Virulence-mediated infectiousness and activity trade-offs and their impact on transmission potential of patients infected with influenza

2019-08-18

# Authors

* Brian McKay1, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA
* Mark Ebell, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA
* Ariella Perry, Colorado Department of Public Health and Environment, Denver, CO, USA
* Ye Shen, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA
* Andreas Handel1, Department of Epidemiology and Biostatistics, The University of Georgia, Athens, GA, USA

1 Corresponding Authors: Brian McKay and Andreas Handel

Address: 101 Buck Rd, Miller Hall, Athens, Georgia 30606

Email: [bmckay52@uga.edu](mailto:bmckay52@uga.edu) or [ahandel@uga.edu](mailto:ahandel@uga.edu)

#Abstract

Most communicable diseases have some amount of virulence that induces infectiousness-enhancing symptoms. However, too much virulence can cause host morbidity and a reduction in transmission potential. For human diseases, the reduction in transmission opportunities is commonly caused by reduced activity. There is limited data regarding the potential impact of virulence on transmission potential. We analyzed data of 326 influenza patients at a university health center during the 2016/2017 influenza season. We classified symptoms as infectiousness-related or morbidity-related and calculated two scores. The scores were used to explore the relationship between infectiousness, morbidity, and activity levels. We found a decrease in activity levels with increasing morbidity scores. There was no consistent pattern between activity level and infectiousness score. We also found a positive correlation between the morbidity and infectiousness scores. Our results provide evidence that for influenza, increasing virulence leads to increased infectiousness and reduced activity. This trade-off determines the transmission potential. Our findings suggest that a reduction of systemic symptoms may increase host activity without reducing infectiousness. Therefore interventions should target both systemic and infectiousness related symptoms to reduce overall transmission potential. Our findings can also inform simulation models to investigate the impact of different interventions on transmission.

*Keywords*: Infectious diseases; Influenza; Trade-off; Transmission; Virulence

# Introduction

Many infectious diseases cause symptoms in at least some of their hosts. Often, those symptoms increase the host’s infectiousness and facilitate the transmission of the pathogen (1–3). Coughing and sneezing for respiratory infections are prime examples. On the other hand, symptoms that are too severe may reduce host activity or in extreme cases cause host death, reducing transmission opportunities. The trade-off hypothesis describes the relationship between virulence and transmission potential (10, p.@Bull1994,4,5,6,7,8,9) and predicts that an intermediate level of virulence leads to maximum fitness (usually quantified by the reproductive number) for the pathogen. At such an optimal level of virulence, the pathogen maximizes transmission by inducing symptoms that increase a host’s infectiousness, while minimizing transmission-reducing morbidity symptoms. The optimal virulence level can depend on both population-level and within-host level processes, the implications of which have been theoretically explored previously (4,5,12–21).

The most commonly discussed and studied trade-off is between increasing transmission potential due to increased host infectiousness and decreasing transmission potential due to host mortality (5). While, this likely applies to many animal diseases and some human diseases (e.g., viral hemorrhagic diseases (22)), for most human pathogens mortality is low, and it is more likely that increased virulence leads to reduced host activity and thus reduced transmission opportunities. Sub-lethal impacts such as weight loss and effects on host fitness have been suggested (5,6,23,24), and interactions between symptoms, activity, and transmission potential have been recognized (25). Despite this, there is very little data available for human pathogens. One study on *Plasmodium falciparum* infections in humans showed an increase in transmission potential as virulence, quantified by mortality, increased, with no apparent trade-off (26). A study in HIV infected individuals showed a negative relationship between duration of asymptomatic infection and viral load and a positive relationship between infectiousness and viral load with optimal transmission potential occurring at an intermediate viral load (27). As far as we are aware, no studies for any other human pathogens have examined data to directly determine the relationship between virulence and transmission.

Here, we investigate this relationship for influenza. Influenza induces symptoms in around 84% of infected individuals (28). Some of the symptoms, such as coughing and sneezing, likely enhance transmission by increasing infectiousness of a host. A recent study provided estimates for the transmission potential of symptomatic versus asymptomatic individuals and found that individuals with symptomatic infections are about 3-12 times as infectious as persons with asymptomatic infections (29). Other symptoms, such as fever, body aches, and general malaise are more likely to lead to a reduction in transmission by reducing host activity. A previous study on influenza in 146 adults and children in the United Kingdom found that healthy individuals had a mean of 12.72 contacts per day, while sick individuals only had a 3.58 (1). The study also showed that the number of contacts decreased as the number of symptoms increased. These studies suggest that there might be a trade-off between infectiousness and activity for influenza, which together determines overall transmission. In this study, we investigate this relationship.

# METHODS

## Data Collection

Our patient population consisted of students who made an appointment at the university health center of a large research university during the 2016/2017 flu season. All patients with a primary complaint of respiratory infections were required to fill out an electronic questionnaire. The questionnaire collected data about their current symptoms and activity level. Patients were required to respond to all symptom-related questions when they scheduled their appointments. We included all symptoms collected by the questionnaire in this analysis. The complete questionnaire is available in the supplementary material.

For the symptoms of weakness and body aches, the patient graded the severity of the symptom as none, mild, moderate, and severe. The patient recorded all other symptom data as present or absent. The patient also reported any changes in their normal behavior. Patients describe their activity level as a number between 0 and 10, with 10 indicating no change in regular activity and 0 being bedridden.

The study population includes all patients with a diagnosis of influenza. The data and results presented in the main text includes patients diagnosed with a rapid antigen or rapid PCR test. To address the impact of the influenza diagnosis method we performed the same analyses for all patients diagnosed with influenza regardless of the method used. The results are in the supplementary material.

The institutional review board approved the study protocol. Data on PCR results for patients is from a study funded by Roche Diagnostics.

## Data Cleaning

We cleaned the data to format the variables and to check for variables with potential errors or missing entries. During the cleaning process, we removed uninformative variables which we defined as any symptoms found to occur in less than 5% of patients. The symptoms of blurred vision and hearing loss both had a prevalence of less than 5%, so they were not considered for further analysis. To allow easy comparison of all symptom variables, we dichotomized weakness and body aches to “absent” or “present”.

## Analysis

We assessed the univariate relationships between activity and each symptom using linear regression treating activity level as a continuous variable. We also performed multiple linear regression. We determined the variables to include in our final model with a sequential forward floating selection, minimizing the root mean square error (RMSE) on test data through a 5-fold cross validation (20 times repeated) (30).

Next, we constructed two cumulative scores, one for overall infectiousness and one for overall morbidity. To that end, we divided all symptoms into those related to infectiousness and those related to morbidity. We defined morbidity symptoms as symptoms that influence overall feelings of well-being but are not associated with infectiousness. Infectiousness symptoms are any symptoms that could plausibly contribute to passing the virus from an infected host to another. Importantly, the grouping of variables to either one of these categories and inclusion of symptoms in the scores was based on *a priori* medical and biologic considerations, independently of any observed correlation with activity level. Doing so prevents any circular reasoning since only including symptoms correlated with activity would, of course, generate a score which would match the impact on activity level.

To prevent redundant variables from being included in the score, we calculated Yule’s Q between symptoms within each category (31). Only one of a pair of symptoms was incorporated in the score if the correlation coefficient was higher than 0.9 (32). We also performed a sensitivity analysis using 0.75 as the cut off for identifying redundant symptoms. The results of this sensitivity analysis is in the supplementary material.

For highly correlated symptom pairs, we included the one in the score with the best balance (closest to 50%) of symptom presence or absence. We summed the symptoms in each category based on absence or presence, creating two scores. Correlations between the infectiousness score, morbidity score, and activity were assessed using Spearman correlation (33,34) and the generalized Mantel–Haenszel procedure (35,36). Linear regression lines are included in the plots to help visualize the relationships.

All analyses were completed using R (version 3.5.3) (37). We used the mlr package for cross-validation (38), vcdExtra to compute Yule’s Q and the CHM trend test (39), DescTools to compute Spearman’s rank correlation coefficient and corresponding confidence intervals (40). The supplementary material includes all of the code and data required to reproduce the results.

# RESULTS

##Study population During the study period, 2326 patients had a respiratory complaint and filled out the questionnaire. Among those, 326 had a lab-based diagnosis of influenza (PCR or rapid antigen). The following analyses focus on those patients since they are most likely to actually be infected with influenza. For analyses of patients who received a flu diagnosis with either the tests or empirically from a physician, see the supplemental material.

Those patients with influenza reported activity levels ranging from 0 to 10 with a median of 4 (SM Figure 1). All of the patients reported symptoms, with only 14% reporting 10 or fewer (out of a total of 25). The most common symptom was coughing and the least common was abdominal pain (Table 1).

Table 1: **Out of the 326 patients included the table shows the number of patients who reported having the following symptoms and the corresponding percentage.**

|  |  |
| --- | --- |
|  | Overall |
| n | 326 |
| Abdominal Pain = Yes (%) | 39 (12.0) |
| Breathlessness = Yes (%) | 131 (40.2) |
| Chest Congestion = Yes (%) | 197 (60.4) |
| Chest Pain = Yes (%) | 110 (33.7) |
| Chills/Sweats = Yes (%) | 287 (88.0) |
| Cough = Yes (%) | 308 (94.5) |
| Diarrhea = Yes (%) | 40 (12.3) |
| Ear Pain = Yes (%) | 59 (18.1) |
| Eye Pain = Yes (%) | 47 (14.4) |
| Fatigue = Yes (%) | 304 (93.3) |
| Headache = Yes (%) | 272 (83.4) |
| Itchy Eyes = Yes (%) | 73 (22.4) |
| Myalgia = Yes (%) | 290 (89.0) |
| Nasal Congestion = Yes (%) | 257 (78.8) |
| Nausea = Yes (%) | 119 (36.5) |
| Runny Nose = Yes (%) | 235 (72.1) |
| Sleeplessness = Yes (%) | 183 (56.1) |
| Sneeze = Yes (%) | 179 (54.9) |
| Sore Throat = Yes (%) | 268 (82.2) |
| Subjective Fever = Yes (%) | 242 (74.2) |
| Swollen Lymph Nodes = Yes (%) | 131 (40.2) |
| Tooth Pain = Yes (%) | 60 (18.4) |
| Vomiting = Yes (%) | 44 (13.5) |
| Weakness = Yes (%) | 307 (94.2) |
| Wheezing = Yes (%) | 106 (32.5) |

##Univariate and subset selection We assessed correlations between activity level and each symptom in a univariate linear analysis (Table 2). All of the statistically significant symptoms had a negative correlation with activity level (Table 2). Next, we considered a multi-variable regression model and performed variable selection based on cross-validated minimization of RMSE. We found that the best performing model was one that included chest congestion, headache, sleeplessness, subjective fever, vomiting, and weakness (Table 2). While vomiting is not a common symptom of influenza, in those patients who did report vomiting it lead to major reductions in their activity.

Table 2: **Results of the univariate and multivariate linear regression of symptoms and activity. The coefficients are the estimated effect on activity when the symptom is present. The multivariate model was selected with a sequential forward floating selection, minimizing the RMSE on test data through a 5-fold cross validation (20 times repeated). 95%CI = The 95% confidence interval for the coefficient.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dependent: Activity Level |  | Mean (sd) | Coefficient (univariable) | Coefficient (multivariable) |
| Abdominal Pain | No | 4.4 (2.6) | - | - |
|  | Yes | 3.4 (2.6) | -1.02 (-1.91 to -0.14, p=0.023) | - |
| Breathlessness | No | 4.3 (2.7) | - | - |
|  | Yes | 4.1 (2.6) | -0.22 (-0.81 to 0.37, p=0.466) | - |
| Chest Congestion | No | 4.7 (2.9) | - | - |
|  | Yes | 4.0 (2.5) | -0.72 (-1.31 to -0.14, p=0.016) | -0.54 (-1.08 to 0.01, p=0.052) |
| Chest Pain | No | 4.4 (2.7) | - | - |
|  | Yes | 4.0 (2.6) | -0.43 (-1.05 to 0.18, p=0.162) | - |
| Chills/Sweats | No | 5.7 (2.8) | - | - |
|  | Yes | 4.1 (2.6) | -1.66 (-2.53 to -0.78, p<0.001) | - |
| Cough | No | 4.2 (3.2) | - | - |
|  | Yes | 4.3 (2.6) | 0.10 (-1.17 to 1.37, p=0.877) | - |
| Diarrhea | No | 4.3 (2.7) | - | - |
|  | Yes | 3.6 (2.6) | -0.72 (-1.60 to 0.15, p=0.106) | - |
| Ear Pain | No | 4.4 (2.6) | - | - |
|  | Yes | 3.7 (2.7) | -0.69 (-1.44 to 0.06, p=0.070) | - |
| Eye Pain | No | 4.2 (2.6) | - | - |
|  | Yes | 4.4 (2.9) | 0.17 (-0.66 to 0.99, p=0.689) | - |
| Fatigue | No | 5.8 (2.4) | - | - |
|  | Yes | 4.1 (2.6) | -1.67 (-2.81 to -0.53, p=0.004) | - |
| Headache | No | 5.6 (2.7) | - | - |
|  | Yes | 4.0 (2.6) | -1.57 (-2.33 to -0.81, p<0.001) | -1.15 (-1.89 to -0.42, p=0.002) |
| Sleeplessness | No | 4.9 (2.7) | - | - |
|  | Yes | 3.7 (2.5) | -1.17 (-1.74 to -0.60, p<0.001) | -0.93 (-1.47 to -0.40, p=0.001) |
| Itchy Eyes | No | 4.4 (2.7) | - | - |
|  | Yes | 3.7 (2.6) | -0.74 (-1.43 to -0.05, p=0.035) | - |
| Myalgia | No | 5.4 (2.9) | - | - |
|  | Yes | 4.1 (2.6) | -1.24 (-2.15 to -0.32, p=0.008) | - |
| Nasal Congestion | No | 4.4 (2.7) | - | - |
|  | Yes | 4.2 (2.7) | -0.24 (-0.95 to 0.47, p=0.507) | - |
| Nausea | No | 4.6 (2.7) | - | - |
|  | Yes | 3.6 (2.4) | -1.06 (-1.65 to -0.47, p<0.001) | - |
| Sore Throat | No | 4.6 (2.6) | - | - |
|  | Yes | 4.2 (2.7) | -0.37 (-1.13 to 0.38, p=0.330) | - |
| Runny Nose | No | 4.7 (2.6) | - | - |
|  | Yes | 4.1 (2.7) | -0.55 (-1.20 to 0.09, p=0.091) | - |
| Sneeze | No | 4.7 (2.7) | - | - |
|  | Yes | 3.9 (2.6) | -0.71 (-1.29 to -0.14, p=0.015) | - |
| Subjective Fever | No | 5.2 (2.5) | - | - |
|  | Yes | 3.9 (2.6) | -1.32 (-1.96 to -0.67, p<0.001) | -0.93 (-1.56 to -0.30, p=0.004) |
| Swollen Lymph Nodes | No | 4.5 (2.7) | - | - |
|  | Yes | 3.9 (2.5) | -0.54 (-1.13 to 0.05, p=0.073) | - |
| Tooth Pain | No | 4.3 (2.6) | - | - |
|  | Yes | 4.0 (2.7) | -0.28 (-1.03 to 0.47, p=0.463) | - |
| Vomiting | No | 4.5 (2.7) | - | - |
|  | Yes | 2.8 (2.1) | -1.67 (-2.49 to -0.84, p<0.001) | -1.46 (-2.24 to -0.68, p<0.001) |
| Weakness | No | 6.6 (2.3) | - | - |
|  | Yes | 4.1 (2.6) | -2.46 (-3.67 to -1.26, p<0.001) | -1.40 (-2.57 to -0.23, p=0.019) |
| Wheezing | No | 4.4 (2.7) | - | - |
|  | Yes | 3.9 (2.5) | -0.54 (-1.16 to 0.08, p=0.085) | - |

##Computation of infectiousness and morbidity scores We divided symptoms into infectiousness-related and morbidity-related and used them to construct an infectiousness and morbidity score. To prevent circular reasoning regarding associations between those scores and activity, the division and potential inclusion of symptoms into each score was done based purely on biological considerations, without regard for any associations with activity found in the previous analysis. We classified coughing, chest congestion, sneezing, runny nose, and nasal congestion as infectiousness related symptoms. The symptoms of subjective fever, having chills and or sweats, body aches, weakness, headache, fatigue, sleeplessness, breathlessness, wheezing, chest pain, sore throat, abdominal pain, diarrhea, nausea, vomiting, ear pain, tooth pain, eye pain, itchy eyes, and swollen lymph nodes were classified as morbidity related symptoms.

Among the symptoms related to infectiousness only cough and chest congestion correlated with each other at a level of greater than 0.9 (SM Figure 2). We kept chest congestion since it was more balanced then cough, which was present in 94% of patients. Among the morbidity symptoms, only vomiting and weakness correlated greater than 0.9 (SM Figure 3). Vomiting was included in the score since it was more balanced then weakness, which was present in 94% of patients. For the results of the sensitivity analysis using 0.75 as the cut off for identifying redundant symptoms, see the supplementary material.

The infectiousness score included all the candidate symptoms except cough, and the morbidity score included all the candidate symptoms except weakness. Each symptom present in a patient, contributed one point to its respective score. The calculated infectiousness score had a possible range of 0 to 4, and the morbidity score had a possible range of 0 to 19.

The median infectiousness score was 3. Only 17 patients had an infectiousness score of 0, 39% had a score of 2 or less, and 29% of patients had the maximum possible score of 4 (Figure 1A).

The mean morbidity score was 8.6, and no patients had a morbidity score of 0, 1, 18 or 19 (Figure 1B). The centered distribution was expected since all the patients felt sick enough to seek medical care, but none were sick enough to require urgent care or hospitalization.

Figure 1: (A) The distribution of infectiousness score with counts for each level. (B) The distribution of the morbidity score with counts for each level. There are no patients with a score of 0, 1, 18, and 19.

Figure 1: (A) The distribution of infectiousness score with counts for each level. (B) The distribution of the morbidity score with counts for each level. There are no patients with a score of 0, 1, 18, and 19.

##Impact of infectiousness score on activity Analysis of the association between the infectiousness score and the patient’s self-reported activity level suggests that the value of this score has a small impact on the activity level of a patient, with higher infectiousness correlating with reduced activity. Spearman’s rank correlation indicates negative relationship ( -0.18 (95% CI: -0.28, -0.07)) and the Cochran-Mantel-Haenszel trend test is statistically significant ( 8.56, 1, < 0.01) (Figure 2). Note however that the data suggest that the relationship between infectiousness and activity is not linear, but instead curved, with lower activity at both the low and high infectiousness score and maximum activity at intermediate infectiousness. We cannot think of a biological mechanism that might lead to this pattern. The reason the overall trend is negative is likely due to the larger sample sizes for infectiousness scores 2-4. Given that the observed negative trend is small and doesn’t show a monotone decline, it is most reasonable to assume based on this data that there is no meaningful relationship between infectiousness score and activity level.

Figure 2: Activity level for each level of the infectiousness score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

Figure 2: Activity level for each level of the infectiousness score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

##Impact of morbidity score on activity Analysis of the association between the morbidity score and the patient’s self-reported activity level suggests that higher morbidity score is associated with reduced activity levels. Spearman’s rank correlation indicates negative relationship ( -0.33 (95% CI: -0.42, -0.23)) and the Cochran-Mantel-Haenszel trend test is statistically significant ( 39.34, 1, < 0.01) (Figure 3). The observed pattern is consistent and clear, with a reduction of 85% in mean activity level going from the lowest to the highest morbidity score.

Figure 3: Activity level for each level of the morbidity score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

Figure 3: Activity level for each level of the morbidity score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

##Impact of morbidity score on infectiousness score  
Analysis of the relationship between the morbidity and infectiousness scores show a positive correlation. Spearman’s rank correlation indicates positive relationship ( 0.28 (95% CI: 0.17, 0.37)) and the Cochran-Mantel-Haenszel trend test is statistically significant ( 24.45, 1, < 0.01) (Figure 4). Apart from the mean activity levels for very low morbidity score values (with very small sample sizes), the pattern is consistent and clear, with an increase of 33% in the mean infectiousness score going from the lowest to the highest morbidity score.

Figure 4: Infectiousness score for each level of the morbidity score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

Figure 4: Infectiousness score for each level of the morbidity score. Red diamonds indicate the mean. The solid blue line is the linear regression fit. The shaded area is the 95% confidence interval for the linear regression

## Conceptualizing our results

The hypothesis of virulence-transmission trade-off as explained in the introduction assumes that increasing levels of virulence lead initially to an increase in transmission-enhancing symptoms, but at some point, virulence leads to transmission-reducing symptoms, with an optimum for the pathogen at some intermediate level. One can quantify this by considering overall transmission potential, , to be proportional to the product of per-contact transmission potential, , contact-rate among infected and susceptible, , and the duration of infectiousness . All 3 quantities can potentially be impacted by virulence, i.e. . Unfortunately, for our study we do not have information on the duration of infectiousness. While it is quite likely that virulence can impact the duration of infectiousness, for the following discussion we assume to be constant. In that case we have . Overall transmission potential is optimized when is maximized. Figure 5 illustrates graphically a relation for contact rate, per-contact transmission potential and overall transmission potential as a function of virulence.

We can map our data onto this conceptual figure if we consider the morbidity score as a proxy of virulence, , the infectiousness score as a proxy of per-contact transmission potential, , and patient-reported activity as a proxy for the contact rate, . Since our data is measured in units with indirect and uncertain mapping to actual per-contact transmission potential and actual contact rate, we standardize the data and manually place it on top of the conceptual lines, this should not be considered a quantitative mapping.

Our study population consisted of individuals who felt sick enough to seek medical care, but none were ill enough to require emergency care. We thus expect them to be in the middle of the virulence spectrum. Based on our data, this range is characterized by infectiousness levels that are close to their maximum value (at least as measured with our absence/presence scale) and are not impacted much by changes in morbidity/virulence. In contrast, activity is more strongly impacted by morbidity/virulence. We would expect that a study population that included asymptomatic and mildly symptomatic infected persons would be on the left side of our data, where increases in virulence would lead to increases in infectiousness, with small initial decreases in activity, while severely ill and hospitalized individuals would fall to the right side of our data in the figure.

Figure 5: This figure illustrates conceptually the hypothetical impact of virulence on total transmission potential (T) resulting from a trade-off between per-contact transmission potential (p) and contact-rate (c). The lines are for illustrative purposes only and not fitted to the data. We are using morbidity as a proxy for virulence. We placed our data in the middle of the full virulence scale since we did not capture anyone not sick enough to seek care nor did we capture anyone who was so ill they were hospitalized or died. The values for infectiousness and activity are re-scaled to allow better visualization. The actual mapping between our measured quantities and the theoretical contact rate and per-contact infectiousness are not known, but based on past research it is feasible to expect that a proportional relationship exists.

Figure 5: This figure illustrates conceptually the hypothetical impact of virulence on total transmission potential () resulting from a trade-off between per-contact transmission potential () and contact-rate (). The lines are for illustrative purposes only and not fitted to the data. We are using morbidity as a proxy for virulence. We placed our data in the middle of the full virulence scale since we did not capture anyone not sick enough to seek care nor did we capture anyone who was so ill they were hospitalized or died. The values for infectiousness and activity are re-scaled to allow better visualization. The actual mapping between our measured quantities and the theoretical contact rate and per-contact infectiousness are not known, but based on past research it is feasible to expect that a proportional relationship exists.

# DISCUSSION

We believe that this is the first study that investigates a trade-off between contact-rate and per contact transmission potential for influenza in humans (5,6,41,42). We showed that for our population, activity decreased as both morbidity and infectiousness scores increased, and we found a positive association between morbidity and infectiousness symptoms.

Limitations of the study include not knowing the flu sub-type for those infected. The type and sub-type of the virus can affect the epidemiological features of the disease (43). Based on influenza surveillance data for the 2016/17 season only 22.1% was influenza B with 77.9% influenza A with the subtype H3N2 making up 97.2% with H1N1 making up the remaining 2.8% (44). Additionally, we only collected data on individuals who were experiencing symptoms severe enough to seek care. As a result, we do not have data on individuals with low virulence infections. As explained above, such data would allow for a complete exploration across the full range of virulence and to determine relationships between transmission, morbidity, and infectiousness. Finally, our study population was made up of college students, i.e., generally young and healthy individuals. As such their symptoms, infectiousness, and activity behavior distributions might not fully apply to a more general population.

Despite these potential limitations, our study provides valuable information that can be useful to inform current and future interventions targeting influenza. For example, our results suggest that a treatment that only reduces those symptoms that are part of our morbidity score, without affecting symptoms that make up our infectiousness score, could lead to increased transmission. While from the perspective of a patient or clinician a reduction in any symptom may be viewed as a positive, such an intervention might lead to worse outcomes on the population level. Current FDA approval of anti-influenza drugs rely on showing an impact on the symptoms, with a focus on more severe and systemic (i.e., morbidity) symptoms (45–47). From a population perspective, it is essential that such drugs also reduce host infectiousness (47–49). Some evidence for this has been found in previous studies(50–52).

Population-level control of infectious diseases makes increasing use of mathematical models (53). The need for these models to be accurate is critical. Researchers have increasingly recognized that capturing human behavior changes during an infectious disease outbreak, both for uninfected and infected individuals is relevant (54,55). As far as we are aware, only one previous modeling study for influenza has tried to capture the impact of infection on behavior (56). Previous studies have shown that symptoms aid infectiousness and impact the number of contacts (1,29). In our analysis, we found an 85% reduction in median activity as a result of increased morbidity. Using data from our study and past studies (1,29) is a starting point for future models that can explore the impacts of infectiousness and contact behavior of infected hosts (25).

# References

1. Eames K, Tilston N, White P, et al. The impact of illness and the impact of school closure on social contact patterns. *Health technology assessment (Winchester, England)*. 2010;14(34):267–312.

2. Antolin MF. Unpacking : Within-host dynamics and the evolutionary ecology of pathogen transmission. *Annual Review of Ecology, Evolution, and Systematics*. 2008;39:415–437.

3. Brown NF, Wickham ME, Coombes BK, et al. Crossing the line: Selection and evolution of virulence traits. *PLoS pathogens*. 2006;2(5):e42.

4. Anderson RM, May R. Coevolution of hosts and parasites. *Parasitology*. 1982;85(2):411–426.

5. Alizon S, Hurford A, Mideo N, et al. Virulence evolution and the trade-off hypothesis: History, current state of affairs and the future. *Journal of evolutionary biology*. 2009;22(2):245–259.

6. Cressler CE, McLEOD DV, Rozins C, et al. The adaptive evolution of virulence: A review of theoretical predictions and empirical tests. *Parasitology*. 2016;143(7):915–930.

7. Bull JJ, Lauring AS. Theory and empiricism in virulence evolution. *PLoS pathogens*. 2014;10(10):e1004387.

8. Ewald PW. Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics*. 1983;14(1):465–485.

9. Levin S, Pimentel D. Selection of intermediate rates of increase in parasite-host systems. *The American Naturalist*. 1981;117(3):308–315.

10. Levin BR. The evolution and maintenance of virulence in microparasites. *Emerging infectious diseases*. 1996;2(2):93.

11. Bull JJ. Virulence. *Evolution*. 1994;48(5):1423–1437.

12. Anderson RM, May RM. Population biology of infectious diseases: Part i. *Nature*. 1979;280(5721):361.

13. Antia R, Levin BR, May RM. Within-host population dynamics and the evolution and maintenance of microparasite virulence. *The American Naturalist*. 1994;144(3):457–472.

14. Sofonea MT, Alizon S, Michalakis Y. Exposing the diversity of multiple infection patterns. *Journal of theoretical biology*. 2017;419:278–289.

15. Coombs D, Gilchrist MA, Percus J, et al. Optimal viral production. *Bulletin of mathematical biology*. 2003;65(6):1003–1023.

16. Coombs D, Gilchrist MA, Ball CL. Evaluating the importance of within-and between-host selection pressures on the evolution of chronic pathogens. *Theoretical population biology*. 2007;72(4):576–591.

17. Lipsitch M, Moxon ER. Virulence and transmissibility of pathogens: What is the relationship? *Trends in microbiology*. 1997;5(1):31–37.

18. Brown SP, Cornforth DM, Mideo N. Evolution of virulence in opportunistic pathogens: Generalism, plasticity, and control. *Trends in microbiology*. 2012;20(7):336–342.

19. Gilchrist MA, Coombs D. Evolution of virulence: Interdependence, constraints, and selection using nested models. *Theoretical population biology*. 2006;69(2):145–153.

20. Mideo N, Alizon S, Day T. Linking within-and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends in ecology & evolution*. 2008;23(9):511–517.

21. Gandon S, Mackinnon MJ, Nee S, et al. Imperfect vaccines and the evolution of pathogen virulence. *Nature*. 2001;414(6865):751.

22. Sofonea MT, Aldakak L, Boullosa LV, et al. Can ebola virus evolve to be less virulent in humans? *Journal of evolutionary biology*. 2018;31(3):382–392.

23. Thomas SR, Elkinton JS. Pathogenicity and virulence. *Journal of invertebrate pathology*. 2004;85(3):146–151.

24. Alizon S, Michalakis Y. Adaptive virulence evolution: The good old fitness-based approach. *Trends in ecology & evolution*. 2015;30(5):248–254.

25. Handel A, Rohani P. Crossing the scale from within-host infection dynamics to between-host transmission fitness: A discussion of current assumptions and knowledge. *Phil. Trans. R. Soc. B*. 2015;370(1675):20140302.

26. Mackinnon MJ, Gandon S, Read AF. Virulence evolution in response to vaccination: The case of malaria. *Vaccine*. 2008;26:C42–C52.

27. Fraser C, Hollingsworth TD, Chapman R, et al. Variation in hiv-1 set-point viral load: Epidemiological analysis and an evolutionary hypothesis. *Proceedings of the National Academy of Sciences*. 2007;104(44):17441–17446.

28. Leung NH, Xu C, Ip DK, et al. The fraction of influenza virus infections that are asymptomatic: A systematic review and meta-analysis. *Epidemiology (Cambridge, Mass.)*. 2015;26(6):862.

29. Van Kerckhove K, Hens N, Edmunds WJ, et al. The impact of illness on social networks: Implications for transmission and control of influenza. *American journal of epidemiology*. 2013;178(11):1655–1662.

30. Kohavi R, John GH. Wrappers for feature subset selection. *Artificial intelligence*. 1997;97(1-2):273–324.

31. Yule GU. An introduction to the theory of statistics. C. Griffin, limited; 1919.

32. Warrens MJ. On association coefficients for 2 2 tables and properties that do not depend on the marginal distributions. *Psychometrika*. 2008;73(4):777.

33. Hollander M, Wolfe DA. Nonparametric statistical methods john wiley & sons. *Inc. New York*. 1973;

34. Conover W. Practical nonparametric statistics, john wiley & sons. 1999.

35. Mantel N. Chi-square tests with one degree of freedom; extensions of the mantel-haenszel procedure. *Journal of the American Statistical Association*. 1963;58(303):690–700.

36. Kuritz SJ, Landis JR, Koch GG. A general overview of mantel-haenszel methods: Applications and recent developments. *Annual review of public health*. 1988;9(1):123–160.

37. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.(<https://www.R-project.org/>)

38. Bischl B, Lang M, Kotthoff L, et al. mlr: Machine learning in r. *Journal of Machine Learning Research* [electronic article]. 2016;17(170):1–5. (<http://jmlr.org/papers/v17/15-066.html>)

39. Meyer D, Zeileis A, Hornik K. Vcd: Visualizing categorical data. 2017.

40. Signorell A. DescTools: Tools for descriptive statistics. 2019.(<https://cran.r-project.org/package=DescTools>)

41. Ebert D, Bull JJ. Challenging the trade-off model for the evolution of virulence: Is virulence management feasible? *Trends in microbiology*. 2003;11(1):15–20.

42. Geoghegan JL, Holmes EC. The phylogenomics of evolving virus virulence. *Nature Reviews Genetics*. 2018;1.

43. Cowling BJ, Chan KH, Fang VJ, et al. Comparative epidemiology of pandemic and seasonal influenza a in households. *New England journal of medicine*. 2010;362(23):2175–2184.

44. Blanton L, Alabi N, Mustaquim D, et al. Update: Influenza activity in the united states during the 2016–17 season and composition of the 2017–18 influenza vaccine. *MMWR. Morbidity and mortality weekly report*. 2017;66(25):668.

45. Ebell MH, Call M, Shinholser J. Effectiveness of oseltamivir in adults: A meta-analysis of published and unpublished clinical trials. *Family practice*. 2012;30(2):125–133.

46. Hayden FG, Sugaya N, Hirotsu N, et al. Baloxavir marboxil for uncomplicated influenza in adults and adolescents. *New England Journal of Medicine*. 2018;379(10):913–923.

47. Jefferson T, Deeks J, Demicheli V, et al. Amantadine and rimantadine for preventing and treating influenza a in adults. *Cochrane Database of Systematic Reviews*. 2004;(3).

48. Nishiura H, Oshitani H. Household transmission of influenza (h1n1-2009) in japan: Age-specificity and reduction of household transmission risk by zanamivir treatment. *Journal of International Medical Research*. 2011;39(2):619–628.

49. Goldstein E, Cowling BJ, O’Hagan JJ, et al. Oseltamivir for treatment and prevention of pandemic influenza a/h1n1 virus infection in households, milwaukee, 2009. *BMC infectious diseases*. 2010;10(1):211.

50. Halloran ME, Hayden FG, Yang Y, et al. Antiviral effects on influenza viral transmission and pathogenicity: Observations from household-based trials. *American journal of epidemiology*. 2006;165(2):212–221.

51. Yang Y, Halloran ME, Longini Jr IM. A bayesian model for evaluating influenza antiviral efficacy in household studies with asymptomatic infections. *Biostatistics*. 2009;10(2):390–403.

52. Tsang TK, Lau LL, Cauchemez S, et al. Household transmission of influenza virus. *Trends in microbiology*. 2016;24(2):123–133.

53. Lessler J, Cummings DA. Mechanistic models of infectious disease and their impact on public health. *American journal of epidemiology*. 2016;183(5):415–422.

54. Funk S, Bansal S, Bauch CT, et al. Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics*. 2015;10:21–25.

55. Carrasco LR, Jit M, Chen MI, et al. Trends in parameterization, economics and host behaviour in influenza pandemic modelling: A review and reporting protocol. *Emerging themes in epidemiology*. 2013;10(1):3.

56. Handel A, Longini Jr IM, Antia R. Neuraminidase inhibitor resistance in influenza: Assessing the danger of its generation and spread. *PLoS Computational Biology*. 2007;3(12):e240.