

Network position and health care worker infections

Troy Tassier¹ · Philip Polgreen² · Alberto Segre³

Received: 16 July 2013 / Accepted: 30 August 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract We use a newly collected data set coupled with an agent-based model to study the spread of infectious disease in hospitals. We estimate the average and marginal infections created by various worker groups in a hospital as a function of their network position in order to identify groups most crucial in a hospital-based epidemic. Surprisingly, we find that many groups with primary patient care responsibilities play a small role in spreading an infectious disease within our hospital data set. We also demonstrate that the effect of different network positions can be as important as the effect of different transmission rates for some categories of workers.

Keywords Infectious disease · Social networks · Vaccine subsidies · Economic epidemiology

This work was supported in part by a pilot grant from the University of Iowa's Carver College of Medicine and the University of Iowa's College of Public Health, and two grants from the National Institutes of Health (R21 AI081164; K01 AI075089). We thank Sriram Pemmaraju, Jason Barr, Mary Beth Combs, and Leanne Ussher for excellent comments and suggestions on the project and Jenifer Kuntz for help with data collection and organization. The paper also has benefitted from comments by seminar participants at Fordham University, Binghamton University, and the New York City Computational Economics and Complexity Workshop as well as from detailed comments from three anonymous reviewers.

✉ Troy Tassier
tassier@fordham.edu

¹ Department of Economics, Fordham University, E528 Dealy Hall, Bronx, NY 10458, USA

² Department of Internal Medicine, Carver College of Medicine, University of Iowa, Iowa City, IA, USA

³ Department of Computer Science, University of Iowa, Iowa City, IA, USA

1 Introduction

Vaccines are a primary way to stop or slow the spread of many infectious diseases, perhaps most notably, influenza. The lack of appropriate vaccination levels is a major health problem. For instance, influenza is a major cause of morbidity and mortality throughout the world despite the availability of a highly effective and inexpensive vaccine. In the US alone, influenza causes an estimated 36,000 deaths and 120,000 hospitalizations annually yet only around 1/3 of healthcare workers are vaccinated each year (Thompson et al. 2003). Efficient provision of vaccinations poses a difficult problem in that the positive externality associated with a vaccination is the product of the probability of infection, the cost of the infection, and the marginal infections generated by an agent if infected (all of which may vary across agents).

There is great concern over the spread of infectious diseases in hospitals, but little knowledge is available to identify healthcare workers who are most likely to acquire and transmit infectious diseases in hospitals. The problem is especially difficult because the transmission of many infectious diseases is not observable. For instance if someone in your household acquires influenza, you likely do not know which of the potentially hundreds of people you come in contact with each day that may have caused the infection. Further if a vaccine is available for an infectious disease and the vaccine is in short supply or is expensive, it is imperative to know which individuals should have the highest priority in vaccine campaigns. In this paper we use a newly collected data set on hospital worker contacts in order to identify hospital worker groups that have the potential to create the largest number of infections based on their location in a hospital contact network.

To achieve this goal, we have collected person-to-person contact information on 140 individuals belonging to one of 15 types of healthcare workers at the University of Iowa Hospitals and Clinics (UIHC). The data contain information on the contacts between healthcare workers and patients and between healthcare workers and other healthcare workers at the hospital. With this information we develop a network model describing the spread of an infectious disease in a hospital. We estimate, using an agent-based model, the effect of network position of different hospital worker groups on the spread of infectious diseases in a hospital. Through this model we are able to identify the hospital worker groups that create the largest externality if removed from the network (perhaps through a vaccination or a quarantine). We argue that methods such as those used in this paper can help hospitals, health care professionals, and epidemiologists to design efficient programs for healthcare worker vaccinations.

Of importance, we note that we only study the externality in terms of network position within the hospital. In this paper we do not consider other potential heterogeneity among agents such as differences in transmission rates across workers, or differences in behavior outside the hospital that may play a role.¹ While these effects are also important, the large differences across classes of workers shown below are worthy of independent study.

¹ In the final part of the results section we do offer some insight as to the size of the network based effects relative to transmission rates.

This is the first paper to use specific micro-level contact data within a hospital that can be used to help guide policy makers and public health officials in the problem of efficiently allocating vaccines within hospitals. The data used in this paper is unique and detailed in comparison to other studies. The data consists of shadow data (where a research assistant follows a specific, randomly chosen hospital worker for an entire shift) for the 15 major groups of healthcare workers at the UIHC, a 700-bed major medical center. This results in over 600 h of direct hospital worker observations and the notation of over 6500 specific worker to worker or worker to patient contacts throughout the hospital. To the best of our knowledge, the data that we have collected comprises the most detailed micro-level healthcare worker contact data set in existence. As a comparison, [Ueno and Masuda \(2009\)](#) collect data on contacts between doctors, nurses, and patients. Their data is based on two calendar days from a small, 129 room, community hospital in Tokyo. They model contacts between nurses, physicians and patients. Their data does not consider contacts with and between other healthcare worker groups (other than nurses and doctors). Based on our data at the UIHC, these assumptions would ignore over 60% of hospital staff, including most of the groups we identify as most crucial to the spread of an infection disease.

We begin by discussing the background and motivation for our study and then move to develop a simple model of infectious disease transmission. In the model, we initially assume homogeneous contacts as in traditional epidemiological models. We then discuss a similar model with heterogeneous contacts and discuss the difficulties of achieving efficient vaccination levels. Following the theoretical discussion, we use our newly collected data on healthcare worker and patient contacts to model the spread of an infectious disease in a hospital setting. The model allows us to identify the healthcare worker groups that would be expected to play the largest role in the spread of infectious diseases, in terms of network position, in this hospital setting.

1.1 Background and motivation

Traditionally, epidemiology research has focused on well-mixed (randomly mixed) populations where agent contacts are homogeneous. In these models, every agent in a population may “bump into” any other agent with equal probability, much as a gas molecule may bump into any other gas molecule with an equal probability over a fixed time period. Only recently have epidemiologists and other researchers begun to study the heterogeneous contact structures between people over which infectious diseases spread (early studies include [Comins et al. 1992](#); [Grenfell and Harwood 1997](#); [Wallinga et al. 1999](#)).

We focus this study on healthcare workers and a particular class of infectious disease, of which influenza is an example. Healthcare workers are at especially high risk of contracting influenza. One study of healthcare workers with a low rate of influenza vaccination demonstrated that 23% of healthcare workers had evidence of influenza infection during a single influenza season ([Elder et al. 1996](#)). Two features of influenza make its spread difficult to control in hospitals. First, people with influenza are infectious 1–4 days before the onset of symptoms. Thus, they can spread the virus when they are still feeling well and are unaware of their own infectious state. Second,

only about 50% of infected persons develop classic influenza symptoms (CDC 2002, 2003). Consequently, restricting healthcare workers with influenza-like symptoms from the workplace will not completely prevent transmission of influenza because healthcare workers with atypical symptoms could continue working and spreading the virus. Furthermore, studies show that healthcare workers are more likely than other workers to return to work early or to keep working when they have influenza-related symptoms (Weingarten et al. 1989).

Because of the ease with which influenza may be contracted and spread by healthcare workers, the Centers for Disease Control and Prevention (CDC) have, for the past two decades, recommended influenza vaccination for all healthcare workers. Yet, in the US, only 36% of workers who have direct patient contact are immunized against influenza annually (Smith and Bresee 2006).

Outside of concerns about traditional influenza, there are additional reasons to study the spread of infectious diseases in hospitals. First, healthcare-associated infections affect about 2 million patients in US hospitals each year (Jarvis 1996). Second, there is a growing fear that hospitals could become a breeding ground for new strains of influenza such as the recent H1N1 influenza outbreak, the potential emergence of person-to-person transmission of avian flu, or other “new viruses.” Much as SARS spread widely in hospitals to begin the SARS epidemic in Toronto (Chowell et al. 2004), person-to-person transmission of avian flu may start in hospitals as well, and, if a more lethal version of H1N1 were to develop, hospitals again could be a breeding ground for new infections. This last point is of particular salience. With the recent H1N1 outbreak and the subsequent work to develop a vaccine, controversies arose concerning which individuals to vaccinate first. Healthcare workers were high on the list. But, as we show below, not all healthcare workers are equal in terms of their importance in spreading infectious diseases. Thus, one primary focus of this paper is identifying the individual hospital workers who are most important to vaccinate should a similar outbreak occur again.

There is a growing literature in economics on the vaccination choices of individuals and of the externalities associated with vaccinations. But scant attention is paid to the network effects determined by heterogeneous contacts that we focus on in this paper. For example, Francis (2004) solves for the optimal tax/subsidy policy for influenza in an SIR model with a constant contact rate and random mixing among the population. Geoffard and Philipson (1997) examine how the individual incentives for vaccination decrease as disease incidence decreases and thereby argue that relying exclusively on private markets is unlikely to lead to disease eradication. Boulier et al. (2007), the most similar paper to ours, investigate the magnitude of the externality associated with a vaccination as a function of the number of vaccinations in the population, the transmission rate of the disease, and the efficacy of the vaccination. They find non-monotonic relationships between each of these items and the magnitude of the vaccine externality. More specifically, the externality and the number of infections prevented per vaccination initially increases before eventually decreasing. However, like Francis, they do not consider the case of heterogeneous contact number or heterogeneous sorting among the population. Finally, much of the recent literature on the economics of infectious disease is summarized in Philipson et al. (2000).

2 A simple model

We begin by describing a simple model where agents in a population have contacts with each other with a uniform probability. This is the traditional random mixing model in epidemiology pioneered by [Kermack and McKendrick \(1927\)](#). The important results in this paper describe exceptions to this homogeneous contact assumption, but we use the simplified model to develop intuition before describing a richer model with heterogeneous contacts. In this simple model, we assume that all agents are homogeneous in that all agents have the same number of contacts with other agents and that these contacts are randomly drawn with uniform probability from the population at large.

Suppose that agents are assigned to one of three states: Susceptible (S), Infected (I), or Recovered (R). A susceptible agent can transition to being infected with probability α if she is in contact with an infected agent. Once infected, an agent transitions to the recovered state according to a parameter κ . Once recovered, the agent is immune to the possibility of future infection. This is a classic Susceptible—Infected—Recovered (SIR) model for infectious diseases such as influenza ([Kermack and McKendrick 1927](#)). The description and parameters yield the following differential equations describing the flows of agents among the various states, assuming a constant population of size N and contact rate of γ . Each equation describes the rate of growth for one of the three populations in the SIR model.

$$\frac{dS_t}{dt} = -\alpha\gamma S_t \frac{I_t}{N} \tag{1}$$

$$\frac{dI_t}{dt} = \alpha\gamma \frac{S_t}{N} I_t - \kappa I_t \tag{2}$$

$$\frac{dR_t}{dt} = \kappa I_t \tag{3}$$

Equation 1 describes how susceptible agents contact γ other agents in the population, of which I_t/N are infected, and how, of these contacts with infected agents, a percentage α cause the susceptible agent to transition to being infected. Equation 2 describes the previously mentioned flows from susceptible into infected and that each infected agent moves to being recovered at rate κ . Finally, Eq. 3 describes the flows from infected to recovered.

We can write these equations in terms of population shares by dividing each of the above equations by the population size N and using lower case letters to denote these population shares, $s_t = S_t/N$, $i_t = I_t/N$, and $r_t = R_t/N$, yielding the following population share equations:

$$\frac{ds_t}{dt} = -\alpha\gamma s_t i_t \tag{4}$$

$$\frac{di_t}{dt} = \alpha\gamma s_t i_t - \kappa i_t \tag{5}$$

$$\frac{dr_t}{dt} = \kappa i_t \tag{6}$$

The number of infected agents will increase in the population if the flows into the infected state from the susceptible state exceed the flows out of the infected state into the recovered state, $\frac{di_t}{dt} > 0$. This condition is equivalent to $\alpha\gamma s_t > \kappa$ or $s_t > \frac{\kappa}{\alpha\gamma}$. If this inequality holds then we say that the population is above the *epidemic threshold*. Note that we cannot remain above the epidemic threshold forever without an introduction of new susceptible agents since $\frac{ds_t}{dt} < 0$: eventually the population will run out of susceptible agents to infect unless the susceptible population is replenished at a sufficient rate.

One goal of healthcare policy is to attempt to place a population below the epidemic threshold so that the number of infectious agents in a population does not grow subject to some cost constraint. A population is most vulnerable to being above the epidemic threshold when the infectious disease first enters a population because $s_t \approx 1$. This implies that each infected agent infects approximately $\frac{\alpha\gamma}{\kappa}$ new agents in the population. This fraction is sometimes referred to as the initial *reproduction number* in the population and is commonly denoted as $R_0 \equiv \frac{\alpha\gamma}{\kappa}$. Without new individuals entering a population in the susceptible state this reproduction number can only decline as the infectious disease spreads.

2.1 Vaccinations

A successful vaccination moves an agent from state S to state R without incurring the costs of infection. If we reduce the initial population of susceptible individuals, s_0 , by enough we can push the population below the epidemic threshold. Specifically, if $s_0 < \frac{\kappa}{\alpha\gamma}$ then the infectious disease dies out of its own accord without further action. Thus an epidemic is prevented whenever $s_0 < \frac{\kappa}{\alpha\gamma}$ which occurs when $(1 - \frac{\kappa}{\alpha\gamma})N$ agents are successfully vaccinated. Vaccinating enough agents to produce this effect is called *herd immunity* (Smith 1970); once enough people are vaccinated, the entire population (herd) is effectively protected without everyone being vaccinated. The question then becomes, given a cost of vaccination, $c(v)$, what is the efficient level of vaccinations to provide in a population and how do we obtain this efficient level? We begin to approach this question by introducing standard value function notation.

In this initial model, once an agent enters state R she remains there forever. Thus the value of being in state R is simply the lifetime discounted utility received in state R. We also introduce the possibility of having heterogenous contacts at this stage by indexing agent j 's contacts (γ_j) and other terms that we allow to vary across agents.

$$V_j(R) = \int_{t=0}^{\infty} \beta^t U_j(R) \quad (7)$$

where $U_j(\cdot)$ = utility of agent j from the specified state and β is the discount rate.

If an agent is in state I, she will remain in state I for $1/\kappa$ periods, on average, until recovered and then enter state R.

$$V_j(I) = \int_{t=0}^{1/\kappa} \beta^t U_j(I) + \int_{t=1/\kappa}^{\infty} \beta^t U_j(R) \quad (8)$$

If an agent is in state S , she receives the same utility as she would if she was recovered, unless she becomes infected. The value to an agent of being in state S is the value of being in state R less the product of the probability that the agent becomes infected and the cost of the infected period.

$$V_j(S) = V_j(R) - \pi(\gamma_j, \alpha, i)c_j \quad (9)$$

where $\pi(\gamma_j, \alpha, i)$ is the probability of becoming infected over the course of the epidemic as a function of the contacts of the agent and the transmission rate of the infectious disease. The cost of being infected is the difference in utility between states S and I during the time spent in state I , $c_j = [U_j(S) - U_j(I)](1/\kappa)$, with $U_j(I) < U_j(R)$.

With the value functions specified we can now specify the vaccination problem for the individual and the social planner.

2.1.1 The individual vaccination problem

To simplify the vaccination choice of individuals we assume that an agent can only be vaccinated at time period 0. At this time the agent will choose to be vaccinated if the value of being in the recovered state less the cost of the vaccination is greater than the value of being in the susceptible state. Thus the agent will choose to be vaccinated if

$$V_j(R) - c(v) > V_j(S) = V_j(R) - \pi(\gamma_j, \alpha, i)c_j \quad (10)$$

which implies the agent will choose to be vaccinated if $c(v) < \pi(\gamma_j, \alpha, i)c_j$.

2.1.2 The social planner's vaccination problem

The social planner's vaccination problem is more difficult than the individual vaccination problem. Essentially, the social planner's problem is to vaccinate agent j if the cost of the vaccination is less than the expected dis-utility of the increase in infections created by agent j if agent j is infected weighted by the probability that agent j is infected. We define the term "marginal infections" of agent j to be the additional infections that occur if j is infected that would not occur if j was not infected. Note that this is not simply the number of infections that agent j creates. As an example, agent k maybe infected by agent j . But, if k would have been infected by another agent had she not been infected by j then this would not be a "marginal infection" of agent j . k will be infected regardless of the vaccination choice of agent j . Along these lines of thinking, measuring marginal infections is a difficult problem for epidemiologists for at least three reasons:

1. As mentioned earlier, many infectious disease transmissions are not observable. Thus it is not easily known how many infections a given agent causes.
2. Even if transmissions are observable, the marginal infections created by agent j are not simply the number of other agents that j infects. This is because some agents that j infects may get infected by agents other than j even if j does not infect

them herself. Further one needs to know how many additional agents are infected by those that j infects and any additional infections created by these agents and so forth. Thus one needs to know information on the dynamics of the entire epidemic to measure the true marginal infections of a given agent.

3. The marginal value of vaccinating an agent depends on the behavior and vaccination choices of other agents. It eventually must be decreasing in the number of other vaccinations that are performed. In the extreme, if there are enough vaccinations in the population to produce herd immunity the marginal value of vaccinating an additional agent only involves the probability that the agent is infected from outside the agent population. In effect, the only value is preventing a single agent from infection because she will not, on average, infect anyone else.

Because of these difficulties we use a simulation approach to help us measure the average and marginal effects of individuals belonging to different worker groups in our hospital contact data. With simulations one can monitor the various infections that occur and also perform controlled experiments to sort out the effects of various groups on potential hospital epidemics.

Define $m_j(\gamma_j, \alpha, \kappa, i, v)$ as the true marginal infections created by agent j if infected, where v is the number of agents vaccinated in the population. For the majority of the paper we will suppress the notation that does not differ across agents and simply refer to marginal infections as $m_j(\gamma_j)$ since the primary focus of the paper will be on the effect of heterogeneous contacts on the spread of infectious diseases. As shown in Boulier et al. marginal infections may be increasing in v for sufficiently small v .² But, marginal infections must eventually decrease in v ; at the extreme, marginal infections are 0 for any level of v above the point at which herd immunity is reached. Thus marginal infections may be increasing or decreasing in the number of vaccinations depending on the specifics of disease transmission, contact patterns, and the number of vaccinations performed.

We can now state the social planner's vaccination problem.

Vaccinate agent j if:

$$c(v) < \pi(\gamma_j, \alpha, i)[1 + m_j(\gamma_j)]c_j(i) \quad (11)$$

Note that the individual and social planner's vaccination problems differ by the term $\pi(\gamma_j, \alpha, i)m_j(\gamma_j)c_j(i)$. This is the positive externality created by a vaccination when a vaccinated agent j protects other agents which he contacts from acquiring the disease from him.

The key terms to investigate in this externality are the probability agent j gets infected, and the marginal infections created by agent j if infected. Note that if these marginal infections, $m_j(\gamma_j)$, go to 0, the social planner's problem and the individual problem converge, and the externality is removed. Similarly the externality is removed

² As brief intuition for this result, the first few vaccinations do very little to prevent an epidemic when the susceptible population is large. If the first vaccinated agent doesn't infect you, someone else will. But, once the vaccinations reach a sufficiently large number, the positive feedbacks associated with a smaller epidemic and a smaller number of susceptible agents decrease marginal infections.

if enough vaccinations are performed to reach herd immunity, again because $m_j(\gamma_j) = 0$ when herd immunity is achieved because the epidemic never occurs.

2.2 Heterogeneous contacts

The main focus of this paper concerns the network position of hospital workers. As such, we assume throughout the paper that the cost of an infection is equal across all agents, $c_j(i) = c_k(i)$ for all j and k . Further, without loss of generality, we can normalize $c_j(i) = 1$. Thus the externality above is the product of the probability of infection and the marginal infections produced by the agent if infected. One question then emerges: how are $\pi(\gamma_j)$ and $m(\gamma_j)$ related? At a simple level, if contacts only vary in degree, that is if the only difference in contacts between two agents is the number of contacts and not other, qualitative, aspects of the contacts, then you would expect $\pi(\gamma_j)$ and $m(\gamma_j)$ to be highly correlated. If an agent has a high likelihood of being infected because he has many contacts then he also has many contacts to pass on the infection.

2.2.1 Example 1: Uniform random contacts with low connectivity

Suppose that each agent pair is connected with a fixed probability. Then, by chance, some agents will have a higher than average number of contacts and some a lower than average number of contacts. In this case, any agent who has a large number of contacts will also generate a large number of secondary infections since there is a lack of structure within the network population. Thus any agent with a high probability of infection, $\pi(\gamma_j)$ will also be expected to generate a high level of marginal infections $m(\gamma_j)$.

Example one is fairly direct. However, various relationships are possible as we show below.

2.2.2 Example 2: Fully connected graphs

In this case each agent in a population is directly connected to every other agent in the population. If this is the case, then any agent that becomes infected is directly tied to all other agents and can infect anyone in the population. Thus once someone is infected each agent has a high probability of becoming infected (either from the original agent or from secondary infections). But, since the first agent contacts everyone in the population, and many agents will be infected from him, the other agents in the population may have a low $m(\gamma_j)$. Thus it is possible to have a high probability of infection $\pi(\gamma_j)$ and low marginal infections generated $m(\gamma_j)$ from the same agent.

2.2.3 Example 3: A bridge between two separate fully connected graphs

Imagine that there are three groups of agents in a population. Two of these groups, call them A and B , are separate fully connected graphs containing equal numbers of agents, who do not have any connections to the other group. In other words an agent

$a \in A$ is connected to every agent $a' \in A$ but no agent $a \in A$ is connected to any agent $b \in B$. Suppose that group B is formed in a similar manner. The third group is composed of one agent, j . Agent j has only two contacts: one to agent $x \in A$ and one to agent $y \in B$. In this example it may be unlikely that agent j gets infected, especially if there is a low transmission rate, because he only has two contacts in the populations. But, if agent j is infected, he may be integral to spreading the disease to the second fully-connected group. Suppose an agent in A becomes infected and subsequently infects agent x (or any of the other agents in A) as well as several other agents in A . Agents in group B are safe from infection as long as agent j is not infected. But, if j becomes infected, then it is possible that a large fraction of agents in B may become infected as well. Thus agent j may have a low probability of being infected, $\pi(\gamma_j)$, but create a large number of marginal infections, $m(\gamma_j)$, if he does become infected.

Note that each of these three examples offer different implications for public policy approaches to encouraging vaccinations. In the first example, each agent has a probability of being infected that is in line with the number of marginal infections generated. In the other two examples, the infection rate and the number of marginal infections generated may not have a simple relationship with each other. This is an important observation if one considers using subsidies or other means to encourage increased vaccination rates.

3 Data

In the remaining portion of the paper we examine a data set on contacts of and between healthcare workers and patients in a large hospital on the University of Iowa campus. We discuss the data and use agent-based models to identify the healthcare workers whose position in the hospital contact network has the potential to create large numbers of infections in the hospital.

Observational data on contacts between healthcare workers and patients was collected during the winter and early spring of 2006–2007 (the 2006–2007 “flu season”) at the University of Iowa Hospitals and Clinics (UIHC). The UIHC is a 700-bed comprehensive academic medical center and regional referral center in Iowa City. Data were collected by randomly selecting UIHC employees from each of 15 job classifications (specified below) and then using research assistants to “shadow” the 140 selected employees. The research assistants manually recorded every human contact of the subjects (within approximately three feet) over a work shift. Note that these observed contacts include anyone contacted within the hospital, not just with the other workers in the shadow sample. This resulted in a total of 6654 recorded contacts over 640 h of observation. Additionally, the RA recorded the worker or patient group category for each observed contact (patient or category of healthcare worker) in our data set, and the location in the hospital where the contact occurred.³

The job categories and number of observed subjects in the data set are as follows: Floor Nurse (8), Food Service (11), Housekeeper (8), Intensive Care Nurse (8), Nurse

³ The data collection was approved by the Institutional Review Board at the University of Iowa.

Assistant (10), Pharmacist (8), Phlebotomist (10), Physical/Occupational Therapist (9), Resident/Fellow/ Medical Student (8), Respiratory Therapist (11), Social Worker (8), Staff Physician (11), Transporter (7), Unit Clerk (9), and X-Ray Technician (14). The data for each group contain approximately 40h of shadowing. The data were summarized into tallies of contacts over 30-min intervals and then aggregated into contacts per 8 h shift by the authors.

Table 1 lists the average number of non-repeated contacts⁴ per 8 h that occur between the various worker (and patient) categories. Note that we were not allowed to choose patients as subjects in our shadow data directly because of privacy concerns. We were only able to observe patient contacts as a result of shadowing other members of the hospital. Thus they do not appear as a row in the table. We use this contact data to model the spread of an infectious disease across the UIHC hospital.

With this data we create a probabilistic contact network for the hospital worker groups. The network is constructed to match the distribution of worker groups at the University of Iowa Hospitals and Clinics. This totals 5232 employees. The distribution of workers across the categories is given in Table 2.

We create a contact network among these agents. In the model, each worker in a given group connects to other workers according to the rates observed in our shadowed subjects given in Table 1. As an example, all floor nurses in the model create 11 contacts to other randomly selected floor nurses on average, 0 contacts to food service workers, etc. We assume that the contacts are symmetric in our model, that is, a contact from a given floor nurse to a given housekeeper is also a contact from the housekeeper to the floor nurse. There are at least two reasons for this assumption. First, if our subject is in close enough proximity to pass on the influenza virus to a second agent, the second agent is also within close enough proximity to pass on the influenza virus to our subject. Thus the ability to acquire or to pass on virus is a symmetric relationship. Second, the reader may note in the table that the matrix of observed contacts is not symmetric because of randomness in the observation of subjects. For instance one notices that the subject floor nurses were not observed to contact food service workers, but a small number of food service worker to floor nurse contacts were observed. Thus by assuming that all contacts that occur in the matrix are undirected, we create a symmetric contact matrix where the total number of contacts from a member of group x to group y (and from group y to group x) is one-half the sum of the observed average contacts from group x to group y and from group y to group x .

We create the contacts in a uniform random manner within groups. Let ρ_{ij} be the ratio of the average contacts between a member of group i and j (taken from Table 1) to the total number of group j employees (taken from Table 2). We then take each pair of employees across each group and create a contact with probability ρ_{ij} . Specifically, let agent a be a member of group i and agent b be a member of group j . Then the probability that a and b are connected is $\rho_{ij} = C_{ij}/N_j$, where C_{ij} is the

⁴ A non-repeated contact is a contact between two workers or a worker and a patient that has not been observed before.

Table 1 Average contacts between worker categories per 8 h

	FN	Food	HK	IC N	N Asst	Pat	Pharm	Phleb	P Ther	Res	R Ther	SW	Stf Phys	Trans	UC	X-ray
FN	11	0	1	1	5	11	1	0	1	3	1	1	2	1	2	1
Food	3	5	1	1	2	14	1	1	0	1	0	1	1	0	4	0
HK	4	0	3	1	2	2	0	0	1	1	0	0	1	0	1	0
IC N	1	0	1	8	1	3	1	0	1	1	2	1	2	0	1	1
N Asst	5	1	1	2	2	10	0	0	1	1	1	1	0	1	2	0
Pharm	4	0	0	1	1	3	2	0	0	3	0	1	2	0	1	0
Phleb	3	0	0	0	1	24	1	2	0	1	1	0	0	0	1	0
P Ther	7	1	1	3	2	10	1	1	3	1	1	1	1	1	2	0
Res	3	1	0	3	1	6	1	0	0	5	1	1	2	0	1	1
R Ther	4	1	1	7	1	6	1	0	1	3	10	0	1	0	1	2
SW	17	0	1	1	1	4	1	0	2	6	1	2	3	0	3	0
Stf Phys	3	0	0	3	1	9	1	0	1	8	0	1	2	0	1	1
Trans	2	0	1	0	1	14	0	1	1	1	1	0	1	4	1	0
UC	32	1	1	8	9	3	1	0	3	10	1	2	3	1	2	1
X-ray	2	0	1	3	1	11	0	0	0	1	0	0	1	1	1	9

Table 2 Employees at the UIHC

Floor nurse	804
Food service	456
Housekeeper	356
IC nurse	190
Nurse Asst	386
Patient	482
Pharmacist	276
Phlebotomist	74
Physical/occupational therapist	90
Residents/fellows/Med students	666
Respiratory therapist	108
Social worker	114
Staff physician	760
Transporter	108
Unit clerk	126
X-ray technician	236
Total	5232

average number of contacts observed between members of groups i and j and N_j is the number of employees belonging to group j .⁵

3.1 Data limitations

Before moving to the computational model, we mention two limitations of our data set. Human contact networks frequently have a small number of individuals with a much larger than average number of contacts, perhaps differing by orders of magnitude. These individuals are often called *hubs*. These hubs have the potential to significantly influence disease transmission because they are highly likely to be infected, and if so, to pass on infections to a large number of individuals. Because our sample includes approximately 2.5% of the hospital worker population, we may be missing hubs in our sample if they exist in this setting. However, we note two things in relation to this: First, the relatively homogeneous workday responsibilities of workers within categories, likely limit the variation of contacts within a category. For instance two physicians or two floor nurses are likely constrained to see a relatively similar number of patients each day. This is unlike many other social network data sets, (such as general friendship or online networks) where there are not these work responsibility constraints on individual contacts. Thus the possibility of hubs with orders of

⁵ Recall that the observed contacts of our subjects may be anyone in the hospital not just one of the 140 shadowed subjects. Thus we divide by the total employees in the worker category not the number of shadow subjects in the worker category.

magnitude differences in numbers of contacts is more limited in our data context. Second, our data set is more comprehensive than any other within hospital contact data set in existence in terms of the worker categories included. Recall from earlier in the paper that the Ueno and Masuda data set only includes physicians, nurses, and patients in a hospital much smaller than is studied here. Our results below suggest that many of the most important groups in the hospital are not included in their study. Thus, at a minimum, our study highlights the importance of funding for studies that aim to collect even more comprehensive data sets that include individuals with non-patient care responsibilities and more comprehensively cover a larger share of hospital workers.

In addition, we note that we consider all contacts in our data set to be equivalent, or “equally weighted.” There may be some concern that not all contacts are created equal in our context. For instance, a contact between a physician and a patient that occurs during a physical examination, may be more likely to spread an infectious disease than other contacts in our data set. We attempt to control for this possibility later in the paper when we consider heterogeneous transmission rates and repeated contacts.

4 Computational model

As mentioned above, transmissions of infectious disease are not usually observable. Thus studying infectious disease transmission using an agent-based model can be a useful tool. In the remainder of the paper we model the spread of an infectious disease across the simulated hospital contact networks described above.

Once created, we use the contact network in a model of the spread of an infectious disease in the hospital as follows: Agents can be in one of three states, susceptible to being infected (S), infected and able to infect others (I), or recovered (and therefore immune) (R). We assume that each infected agent recovers after 5 periods which is in-line with the infectious period for influenza. Once recovered the agent enters state R and is therefore immune to further transitions to the infected state.

Initially, all agents in the model are in state S. Agents may be vaccinated against infection. Vaccinations occur only in the initial period of the model. Once vaccinated, an agent moves immediately from state S to state R and is thus immune for the remainder of the model.

In the initial period of the model, each agent in state S (all agents that have not been vaccinated) is subject to infection with probability $\alpha_0 = 0.005$. These are the agents of our model that seed the potential epidemic. Once these initial infections occur we assume that each contact in our network occurs once in each subsequent period of the model. If a contact occurs between a state I and a state S agent, the state S agent transitions to state I with probability α , which we vary across experiments. We continue the model until no agents remain in state I. In each period of the model we calculate the fraction of agents in each worker category in each state (S, I, or R).

For each of the results reported below we run 100 replications of each parameter set or computational experiment reported. The results reported are averages over these replications. In addition to the results reported below, we have studied a wide range of

parameters for our model and find the results reported below to be robust to changes in all of the parameters within reasonable bounds.⁶

4.1 Computational experiments

The purpose of the model is to estimate $m(\gamma_j)$ and the externality generated for the network of contacts in the UIHC shadow data and, in turn, to identify the classes of workers most important to vaccinate. This is a two step process. First we perform a series of base-line models as described above with none of the healthcare workers and patients vaccinated. From this baseline, we observe the rate of infection for each class of agents in the hospital population (15 worker groups and patients for a total of 16 groups). We denote the infection rate of group k in the *base* model as a function of the transmission rate α as $\pi_b^k(\alpha)$ and the overall infection rate in the entire population as a function of α as $\pi_b^0(\alpha)$. Second, we want to calculate the average and marginal infections generated by each group. To do so, we change the vaccination rate for each group, one group at a time, and re-run the model. As an example, we run the model with all floor nurses vaccinated and no other vaccinations and observe $\pi_1^0(\alpha)$. Then we run the model with all housekeepers vaccinated and no one else and observe $\pi_2^0(\alpha)$ and so on for each group. We then compare the change in the average number of infections between the models, $\delta(b, k) = (\pi_b^0(\alpha) - \pi_k^0(\alpha))N$, which is the difference between the overall infection rate in the base model with no vaccinations and the overall infection rate in the model with all of group k vaccinated, multiplied by the total population size N . Now, using the notation described above, the change in the number of infections $\delta(b, k)$ is equal to the change in number of people vaccinated, N^k , multiplied by the probability that each of these agents becomes infected if not vaccinated, multiplied by the number of additional infections each infected agent would generate. Simplifying, if we assume each agent infects the same number of individuals, we can write the average number of secondary infections generated per infected person in group k , as A^k and write:

$$\delta(b, k) = N^k(\pi_b^k(\alpha))(A^k) \tag{12}$$

One can then find an estimate of the average secondary infections per infected person as:

$$\hat{A}^k = \frac{\delta(b, k)}{\pi_b^k(\alpha)N^k} \tag{13}$$

Effectively, this process removes each group from the hospital, one at a time. We then can observe the effect of each individual worker group on the size of the modeled epidemic.

Instead of vaccinating all agents of a group at once, we can vaccinate a fraction of a group. As we change this fraction at intervals ΔN^k we can view the effect of increases

⁶ Additional results and details are available from the corresponding author.

in vaccination rates for each group (one at a time). We then have an estimate of the marginal infections prevented per vaccination as:

$$\hat{m}^k = \frac{\delta(b, k)}{\pi_b^k(\alpha) \Delta N^k} \quad (14)$$

We now proceed in two steps. First we investigate the effect of each group in total on the epidemic process by vaccinating an entire group one at a time and calculating \hat{A}^k for each group. We then choose a sample of interesting groups and study the details of the epidemic process as we vary the number of vaccinations performed in each of these groups, at specified intervals between 0 and 100%. Interestingly, as we vary the percent of each group vaccinated, we will see that there are different outcomes across these different groups in terms of marginal infections generated, probability of infection and the overall effect of a vaccination (in terms of reducing the number of infections).

4.1.1 Base experiments

We begin by varying the transmission rate, α , over the range [0.0004, 0.006] and observing the base infection rate $\pi_b^0(\alpha)$. The results are displayed in Fig. 1. As one can see in the figure, a sufficiently large transmission rate is needed to generate an epidemic of reasonable size. Further, as expected, the number of infections generated monotonically increases as a function of the transmission rate, α . Our primary interest is in intermediate ranges of epidemic outbreaks. If the transmission rate is too high then almost everyone in a population needs to be vaccinated in order to reach herd immunity. And, if the transmission rate is too low, then there is not a large need to worry about vaccine priority. Thus, we concentrate on two intermediate levels of the transmission rate $\alpha = 0.0030$ and $\alpha = 0.0035$. With no vaccinations, these levels yield an epidemic where between one-third and one-half of the population is infected over the course of the epidemic.

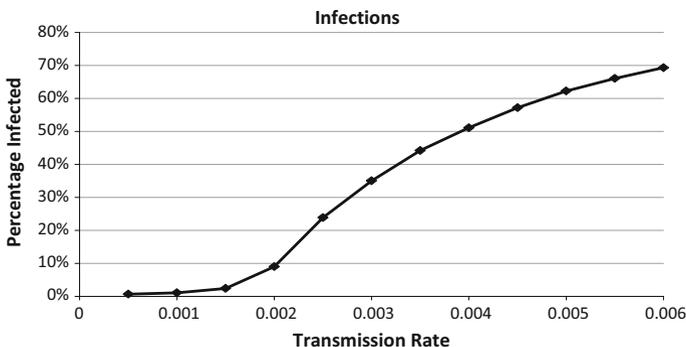


Fig. 1 Infection rate as a function of transmission rate

Table 3 Estimate of average infections prevented per vaccination, $\alpha = 0.0030$

Group	Secondary infections per person	Decrease in infections per vaccination
Floor nurse	3.21	1.91
Food service	1.56	0.26
Housekeeper	1.46	0.17
IC nurse	1.75	0.55
Nurse Asst	1.71	0.53
Patient	2.81	1.41
Pharmacist	1.41	0.26
Phlebotomist	4.57	3.24
Physical/occupational therapist	3.57	2.27
Residents/fellows/Med students	1.68	0.50
Respiratory therapist	3.02	1.50
Social worker	4.30	3.10
Staff physician	1.59	0.29
Transporter	2.60	1.21
Unit clerk	8.35	7.35
X-ray technician	1.78	0.44

4.1.2 Average effect of vaccinations

We now find the average effect of vaccinations across the hospital worker groups using the procedure described above for $\alpha = 0.0030$ and $\alpha = 0.0035$. We present results for the average “secondary infections” generated per infected person, \hat{A}^k , and the product of average infections and probability of infection, which yields the “decrease in infections per vaccination,” $\frac{\delta(b,k)}{N^k}$ in Tables 3 and 4.

From the decrease in infections per vaccination we have an indication of how much the vaccination of an individual group member is contributing to preventing the spread of an epidemic. The results of this experiment suggest where efforts should be directed in the event of an influenza vaccine shortage or in the event of the development of new disease for which a vaccine may be developed (e.g., avian flu, swine flu, etc.) but is initially in short supply until mass quantities may be made available. Note that some of the groups have vaccinations that prevent less than one infection per vaccination. This occurs because these groups have a low probability of infection and sufficiently low number of average infections that each member generates if infected. Groups with large decreases in infections per vaccination are the ones to prioritize in times of a vaccine shortage, assuming equal costs of infection.

In these experiments we see three clear groups that stand above the others in terms of the effect of vaccinations. For the parameters of the experiments, each vaccination

Table 4 Average infections prevented per vaccination, $\alpha = 0.0035$

Group	Secondary infections per person	Decrease in infections per vaccination
Floor nurse	2.73	1.91
Food service	1.27	0.29
Housekeeper	1.23	0.19
IC nurse	1.63	0.68
Nurse Asst	1.53	0.60
Patient	2.61	1.63
Pharmacist	1.22	0.29
Phlebotomist	3.71	3.07
Physical/ occupational therapist	2.97	2.24
Residents/ fellows/Med students	1.44	0.54
Respiratory therapist	2.26	1.43
Social worker	3.38	2.81
Staff physician	1.37	0.34
Transporter	2.06	1.17
Unit clerk	6.64	6.27
X-ray technician	1.31	0.44

of a unit clerk, social worker, and phlebotomist, results in a decrease of 3.1 infections or more on average for $\alpha = 0.0030$ and of 2.2 or more for $\alpha = 0.0035$.⁷ In addition vaccinating unit clerks is extremely effective; each unit clerk vaccination results in a decrease of over 7 infections for $\alpha = 0.0030$ and over 6 infections for $\alpha = 0.0035$. Somewhat surprisingly, some of the groups that are seen as central to the functioning of a hospital play a very small or moderate role in spreading an infectious disease. Vaccinating staff physicians results in a lower than average decrease in infections. We revisit this result in our discussion of transmission rates later in the paper.

Also of note, as one would expect, as the transmission rate increases, the probability of infection increases. But, this has the effect of making individual vaccinations less beneficial on average. Note that \hat{A}^k is smaller for each group for a higher transmission rate. This has the effect of lowering the variance of average infections. For the $\alpha = 0.0030$ case above the standard deviation is 1.79, and for the $\alpha = 0.0035$ case, the standard deviation is 1.41. As the transmission rate increases, a larger fraction of individuals are infected throughout the population. Thus there are more opportunities for each individual to be infected if she has not already been infected. Vaccinating a given person in the population will only prevent one of these multiple channels for infection. So, as the infection rate increases, the effectiveness of a vaccination becomes

⁷ We discuss characteristics of the network contacts of these groups below.

more uniform across the groups. This has direct policy applications. An infectious disease that is highly contagious could best be met with a uniform vaccination strategy since each individual in the population will create a similar level of infections on average. But an infectious disease with a low level of contagiousness could most effectively be met with a targeted vaccination campaign (Bansal and Pourbohloul 2007).

4.1.3 Marginal effect of vaccinations

We next look at the marginal effect of a vaccination as the number of vaccinations increase. We present results in Figs. 2, 3, 4, 5, 6 and 7 for five interesting worker group categories for the same two transmission rates discussed above. Unit Clerks, Social Workers and Phlebotomists are chosen because of their large number of secondary infections generated. We also choose Floor Nurses and Staff Physicians because of interest in the effect of worker groups with primary patient care responsibilities. In Fig. 2 we plot the marginal infections prevented per vaccination as a function of the number of vaccinations performed for the five groups. Recall that the number of marginal infections is the additional number of infections that an agent generates if the agent becomes infected. In Fig. 3 we plot the probability of infection for the five groups. And, in Fig. 4 we plot the product of marginal infections and probability of infection which yields the total number of infections prevented per vaccination. These figures all consider a transmission rate of 0.0030. Figures 5, 6, and 7 plot the same relationships for a transmission rate of 0.0035.

We begin with marginal infections. Recall that marginal infections may be increasing or decreasing in the number of vaccinations performed for small numbers of vaccinations (Boulier et al. 2007). (Here the number of vaccinations performed is small relative to the entire population as we are only vaccinating some members of one of the 15 groups. So, even if we vaccinate an entire group, this is a small number relative to the entire population.) Here, we see two interesting outcomes. First, we see that marginal infections for both Unit Clerks and Floor Nurses increase as more vaccinations are performed. For these two groups, each additional vaccination prevents a larger and larger number of infections. This is particularly extreme for a transmission rate of 0.0030 and the case of Unit Clerks where a small number of vaccinations results

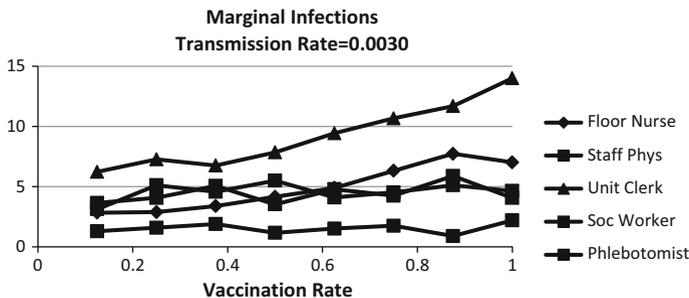


Fig. 2 Marginal infections per vaccination. Transmission rate = 0.0030

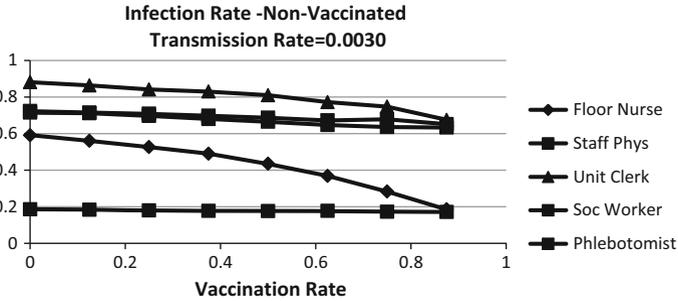


Fig. 3 Infection rate for non-vaccinated individuals. Transmission rate = 0.0030

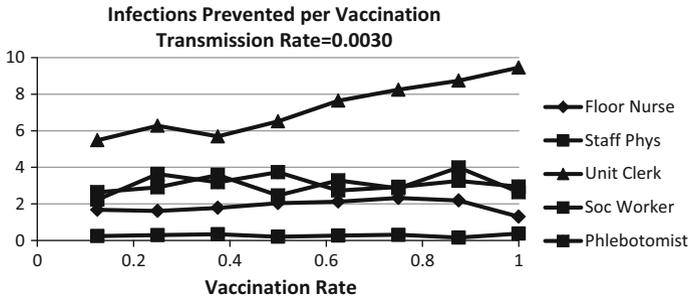


Fig. 4 Infections prevented per vaccination. Transmission rate = 0.0030

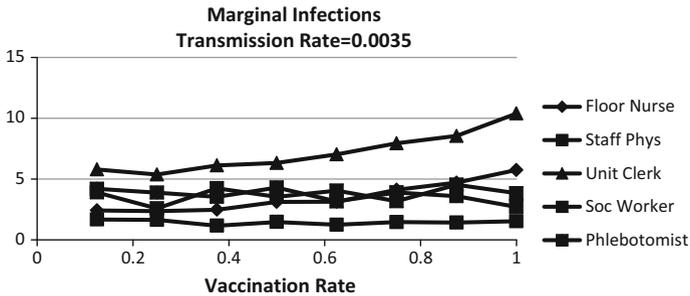


Fig. 5 Marginal infections per vaccination. Transmission rate = 0.0035

in about 6 marginal infections prevented and the last vaccination of a Unit Clerk results in about 14 marginal infections prevented. We see the same increasing relationship for a transmission rate of 0.0035 but to a smaller extent. (Recall that the effect of a vaccination becomes more homogeneous as the transmission rate increases.) We see a similar increasing relationship for Floor Nurses as well. However the increase is less pronounced. The other three groups have relatively flat plots of marginal infections. Thus for these three groups there is little difference between marginal infections generated and average infections generated.

Figures 3 and 6 show the change in infection rate for the non-vaccinated agents in the model. As the vaccination rate of a group increases all groups show at least a small

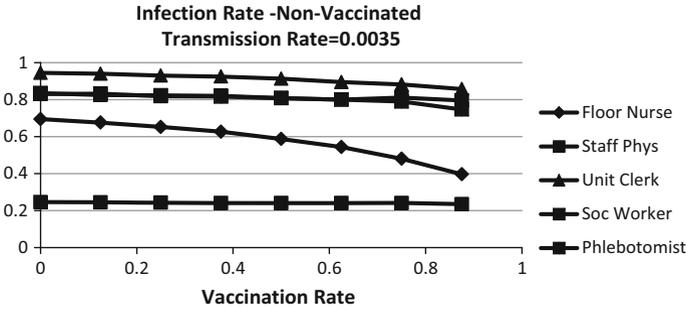


Fig. 6 Infection rate for non-vaccinated individuals. Transmission rate = 0.0030

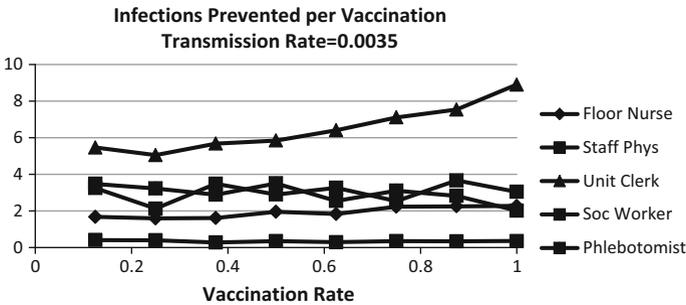


Fig. 7 Infections prevented per vaccination. Transmission rate = 0.0030

decrease in infection rate. For the Floor Nurse group this decrease is large, dropping from over 60 to under 20 % at a transmission rate of 0.0030 and from nearly 70 to 40 % at a transmission rate of 0.0035. The other groups show a much more modest decrease in infection rate.

The product of the two previously discussed statistics yields the number of infections prevented per vaccination. Again, Unit Clerks display a unique relationship in that the number of infections prevented per vaccination dramatically increases in the number of vaccinations performed. The other four groups result in much flatter plots. Thus again there is little difference between the marginal and the average for these groups.

There are two interesting points to be made from these results. The first is that the effect of vaccinating Unit Clerks in our data is most important both from a marginal and an average perspective regardless of how many vaccinations have been performed. Particularly, the marginal effect of vaccinating a Unit Clerk increases at a greater rate than the probability of infection for a non-vaccinated unit clerk falls. Thus, it is always more beneficial to vaccinate one more unit clerk as opposed to a worker from another group (assuming transmission rates are equal). The second is that there is little difference between the average and the marginal effect of a vaccination for the other groups considered here. This second point can be interpreted as good news from a policy making perspective in the sense that the optimal allocation of vaccinations does not switch as more of a group is vaccinated. In other words it is not the case that Group A is the optimal group to target up to some vaccination percentage, after

which Group B should be targeted. Switching such as that would indicate a much more complicated solution to the optimal vaccine allocation problem. Here, because there is little difference between the marginal and the average effect of a vaccination for most groups, one can pragmatically target the groups with the largest average effect of a vaccination.

4.2 Network characteristics of most important groups

We now move to discuss the important features of the contact network that creates the externality. As we will see below, it is not just the number of contacts that an agent has but also which specific agents and groups the agent contacts, as well as who the agent's contacts connect to in turn.

We begin by looking at some basic statistics of the contacts in our data in Table 5. For each group, the table displays the total number of contacts, the percentage of total contacts that are with members of an agent's own group, and the number of patient contacts. Total contacts and contacts with patients could be directly correlated with the likelihood of being infected and with passing on infections. The percentage of contacts within one's own group can indicate how varied one's network is and how widespread one's connections are. For instance having few contacts within one's own group provides the possibility of introducing an infection to other groups within the hospital. In addition the table also contains a common network characteristic measure, *betweenness centrality*, which we discuss below.

On a network, the geodesic distance, $g_{a,b}$, between nodes a and b is the length of the shortest path between the two nodes that traverses connections on the network measured as the number of connections traversed (or "hops" required) to reach node b from node a . As one measure of the *centrality* of a node on a graph one can calculate the average distance to all other nodes on the graph. Thus if a graph \mathcal{G} is composed of N nodes, average distance for node a is calculated as:

$$\bar{g}_a = \frac{\sum g_{a,b}}{N-1} \quad \forall b \neq a \in \mathcal{G} \quad (15)$$

A short average distance indicates that a node is close to other nodes on average and thus may be likely to be infected and to pass on infections. Thus it is sometimes considered a measure of the centrality or importance of a node in a network.

Betweenness is another measure of the centrality of a node in a network. Let C_{jk}^i be the proportion of all geodesics linking node j and node k which pass through node i . Let C^i be the sum of all C_{jk}^i for $i \neq j \neq k$. Let \bar{C}^i be the maximum possible value for C^i . (Normalized) *Betweenness* for node i , B^i is then:

$$B^i = C^i / \bar{C}^i \quad (16)$$

Betweenness for node i is therefore a measure of the proportion of shortest paths between nodes that go through node i .⁸

⁸ For more discussion of network centrality measures see Jackson (2008).

Table 5 Contact characteristics

Group	Total contacts	% Contacts in own group	Patient contacts	Betweenness
Floor nurse	214	20	36	0.031
Food service	40	30	14	0.008
Housekeeper	21	34	2	0.004
IC nurse	134	31	23	0.029
Nurse Asst	90	8	25	0.121
Pharmacist	34	12	3	0.008
Phlebotomist	51	10	40	0.036
Physical Ther	57	16	15	0.089
Resident/fellow	144	29	11	0.018
Respiratory Ther	85	38	12	0.062
Soc worker	50	8	4	0.113
Staff Phys	77	6	11	0.009
Transporter	32	25	14	0.026
Unit clerk	111	4	3	0.248
X-ray tech	70	53	19	0.014
Average	81	21	15	0.0544

Columns 1–3 taken from data. Column 4 taken from simulated networks

In the table below we report group level values for betweenness centrality. The values are created in the following manner. First we create a network using the same methods as described above. Second, we then calculate the betweenness centrality measurement for each agent in the simulated network. Third, we calculate the average value for each hospital worker group and report the result in the table below. This network variable is likely to be an indicator of importance for disease transmission in the network. If a hospital worker group has a high average betweenness value, then the nodes in this group are potentially important in passing infections on to other nodes as it plays a crucial role in location along many of the shortest paths between nodes in a network. As such, this measure should be closely related to the marginal infections that a group generates.

What is most interesting in Table 5 is the lack of a clear relationship between any of the variables in the first three columns and our previously-listed most important groups (highlighted in bold). The only measure that consistently aligns well is the betweenness measure. For a moment concentrate on the values in the first three columns. Each of these three plausibly important characteristics fail to display a meaningful relationship with the average or marginal infections generated. If we concentrate on the top three most important groups, some have relatively large numbers of contacts (unit clerks), although not the largest, while others have contacts significantly below the average (phlebotomists). Some have large numbers of patient contacts (phlebotomists) while others have some of the smallest number of patient contacts (unit clerks and social workers). One interesting thing that appears in the table is that phlebotomists have almost all of their contacts with patients and other phlebotomists. Thus the network

position of phlebotomists, is in some sense, very similar to that of patients. Overall, what these observations imply is that there is not likely to be a simple relationship (or a small set of simple relationships) indicating which individuals are most important to vaccinate by looking at easily observed worker interaction patterns. Instead the relationship depends on the intricate and complex web of relationships that make up the entire contact network of the hospital. This is what is captured in the betweenness centrality measure. Betweenness measures the percentage of shortest paths in the entire network on which an agent is located. If we remove agents with high betweenness measures (by vaccinations) from the disease propagation network, we disrupt the flow of an epidemic. For concreteness, the correlation between the betweenness measure and the average infections generated is about 0.84 for both $\alpha = 0.0030$ and $\alpha = 0.0035$.

4.3 Heterogeneous transmission rates and repeated contacts

As mentioned above, the primary focus of this paper concerns the effect of network position on the marginal infections generated within a hospital. Here, we consider two robustness checks to the results presented above. First, we discuss a comparison of the magnitude of the above described network effects and the effect of transmission rates on marginal infections for an interesting example group, staff physicians. Second, we do an additional set of experiments using an additional data set that includes observed repeated contacts. We perform these robustness checks for two main reasons: First, because of the job responsibilities of different worker groups, the different worker groups may have different transmission rates, durations of contacts, or frequency of contacts creating another source of heterogeneity in infections. As a few examples, a staff physician may be more likely to transmit an infection over the course of a patient exam that includes a series of physical contacts, compared to a nurse who has a brief arm's-length conversation with a hospital transporter. A floor nurse may have multiple interactions with the same patient during the course of the day.⁹

We begin this analysis by varying the transmission rate of staff physicians. (Recall that staff physicians created a lower than average number of infections in our earlier model.) Again, this is primarily to account for the fact that physicians may have longer duration contacts with patients and the contacts may more frequently involve physical touch. We now assign a special transmission rate to staff physicians that we denote as α_p . This will be the transmission rate for any contact between a physician and another agent where one of the agents is infected. We use $\alpha = 0.0030$ for all other contacts in the population and vary α_p from $\alpha_p = 0.0030$ and $\alpha_p = 0.0200$. As we do this we again measure average infections as described earlier and show the resulting average infections for unit clerks and staff physicians in Fig. 8.

Before we present the results we note that there are alternative ways to model this scenario. For instance, we could re-weight our contact matrix. If we knew, for instance, that physician to patient contacts lasted twice as long as other pairs of contacts in

⁹ As an alternative, instead of changing transmission rates, one could also consider re-weighting the contact matrix. For instance, again, some contacts, say between a physician and a patient may have a long duration. Thus one could more heavily weight the edges of the contact matrix that are of long duration or are of a type that could lead to a higher probability of transmission.

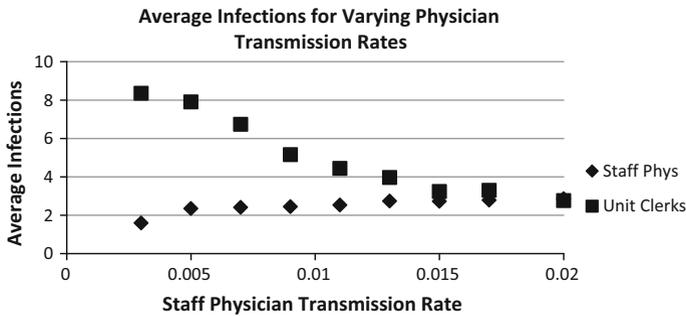


Fig. 8 Average infections as a function of physician transmission rate holding all other transmission rates constant at 0.0030

the hospital worker population, we could re-weight each physician to patient contact by a factor of two. But note that, on average, this is equivalent to increasing the transmission probability by a factor of two because the expected number of new infections is the number of susceptible to infected contacts in a period multiplied by the transmission rate. A doubling of the number of contacts is equivalent to a doubling of the transmission rate.¹⁰

Recall from our results above that when the transmission rates are equal the average secondary infections created by unit clerks are slightly greater than 8 and slightly greater than 1.5 for staff physicians. As we increase the transmission rate for staff physicians we see several things: First, as you would expect, the average infections for staff physicians increases. But the change is not large. For $\alpha_p = 0.0200$ the average secondary infections created by staff physicians is 2.86, slightly less than a two-fold increase and still well below the level of average infections noted for unit clerks when the transmission rates are equal. For unit clerks, as α_p increases, the average infections of unit clerks drops rapidly. This occurs for at least two reasons: One is that unit clerks become less important relative to staff physicians as α_p increases. Another is that, as α_p increases, the overall infection rates increase, and as we reported earlier, this causes average infections to become more uniform across groups. Overall, the level of average infections between these two groups does not become similar until the α_p increases to about 0.0150, a five-fold increase in the staff physician transmission rate. And, the average infections of staff physicians does not become greater than that of unit clerks until $\alpha_p = 0.0200$.

Most interesting about these results is the magnitude of the network effects relative to the magnitude of the transmission rates. In the case of staff physicians and unit clerks it takes a 5–6 times increase in the transmission rate of staff physicians to “make up” the difference in network position. This suggests that the network effect differences in average infections are very important in understanding overall transmission patterns.

¹⁰ Note that this ignores some subtleties such as one susceptible individual having multiple infected contacts in the same time period. But, for a relatively large population with a small per period incidence of infection, as we have here, these issues are negligible.

As a second robustness check we use an additional cut of our data that includes observed repeated contacts. In some instances members of the hospital population were observed to have come in contact more than once during the observation period. We use this data as one way to include the weighting of contacts mentioned above into the analysis. Table 6 displays the observed contacts that were observed and had occurred previously in our data. As you see in the table many of these contacts involved repeated interactions with patients (often by members of the nursing staff). We re-perform the analysis above with these additional contacts added into the data. The only additional change is that we modify the transmission rate to $\alpha = 0.0020$ so that the total number of infections in the population remains nearly constant compared to the non-repeat contact data. (Recall from earlier in the paper that a larger epidemic smooths out differences in the population groups and makes the average and marginal values more similar across groups. Thus we control for epidemic size by varying the transmission rate.)

We present the results in Table 7. You will notice in the table that most of the groups that previously had the largest effect still do. Unit clerks are still the most important group to vaccinate, but the difference in magnitude between unit clerks and other important groups is less than in the the previous model. Further, as one would expect, groups that have more repeated contacts, such as all types of nurses, become more important. The important point though, is that the relative ranking of most of the groups changes very little. Of course this is partially due to there being relatively few repeated contacts in the data set.

Taking these two robustness checks in combination, they demonstrate two important things. First, network structure is at least as important as transmission rate in determining the course of an epidemic in our data set. Second, while the data we have collected is not perfect in terms of comprehensiveness, the relative ranking of group importance appears relatively robust to alternative measures of network structure. A minimum, taken in combination, these results suggest a need for a greater emphasis on network based data collection in order to better understand both micro level and macro level epidemiology.

4.4 Protecting patients and physicians

We now make a small shift in focus. Generally, a hospital's primary goal is to restore or improve patient health. Thus prioritizing healthcare worker vaccinations so as to best protect patients may be a legitimate goal of a hospital. In other words hospital administrators may care about protecting patients from infection as much, or more, than they do about protecting the entire hospital population from infection. Of course these may be closely related goals. In addition, in a large scale epidemic, it may be of great importance to have a healthy staff of physicians to treat patients. Thus protecting physicians may be another important goal in vaccine priority within a hospital. In Tables 8 and 9 we display the same relationships as displayed in our initial results section above but this time only with regard to *patient* infections generated and *staff physician* infections generated, not infections in the entire hospital.

Table 6 Additional repeated contacts between worker categories per 8 h

	F N	Food	HK	IC N	N Asst	Pat	Pharm	Phleb	P Ther	Res	R Ther	SW	Stf Phys	Trans	UC	X-Ray
F N	10	0	0	0	6	25	0	0	1	1	0	0	0	0	5	0
Food	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
HK	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
IC N	0	0	0	13	0	19	0	0	0	2	1	0	2	0	1	0
N Asst	6	0	0	3	2	15	0	0	0	0	0	0	0	0	4	0
Pharm	3	0	0	1	0	1	0	0	0	4	0	0	2	0	1	0
Phleb	0	0	0	0	0	17	0	0	0	0	0	0	0	0	0	0
P Ther	3	0	0	1	1	5	0	0	2	0	0	0	0	0	1	0
Res	2	0	0	6	1	5	1	0	0	16	1	0	6	0	0	0
R Ther	2	0	0	9	0	6	0	0	0	2	6	0	1	0	0	0
SW	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Stf Phys	0	0	0	1	0	1	0	0	0	18	0	0	1	0	0	0
Trans	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
UC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
X-Ray	2	0	0	2	0	9	0	0	0	0	0	0	0	0	0	9

Table 7 Estimate of average infections prevented per vaccination including repeated contacts, $\alpha = 0.0020$

Group	Secondary infections per person	Decrease in infections per vaccination
Floor nurse	2.94	1.60
Food service	1.30	0.17
Housekeeper	1.64	0.14
IC nurse	2.90	1.53
Nurse Asst	2.12	0.77
Patient	3.76	2.26
Pharmacist	1.45	0.30
Phlebotomist	5.00	3.86
Physical/occupational therapist	3.39	2.12
Residents/ fellows/Med students	1.93	0.68
Respiratory therapist	3.63	2.15
Social worker	3.59	2.07
Staff physician	1.79	0.37
Transporter	2.09	0.79
Unit clerk	5.97	4.57
X-ray technician	1.92	0.61

In this analysis we see very similar results to the overall population results. Beginning with patients, the four groups (unit clerk, social worker, physical and occupational therapist, and phlebotomist) that played the most important role in transmitting to the hospital population as a whole also play an important role in their effect on patients specifically. However, we see some difference in the two groups of models. First, groups that have more direct patient contacts increase in importance. For instance, phlebotomists replace unit clerks as the most important group. Second new groups emerge as important for transmitting to patients. For instance, hospital transporters are among the top four groups in transmitting to patients but are significantly below the average in terms of transmissions to the general population. With this in mind it seems that giving vaccination priority to health care workers with direct patient contacts is more important for protecting patients than it is for protecting the general population, as one would expect. But still, some of the groups with the largest impact on infecting patients have few direct patient contacts (unit clerks and social workers, for example).

For staff physicians we see similar results. Again, the four most important groups remain important, but other groups such as floor nurses increase in importance when considering staff physicians specifically.

To summarize the results of this section, the same groups that create infections in the general population also create infections in the patient and staff physician population. But, groups that have direct patient or staff physician contacts have increased importance. Still, one should not ignore other groups central to the network of the hos-

Table 8 Estimate of average patient infections prevented per vaccination, $\alpha = 0.0030$

Group	Secondary patient infections per person	Decrease in patient infections per vaccination
Floor nurse	0.34	0.20
Food service	0.14	0.02
Housekeeper	0.05	0.01
IC nurse	0.06	0.02
Nurse asst	0.12	0.04
Pharmacist	0.00	0.00
Phlebotomist	1.25	0.89
Physical/ occupational therapist	0.47	0.30
Residents/ fellows/ Med students	0.07	0.02
Respiratory therapist	0.29	0.14
Social worker	0.32	0.23
Staff physician	0.07	0.01
Transporter	0.50	0.23
Unit clerk	0.68	0.60
X-ray technician	0.17	0.04

pital that have only few direct patient or staff physician contacts (e.g., social workers and unit clerks).

5 Conclusion

We utilize a newly collected data set on contacts of health care workers at a large university hospital to estimate network effects for infectious disease transmission. Interestingly the most important groups to vaccinate tend to have heterogeneous contacts throughout the hospital. Groups such as social workers and unit clerks are very important to vaccinate even though they have been given low priority in past vaccine campaigns because of their relatively limited number of patient contacts. For instance, the CDC recommended in their Interim Influenza Vaccination Recommendations in 2004–2005 that influenza vaccine priority be given to “health-care workers involved in direct patient care” and further stated, “Persons who are not included in one of the priority groups described above should be informed about the urgent vaccine supply situation and asked to forego or defer vaccination.” This mismatch of scientific results and past policy decisions suggests that future research in this area is warranted especially when one considers the public health dangers associated with the emergence of avian flu, a more lethal version of swine flu, or recent dangers such as SARS. With

Table 9 Estimate of average staff physician infections prevented per vaccination, $\alpha = 0.0030$

Group	Secondary physician infections per person	Decrease in physician infections per vaccination
Floor nurse	0.23	0.14
Food service	0.04	0.01
Housekeeper	0.00	0.00
IC nurse	0.14	0.04
Nurse Asst	0.03	0.01
Patient	0.18	0.09
Pharmacist	0.05	0.01
Phlebotomist	0.20	0.14
Physical/ occupational therapist	0.19	0.12
Residents/ fellows/ Med students	0.14	0.04
Respiratory therapist	0.22	0.11
Social worker	0.38	0.27
Transporter	0.04	0.02
Unit clerk	0.61	0.54
X-ray technician	0.05	0.01

that stated, we want to be careful to recognize that one reason to vaccinate primary care providers is to assure individuals are available to care for the sick. This important incentive is outside the scope of our model.

The results of this paper lead to important public policy considerations. Specifically, hospital workers with a low probability of infection may be likely to ignore recommendations for vaccination even if they are central to the spread of an infectious disease. One way to increase the overall vaccination level is with a subsidy program. But, as the results in this paper show, not all hospital workers are equal in terms of the positive externality generated by a vaccination. Because of the heterogeneous contacts throughout the hospital, some workers are more important to the spread of an infectious disease than others. Thus if hospitals and other public health organizations want to efficiently distribute vaccines they need to target specific worker groups, perhaps by allocating subsidies, on the basis of discrepancies in probability of infection and marginal infections generated. This paper is the first to use specific micro-level contact data within a hospital to guide policy makers and public health officials in this endeavor.

To be clear, these results are not meant to be specifically calibrated to measure the exact effect of vaccinations in these groups. Instead our hope is that the orderings of the hospital worker groups (which are robust across the parameters that we have explored) indicate where public health officials can effectively intervene in order to prevent

widespread epidemics within hospitals. And these experiments reveal interesting and surprising groupings. Prior to this study, as quoted above, it had been argued that groups like unit clerks be excluded from influenza vaccine campaigns, in times of vaccine shortages, because of their minimal patient contacts. The results of this study suggest that decisions such as these need to be more fully explored.

References

- Bansal S, Pourbohloul B, Meyer LA (2007) A comparative analysis of influenza vaccination programs. *PLoS Med* 3:e387
- Boulter BL, Datta TS, Goldfarb RS (2007) Vaccination externalities. *B E J Econ Anal Policy* 7. Article 23
- CDC (2002) Epidemiology and prevention of vaccine-preventable diseases, 7th edn. Public Health Foundation, Washington
- CDC (2003) Prevention and control of influenza: recommendations of the advisory committee on immunization practice (ACIP). *Morb Mortal Wkly Rep* 52:1–36
- Chowell G, Castillo-Chavez C, Fenimore PW, Kribs-Zaleta CM, Arriola L, Hyman JM (2004) Model parameters and outbreak control for SARS. *Emerg Infect Dis* 10:1258–1263
- Comins HN, Hassell MP, May RM (1992) The spatial dynamics of host parasitoid systems. *J Anim Ecol* 61:735–748
- Elder AG, O'Donnell B, McCrudden EA et al (1996) Incidence and recall of influenza in a cohort of glasgow healthcare workers during the 1993–94 epidemic: results of serum testing and questionnaire. *Br Med J* 313:1241–1242
- Francis PJ (2004) Optimal tax/subsidy combinations for flu season. *J Econ Dyn Control* 28:2037–2054
- Goeffard P-Y, Philipson T (1997) Disease eradication: private versus public vaccination. *Am Econ Rev* 87:222–230
- Grenfell B, Harwood J (1997) (Meta)Population dynamics of infectious diseases. *Trends Ecol Evol* 12:395–399
- Jackson MO (2008) Social and economic networks. Princeton University Press, Princeton
- Jarvis WR (1996) Selected aspects of the socioeconomic impact of nosocomial infections: morbidity mortality, cost, and prevention. *Infect Control Hosp Epidemiol* 17:552–557
- Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A* 115:700–721
- Philipson T (2000) Epidemiology economic, diseases infectious. In: Culyer AJ, Newhouse JP (eds) *Handbook of health economics*, vol 1B. Elsevier, NY, pp 1761–1797
- Smith CEG (1970) Prospects of the control of disease. *Proc R Soc Med* 63:1181–1190
- Smith NM, Bresee JS (2006) Prevention and control of influenza: recommendations of the advisory committee on immunization practices, morbidity and mortality weekly report 55, Centers for Disease Control and Prevention, pp 1–42
- Thompson WW, Shay DK, Weintraub E et al (2003) Mortality associated with influenza and respiratory syncytial virus in the United States. *J Am Med Assoc* 289:179–186
- Ueno T, Masuda N (2009) Controlling nosocomial infection based on the structure of hospital social networks. *J Theor Biol* 254:655–666
- Wallinga J, Edmunds WJ, Kretzschmar M (1999) Perspective: human contact patterns and the spread of airborne infectious diseases. *Trends Microbiol* 7:372–377
- Weingarten S, Riedinger M, Bolton LB et al (1989) Barriers to influenza vaccination acceptance. A survey of physicians and nurses. *Am J Infect Control* 17:202–207