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

Brian McKay¹, Mark Ebell¹, Ariella Perry Dale³, Ye Shen¹ and Andreas Handel²

¹Department of Epidemiology and Biostatistics, and ²Department of Epidemiology and Biostatistics and Health Informatics Institute and Center for the Ecology of Infectious Diseases, The University of Georgia, Athens, GA, USA
³Colorado Department of Public Health and Environment, Denver, CO, USA

Communicable diseases are often virulent, i.e. they cause morbidity symptoms in those infected. While some symptoms may be transmission-enhancing, other symptoms are likely to reduce 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 performed an exploratory data analysis of 324 influenza patients at a university health centre 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 (virulence), and activity level. We found a decrease in the activity level with increasing morbidity scores. There was no consistent pattern between an activity level and an infectiousness score. We also found a positive correlation between morbidity and infectiousness scores. Overall, we find that increasing virulence leads to increased infectiousness and reduced activity, suggesting a trade-off that can impact overall transmission potential. Our findings indicate 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 that investigate the impact of different interventions on transmission.

1. Introduction

Many infectious diseases cause symptoms in at least some infected hosts. Often, those symptoms increase the host’s infectiousness and facilitate the transmission of the pathogen [1,2]. Coughing and sneezing for respiratory infections are prime examples. However, symptoms that are too severe may reduce host activity or, in extreme cases, cause host death, reducing transmission opportunities. The induction of symptoms by a pathogen is often called virulence, and the idea that some virulence is needed for efficient transmission but too much virulence reduces transmission is often called the trade-off hypothesis and has been studied extensively [3–5]. The trade-off hypothesis predicts that an intermediate level of virulence leads to maximum pathogen fitness (usually quantified by the reproductive number). The virulence level which optimizes fitness depends on both population-level and within-host-level processes, the implications of which have been theoretically explored previously [2,3,6–13].

The most studied trade-off is between increasing transmission potential due to increased host infectiousness and decreasing transmission potential due to host mortality [3]. For humans, malaria might fall into this category. It has been shown that as malaria parasite density increased within a host (considered a measure of virulence), per-contact transmission potential (quantified by
gametocyte density) increased, as did the duration of infection and mortality [14, 15]. The former two quantities are assumed to increase transmission potential, while the latter reduces it. It is, however, unclear if the increase in mortality is a strong enough effect to play an important role in the virulence trade-off in malaria. Host death is likely an important determining factor of transmission potential for many animal diseases and some human diseases (e.g. viral haemorrhagic diseases [16]). However, for many human pathogens, mortality is low, and it is more likely that increased virulence leads to reduced host activity and thus reduced transmission potential. Transmission potential, $T$, is proportional to the product of per-contact transmission potential (i.e. infectiousness), $p$, the rate at which infected hosts have contact with other hosts, $c$, and the duration of the infectious period, $d$.

All three quantities can potentially be impacted by virulence, $v$, leading to $T \sim p(v) \times c(v) \times d(v)$.

A previous study for HIV infection in humans showed that $d$ and $p$ are both impacted by virulence, namely increased virulence (quantified by increased pathogen load), was shown to increase per-contact transmission potential, $p$, and reduce the duration of infectiousness (i.e. time to AIDS/death), $d$. Maximum transmission potential was found to be at an intermediate level of pathogen load (virulence) [17, 18]. A study of dengue infection showed similar results [19]. The viral production rate, which was considered as a proxy for virulence, led to both increased peak viral load (which is correlated with increased transmission potential) as well as increased clearance rate, i.e. shorter duration of infection. This led to a similar relationship as found for HIV, namely an intermediate peak viral load resulted in the highest transmission potential.

While there is some evidence that the virulence trade-off applies to other human diseases [20–23], there are, as far as we are aware, no other studies of human pathogens that directly address this topic. In addition, the studies we just described focused on the duration, $d$, and per-contact transmission potential, $p$. None were able to investigate the impact of virulence on contact behaviour, $c$.

Here, we investigate the trade-off hypothesis using data from influenza infections in humans, focusing on transmission potential ($p$) and contact rate ($c$). Influenza induces symptoms in around 84% of infected individuals [24]. Some of the symptoms, such as coughing and sneezing, likely enhance transmission by increasing the infectiousness of a host ($p$). 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 [25]. Other symptoms, such as fever, body aches, and general malaise, are more likely to lead to a reduction in transmission by reducing host activity ($c$). A previous study on influenza in 146 adults and children in the UK found that individuals without symptoms had a mean of 12.72 contacts per day, while those individuals with symptoms only had 3.58 [26]. The same study 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 are important factors in the overall transmission potential. Here, we investigate this relationship further using individual patient data. While it is quite likely that virulence can impact the duration of infectiousness for influenza, we, unfortunately, do not have data for this component and thus focus on the other two quantities. We expect that virulence leads to increasing $p$ while negatively affecting $c$. The transmission potential due to $p$ and $c$ can be considered as proportional to the product of those quantities, i.e. $T_{pc} \sim p(v) \times c(v)$. The lines in figure 2 illustrate our hypothesized relation for contact rate, per-contact transmission potential, and overall transmission potential, $T_{pc}$, as a function of virulence. Our data (symbols in figure 2, discussed in detail next) show overall support for this relation.

2. Material and methods

(a) Data collection

Students with a primary complaint related to a respiratory infection who made an appointment at the health centre of a large research university from December 2016 to February 2017 filled out an electronic questionnaire (see [27] for more details). The questionnaire collected data about their current symptoms and activity level. A response was required for 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 electronic supplementary material.

For the symptoms of cough, 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 behaviour. Patients described 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 include patients diagnosed with a rapid antigen or rapid polymerase chain reaction (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. These results are shown in the electronic supplementary material.

The institutional review board approved the study protocol.

(b) 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 morbidity symptoms, we dichotomized weakness and body aches to ‘absent’ or ‘present’. Patients reported cough as present or absent as well as severity, but only cough absent or present was considered for the main analysis.

(c) Analysis

We assessed the univariate relationships between activity and each symptom using the linear regression-treating activity level as a continuous variable. We also performed multiple linear regressions. 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 fivefold cross-validation (20 times repeated) [28].

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 (electronic supplementary material, table S1). 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 a susceptible host [29,30]. Importantly, the grouping of variables into either the morbidity or infectiousness symptom categories was based on a priori medical and biological considerations, independently of any observed correlation with the 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 the activity level. These scores are similar to systemic and respiratory scores used in past studies of influenza infection [29,31,32]. The scores are computed as a sum of the symptoms that are present. Our data did not allow us to take into account symptom severity, though a comparison of our scores with cough, weakness, and body aches, for which we have severity, shows that there is a good positive correlation between severity and the number of symptoms (see electronic supplementary material). To test the robustness of our results, we completed several sensitivity analyses using different approaches to generate the scores. Results from these analyses are shown in the electronic supplementary material.

Correlations between the infectiousness score, morbidity score, and activity were assessed using Spearman correlation [33] and the generalized Mantel–Haenszel procedure [34]. Linear regression is used to estimate the average change in activity, and the lines are included in the plots to help visualize the relationships. All analyses were completed using R (v: 3.5.3) [35]. We used the mlr package for cross-validation [36], vcdExtra to compute Yule’s Q, and the Cochran-Mantel-Haenszel trend test [37] and DescTools to compute Spearman’s rank correlation coefficient and corresponding confidence intervals [38]. All of the code and data required to reproduce the results are available through Dryad.

3. Results

(a) Study population
During the study period, 2380 patients had a respiratory complaint and filled out the questionnaire. Among those, 324 had a lab-based diagnosis of influenza (PCR or rapid antigen). The following analyses focus on those patients since they are most likely infected with influenza. For analyses of patients who received a flu diagnosis based on lab tests or empirically from a physician, see the electronic supplementary material. Patients with influenza reported activity levels ranging from 0 to 10, with a median of 4 (electronic supplementary material, figure S1). All of the patients reported symptoms, with only 14% reporting 9 or fewer (maximum possible 25). The most common symptoms were coughing and weakness, while the least common was abdominal pain (electronic supplementary material, table S1).

(b) Univariate and subset selection
We assessed correlations between the activity level and each symptom in a univariate linear analysis (electronic supplementary material, table S2). All statistically significant symptoms had a negative correlation with the activity level (electronic supplementary material, table S2). Next, we considered a multivariable regression model and performed a variable selection based on the cross-validated minimization of RMSE. We found that the best performing model was one that included chills/sweats, subjective fever, headache, weakness, sleeplessness, and vomiting (electronic supplementary material, table S2). While vomiting is not a common symptom of influenza, in those patients who did report vomiting it led to major reductions in their activity.

(c) Computation of infectiousness and morbidity scores
For the main analysis, we classified five symptoms as infectiousness-related: coughing, sneezing, runny nose, nasal congestion, and chest congestion. Twenty symptoms were classified as morbidity-related: subjective fever, having chills and or sweat, 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. Each symptom present in a patient contributed one point to their respective scores. For those symptoms for which we had severity, we investigated correlations with the number of symptoms and found total number of symptoms to be a good proxy (see the electronic supplementary material). Analyses using several alternative approaches for computing the score are shown in the electronic supplementary material and summarized below in the Sensitivity Analysis section.

The infectiousness score had a possible range of 0–5, and the morbidity score had a possible range of 0–20. The median infectiousness score was 4. Only two patients had an infectiousness score of 0, 20% had a score of 2 or less, and 28% of patients had the maximum possible score of 5. The mean morbidity score was 9.574, and no patients had a morbidity score of 0, 1, 19, or 20. The centred distribution is assumed to be the result of how patients were included in the study, that is all the patients felt sick enough to seek medical care, but none were sick enough to require urgent care or hospitalization. Results presented in the electronic supplementary material show plots of the score distributions (electronic supplementary material, figure S2A,B).

(d) 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 a negative relationship ($r = -0.15$ (95% CI: $-0.25$, $-0.04$)), and the Cochran–Mantel–Haenszel trend test is statistically significant ($\chi^2 = 6.363$, d.f. = 1, $p < 0.01$) (figure 1a). Note, however, that the data suggest a nonlinear relationship between infectiousness and activity. We cannot think of a biological mechanism that might lead to this pattern. Given that the observed negative trend is small and does not show a monotone decline, it is most reasonable to assume based on this data that there is no meaningful relationship between the infectiousness score and the activity level. Results presented in the electronic supplementary material show that the overall pattern remains the same if details of the analysis approach are changed (electronic supplementary material, figures S6–S8B and S13).

(e) Impact of morbidity score on activity
Analysis of the association between the morbidity score and the patient’s self-reported activity level suggests that the
higher morbidity score is associated with reduced activity levels. Spearman’s rank correlation indicates a negative relationship ($r = -0.32$ (95% CI: $-0.42$, $-0.22$)), and the Cochran–Mantel–Haenszel trend test is statistically significant ($\chi^2 = 38.577$, d.f. = 1, $p < 0.01$) (figure 1b). The observed pattern is consistent and clear, with a mean 4.7-fold reduction in the activity level going from the lowest to the highest morbidity score. The strong negative relationship is preserved if details of the analysis approach are changed and results are presented in electronic supplementary material, figures S9B, S10B and S14.

(f) Impact of morbidity score on the infectiousness score

Analysis of the relationship between the morbidity and infectiousness scores shows a positive correlation. Spearman’s rank correlation indicates a positive relationship ($r = 0.28$ (95% CI: $0.18$, $0.38$)), and the Cochran–Mantel–Haenszel trend test is statistically significant ($\chi^2 = 26.093$, d.f. = 1, $p < 0.01$) (figure 1c). Apart from the mean activity levels for very low morbidity score values (with very small sample sizes), the pattern is consistent and clear, with a mean 1.8-fold increase in the infectiousness score going from the lowest to the highest morbidity score. A positive relationship is observed regardless of how the infectiousness or morbidity scores are calculated (electronic supplementary material, figures S6–S10C) and in the analysis of empirically diagnosed patients (electronic supplementary material, figure S15).

(g) Sensitivity analysis

We performed several sensitivity analyses of our results. In one type of analysis, we computed both the morbidity and infectiousness scores in different ways and showed that overall results remain the same. In a different analysis, we considered all patients diagnosed with influenza, not just those that had a positive lab test. Again, overall results remained robust. Details of all these analyses are presented in the electronic supplementary material.

(h) Conceptualizing our results

We can place our data into the conceptual framework introduced in the introduction with the morbidity score as a proxy for virulence, $v$, the infectiousness score as a proxy of per-contact transmission potential, $p$, and patient-reported activity as a proxy for the contact rate, $c$ (figure 2). Since our data are measured in units with indirect and uncertain mapping to actual per-contact transmission potential and actual contact rate, we standardized the data and manually placed 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. It is thus a reasonable assumption to expect them to be somewhere in the middle of the virulence spectrum. Based on our data, this range is characterized by infectiousness levels that are increasing as morbidity (virulence) increases (figure 2). Activity is more strongly impacted and decreases as morbidity increases. We speculate that a study population that included
asymptomatic and mildly symptomatic-infected persons would be on the left side of our data, while severely ill and hospitalized individuals would fall to the right side of our data.

4. Discussion

We believe this is the first study that investigates a trade-off between contact rate and per-contact transmission potential for influenza in humans. 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. This is a secondary analysis of data that was not specifically collected to address our question. As such, results should be considered exploratory instead of confirmatory. All data were self-reported, and most symptoms were only reported as present or absent. Patient-reported symptoms can be very subjective [39]. Due to limitations in the available data, we had to make certain assumptions. We only used the absence or presence of symptoms for our scores. More detailed data would have allowed for a more accurate score taking into account severity/frequency of symptoms. We only collected data on individuals who were experiencing enough symptoms to seek care, but not symptoms so severe that emergency care was required. As a result, we do not have data on individuals with low or high virulence infections. Such data would allow for a complete exploration across the full range of virulence and 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 behaviour distributions might not fully apply to a more general population. Our data also does not contain information on the influenza strain with which individuals were infected, which may impact disease outcomes [40].

The mapping of our infectiousness score on actual transmission potential is unclear. Different symptoms likely contribute to the infectiousness of a patient in different amounts. Coughing and sneezing create both droplets and aerosols that have varying concentrations of virus and environmental persistence and can transmit the disease directly and indirectly [41–47]. While a runny nose does not create aerosols or droplets without proper hygiene, it can lead to contaminated objects or self-inoculation of contacts [42,48]. Our score is limited but similar to scores used in previous studies to quantify infectiousness [29,31,32].

Despite these potential limitations, our study can help inform current and future interventions targeting influenza. For example, if our results hold up, it suggests that any anti-influenza treatment that reduces morbidity-related symptoms, without affecting infectiousness-related symptoms, could lead to increased transmission. While from the perspective of a patient or a clinician, a reduction in any symptom may be viewed as a positive, such an intervention might lead to worse outcomes on the population level [49]. The current Food and Drug Administration (FDA) approval of anti-influenza drugs relies on showing an impact on the symptoms, with a focus on more severe and systemic (i.e. morbidity) symptoms [50–52]. From a population perspective, it is essential that such drugs also reduce host infectiousness [52–55]. Some evidence for this has been found in previous studies [54,56–58] as well as being explored in mathematical models [49,59,60].

Population-level control of infectious diseases makes increasing use of mathematical models [61]. The need for these models to be accurate is critical. Researchers have increasingly recognized that capturing human behaviour changes during an infectious disease outbreak, both for uninfected and infected individuals, is relevant [62,63]. As far as we are aware, only a few previous modelling studies for influenza have tried to capture the impact of infection on contact rates [64–67]. Previous studies have shown that symptoms aid infectiousness and impact the number of contacts [25,26]. In our analysis, we found a strong reduction in mean activity as a result of increased morbidity. Using data from our study and past studies [25,26] can be a starting point for future models to explore the impacts of infectiousness and contact behaviour of infected hosts [68].

Ethics. The University of Georgia’s Institutional Review Board (IRB) approved the study protocol.

Data accessibility. All of the data and code required to complete the analysis are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.51c9zv4v [69].

Authors’ contributions. B.M. aided in data collection, data analysis, and drafting the manuscript; M.E. designed and facilitated the data collection, reviewed and revised the manuscript; A.P.D. aided in data collection and reviewed the manuscript; Y.S. helped develop the analytical strategy and the methods and results section of the manuscript; A.H. was responsible for the conceptual design, analysis, and revised the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. A.H. and B.M. were partially supported by the NIH grant no. U19AI117891. This is a secondary analysis of data that had been previously collected with support from Roche diagnostics.

References


2. Lipsitch M, Moxon ER. 1997 Virulence and previous studies to quantify infectiousness [29,31,32].


