model is implemented in the statistical programming language R (R Core Team 2012). Our software can be used to forecast incidence under various scenarios such as weather events and vaccination strategies. However, as not all persons are familiar with R, we further develop a web-based tool to run our model. Development of this tool is Aim 4 of this dissertation. The online tool will allow epidemiologists and public health researchers to have the ability to forecast coccidioidomycosis incidence by simply uploading files and clicking a simulate button. We use HTML and CSS for

design and display, JavaScript for asynchronous and client-side interactions, and PHP (PHP Group 2013) for dynamic content processing and to interface with R. The directory listing for a virgin install of our web-based tool is shown in Table 10.

Table 10. Directory listing of web-based tool.

Directory Listing	License
/css	
apprise.css	MIT License
cocci.main.css	
mit.license.apprise.txt	
/data	
/default	
/inputfiles	
alpha.birthrate.2009.txt	
beta.birthrate.2009.txt	
env.2009.txt	
env.2010.txt	
input.pop.2009.txt	
isr.txt	
mortrate2009.txt	
p.diabetes.txt	
p.dust.txt	
p.hiv.txt	
p.params.txt	
p.smoke.txt	
p.vacc.txt	
ppi.txt	
prec2009.txt	
prec2010.txt	
temp2009.txt	
temp2010.txt	
batch.R	
model.R	
script.R	

Table 10 (continued).

/images		
	about.gif	
	browse.png	
	content.png	
	download.png	
	favicon.ico	
	footer.png	
	loading.gif	Open Source
	os.license.upload.txt	
	pattern.png	
	pd.license.spherule.txt	
	restart.png	
	simulate.png	
	spherule.jpg	Public Domain
	upload.png	
/js		
	ajaxfileupload.js	Open Source
	apprise-1.5.full.js	MIT License
	cocci.main.js	
	jquery-cookie.js	MIT License
	jquery-1.8.3.min.js	MIT License
	mit.license.apprise.txt	
	mit.license.cookie.txt	
	mit.license.jquery.txt	
	os.license.upload.txt	
config.php		
footer.php		
functions.php		
header.php		
index.php		
noscript.php		
upload.php		

As shown by the license column, there is one CSS file, two image files, and four JavaScript files (all with permissive free software licenses) that are incorporated into our web-based tool. We create one CSS file, ten image files, one JavaScript file, two R files, and seven PHP files for our

web-based tool. One R file and the 18 data files were created for our stochastic epidemiological agent-based model and incorporated into the web-based tool. We also include seven permissive free software licenses (two are duplicates) in our web-based tool for the four JavaScript libraries (two with two files) and one image we incorporate into our tool. Further files and directories will be automatically created by our web-based tool as needed to store content.

5.0 DISCUSSION

We conduct an extensive review and summarization of the literature on coccidioidomycosis. Using published reports and analyses regarding coccidioidomycosis, we construct an agent-based model. This model has been developed using incidence data through 2009 and published analyses of data through 2009. In addition, population, natality, mortality and comorbidity data were used. Analyses that use any incidence data past 2009, even if for other jurisdictions, were excluded for use in model development.

5.1 MODEL VALIDATION

For Aim 2, we assess the validity of the model in two ways. First, we assess the face validity (Objective 2a) by examining whether the outputs from the model match with the relative risks and other characteristics that were used to construct the model. Second, we use the model to forecast incidence data and compare the results to actual counts of reported cases in an independent test of the model (Objective 2b). In our assessment of face validity, we do not expect perfect agreement because we only examine marginal probabilities and marginal relative risks. In addition, differences between the population of Maricopa County in 2010 and the reference population of Kern County, California in 1995-1996 could dramatically affect the marginal probabilities and marginal relative risks.

Most (10/14) of the model inputs and the corresponding estimates from the simulations agreed to within 10%. Four of the parameters exhibited more significant differences: the percent of symptomatic infections with severe pulmonary symptoms (model input was 20.3% versus estimate of 27.2%), and the relative risks of severe pulmonary symptoms for diabetics versus non-diabetics (2.81 versus 3.43), immunocompromised versus immunocompetent (5.26 versus

3.76), and smokers versus non-smokers (2.34 versus 1.91). It is noteworthy that each of these four parameters concern severe pulmonary symptoms. The model inputs were all based on the reference population of Kern County, and the individual probabilities of developing severe pulmonary symptoms were further influenced by individual risk factors such as smoking and comorbidity. The median age for symptomatic individuals in the simulations as 43 years, compared to 34 years in the Kern County patient pool. As age is a risk factor for severe pulmonary symptoms, this difference in age distribution between the two counties may partially account for the difference in the percentage of patients with severe pulmonary symptoms and the relative risks associated with co-morbidity and smoking.

We compare the proportion of actual reported cases in Maricopa County in 2009 that fell into different demographic groups to the proportions of forecasted infections and forecasted reported infections in those same demographic groups. We observe very close agreement between the proportions of reported infections. We also observe some differences between the proportions of forecasted infections and forecasted case reports of certain demographic groups. In particular, in younger individuals we observe that the proportion of forecasted infections was greater than the proportion of forecasted reported infections. Similarly, in older individuals the proportion of forecasted infections was less than the proportion of forecasted case reports. This difference is due to the greater disease severity in older individuals, and in our model individuals with complications are more likely to generate a case report. We also observe a similar effect for African-Americans, whose proportion of infections was less than the proportion of reported infections. As with age, the difference is attributable to the greater risk for African-Americans to develop severe pulmonary symptoms and disseminated disease.

We also examine the temporal trends in total infections, symptomatic infections, and in reported infections. The number of total infections was strongly correlated with PM10 with zero lag as would be expected from the model. We observe a strong correlation between the numbers of cases by date of infection and the numbers of cases by date of symptom onset with a lag of 2 weeks, which corresponds to the typical incubation period for coccidioidomycosis. The number of reported infections for 2010 was consistent with increased levels of dust in the latter half of 2009 and the lack of spikes in reported infections was consistent with little dust activity throughout 2010. We also observe a broad peak between 3 and 9 weeks in the cross-correlation between the date of case report and date of symptom onset. This lagged correlation is consistent

with the median time from symptom onset to case report of 55 days used in the model. We further examine whether the model could correctly forecast future incidence. For nearly all demographic groups, the actual number of reported cases fell within our forecast intervals for those demographic groups. While we are not able to assess the model's ability to forecast incidence across race strata due to a large number of reported cases with unknown race, the agreement across age and gender groups is very strong.

5.2 SENSITIVITY ANALYSIS

Objective 2c asserted that the model would respond predictably to changes in key parameter values. The sensitivity analysis for $p_{S|I}$ showed that the value of $p_{S|I}$ has a large impact on the total number of infections, but does not affect the number of symptomatic infections or the number of reported infections. This result, while a bit counter-intuitive, makes sense in the context of the model. A key input into the model is the probability of reported infections by strata. Given this input, if the probability of symptoms is decreased, then to yield the correct rate of reported infections the number of total infections must be increased. Similarly, if the probability of symptoms is increased, the total number of infections must be decreased to keep the number of reported infections approximately constant within strata. The implication is that the probability of symptoms given an infection does affect estimates of the total number of infections. However, it does not affect estimates of the public health burden from coccidioidomycosis as the number of symptomatic infections remains unaffected.

Similarly, the sensitivity analysis for the probability of report given the four symptom/outcome groups dramatically affects estimates of the total number of infections and the number of symptomatic infections, but does alter estimates of the numbers of reported infections. Again, this counter-intuitive result does make sense in the context of the model. If our estimate for the rate of reporting is increased from our initial base case, but the number of reports remains unchanged then it follows that there must be fewer symptomatic infections. This model behavior is reassuring and suggests that the model is operating as it should.

Furthermore we conduct a sensitivity analysis for the relative risk of infection for individuals in the high versus low dust exposure group. Decreasing the relative risk from 1.64 to

1.34 decreased the total number of infections by approximately 1,000 and the number of reported infections by approximately 200. Increasing the relative risk from 1.64 to 1.92 increased the number of infections by approximately 1,200 and the number of case reports by 135. These changes are not outside of the simulation error bounds from ten model runs, and the direction of the changes makes sense. The implication is that the model results are insensitive to changes in the relative risk for dust exposure group. This is likely due to the low proportion of individuals in the high dust exposure group – approximately 5% in our simulation.

5.3 PUBLIC HEALTH IMPLICATIONS

For Aim 3 of this dissertation, we investigate the public health impact of vaccination strategies, vaccination coverage, and a weather event on forecasted incidence. We first use our model to examine the impact on public health from the availability of a vaccine. We investigate three vaccination strategies: all individuals could receive vaccination, only children could receive vaccination, and only adults could receive vaccination. We find that if a vaccine were available with 75% effectiveness, and all individuals were vaccinated at influenza rates, annual morbidity would decrease by nearly 6,000 cases and annual mortality by 70 cases in Maricopa County alone. Vaccination of adults only would achieve nearly the same benefit.

In contrast, vaccination of children only would reduce annual morbidity by 374 cases and mortality by five. While there would still be positive benefit to such a strategy, vaccination of all individuals would provide a greater reduction in morbidity and mortality. Further, this analysis is only applicable over a relatively short time interval of one year. It is possible that vaccination of children would have a much greater benefit than indicated in this analysis because children would be less likely to have prior exposure to coccidioidal fungi than adults. Our model is limited in that it does not account for date of first exposure to coccidioidomycosis, and this limitation may in particular affect conclusions regarding vaccination of children. In addition, as discussed in Section 3.11.1, we assume that the proportion of protected individuals equals the vaccination rates for influenza. If a single inoculation provides lifelong protection, the proportion of protected individuals might be considerably greater and the above estimates for the reduction in morbidity and mortality would be conservative.

We also examine the public health impact of different vaccination coverage rates. With availability of a vaccine, a public health policy question could arise regarding how actively to promote vaccination. In the case of influenza, a significant effort is put forth to encourage students, workers, and the elderly to be vaccinated. As stated above, we find that annual morbidity would decrease by approximately 6,000 cases and mortality by 70 cases if children and adults were vaccinated at influenza rates. If instead, the odds of individuals receiving vaccination were half those of influenza, morbidity would decrease by approximately 3,800 and mortality by 40 cases. At one-quarter influenza rates, the reduction is approximately 2,400 symptomatic cases and 30 deaths. Hence, even with low vaccination rates, availability of an effective vaccine would provide significant reductions in morbidity and mortality in Maricopa County, Arizona and elsewhere coccidioidomycosis is endemic.

Finally, we examine the public health impact of weather events and develop an online tool. For instance, we find that a moderate-sized weather event such as a dust storm could increase morbidity by nearly 4,700 cases, generate more than 2,000 additional case reports and increase mortality by 42 cases. This suggests that public health researchers might strive to achieve greater awareness of the dangers of exposure to dust during such weather events in hopes of reducing infections. It may also impact physician awareness campaigns and increase physician knowledge of the disease, leading to improved diagnosis and earlier treatment as needed. Further, our online tool can be used by public health officials to examine the impact of various alternative policies and thus improve policy decisions. Specifically, officials can forecast morbidity and mortality across strata under different hypothetical scenarios such as vaccination strategies and use this information in conjunction with costs and other key inputs to guide decision-making. More information about our online tool is given in Section 5.4.

5.4 ONLINE TOOL

We develop an online tool pursuant to Aim 4 of this dissertation. Our web-based tool runs on a Linux server, and we estimate that results of a partial run would be available within the first several hours of the requested analysis at which point results would be emailed to the end user.

Total computation time will depend on several factors including server utilization, population size, and the number of requested simulations. Run time is constrained to one year to ensure the validity of incidence forecasts and the timely processing of the requested simulations. Further, given the reliance on asynchronous JavaScript, if an end user has JavaScript disabled in her browser, she will be prompted to enable it. Should an end user wish to install our web-based tool rather than use it at its hosted location, she would need a Linux server configured like the hosted server, have the ability to FTP files, but no file edits to the web-based tool should be necessary.

5.5 FUTURE POTENTIAL USES

In this dissertation we examine the public health impact of vaccination strategies, vaccination coverage, and weather events. However, our model could also be used to examine the public health impact of other decisions and events. For example, we could use the model to assess the impact of more frequent or mandatory wearing of masks by construction and agricultural workers (Grefenstette and Talbott 2013). If mask-wearing were mandatory, we could use the model to assess the impact by determining the modified relative risk for the high dust exposure group and comparing the model output to that found using the unmodified relative risk. If mask-wearing were not mandatory but instead encouraged via an awareness campaign, we would need to estimate compliance and make revisions to the model to support the additional complexity. The model could also be used to assess the impact of changes in regional employment, such as caused by economic up-turns and down-turns, housing booms and busts, etcetera. This impact could be investigated by altering the proportion of individuals in affected industries.

Additionally, watering down of open lands can reduce the blowing of soil and hence reduce the ability of *Coccidioides* species to spread. While dust control measures are required by the Maricopa County Air Quality Department when it issues permits for soil disruption (Maricopa County Air Quality Department 2013), the model could be used to examine the public health impact of these requirements, or the impact of greater compliance, as well as to examine the effect of greater dust control in the regions surrounding Maricopa County (Lake and Talbott 2013). Of particular concern are tribal communities on the border of Maricopa County that are

sovereign nations and may have less strict dust control requirements than Maricopa County. However to examine this impact, we may need to construct a model for the effect of such changes on dust levels in Maricopa County. Further, while Maricopa County residents tend to go indoors if accessible during dust storms, a public health awareness campaign could be conducted to educate individuals about the importance of being indoors during a dust storm, and the model could be adopted to examine the public health impact of such a campaign (Talbott 2013). While many approaches exist to modeling the impact of such a campaign, one approach would be to reduce the relative risk associated with PM10 whenever particulate matter exceeds some value such as 150 μg/m3 when residents would presumably go indoors if possible.

In addition, the model could also be adapted to examine the public health impact of migration. The present version of the model only considers residents of Maricopa County, as we assume a closed population. However non-residents, in particular the many seasonal visitors who spend their winter months in the southwestern United States, may acquire coccidioidomycosis during their overwinter in endemic areas. Such individuals may become infected in Maricopa County and hence present an additional health care burden in the County, or they may return to their home location with disease that is less likely to be properly diagnosed. Hence the model could be used to assess the effect that seasonal migration has on the public health burden from coccidioidomycosis in Maricopa County and across the United States.

6.0 CONCLUSIONS

In this dissertation we develop an agent-based model of coccidioidomycosis that combines expert knowledge from the literature and real historical data (Aim 1). We assess the face validity of the model by comparing its literature-based inputs to the outputs from the model and find that the model is behaving as specified (Aim 2: Objective 2a). We also use the model to construct independent forecasts of incidence for Maricopa County in 2010 and find that the model correctly forecasts incidence (Aim 2: Objective 2b). For most of the demographic groups that we could examine, the actual number of reported cases fell within our forecast intervals for those demographic groups. From our sensitivity analyses (Aim 2: Objective 2c), we find the number of reported cases is generally insensitive to changes in parameter values because rates of reported cases are typically available and used by our agent-based model.

We use our model to assess the public health impact of vaccination strategies, vaccination coverage, and weather events (Aim 3). These assessments illustrate the utility of such an agent-based model. Using the model we can examine scenarios of public health importance in order to gain insight that can inform and affect decision-making. In order to allow wide access to our model and enable epidemiologists and public health researchers to forecast coccidioidomycosis incidence, we create a web-based tool (Aim 4). Investigators can use this tool to conduct simulation studies for public health scenarios and locales of interest. By making this tool publicly available, we place it in the hands of public health decision-makers so they can better serve their jurisdictions and improve the health and well-being of their residents.

6.1 PUBLIC HEALTH SIGNIFICANCE

There are an estimated 150,000 new cases of coccidioidomycosis in the United States each year. Through the development of an agent-based model, the dissertation provides a novel approach to forecast coccidioidomycosis incidence, morbidity, and mortality. This model is accessible as a web-based tool, bringing customizable forecasts to the fingertips of epidemiologists, researchers, and public health decision-makers. Such forecasts support assessment of the present and future public health impact of coccidioidomycosis and can be used to guide decisions on future vaccination strategies and education campaigns.

BIBLIOGRAPHY

Adam, R. D., S. P. Elliott, et al. (2009). "The Spectrum and Presentation of Disseminated Coccidioidomycosis." Am J Med **122**(8): 770-777.

Ajello, L. (1967). "Comparative Ecology of Respiratory Mycotic Disease Agents." Bacteriol Rev **31**(1): 6–24.

Ajello, L., K. Maddy, et al. (1965). "Recovery of Coccidioides immitis from the air." Med Mycol **4**(2): 92-95.

American Red Cross. (2012). "Blood Types." Retrieved 23 October 2012, from http://www.redcrossblood.org/learn-about-blood/blood-types.

Ampel, N. (2000). Coccidioidomycosis. Fungal Diseases of the Lung. G. Sarosi and S. Davies. Philadelphia, Lippincott Williams and Wilkins**:** 59-77.

Arizona Department of Health Services. (2007). "Valley Fever Annual Report 2007." Retrieved 22 October 2012, from http://www.azdhs.gov/phs/oids/epi/disease/cocci/documents/2007-valley-fever-report.pdf.

Arizona Department of Health Services. (2009). "Coccidioidomycosis (Valley Fever) Lab Tests." Retrieved 27 October 2012, from http://www.azdhs.gov/phs/oids/epi/disease/cocci/cocci_lab.htm.

Arizona Department of Health Services. (2012). "Arizona Health Status and Vital Statistics Annual Reports,." Retrieved 22 October 2012, from http://www.azdhs.gov/plan/report/ahs/.

Arora, N. P., V. Taneja, et al. (2012). "Coccidioidomycosis masquerading as malignancy." BMJ Case Rep **25**(10).

Baptista-rosas, R., J. Catalán-dibene, et al. (2012). "Molecular detection of Coccidioides spp. from environmental samples in Baja California: linking Valley Fever to soil and climate conditions." Fungal Ecol **5**(2): 177–190.

Barker, B., J. Tabor, et al. (2012). "Detection and phylogenetic analysis of Coccidioides posadasii in Arizona soil samples." Fungal Ecol **5**(2): 163–176.

Barnato, A. E., G. D. Sanders, et al. (2001). "Cost-effectiveness of a potential vaccine for Coccidioides immitis." Emerg Infect Dis **7**(5): 797-806.

Barrera, C. (1986). "Formation and germination of fungal arthroconidia." Crit Rev Microbiol **12**(4): 271-292.

Benedum, C. and C. A. Tsang. (2012). "Coccidioidomycosis Surveillane in Arizona: A Comparison of 2007 and 2011 Data." Retrieved 27 October 2012, from http://www.azdhs.gov/phs/oids/training/documents/2012/BenedumCorey.pdf.

Bercovitch, R. S., A. Catanzaro, et al. (2011). "Coccidioidomycosis during pregnancy: a review and recommendations for management." Clin Infect Dis **53**(4): 363-368.

Blair, J. E. (2007). "Coccidioidomycosis in Patients Who Have Undergone Transplantation." Ann N Y Acad Sci **1111**(1): 365-376.
Blair, J. E. (2007). "State-of-the-Art Treatment of Coccidioidomycosis." Ann N Y Acad Sci **1111**(1): 411-421.
Blair, J. E., B. Coakley, et al. (2006). "Serologic testing for symptomatic coccidioidomycosis in immunocompetent and immunosuppressed hosts." Mycopathologia **162**(5): 317-324.
Blair, J. E. and J. T. Currier (2008). "Significance of isolated positive IgM serologic results by enzyme immunoassay for coccidioidomycosis." Mycopathologia **166**(2): 77-82.
Blair, J. E. and J. L. Logan (2001). "Coccidioidomycosis in solid organ transplantation." Clin Infect Dis **33**(9): 1536-1544.
Braddy, C. M., R. L. Heilman, et al. (2006). "Coccidioidomycosis after renal transplantation in an endemic area." Am J Transplant **6**(2): 340-345.
Burwell, L. A., B. J. Park, et al. (2009). "Outcomes among inmates treated for coccidioidomycosis at a correctional institution during a community outbreak, Kern County, California, 2004." Clin Infect Dis **49**(11): e113-119.
Cairns, L., D. Blythe, et al. (2000). "Outbreak of Coccidioidomycosis in Washington State Residents Returning from Mexico." Clin Infect Dis **30**(1): 61-64.
Caldwell, J., G. Welch, et al. (1996). The economic impact of coccidioidomycosis in Kern County, California, 1991-1993. Proceedings of the Fifth International Conference on Coccidioidomycosis. H. Einstein and A. Catanzaro. Washington, DC, National Foundation for Infectious Diseases. **88-97**.
Caldwell, J. W., E. L. Arsura, et al. (2000). "Coccidioidomycosis in pregnancy during an epidemic in California." Obstet Gynecol **95**(2): 236-239.
Campbell, C. C. (1980). "(Philosopical) review of air currents as a continuing vector." Ann N Y Acad Sci **353**: 123-139.
Centers for Disease Control and Prevention (1996). "Coccidioidomycosis, Arizona, 1990-1995." Morb Mortal Wkly Rep **45**(49): 1069-1073.
Centers for Disease Control and Prevention. (2010). "Detailed Data on Diagnosed Diabetes." Retrieved March 3, 2013, from http://www.cdc.gov/diabetes/statistics/prev/national/tprevwfage.htm.
Centers for Disease Control and Prevention (2010). HIV Surveillance Report. **17**.
Centers for Disease Control and Prevention. (2011). "Final state-level influenza vaccination coverage estimates for the 2010–11 season–United States, National Immunization Survey and Behavioral Risk Factor Surveillance System, August 2010 through May 2011." Retrieved March 3, 2013, from http://www.cdc.gov/flu/professionals/vaccination/coverage_1011estimates.htm#Table1.
Centers for Disease Control and Prevention. (2011). "National Notifiable Diseases Surveillance System (NNDSS)." Retrieved 27 October 2012, from http://wwwn.cdc.gov/NNDSS/beta/bconditionsummary.aspx?CondID=37.
Centers for Disease Control and Prevention (2011). "Summary of Notifiable Diseases, United States, 2009." Morb Mortal Wkly Rep **58**(53): 1-100.
Centers for Disease Control and Prevention (2012). "Current Cigarette Smoking Among Adults - United States 2011." Morb Mortal Wkly Rep **61**(44): 889-894.
Centers for Disease Control and Prevention. (2012). "Youth and Tobacco Use." Retrieved March 3, 2013, from

http://www.cdc.gov/tobacco/data_statistics/fact_sheets/youth_data/tobacco_use/index.htm#estimates.
Chang, D. C., S. Anderson, et al. (2008). "Testing for Coccidioidomycosis among Patients with Community-Acquired Pneumonia." Emerg Infect Dis **14**(7): 1053-1105.
Charlton, V., K. Ramsdell, et al. (1999). "Intrauterine transmission of coccidioidomycosis." Pediatr Infect Dis J **18**(6): 561-563.
Chiller, T. M., J. N. Galgiani, et al. (2003). "Coccidioidomycosis." Infect Dis Clin North Am **17**(1).
Cohen, I. M., J. N. Galgiani, et al. (1982). "Coccidioidomycosis in renal replacement therapy." Arch Intern Med **142**(3): 489-494.
Cole, G. T., J. M. Xue, et al. (2004). "A vaccine against coccidioidomycosis is justified and attainable." Med Mycol **42**(3): 189-216.
College of Veterinary Medicine. (2010). "Coccidioidomycosis." Retrieved 09 October, 2012, from http://www.cfsph.iastate.edu/Factsheets/pdfs/coccidioidomycosis.pdf.
Comrie, A. C. (2005). "Climate Factors Influencing Coccidioidomycosis Seasonality and Outbreaks." Environ Health Perspect **113**(6).
Comrie, A. C. and M. F. Glueck (2007). "Assessment of climate-coccidioidomycosis model: model sensitivity for assessing climatologic effects on the risk of acquiring coccidioidomycosis." Ann N Y Acad Sci **1111**: 83-95.
Converse, J. L. and R. E. Reed (1966). "Experimental epidemiology of coccidioidomycosis." Bacteriol Rev **30**(3): 678-695.
Cox, R. A. and D. M. Magee (2004). "Coccidioidomycosis: host response and vaccine development." Clin Microbiol Rev **17**(4): 804-839, table of contents.
Crum, N., C. Lamb, et al. (2002). "Coccidioidomycosis outbreak among United States Navy SEALs training in a Coccidioides immitis-endemic area-Coalinga, California." J Infect Dis **186**(6): 865-868.
Crum, N. F. and G. Ballon-Landa (2006). "Coccidioidomycosis in pregnancy: case report and review of the literature." Am J Med **119**(11): 993 e911-997.
Crum, N. F., E. R. Lederman, et al. (2004). "Coccidioidomycosis: a descriptive survey of a reemerging disease. Clinical characteristics and current controversies." Medicine (Baltimore) **83**(3): 149-175.
Cummings, K. C., A. McDowell, et al. (2010). "Point-source outbreak of coccidioidomycosis in construction workers." Epidemiol Infect **138**(4): 507-511.
Deresinski, S. C., D. Pappagianis, et al. (1979). "Association of ABO blood group and outcome of coccidioidal infection." Sabouraudia **17**(3): 261-264.
Dickson, E. C. (1937). ""Valley Fever" of the San Joaquin Valley and Fungus Coccidioides." Cal West Med **47**(3): 151–155.
Dodge, R. R., M. D. Lebowitz, et al. (1985). "Estimates of C. immitis infection by skin test reactivity in an endemic community." Am J Public Health **75**(8): 863-865.
Dykes, T., A. Stone, et al. (2005). "Coccidioidomycosis of the Epididymis and Testis." Am J Roentgenol **184**(2): 552-553.
Einstein, H. E. and R. H. Johnson (1993). "Coccidioidomycosis: New Aspects of Epidemiology and Therapy." Clin Infect Dis **16**(3): 349-356.
Elconin, A., R. Egeberg, et al. (1964). "Significance of Soil Salinity on the Ecology of Coccidioides immitis." J Bacteriol **87**(3): 500–503.

Espinel-Ingroff, A. V. (2003). Medical Mycology in the United States : A historical analysis (1894-1996). Dordrecht, The Netherlands, Kluwer Academic Publishers.
Farness, O. (1941). "Coccidioidomycosis." J Am Med Assoc **116**(16): 1749-1752.
Fienberg, S. (1970). "An Iterative Procedure for Estimation in Contingency Tables." Ann Math Statist **41**(3): 907-917.
Fish, D. G., N. M. Ampel, et al. (1990). "Coccidioidomycosis during human immunodeficiency virus infection. A review of 77 patients." Medicine (Baltimore) **69**(6): 384-391.
Fisher, F. S., M. W. Bultman, et al. (2007). "Coccidioides Niches and Habitat Parameters in the Southwestern United States." Ann N Y Acad Sci **1111**(1): 47-72.
Fisher, M., G. Koenig, et al. (2002). "Molecular and phenotypic description of Coccidioides posadasii sp. nov., previously recognized as the non-California population of Coccidioides immitis." Mycologia **94**(1): 73-84.
Flaherman, V., R. Hector, et al. (2007). "Estimating Severe Coccidioidomycosis in California." Emerg Infect Dis **13**(7): 1087-1090.
Flynn, N. M., P. D. Hoeprich, et al. (1979). "An Unusual Outbreak of Windborne Coccidioidomycosis." N Engl J Med **301**(7): 358-361.
Foote, A. (2011). "Valley Fever infections increase." Retrieved 16 Novemeber 2012, from https://www.vfce.arizona.edu/resources/inthenews/Valley-Fever-infections-increase.pdf.
Fresno County Community Health Division. (2011). "Epidemiology of Coccidioidomycosis in Six California Counties." Retrieved 22 October 2012, from http://www.co.fresno.ca.us/DivisionPage.aspx?id=49446.
Gaidici, A. and M. Saubolle (2009). "Transmission of Coccidioidomycosis to a Human via a Cat Bite." J Clin Microbiol **47**(2): 505–506.
Galgiani, J. N. (1993). "Coccidioidomycosis." West J Med **159**(2): 153-171.
Galgiani, J. N. (2012). "Coccidioidomycosis in Arizona Counties Relation to 2011 Dust Storms." Retrieved March 3, 2013, from https://www.vfce.arizona.edu/resources/CocciStudyGroup/CSG2012Presenterspdf%27s/Coccidioidomycosisin-Arizona-Counties-Relation-to-2011-Dust-Storms.pdf.
Galgiani, J. N., N. M. Ampel, et al. (2005). "Coccidioidomycosis." Clin Infect Dis **41**(9): 1217-1223.
Gehlbach, S., J. Hamilton, et al. (1973). "Coccidioidomycosis-an occupational disease in cotton-mill workers." Arch Intern Med **131**(2): 254-255.
Goegebuer, T., K. Nackaerts, et al. (2009). "Coccidioidomycosis: an unexpected diagnosis in a patient with persistent cough." Acta Clin Belg **64**(3): 235-238.
Grefenstette, J. J. and E. O. Talbott (2013). Personal Correspondence.
Hawksworth, D. (1991). "The fungal dimension of biodiversity: magnitude, significance, and conservation." Mycol Res **95**: 641-655.
Hawksworth, D. (2001). "The magnitude of fungal diversity: The 1.5 million species estimate revisited." Mycol Res **105**: 1422-1432.
Hector, R. (2011). "Fighting Coccidioidomycosis Prevention: Vaccine Treatment: Nikkomycin Z." Retrieved 27 October 2012, from http://www.countyofkings.com/health/forms/Cocci_Hector2.ppt.
Hector, R. F., G. W. Rutherford, et al. (2011). "The Public Health Impact of Coccidioidomycosis in Arizona and California." Int J Environ Res Public Health **8**(4): 1150-1173.

Hector, R. F., B. L. Zimmer, et al. (1990). "Evaluation of nikkomycins X and Z in murine models of coccidioidomycosis, histoplasmosis, and blastomycosis." Antimicrob Agents Chemother **34**(4): 587-593.

Hooper, J. E., Q. Lu, et al. (2007). "Disseminated coccidioidomycosis in pregnancy." Arch Pathol Lab Med **131**(4): 652-655.

Huang, J., B. Bristow, et al. (2012). "Coccidioidomycosis-associated deaths, United States, 1990–2008." Emerg Infect Dis **18**(11): 1723–1728.

Irfan, U. (2012). "Valley Fever on the Rise in U.S. Southwest, with Links to Climate Change " Retrieved 09 October, 2012, from http://www.scientificamerican.com/article.cfm?id=valley-fever-on-the-rise-in-us-southwest.

Jaroszewski, D. E., W. J. Halabi, et al. (2009). "Surgery for Pulmonary Coccidioidomycosis: A 10-Year Experience." Ann Thorac Surg **88**(6): 1765-1772.

Kamei, K., A. Sano, et al. (2003). "The trend of imported mycoses in Japan." J Infect Chemother **9**(1): 16-20.

Kirkland, T. and J. Fierer (1996). "Coccidioidomycosis: a reemerging infectious disease." Emerg Infect Dis **3**(2): 192-199.

Klotz, S., D. Drutz, et al. (1984). "The critical role of CO2 in the morphogenesis of Coccidioides immitis in cell-free subcutaneous chambers." J Infect Dis **150**(1): 127-134.

Kolivras, K., P. Johnson, et al. (2001). "Environmental variability and coccidioidomycosis (valley fever)." Aerobiologia **17**: 31-42.

Kuberski, T., J. Herrig, et al. (2010). "False-positive IgM serology in coccidioidomycosis." J Clin Microbiol **48**(6): 2047-2049.

Kumar, K., A. Narasimhan, et al. (2011). "Coccidioidomycosis in Chennai." J Assoc Physicians India **59**: 122-124.

Lacy, G. and F. Swatek (1974). "Soil Ecology of Coccidioides immitis at Amerindian Middens in California." Appl Microbiol **27**(2): 379-388.

Lake, D. F. (2012). Personal Correspondence.

Lake, D. F. and E. O. Talbott (2013). Personal Correspondence.

Larone, D. (1995). Medically Important Fungi - A Guide to Identification, 3rd ed. Washington, D.C., ASM Press.

Leake, J. A. D., D. G. Mosley, et al. (2000). "Risk Factors for Acute Symptomatic Coccidioidomycosis among Elderly Persons in Arizona, 1996–1997." J Infect Dis **181**(4): 1435-1440.

Lecara, G., R. A. Cox, et al. (1983). "Coccidioides immitis vaccine: potential of an alkali-soluble, water-soluble cell wall antigen." Infect Immun **39**(1): 473-475.

Lee, C. H., L. Wilcox, et al. (2008). "Coccidioides immitis: two cases of misidentified mycosis." Can Respir J **15**(7): 377-379.

Levine, H. B., J. M. Cobb, et al. (1961). "Immunogenicity of spherule-endospore vaccines of Coccidioides immitis for mice." J Immunol **87**: 218-227.

Levine, H. B. and C. E. Smith (1967). The reaction of eight volunteers injected with Coccidioides immitis spherule vaccine: First human trials. Proceedings of the Second Coccidioidomycosis Symposium. L. Ajello. Tucson, AZ, University of Arizona Press**:** 197-200.

Linsangan, L. C. and L. A. Ross (1999). "Coccidioides immitis infection of the neonate: two routes of infection." Pediatr Infect Dis J **18**(2): 171-173.

Logan, J. L., J. E. Blair, et al. (2001). "Coccidioidomycosis complicating solid organ transplantation." Semin Respir Infect **16**(4): 251-256.

Louie, L., S. Ng, et al. (1999). "Influence of host genetics on the severity of coccidioidomycosis." Emerg Infect Dis **5**(5): 672-680.

Madden, L. and M. Wheelis (2003). "The threat of plant pathogens as weapons against U.S. crops." Annu Rev Phytopathol **41**: 155-176.

Maddy, K. and H. Crecelius (1967). Establishment of Coccidioides immitis in negative soil following burial of infected animals and animal tissues. Papers from the Second Symposium on Coccidioidimycosis. L. Ajello. Tucson, AZ, University of Arizona Press**:** 309-312.

Maddy, K. T., H. G. Crecelius, et al. (1960). "Distribution of Coccidiodes immitis determined by testing cattle." Public Health Rep **75**: 955-962.

Maricopa County Air Quality Department. (2013). "Dust Sources: Dust Control." Retrieved 03 April 2013, from http://www.maricopa.gov/aq/divisions/compliance/dust/dust_sources/.

Maricopa County Department of Public Health, Office of Epidemiology. (2011). Maricopa County Arizona Health Status Report 2009. Phoenix, Arizona. http://www.maricopa.gov/publichealth/Services/EPI/pdf/hsr/2009%20Countywide%20Report.pdf.

Maricopa County Department of Public Health, Office of Epidemiology. (2012). Maricopa County Arizona Health Status Report 2010. Phoenix, Arizona. http://www.maricopa.gov/publichealth/Services/EPI/pdf/hsr/2010-Countywide-HSR.pdf.

Mayo Clinic. (2006). "Building Mayo: Mayo Clinic Scottsdale/Phoenix, Ariz." Retrieved 22 October 2012, from http://www.mayoclinic.org/tradition-heritage/arizona-campus.html.

McGinnis, M. (1991). Introduction to mycology. Medical Microbiology. S. Baron. New York, Churchill Livingstone**:** 921–933.

Mendel, E., E. N. Milefchik, et al. (1994). "Coccidioidomycosis brain abscess. Case report." J Neurosurg **81**(4): 614-616.

Merriam, C. H. (1898). Life Zones and Crop Zones of the United States. Washington, D.C., U.S. Department of Agriculture, Division of Biological Survey. **Bulletin 10**.

Miller, M., R. Hendren, et al. (2004). "Posttransplantation Disseminated Coccidioidomycosis Acquired from Donor Lungs." J Clin Microbiol **42**(5): 2347–2349.

Morrow, W. (2006). "Holocene coccidioidomycosis: Valley Fever in early Holocene bison (Bison antiquus)." Mycologia **98**(5): 669-677.

National Climatic Data Center. (2012). "Global Summary of Day,." Retrieved 21 October 2012, from http://www7.ncdc.noaa.gov/CDO/cdoselect.cmd?datasetabbv=GSOD.

Olivere, J. W., P. A. Meier, et al. (1999). "Coccidioidomycosis--the airborne assault continues: an unusual presentation with a review of the history, epidemiology, and military relevance." Aviat Space Environ Med **70**(8): 790-796.

Osaki, T., H. Morishita, et al. (2005). "Pulmonary Coccidioidomycosis That Formed a Fungus Ball with 8-years Duration." Intern Med **44**(2): 141-144.

Pappagianis, D. (1993). "Evaluation of the protective efficacy of the killed Coccidioides immitis spherule vaccine in humans. The Valley Fever Vaccine Study Group." Am Rev Respir Dis **148**(3): 656-660.

Pappagianis, D. (2007). "Coccidioidomycosis in California state correctional institutions." Ann N Y Acad Sci **1111**: 103-111.

Pappagianis, D. and H. Einstein (1978). "Tempest from Tehachapi takes toll or Coccidioides conveyed aloft and afar." West J Med **129**: 527-530.
Pappagianis, D., R. Hector, et al. (1979). "Immunization of mice against coccidioidomycosis with a subcellular vaccine." Infect Immun **25**(1): 440-445.
Pappagianis, D., H. B. Levine, et al. (1967). Further studies on vaccination of human volunteers with killed Coccidioides immitis. Proceedings of the Second Coccidioidomycosis Symposium. L. Ajello. Tucson, AZ, University of Arizona Press**:** 201-210.
Pappagianis, D., R. K. Sun, et al. (1993). "Coccidioidomycosis - United-States, 1991-1992." Morb Mortal Wkly Rep **42**(2): 21-24.
Pappagianis, D. and B. L. Zimmer (1990). "Serology of coccidioidomycosis." Clin Microbiol Rev **3**(3): 247-268.
Parish, J. M. and J. E. Blair (2008). "Coccidioidomycosis." Mayo Clin Proc **83**(3): 343-349.
Park, B. J., K. Sigel, et al. (2005). "An Epidemic of Coccidioidomycosis in Arizona Associated with Climatic Changes, 1998–2001." J Infect Dis **191**(11): 1981-1987.
Petersen, L. R., S. L. Marshall, et al. (2004). "Coccidioidomycosis among workers at an archeological site, northeastern Utah." Emerg Infect Dis **10**(4): 637-642.
Petrini, B., C. M. Skold, et al. (2003). "Coccidioidomycosis mimicking lung cancer." Respiration **70**(6): 651-654.
PHP Group, The. (2013). "PHP: Hypertext Preprocessor." from http://www.php.net/.
Powell, B. L., D. J. Drutz, et al. (1983). "Relationship of progesterone- and estradiol-binding proteins in Coccidioides immitis to coccidioidal dissemination in pregnancy." Infect Immun **40**(2): 478-485.
R Core Team. (2012). "R: A language and environment for statistical computing." from http://www.R-project.org/.
Rempe, S., M. S. Sachdev, et al. (2007). "Coccidioides immitis fungemia: clinical features and survival in 33 adult patients." Heart Lung **36**(1): 64-71.
Rosenstein, N. E., K. W. Emery, et al. (2001). "Risk Factors for Severe Pulmonary and Disseminated Coccidioidomycosis: Kern County, California, 1995–1996." Clin Infect Dis **32**(5): 708-714.
Ruddy, B. E., A. P. Mayer, et al. (2011). "Coccidioidomycosis in African Americans." Mayo Clin Proc **86**(1): 63-69.
Ruggles, D. (2008). "Testicular Coccidioidomycosis." Urol Nurs **28**(2): 113-114.
Rutherford, G. and M. Barrett (1996). "Epidemiology and control of coccidioidomycosis in California." West J Med **165**(4): 221–222.
Saubolle, M. (1996). Life cycle and epidemiology of Coccidioides immitis. Proceedings of the Fifth International Conference on Coccidioidomycosis. H. Einstein and A. Catanzaro. Washington, D.C., National Foundation for Infectious Diseases**:** 1-9.
Saubolle, M. (2000). Mycology and the Clinical Laboratory in the Diagnosis of Respiratory Mycoses. Fungal Diseases of the Lung. G. Sarosi and S. Davies. Philadelphia, Lippincott Williams and Wilkins**:** 1-16.
Saubolle, M., P. McKellar, et al. (2007). "Epidemiologic, Clinical, and Diagnostic Aspects of Coccidioidomycosis." J Clin Microbiol **45**(1): 26-30.
Schmelzer, L. L. and I. R. Tabershaw (1968). "Exposure factors in occupational coccidioidomycosis." Am J Public Health Nations Health **58**(1): 107-113.
Schneider, E., R. A. Hajjeh, et al. (1997). "A Coccidioidomycosis Outbreak Following the Northridge, Calif, Earthquake." JAMA **277**(11): 904-908.

Sharpton, T. J., J. E. Stajich, et al. (2009). "Comparative genomic analyses of the human fungal pathogens Coccidioides and their relatives." Genome Res **19**(10): 1722-1731.
Shubitz, L. (2007). "Comparative Aspects of Coccidioidomycosis in Animals and Humans." Ann N Y Acad Sci **1111**: 395–403.
Sievers, M. L. (1974). "Disseminated coccidioidomycosis among southwestern American Indians." Am Rev Respir Dis **109**(6): 602-612.
Smith, C. E., R. R. Beard, et al. (1946). "Effect of season and dust control on coccidioidomycosis." J Am Med Assoc **132**(14): 833-838.
Sorensen, R. H. (1964). "Survival characteristics of mycelia and spherules of Coccidioides immitis in a simulated natural environment." Am J Hyg **80**(3): 275 - 285.
Spinello, I. M., R. H. Johnson, et al. (2007). "Coccidioidomycosis and pregnancy: a review." Ann N Y Acad Sci **1111**: 358-364.
Sunenshine, R. H., S. Anderson, et al. (2007). "Public Health Surveillance for Coccidioidomycosis in Arizona." Ann N Y Acad Sci **1111**(1): 96-102.
Tabor, J. A. and M. K. O'Rourke (2010). "A risk factor study of coccidioidomycosis by controlling differential misclassifications of exposure and susceptibility using a landscape ecology approach." Sci Total Environ **408**(10): 2199-2207.
Talbott, E. O. (2013). Personal Correspondence.
Tamerius, J. D. and A. C. Comrie (2011). "Coccidioidomycosis Incidence in Arizona Predicted by Seasonal Precipitation." PLoS ONE **6**(6): e21009.
Thompson, G. R. (2011). "Strategies to Combat Coccidioidomycosis: Are We Making Any Progress?" Curr Fungal Infect Rep **5**(4): 215-223.
Tripathy, U., G. L. Yung, et al. (2002). "Donor transfer of pulmonary coccidioidomycosis in lung transplantation." Ann Thorac Surg **73**(1): 306-308.
Tsang, C. A., S. M. Anderson, et al. (2010). "Enhanced Surveillance of Coccidioidomycosis, Arizona, USA, 2007–2008." Emerg Infect Dis **16**(11): 1738-1744.
Ugurlu, S., A. G. de Alba Campomanes, et al. (2005). "Coccidioidomycosis of the eyelid." Ophthal Plast Reconstr Surg **21**(2): 157-159.
United States Bureau of Labor Statistics. (2009). "Employed persons by occupation, race, Hispanic or Latino ethnicity, and sex." Retrieved 23 October 2012, from ftp://ftp.bls.gov/pub/special.requests/lf/aa2009/pdf/cpsaat10.pdf.
United States Bureau of Labor Statistics. (2009). "Employed persons in nonagricultural industries by age, sex, race, and Hispanic or Latino ethnicity." Retrieved 23 October 2012, from ftp://ftp.bls.gov/pub/special.requests/lf/aa2009/pdf/cpsaat14.pdf.
United States Census Bureau. (2010). "Arizona QuickLinks." Retrieved 23 October 2012, from http://quickfacts.census.gov/qfd/states/04000lk.html.
United States Environmental Protection Agency. (2012). "AirData." Retrieved 21 October 2012, from https://ofmext.epa.gov/AQDMRS/aqdmrs.html.
Valley Fever Center for Excellence. (2008). "Current Activities Of The Valley Fever Center." Retrieved 27 October 2012, from https://www.vfce.arizona.edu/resources/pdf/CurrentActivitiesOfTheValleyFeverCenter.pdf.
Valley Fever Center for Excellence. (2010). "Valley Fever (Coccidioidomycosis) " Retrieved 23 October 2012, from https://www.vfce.arizona.edu/valleyfeverinpeople/faqs.aspx.

Valley Fever Center for Excellence. (2010). "Valley Fever in Other Animal Species " Retrieved 08 October, 2012, from https://www.vfce.arizona.edu/ValleyFeverInPets/VFID-other.aspx.

Valley Fever Center for Excellence. (2012). "Canine Nikkomycin Z Study Results." Retrieved 27 October 2012, from https://www.vfce.arizona.edu/resources/pdf/Lay_Summary_Canine_Nik_Z_Study_Results.pdf.

Valley Fever Center for Excellence. (2012). "The Search for the Cure for Valley Fever Nikkomycin Z Development at the University of Arizona." Retrieved 27 October 2012, from https://www.vfce.arizona.edu/resources/pdf/Bio5%20summary%20NikZ%20development%20plan%20March%202012.pdf.

Vikram, H. R. and J. E. Blair (2009). "Coccidioidomycosis in transplant recipients: a primer for clinicians in nonendemic areas." Curr Opin Organ Transplant **14**(6): 606-612.

Wang, C. Y., J. S. Jerng, et al. (2005). "Disseminated coccidioidomycosis." Emerg Infect Dis **11**(1): 177-179.

Wang, Z.-y., S.-l. Wen, et al. (2011). "A case study of imported pulmonary coccidioidomycosis." J Zhejiang Univ Sci B **12**(4): 298–302.

Werner, S. B. and D. Pappagianis (1973). "Coccidioidomycosis in Northern California—An Outbreak among Archeology Students near Red Bluff." Calif Med **119**(3): 16-20.

Werner, S. B., D. Pappagianis, et al. (1972). "An Epidemic of Coccidioidomycosis among Archeology Students in Northern California." N Engl J Med **286**(10): 507-512.

Wilson, J., C. Smith, et al. (1953). "Primary cutaneous coccidioidomycosis; the criteria for diagnosis and a report of a case." Calif Med **79**(3): 233-239.

Wright, P. W., D. Pappagianis, et al. (2003). "Donor-related coccidioidomycosis in organ transplant recipients." Clin Infect Dis **37**(9): 1265-1269.