Determining variable contagiousness of MRSA by setting

A thesis submitted to the University of Arizona College of Medicine --Phoenix
in partial fulfillment of the requirements for the Degree of Doctor of Medicine

Joshua Routh
Class of 2013

Mentor: Sarada Panchanathan, MD, MS
Acknowledgements

Over the last four years I have received support from several wonderful people. Dr. Sarada Panchanathan has been the most supportive and encouraging mentor imaginable. Her guidance turned what could have been a chore into a delightful and enriching experience. Dr. Gerardo Chowell has been instrumental in advising me throughout the course of this study.

I was grateful to receive funding from the Arizona Biomedical Research Commission to conduct this study.
Abstract

Objective and Hypothesis

Methicillin resistant Staphylococcus *aureus* (MRSA) is currently a major cause of skin and soft tissue infections (SSTI) in the United States. In order to characterize the spread of MRSA in the pediatric population we built a probabilistic, discrete-event, individual-based simulation. Specifically, our model looked at the spread of MRSA in households and at schools to determine if there was a difference in communicability between the two settings.

Methods

We developed a probabilistic, discrete-event, individual-based model. This model was validated using insurance billing data for skin and soft tissue infections. The first validation trained the model for two years of data, and validated it with the next two years of data. The second method trained the model in one region and validated it in another. Following the validation, the Poisson-bootstrap resampling method was used to find specific values for a contagiousness factor (CF) in households and schools.

Results

Both methods of validation supported the model with no statistically significant difference. The bootstrap resulted in a CF$_{household}$ of 30.69 (95% CI [29.09, 32.29]) and a CF$_{school}$ of 0.55 (95% CI [0.46 to 0.64]). Effective reproduction number for the school setting was found to be 0.0015 and 0.06 to 3.04 for households of different size.

Conclusion

In this study we characterize a marked difference in communicability in the household and at school, which has not previously been shown. The identification of colonization clusters in households can be used to design strategies reduce the disease burden. The model can be used to simulate and predict responses to different interventions.
Table of contents

Introduction 1
Methods 3
   Study Location 3
   Epidemiologic Data Collection 3
   Model Design 4
   Disease states 6
   Contagiousness Factor 8
   Model transition probabilities 8
   Data Structure 10
Households 12
Schools 12
Time 12
Initialization 13
Training and Validation 13
Results 15
Discussion 21
Future Directions 25
Conclusion 26
References 27
List of figures and tables

Figure 1: Schematic depiction of the disease states with labeled transition states
Figure 2: Schematic of data hierarchy
Figure 3: Time series of monthly MRSA incidence from training and validation by split-time method
Figure 4: Time series of monthly MRSA incidence from training and validation by geographic method
Figure 5: Time series of monthly MRSA incidence from bootstrap optimization. The dotted line indicates the simulated incidence, and the solid line is the observed incidence.
Figure 6: Effective reproduction number in households with introduction of 1 colonized or infected individual

Table 1: Previously validated values for epidemiologic parameters
Table 2: Results of the Poisson-bootstrap resampling analysis
Introduction

Methicillin resistant Staphylococcus aureus (MRSA) is endemic across many US hospitals [1], long-term care facilities [2], and communities [3, 4], with an estimated 125,969 hospitalizations annually in the United States from 1999 to 2000 [5]. Moreover, MRSA-related hospitalizations more than doubled from 1999 through 2005 [1]. While the rates of health care–associated (HA) MRSA infections are declining in the US [6], community acquired (CA) -MRSA infections, predominantly skin and soft tissue infections (SSTI) [7, 8], have increased markedly in the last decade in the United States[3, 9]. Controlling MRSA transmission is hampered by the fact that traditional transmission control measures focus on hospitals settings [11] while CA-MRSA infections are often acquired in the community and treated in emergency rooms and offices [8, 10].

MRSA colonization clusters have been identified in day-care facilities [12], households [13,14], and sports teams [15, 16] through outbreak investigations. There is evidence that these clusters linked to outbreaks can be neutralized with various decolonization regimes [17, 18]. However, the overall impact of these clusters on transmission and the broader population effects of interventions targeting them are poorly understood.

Current community guidelines and recommendations focus on prevention through personal hygiene and appropriate treatment of symptomatic disease [19, 20]. Moreover, pediatric infectious disease guidelines currently recommend decolonization of a patient only after recurrent MRSA infections [21]. Proposed measures of screening and decolonization of the nares with mupirocin or other agents are not backed by sufficient evidence [20]. Similarly household decolonization has been advocated and studied on a small scale, but the impact of individual decolonization and decolonization at the household level on MRSA transmission are still debated [18,22]. For instance, recommendations have been made against mupirocin decolonization in asymptomatic patients to prevent drug resistance [23]. In order to make evidence-based recommendations, a better understanding of transmission dynamics of MRSA
among children in the community is warranted. Specifically, identification of the settings where most of the transmission events occur could lead to cost-effective intervention strategies.

Here we aim at characterizing the transmission dynamics of MRSA at the community level in order to inform public health policy and general pediatric guidelines. Our hypothesis was that, after accounting for time spent in each location, the intimate nature of interactions in a household is likely to result in a greater degree of spread of MRSA within households compared to schools.
Methods

Study location: Maricopa County, Arizona

Maricopa County is the third most populous local public health jurisdiction in the US, behind New York City and Los Angeles County, with a population of 3.8 million comprising 60% of Arizona’s population.

Epidemiologic Data Collection

We obtained our study data from the Center for Health Information Research (CHIR). CHIR is a university-community partnership between ASU and several Arizona providers, insurers and employers. It maintains a research data repository that integrates Arizona-based administrative, clinical and public health data across a large number of sources permitting the conduct of population-based research on residents of Arizona [24]. Specifically, we obtained data on hospitalization and outpatients visits of children and adolescents (age<=18 year) who were continuously enrolled for at least 6 months in the Arizona Health Care Cost Containment System (AHCCCS) program during the period January 1, 2005 to December 31, 2008. The records of all encounters with age<=18 year with skin or soft tissue infection based on ICD 9 codes (680.xx-682.9x) in Maricopa County were extracted. These codes correspond to cellulitis and abscesses, but do not include superficial skin infections such as impetigo. These were clinical case definitions since most often cellulitis is treated without any laboratory testing and test results for abscesses are often unavailable at the presenting visit. We have altogether 51,287 patient encounters including both first-time and recurrent infections during the 4-year period, with the denominator population quite stationary each year, ranging from 287,091 to 293,550.

Information regarding the prevalence of methicillin resistance in staphylococcal isolates in the community was obtained from statistics on wound cultures from 3 geographically separated urban emergency departments in the Greater Phoenix area. A previous study showed the percentage of skin and soft tissue infections caused by MRSA was 47.8, 46, 43.1, and 41.8% respectively in 2005–2008 [24].
Patients were stratified into 5 age groups: 0-2, 3-4 years, 5-9 years, 10-14 years, and 15-18 years. Using the corresponding denominator population obtained from CHIR, we compiled monthly SSTI incidence for each ZIP code in Maricopa County. Population segments with fewer than 50 AHCCCS members in an age group were included in the model, but those particular segments were excluded from the model training and validation phases to eliminate biases based on small sample sizes.

Model Design
An agent-based model was developed describing the spread of MRSA in the pediatric population in Maricopa County. The model used parameter values from a previously validated deterministic model of the spread of MRSA in the same pediatric population [24]. These values are enumerated in Table 1. Rates in the deterministic model were converted into probabilities in the agent based model. This model also included new parameters as described below in order to characterize communicability in households and schools.
Table 1: Previously validated values for epidemiologic parameters [24]

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contagiousness-infected ($\beta$)</td>
<td>0.3</td>
</tr>
<tr>
<td>Contagiousness factor-colonized ($f$)</td>
<td>0.25</td>
</tr>
<tr>
<td>Decolonization fraction ($d$)</td>
<td>0.6</td>
</tr>
<tr>
<td>Time to decolonization ($t_d$) in months</td>
<td>7</td>
</tr>
<tr>
<td>Infection time ($t_i$) in months</td>
<td>1</td>
</tr>
<tr>
<td>Infectivity ($b$)</td>
<td>0.035</td>
</tr>
</tbody>
</table>
Disease States

The previous deterministic model used differential equations and categorized the population into three groups: susceptible, colonized and infected. This model translates compartments from the homogenous population characteristics into individual disease states for simulated agents that represent members of the population on a roughly 1:1 basis. As such, each individual in the simulation has a disease state of susceptible, colonized or infected. Figure 1 illustrates disease states and their transitions. The disease state is hierarchical with infected superseding colonized superseding susceptible. In this way all infected individuals are also considered colonized for the purpose of tracking colonization length and colonization events. Person to person disease spread is modeled by discrete events where individuals transitions from a susceptible state to a colonized state following exposure to one or more infected or colonized individuals. A clinically significant SSTI is modeled by a transition from colonized to infected. Similarly treatment and decolonization are modeled by state transitions. Specific descriptions disease transitions occur are detailed below.
Figure 1: Schematic depiction of the disease states with labeled transition states
Contagiousness Factor
We estimated the relative contagiousness factor (CF) at the household level relative to school settings. CF is a dimensionless multiplier that represents the variation in contagiousness of MRSA in different settings. We estimated CF through an approach of “trajectory matching,” where one searches for the combination of model parameters that produces an epidemic curve most statistically similar to the observed one. Once epidemiological parameters are estimated, these are used to generate estimates of the effective reproduction number for specific settings.

Model transition probabilities
In general, for each discrete event we calculated the probability of an individual progressing to a new state. The probability of becoming colonized depends on the contagiousness for the time period spent in a specific setting (e.g. household vs. school), the number of infected people in the subpopulation, the number of colonized people in the subpopulation, and the relative contagiousness of colonized individuals.

Equation 1 was used to calculate the contagiousness for some defined time step from the monthly contagiousness found in the literature or used in a previous model.

\[
C_{step} = (1 + C_{month})^{t_{step}} - 1
\]

Where \( C_{step} \) is the contagiousness for the desired time step length, \( t_{step} \) is the length of the time step in months and \( C_{month} \) is the contagiousness for a month. Of note, this equation can be used to calculate any rate from incidence per month to incidence per arbitrary simulated time period.

Equation 2 was used to calculate the contagiousness, \( C_{inf} \), for an individual who spends a fraction of his time, \( t \), in a setting, \( f \):

\[
C_{inf} = \left(1 + C_{step}\right)^{t \cdot CF} - 1
\]

Where \( C_{step} \) is the contagiousness for the period, \( t \) is the fraction of time spent in that context and \( CF \) is a contagiousness factor as described above. The CF is included at this stage to
describe its significance in terms of time. For example, one setting having a CF of 3 times that of another setting is analogous to spending three times as long in one setting as spent in a setting with equal risk.

Equation 3 calculates the probability of an individual becoming colonized during a given period of time after encountering a number colonized and infected individuals.

\[
P_{\text{col}}(n_{\text{col}}, n_{\text{inf}}) = (1 - (1 - C_{\text{inf}})^{n_{\text{inf}}} \ast (1 - C_{\text{inf}} \ast R)^{n_{\text{col}}} / N
\]

Where \( P_{\text{col}} \) is the probability that an individual will become colonized during a given period of time, \( C_{\text{inf}} \) is the contagiousness of an infected individual and \( R \) is the relative contagiousness of a colonized individual, \( n_{\text{col}} \) is the number of colonized individuals in the sub population and \( n_{\text{inf}} \) is the number of infected individuals in the sub population. \( C_{\text{inf}} \) here is the contagiousness for the time period and fraction of time spent determined by Eq1 and Eq2. \( N \) is the total number of people in the subpopulation. Using the agent-based model we were able to characterize these clusters. Applying the resultant CF_{school} value of 0.55 to a classroom with one index case, we found a probability of colonization of 0.019% for an infected index case and 0.0049% for a colonized index case. The school setting was found to have an \( R_{\text{eff}} \) of 0.0057 for a colonized individual. Because this is much less than 1, it is unlikely that exposures at school play a significant role in the endemic state of MRSA, other than means of maintaining a steady state by introducing MRSA colonization to households at the same rate as it is eradicated from others.

A similar equation to Eq1 was used to calculate the probability of a colonized individual to transition to the infected state during a period. This probability is assumed to be the same for all colonized individuals and does not depend on their environment or other factors.

Transitions from infected to colonized represent receiving treatment, and transitions from
colonized to susceptible represent spontaneous decolonization. Both are simulated by selecting a wait time by sampling a normal distribution with a mean equal to the literature values and a standard deviation of ¼ that value. Each individual who enters one of these disease states selects a wait time then remains in that state until the correct amount of wait time has passed.

**Data structure**

The experimental model was designed to use distinct data elements to represent individuals in the community. Each modeled individual has data elements representing age and MRSA status. All individuals belong to a household. In addition to containing household members, a household is also associated with a ZIP code. School-age individuals belong to a school district according to their ZIP code, while adult individuals belong to a work group independent of their geographic location. Within their respective school districts, school-age individuals belong to a classroom. The data hierarchy for individuals, households, ZIP codes, and school districts is depicted in Figure 2. Although only one ZIP code is illustrated, each ZIP code belongs to a primary and secondary school district independently.
Figure 2: Data hierarchy
Individuals have the properties of age and MRSA status. Each individual belongs to a household that may contain other individuals. Several households make up a ZIP code. One or more ZIP codes together constitute a school district. ZIP codes do not belong to one school district as depicted, but to one primary school district and one secondary school district.

**Households**

Each household data element was populated using data from the 2000 United States census. The model queried census data for the specific geographic region being modeled. The specific parameters from the census data used in populating the model were the number of children in each age group 18 and under, the total population, the total number of households, the number of households with children and the average household size. It was assumed that all households with children contained exactly 2 adults.

**Schools**

School districts were modeled as a simple collection of discrete classrooms. A cluster of 30 individuals were assigned to each classroom according to age. The cluster size was selected to simulate a class size and the likely number that any child would interact with closely on a regular basis. For individuals age 5-14, the cluster included others individuals of the same age assignment, corresponding to primary school where children generally interact with others in their class. For individuals age 15-18, the cluster included 30 individuals within the whole age range, corresponding to more loosely constrained groupings in secondary school. Primary schools and secondary schools were organized geographically, and it was assumed that children from households in the same geographic area attended the same schools as defined by school district boundaries. We did not model child care for children under the age of 5 because of the large variety of child-care arrangements and the geographical and socioeconomic complexity of such arrangements. It was assumed that children under 5 spend all of their time in the household.
Time
The time assigned to each of two locations varies according to age. While children under age 5 spend 100% of their time in household environment, school-aged children are assigned to spend 1/3 of their time in the school cluster, corresponding to the school day and additional time after school spent with the school cluster group of friends.

Initialization
During the initialization of the model individuals were populated into households within ZIP codes according to census data for population, total number of households, number of households with children and average household size. The newly populated households were assigned to primary school districts and secondary school districts according to approximate mapping of ZIP codes to school district boundaries. Individuals in the population were assigned a susceptible, colonized or infected MRSA status according to the incidence and prevalence derived from the starting point of the CHiR dataset.

Training and Validation
The model was calibrated using the epidemiological time series data through an iterative optimization process. Overall, the goal of this process was to determine the values of a $CF_{household}$ and a $CF_{school}$ that would produce the best model fit to observed time series data. The model goodness of fit was measured by a chi squared goodness-of-fit test statistic for the incidence across pediatric age groups. We generated 5 realizations of our stochastic individual-based model for each parameter set. To find the optimal values for $CF_{household}$ and $CF_{school}$, we first generated epidemic model realizations over a variety of sets of these parameters spanning the complete range of possible. After this initial low detail survey of possible values a modified simulated annealing heuristic optimization was implemented to obtain the best fit model parameters.

We validated our transmission model using two methods that employ different levels of spatial
and temporal information. The first “split-time” method focused on the population living in a single secondary school district, the Phoenix Union High School District (PUHSD), and the corresponding primary school districts. This single district population was initially trained using data from the first two years of our dataset. Then the model’s simulated incidence trajectory for the following two years was compared to the corresponding observed epidemiologic data using a standard chi squared goodness-of-fit calculation. To perform this calculation the incidences for a age-specific incidences were combined across all ZIP codes, weighted according to the local population in each age-group and ZIP code.

In the second, ‘geographic’ method, the model was first trained for the population in the PUHSD using the entire four years of available data. Then our model was validated by using the optimized parameters over the same four years in the combined population of the Mesa and Tempe secondary school districts. A combination of these two school districts was used in order to have a comparable population size to PUHSD.

After the initial validation a bootstrap-Poisson resampling analysis was conducted for the PUHSD population to estimate confidence intervals for the $C_{F_{\text{household}}}$ and $C_{F_{\text{school}}}$. 
**Results**

When the model was trained with the split-time method, the model was optimized to a CF\textsubscript{household} of 41.88 and a CF\textsubscript{school} of 0.04. The chi squared value for the validation years was 25.99 with a P value >0.9999\textsuperscript{1}. Figure 3 compares the optimized simulation to the observer data.

The geographic method optimized to a CF\textsubscript{household} of 47.78 and a CF\textsubscript{school} of 0.01 with a chi squared value of 22.73 and a P value > 0.9999. Results from the validation phase in the Mesa\textbackslash Tempe area had a chi squared of 58.41 and a P value >0.9999. Figure 4 compares the training and validation simulation time series.

The resampling results are listed in Table 2. The model was resampled for 20 iterations. CF\textsubscript{household} had a mean of 30.69 and a 95% CI of [29.09, 32.29]. The CF\textsubscript{school} had a mean of 0.55 and a CI of [0.46 to 0.64]. Figure 3 shows the averaged results 20 simulations with the resulting parameters from the bootstrap approximation. Figure 5 shows the averaged results 20 simulations with the resulting parameters from the bootstrap approximation.

We used simple, single-setting simulations to demonstrate the implications of the CFs found above. Each simulation models the spread of MRSA in a specified group using the parameters found by using the Poisson-bootstrap method. The simulations begins with one index case in the group becoming either colonized or infected, and then track the group in isolation from outside factors until the colonization has cleared. Each simulation was run 1000 times and the results were averaged. Figure 6 shows the effective reproduction number ($R_{eff}$) in the household setting graphed against household size. In the school setting the effective reproduction was found to be 0.0057 and 0.0078 for a colonized and infected index case respectively.

\textsuperscript{1} Low P values are commonly used to show statistical significance of findings; however, here high P values are used to show a statistically insignificant difference between the observed data and the simulation.
Figure 3: Time series of monthly MRSA incidence from training and validation by split-time method. The dotted line indicates the simulated incidence, and the solid line is the observed incidence.
Figure 4: Time series of monthly MRSA incidence from training and validation series by geographic method. The graphs on the left correspond to the training series. Those on the right are validation. The dotted line indicates the simulated incidence and the solid line is the observed incidence.
Table 2: Results of the Poisson-bootstrap resampling

<table>
<thead>
<tr>
<th></th>
<th>$CF_{\text{household}}$</th>
<th>$CF_{\text{school}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>30.69</td>
<td>0.55</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>3.65</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>CI</strong></td>
<td>[29.09, 32.29]</td>
<td>[0.46, 0.64]</td>
</tr>
</tbody>
</table>
Figure 5: Time series of monthly MRSA incidence from bootstrap optimization. The dotted line indicates the simulated incidence, and the solid line is the observed incidence.
Figure 6: Effective reproduction number ($R_{eff}$) in households with introduction of 1 colonized or infected individual. Squares correspond to simulations of household index cases. Diamonds correspond to simulations with colonized index cases.
**Discussion**

We have developed a probabilistic, discrete-event, individual-based model for CA-MRSA in the pediatric population that was trained and validated using local administrative data. Validation was achieved through both testing a model trained with observed data for an initial time period in the same region and testing a model trained with data for one region in another region. A Poisson-bootstrap resampling yielded a $CF_{\text{household}}$ of 30.69 (95% CI [29.09, 32.29]) and a $CF_{\text{school}}$ of 0.55 (95% CI [0.46, 0.64]). The implication of these results is that contact at household is 55.8 (95% CI [46.45, 70.20]) times more likely to result in colonization than the same amount of time spent at school. This supports our hypothesis that contagiousness at household is greater than at school even when accounting for the time spent each place in the calculation.

Several previous studies have identified households as disease clusters and successfully decolonized them [13,18]. One prospective, observational study found that 50% of clinically encountered MRSA infections had at least one other member of their household who was colonized [25]. In total 67% of index patient household members were found to be colonized. Little is known about the prevalence and epidemiologic impact of these clusters.

Using the agent-based model we were able to characterize these clusters. Applying the resultant $CF_{\text{school}}$ value of 0.55 to a classroom with one index case, we found a probability of colonization of 0.019% for an infected index case and 0.0049% for a colonized index case. The school setting was found to have an $R_{\text{eff}}$ of 0.0057 for a colonized individual. Because this is much less than 1, it is unlikely that exposures at school play a significant role in the endemic
state of MRSA, other than means of maintaining a steady state by introducing MRSA colonization to households at the same rate as it is eradicated from others.

The corresponding probabilities in a household with 5 members with a $CF_{\text{household}}$ of 30.69 are 9.6% for an infected index case and 2.4% for a colonized index case. Taking the situation where one index individual in a household is colonized we can illustrate a behavior seen in the model. From the time of the initial colonization, the index case has an average time to decolonization of 7 months [24]. Even if the individual does not become infected, for every discrete week the other four members of the family have a 2.4% chance of becoming colonized. In general, there is only a 6.6% chance that the colonization will clear before the index case has decolonized.

Once another individual becomes colonized the probability colonization further increase and the probability of eradicating MRSA in the household decreases further. A household with 5 members carries an $R_{\text{eff}}$ of 0.93 and will take an average of 54 months to decolonize once one of its members is colonized. This may go undetected by traditional monitoring as it will only result in 1.32 infections over the 54 month event. A reproduction number of 1 is required to maintain an endemic state. Using a regression to approximate a continuous function for the $R_{\text{eff}}$, the household size where $R_{\text{eff}}$ equals 1 is found to be 5.089. It is possible to maintain an endemic in households with 6 or more members.

Potential interventions should be explored as targets for public health policy. Household colonization represents a large part of the disease burden of CA-MRSA. According to this model, larger households have an increased chance of harboring MRSA colonization.
Interventions aimed at schools should be avoided as it is as based on our model, they will have minimal effect on the overall epidemiology. No recommendations about child care centers should be based on this work since they were not included in the model.

Future work should focus on exploring the impact of specific proposed prevention measures on the disease burden. It is likely that identifying and eliminating household clusters will have a significant impact the prevalence of MRSA, but at this time no specific recommendations can be made for preventive measures other than adopting larger family size as a risk factor.

The major strength of this study is the large sample size of the dataset. The large number of patient encounters represented contributes to the narrow confidence interval and the statistical significance of the findings. Furthermore, averaging over several simulations helps limit the impact due to a single random event in the simulation. By using a dataset which includes outpatients we are able to characterize the behavior of MRSA in the general pediatric population. Many other studies have focused on hospital outbreaks or specific foci of infection in the community.

The major weakness of this model is the use of administrative data. We used administrative billing data in combination with the incidence reported by emergency room cultures of skin and soft tissue infections because more direct measures are unavailable for analysis. We chose to use administrative data because most skin and soft tissue infections are currently managed.
clinically without definitive cultures. Even though there are more potential variations in coding, this dataset represents the best approximation available at this time. We also approximate the incidence of MRSA in the general population using data from emergency department incision and drainage cultures. This likely overestimates the incidence because patients in the emergency department are more likely to have abscesses. MRSA is more likely to form an abscess than other causes of SSTI, so the portion of incision and drainage cultures that are MRSA is likely higher than that of SSIs in general. Medicaid billing data was used to train and validate the model because it was the most complete record available; however, because this dataset represents a lower socioeconomic class than the general public, it may have caused us to overestimate the true incidence of MRSA SSTIs. While these factors likely contribute to an overestimation of the incidence of MRSA, this sort of system-wide overestimation is unlikely to affect the ratio of contagiousness at home compared to school.

For the sake of simplicity we assumed that children attended school in the geographic area where they lived. This assumption is not universally true, but the small amount of variation from specific exceptions is unlikely to affect the overall epidemiology. We assumed that children under 5 spent all of their time in the household. This assumption was made because we lacked information about child-care arrangements in the population. In doing so we did not factor in child care centers as potential setting for the spread of MRSA even though previous studies have identified MRSA clusters in child care centers[12]. Including MRSA spread at childcare centers would likely lead to a higher incidence of MRSA in the 0-2 and 3-4 age groups, possibly leading to a closer fit to the observed data(Fig 5). If including child-care arrangements
effectively increases the infection rate in the 0-2 and 3-4 age groups, the model may optimize to a lower contagiousness overall, bringing down the 5-9 and 10-14 incidence for a tighter fit. Future iterations of the model may include child care centers and explore their impact in MRSA epidemiology.
Future Directions

Various features may be added to make the model. Differences in incidence between age groups may be associated with age-related factors such as hygiene. This model is well-suited to explore those differences and their possible impact.

Now that the model has been validated and a target for intervention has been found, attention should be turned to using the model to evaluate potential solutions. The next step will be to explore proposed interventions and analyze their impact using the individual-based model. Individual-based models are uniquely equipped for predicting the impact of proposed changes in behavior.

After an intervention is selected with the help of the local public health department, a program should be developed and implemented to mitigate the disease burden. Future analysis of CHiR data following the intervention will show if the predicted impact was made.
Conclusion

CA-MRSA represents a significant disease burden in the pediatric population. While treatment guidelines for infected individuals are plentiful, limited evidence exists to make informed guidelines for public policy or to endorse preventive measure at the primary care level. Using the model described here we have shown that school contact is not likely a major contributor to the endemic. Exposures in the household, likely due to the more intimate setting, are more likely to lead to colonization even when the difference in time is factored in. Interventions should be explored to stem the re-colonization that takes place in larger households. In the future specific interventions can be analyzed using this model to estimate their impact on the population.
References


