Coughs, colds and “freshers’ flu” survey in the University of Cambridge, 2007–2008

Ken T.D. Eames a, Maria L. Tang a,b,∗, Edward M. Hill b,c, Michael J. Tildesley b,c, Jonathan M. Read b,d, Matt J. Keeling b,c, Julia R. Gog a,b,∗∗

a Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Cambridge, CB3 0WA, UK
b Joint Universities Pandemic and Epidemiological Research, UK
∗ The Zeeman Institute for Systems Biology & Infectious Disease Epidemiology Research, School of Life Sciences and Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK
d Lancaster Medical School, Lancaster University, Lancaster, UK

A B S T R A C T

Universities provide many opportunities for the spread of infectious respiratory illnesses. Students are brought together into close proximity from all across the world and interact with one another in their accommodation, through lectures and small group teaching and in social settings. The COVID-19 global pandemic has highlighted the need for sufficient data to help determine which of these factors are important for infectious disease transmission in universities and hence control university morbidity as well as community spillover. We describe the data from a previously unpublished self-reported university survey of coughs, colds and influenza-like symptoms collected in Cambridge, UK, during winter 2007–2008. The online survey collected information on symptoms and socio-demographic, academic and lifestyle factors. There were 1076 responses, 97% from University of Cambridge students (5.7% of the total university student population), 3% from staff and <1% from other participants, reporting onset of symptoms between September 2007 and March 2008. Undergraduates are seen to report symptoms earlier in the term than postgraduates; differences in reported date of symptoms are also seen between subjects and accommodation types, although these descriptive results could be confounded by survey biases. Despite the historical and exploratory nature of the study, this is one of few recent detailed datasets of influenza-like infection in a university context and is especially valuable to share now to improve understanding of potential transmission dynamics in universities during the current COVID-19 pandemic.

1. Introduction

There is significant potential for spread of infectious respiratory illnesses at universities. In addition to the annual influenza season, the start of which typically coincides with the start of the academic year for UK universities (Mook et al., 2008), the current COVID-19 global pandemic has highlighted the need to understand and control high morbidity of infectious disease amongst university students (Enright et al., 2021). University students represent a unique relatively closed population with extremely high levels of close social mixing, yet where there is a relatively limited level of enforceable control. Outbreaks in student communities also risk spillover into the general population, including from students travelling home in between terms (Mangrum and Niekamp, 2020; Buck et al., 2006), as well as negatively and significantly impacting the health and academic performance of students (Nichol et al., 2005).

Several recorded outbreaks report high influenza-like illness (ILI) attack rates for students (Layde et al., 1980; Pons et al., 1980; Sobal and Loveland, 1982). In the UK, the coined term ‘freshers’ flu’ attributes these outbreaks to first-year students who, at the start of the academic year, may be exposed to new infections from the mass migration of students from across the world through socialising. Higher incidence has been observed for meningococcal disease in first-year students compared to higher years (Bruce et al., 2001; Ala’Aldeen et al., 2011), and for influenza in undergraduate-aged students compared to older students or faculty staff (Pons et al., 1980; Layde et al., 1980).
There are many possible routes of infection within a university setting. Population density in university accommodation can be high, with many students sharing kitchen or bathroom facilities. As such, students living on campus have been observed to have higher attack rates than those living off campus (Pons et al., 1980). Dormitory rooms in particular have been considered to be common settings for likely transmission of respiratory pathogens (Tsuang et al., 2004), as well as dormitory common areas and also small classes (Zivich et al., 2020). Other factors such as catered university accommodation (Neal et al., 1999), mid-term vacations (Sobal and Loveland, 1982) and specific social calendar events (Iuliano et al., 2009) have all seemed to contribute to close-contact infectious disease outbreaks in universities.

Targeted epidemiological data on university students linked with demographic and lifestyle data can help determine the role of different factors impacting the spread of infectious diseases in a university and thus inform suitable preventative and intervention measures. Most available ILI university datasets of this kind to date are from the US (Layde et al., 1980; Pons et al., 1980; Sobal and Loveland, 1982; Tsuang et al., 2004; Iuliano et al., 2009; Guh et al., 2011), few are from recent years (Zivich et al., 2020), and most contain only limited demographic and lifestyle data. After the start of the COVID-19 pandemic, a study at the University of Cambridge has identified university transmission clusters in the setting of national COVID-19 restrictions linked to nightclubs, halls of residence and students in the same course or year (Aggarwal et al., 2021; Warne et al., 2021). Most other COVID-19 university studies are from the US (Teran et al., 2020; Tilley et al., 2020; Wilson et al., 2020; Kianersi et al., 2021; Vang et al., 2021), but universities can differ greatly in structure and student behaviour between countries.

In this paper, we analyse and share the data from a previously unpublished UK university survey of self-reported coughs, colds and influenza-like symptoms in Cambridge, gathered between October 2007 and mid-February 2008, during the 2007–2008 influenza season. The online survey had 1076 responses, mainly from students of the University of Cambridge, and surveyed epidemiological as well as demographic, academic, domestic and social information. Although an imperfect exploratory study, it is valuable as one of the few detailed datasets on recent influenza-like infection in a university setting prior to the COVID-19 pandemic, and especially pertinent to share now because of the need to understand and predict transmission within universities in the current COVID-19 pandemic.

2. Methods

2.1. Data collection

The survey targeted participants who were experiencing or had experienced any kind of self-defined cough, cold or influenza-like symptoms. In this paper, we will refer to these as ILI symptoms in a general sense rather than the clinical definition. Participants were recruited by posters advertising the study in the university colleges, departments and other locations around Cambridge frequented by students. Locations were chosen to try to minimise bias in the demographics of the participants recruited, although this could not be guaranteed. The survey was re-advertised and posters were re-applied with new catchphrases when possible over the duration of the survey. The survey was also circulated to email contacts of each college’s undergraduate and postgraduate student bodies and college nurses in order for them to publicise to their own communities. Participants who opted in were entered into a prize draw. No other interactions were had with survey participants.

The first two terms of the 2007/2008 academic year at the University of Cambridge ran from 2 October to 30 November 2007 and 15 January to 14 March 2008, which is when students were expected to be resident in Cambridge. The survey study period started shortly before the first responses were received on 8 October 2007 and ended shortly after the last recorded responses on 12 February 2008 (the exact dates the survey opened and closed were not recorded).

Participants completed an online survey with the date of their symptom onset, demographic information (including age and gender), details of academic study, college, and social and domestic factors. Participants could submit multiple entries recording separate instances of ILI symptoms, and the number of previous times the survey completed was a survey question. The survey had 1076 responses from 1027 participants. The variables are described in Table 1. To preserve non-identifiability, some variables and options were removed or aggregated (see supplementary Text S1). The study was explained and informed consent obtained from all participants before the survey began. The study was approved by the Human Biology Research Ethics Committee of the University of Cambridge (ref. 2007.06). The anonymous and non-identifiable dataset is provided with this paper.

2.2. Statistical analysis

We fitted a Cox proportional hazards model on the days between the start of the first term and date of symptom onset, and chose models based on Akaile's information criterion (AIC) (Akaile, 1998). We used participants with a symptom onset date during the first term from 2 October–30 November 2007 when students were resident in Cambridge. We fit models with combinations of all variables in Table 1, excluding the predictors about post-symptom-onset factors: duration of symptoms, whether participants thought they could retrospectively identify who infected them, whether behaviour changed whilst ill and the behavioural changes experienced. We also exclude number of Cambridge Facebook friends, due to the amount of missing values. Recurrent events (those with the “previous” variable > 0) were excluded due to dependence. Finally, removing submissions with missing values in any of the factors under investigation gave 723 participants.

The proportional hazards assumption was checked by the Grambsch–Therneau proportional hazards assumption test (Grambsch and Therneau, 1994). Proportional hazards models were fitted in R with the survival package (Therneau, 2020), and proportional hazards models with explicit time-dependent co-efficients with the timereg package (Scheike and Zhang, 2011).

We also investigate how the risk ratio of different subpopulations depends on the unknown reporting rates, the probability that someone who experiences symptoms reports them in our survey. Subpopulation sizes in Cambridge were taken from the academic year 2007–2008 (University of Cambridge, 2008). Further, we approximate the risk ratios using subgroup survey response rates from other surveys of university students. This assumes that survey response rates do not depend on whether someone experiences symptoms or not, although we acknowledge the limitations of this (see Discussion). The surveys chosen were limited to those with available within-group survey response rates and chosen to have university student populations and the subject of ILIs where possible (Table S1), in an attempt to give comparable response rates to our survey’s unknown rates.

3. Results

3.1. Study population

The survey had 1076 responses, mainly from University of Cambridge students (97%), but also university staff (3%) and other participants (<1%) (Fig. 1(a)). The student participants make up 5.7% of the university population (6.2% of the undergraduate population...
and 4.6% of the postgraduate population) in the 2007–2008 academic year (University of Cambridge, 2008). The majority of survey participants were younger than 25 (Fig. 1(b)) and the survey over-sampled female students compared to the student population (Fig. 1(d)). The distribution of students among colleges in the survey also differed from that of the university (Fig. 1(c)).

For undergraduates, the survey over-sampled first-years and underlined second and third years, which was also reflected in the undergraduate age distribution (Figure S1) as most first-year undergraduates were aged under 20 (Fig. 1(e)). Undergraduates studying Natural Sciences were also over-sampled (Figure S1d).

Postgraduates were under-sampled, compared to undergraduates, but the sample postgraduate age distribution was similar to that of the university postgraduate population (Figure S1). These differences could reflect non-uniform illness incidence between groups, but could also be confounded by sampling biases, such as from the location of advertisements or differences in word-of-mouth advertisement.

### 3.2. Population-level submission summary

In the rest of this results section, the survey output is described as raw data without comparison to the student population, so differences in reported ILI could be a result of sampling biases. The survey ran over the first two terms of the 2007–2008 academic year and gathered 1076 responses between 8 October 2007 and 12 February 2008 (Fig. 2(a)). We observed separate waves during the two terms, with the first wave significantly bigger (Fig. 2(b)). Participants also retrospectively reported illness from September before term started and the Christmas holidays, however the observed reduction in reported illness outside of term time could be impacted by reduced reporting. 90% of participants answered the survey within four weeks of their reported symptom onset date, with 40% answering within a week (Figure S2).

#### 3.3. Demographic factors

There was little difference in timing of reported illness onset between male and female participants (Figure S3a–S3b). However, the peak number of undergraduates reporting symptoms in this survey was more than double the postgraduate peak for both waves (Fig. 3(a)). The first wave was also seen to be faster in the undergraduate population compared to the postgraduate population (Fig. 3(b)), but the pattern was less clear for the smaller second term wave. Although the total number of first-year undergraduates with symptoms was larger, there was no evidence that first-year undergraduates became ill before other years, but fourth-year undergraduates were seen to become ill later than undergraduates in other years, at a similar rate to postgraduates.
Fig. 1. Distribution of participants in the survey by (a) academic status and (b) age. Distribution of students in the survey compared to the student population by (c) college (colleges ordered by most over-represented to most under-represented), and (d) gender. (e) Age distribution of first-year undergraduates and postgraduates admitted to the University of Cambridge in October 2007. *First-year postgraduates aged under 21 total 0.5%. Student population statistics are of full-time undergraduate and postgraduate students in the academic year 2007–2008 from University of Cambridge (2008).
3.4. Academic factors

Out of undergraduate survey participants, the modal response for subject of study was Natural Sciences (Figure S1d), the undergraduate course with the largest cohort (University of Cambridge, 2008). Natural Sciences students were heavily over-sampled, and students studying Medicine and Other Sciences were also slightly over-sampled. This bias could be due to the relevance of the study to these subjects and hence stronger motivation of these students to contribute to the survey. There is some evidence during the first wave of an earlier infection peak in undergraduate students studying Medicine and a later postgraduate peak compared to all other undergraduate subjects (Figs. 4(a)–4(b)), but this could be confounded by sampling biases (including word-of-mouth advertisement between students on the same course).

The number of hours of lectures, practicals and supervisions per week can vary greatly between courses and years, as well as individual choices. The size of lecture audiences also varies from a few hundred for courses with compulsory modules and large year groups, to less than ten for courses with many optional modules. Supervisions consist of very small-scale teaching, usually with one to three students per supervisor. The number of lectures and supervisions per week throughout term will vary between courses and years, but lectures tend to start halfway into the first week of term, whereas supervisions in the first term tend to start a few weeks later. Figs. 4(c) and 4(d) show the relationship between the date of symptom onset and the hours of lectures and supervisions in the seven days beforehand, with no obvious trend and large variances during periods of low total reported cases in September and from mid-November. Some participants with a symptom onset date during the holidays report non-zero hours of lectures and supervisions which may suggest a misunderstanding, such as mistakenly reporting lectures and supervisions from the seven days previous to submitting the survey or from the last seven days of the previous term.

3.5. Domestic factors

The University of Cambridge has a collegiate system which is shared by only a few other universities in the UK. In Cambridge, students are generally grouped by college for accommodation and, depending on the course, supervisions (small-scale teaching). Each college offers accommodation, mainly to undergraduates but also postgraduates and some staff, their own centralised catering, and holds social events for college members. Our survey had participants from each of the 31 colleges of the University of Cambridge (1c). There is some evidence of differences in peak height and timing by college (Figure S4), with most first wave peaks occurring before mid-October, but this could be confounded by sampling biases.

Undergraduates are usually expected to live in college accommodation for most of the duration of their degree, which may be on the main college site or elsewhere in the city. Postgraduates are often offered college accommodation for at least the first few years of their degree, which is likely to be off the main college site. These relationships are seen in our survey data (Fig. 5(a)). The majority of reported cases were in halls and the first wave seems to be very slightly earlier in halls than other types of accommodation (Figs. 5(b) and 5(c)). There is little evidence of a relationship between number of people sharing kitchens with date of symptom onset in term time (Figure S5b). Between terms, most undergraduate students do not reside in their term-time accommodation.
Fig. 3. Distribution of weekly reported cases and cumulative distribution of reported cases over the first wave by (a, b) academic status, (c, d) undergraduate year of course and (e, f) age. University of Cambridge academic terms are highlighted in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
3.6. Social factors

In the first term, there is some indication that students reporting symptoms earlier in the term also reported more evenings out in the week previous to their symptom onset (Fig. 6(b)). However, it might be expected that students go on more evenings out at the beginning of the first term when there are many social events for ‘Freshers’ week’, and evenings out may reduce over the term as workload increases. The increase in evenings out reported in the last seven days of term might also be expected as part of an increase in socialising at the end of term. A definition of ‘evenings out’ was not specified. Outside of term time, university students might display different social behaviour compared with the term time behaviour reported in the survey, so any analysis should consider time in and out of term separately. This may particularly be true for undergraduates compared to postgraduates, as undergraduates have a longer Christmas vacation and are more likely to leave their university accommodation outside of term time than postgraduates.

Facebook was a very popular social media platform at the time of survey for university students — a number of surveys in various universities in the US have found that very high proportions of the student population had a Facebook account (94% in 2007 (Valenzuela et al.,
Fig. 5. (a) Distribution of reported cases by type of accommodation and academic status. (b) Distribution of weekly reported cases by type of accommodation. (c) Cumulative distribution of reported cases over the first wave by type of accommodation. University of Cambridge academic terms are highlighted in orange. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In our survey, 86% of student participants answered the question on Facebook friends at the University of Cambridge with an exact or estimated number, suggesting they had a Facebook account. In 2007, the number of Facebook friends at the same university as you was clearly displayed on your profile. Participants who answered with an exact number reported between 0 and 781 Cambridge Facebook friends, with an overall mean of 119 (s.d. = 110) (Fig. 6(c)). Undergraduates had a mean of 132 (s.d. = 115) and postgraduates had fewer Cambridge Facebook friends on average than undergraduates, with a mean of 93 (s.d. = 93). Other surveys of undergraduates in the US report a similar positive skew in the distribution of all Facebook friends (not limited to the same university) and higher means of 150–200 in 2006 (before Facebook opened to everyone without a school or university affiliation) (Ellison et al., 2007) and 395 in 2008 (Tong et al., 2008).

Number of Facebook friends has been seen to significantly and positively relate to the rate of upper respiratory infections in university students (Campisi et al., 2012). It has also been thought to be an indicator of sociability (since seen to be positively related with extroversion (Chen, 2014) and negatively related to anxiety about face-to-face communication (Orr et al., 2009)) which could correspond to in-person interactions in which infection could spread. In our survey, looking in particular at the number of Facebook friends also attending the University of Cambridge could hence correlate with number of face-to-face interactions within the university community. Fig. 6(d) shows
3.7. Symptoms and infection

Most participants reported showing symptoms at the time of answering the survey, but for those whose symptoms had ended, the modal duration of symptoms was more than seven days (Fig. 7(a)). The majority of participants (95%) reported only one incidence of illness in the survey, and less than 5% completed the survey at least twice (Fig. 7(b)).

Participants also reported whether they experienced behavioural changes while ill and if so, which ones (Figs. 7(c), 7(d)). 67% reported a change in behaviour as a result of their symptoms. The most common behavioural changes were missing social activities (reported by 53% of all participants) and staying in bed (39%). Participants also missed meals (27%), missed work (24%) and missed lectures (19%), but few missed supervisions (4%).

the distribution of Cambridge Facebook friends by date of symptom onset, with no obvious trend across the first term. After Christmas, there is some indication of a positive trend in Cambridge Facebook friends with symptom onset date, but this may reflect a general trend of adding more friends throughout the academic year, particularly to keep in contact with over the vacation.
23% of participants thought they could identify the person who infected them. Whether they were correct would be confounded by many personality and behavioural traits, but a crude method for targeted backwards contact tracing could still be useful to identify likely settings or relationships for transmission.

### 4. Statistical analysis

#### 4.1. Proportional hazards model

Given the available data, the three most preferred proportional hazards models by AIC in Table 2 agreed that the number of evenings out during the previous week increased the risk of symptom onset at any time during the first term by a factor of 10% per evening out. No significant difference was found between participants living in halls and in college houses, but those not living in college accommodation were found to be around 30% more at risk than those in halls, although this was not highly significant ($p$-value 0.056). This may also be influenced by interactions with other variables; undergraduates living in halls report infection marginally earlier than those outside college accommodation, but the few postgraduates who live in halls report infection later than those outside college accommodation (Figure S6).

The models also consistently found undergraduates in all subjects to be at higher risk than postgraduates, but this risk varied according to...
subject (Table 2). Compared to postgraduates, undergraduates studying medicine were found to be at the most risk of symptom onset, on average over 400% more. Natural Sciences undergraduates had an increased risk of over 350%. Social Sciences or Other Sciences undergraduates had a 150%–250% increased risk and Arts and Humanities undergraduates had the smallest increase in risk of over 100%. There were not enough (21) in the sample to give a significant result for postdoctoral researchers and staff. Third-year and fourth-year undergraduates were found to have lower risk than first-year undergraduates by around 40% and 50% respectively, whereas second-year undergraduates were not significantly different to first-year undergraduates with a p-value of 0.6.

All models in Table 2 saw that the number of lectures and supervisions in the week before symptom onset significantly decreased the risk of symptom onset at a significance level of 0.05 (supervisions \( p = 0.087 \), lectures \( p = 0.090 \) respectively). Explicitly modelling the time-dependence of the supervisions and lectures variables instead gives estimates for all other parameters broadly consistent with the time-independent model (Table S3). When instead taking the supervisions and lectures predictors as categorical variables in Table S4, study participants with any number of supervisions were found to be at least as high risk of symptom onset than those with no supervisions. The more lectures participants had, the lower their risk compared to those with no lectures, with the exception of those with 1 lecture who may be more at risk. Other estimates were consistent with the numeric model.

### 4.2. Risk ratios for different reporting rates

Raw numbers of reported illness within different student subpopulations were highly dependent on both the subpopulation’s size and could also be dependent on the survey reporting rate. For example, the number of reported cases of symptoms was 36 times more for undergraduates in halls than undergraduates in other accommodation, which was skewed by the majority of the Cambridge undergraduate population living in halls, and could also have been affected by different survey reporting rates between students living in different types of accommodation. Therefore, we explored how reporting rates within different subpopulations can influence risk ratios for this study and give example risk ratios using response rates from other surveys.

Risk ratios depend linearly on the ratio of the subpopulations’ reporting rates, as approximately do the bounds on the region of 95% significance (Fig. 8). Hence risk ratios estimated using the unequal response rates from the surveys in Table S1 differed from the risk ratios obtained from assuming equal reporting rates (Table 3). The estimated risk ratio of male and female students depended on the survey used; response rates from Nichol et al. (2005) and Iuliano et al. (2009) suggested respectively greater risk of symptom onset in male than female students, or an insignificant difference in risk (Fig. 8(a)). However, the response rates from three surveys all estimated first-year undergraduates to have higher risk, between two to three times higher,

<table>
<thead>
<tr>
<th>Predictor</th>
<th>HR (95% CI)</th>
<th>p-value</th>
<th>HR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evenings out</td>
<td>1.10 (1.06, 1.15)</td>
<td>&lt;0.005</td>
<td>1.10 (1.05, 1.15)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Supervisions</td>
<td>0.79 (0.74, 0.84)</td>
<td>&lt;0.005</td>
<td>0.79 (0.74, 0.84)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Lectures</td>
<td>0.95 (0.94, 0.96)</td>
<td>&lt;0.005</td>
<td>0.95 (0.93, 0.96)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Type of accommodation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halls (ref)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>College house</td>
<td>1.00 (0.81, 1.23)</td>
<td>0.97</td>
<td>1.00 (0.81, 1.23)</td>
<td>0.97</td>
</tr>
<tr>
<td>Other accommodation</td>
<td>1.31 (0.99, 1.73)</td>
<td>0.056</td>
<td>1.31 (0.99, 1.73)</td>
<td>0.056</td>
</tr>
<tr>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postgraduate (ref)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Arts and Humanities</td>
<td>2.27 (1.66, 3.11)</td>
<td>&lt;0.005</td>
<td>2.09 (1.59, 2.76)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Medicine</td>
<td>5.47 (3.69, 8.11)</td>
<td>&lt;0.005</td>
<td>5.04 (3.51, 7.22)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Natural Sciences</td>
<td>4.99 (3.60, 6.90)</td>
<td>&lt;0.005</td>
<td>4.60 (3.45, 6.15)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Social Sciences</td>
<td>3.12 (2.25, 4.33)</td>
<td>&lt;0.005</td>
<td>2.86 (2.14, 3.83)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Other Sciences</td>
<td>3.31 (2.44, 4.49)</td>
<td>&lt;0.005</td>
<td>3.06 (2.33, 4.02)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Postdoc/staff/none</td>
<td>1.00 (0.62, 1.60)</td>
<td>0.99</td>
<td>1.11 (0.70, 1.77)</td>
<td>0.66</td>
</tr>
<tr>
<td>Year of course</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st year undergraduate (ref)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2nd year undergraduate</td>
<td>1.06 (0.84, 1.35)</td>
<td>0.62</td>
<td>1.06 (0.84, 1.33)</td>
<td>0.61</td>
</tr>
<tr>
<td>3rd year undergraduate</td>
<td>0.61 (0.47, 0.77)</td>
<td>&lt;0.005</td>
<td>0.60 (0.47, 0.77)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>4th year undergraduate</td>
<td>0.51 (0.35, 0.75)</td>
<td>&lt;0.005</td>
<td>0.53 (0.36, 0.78)</td>
<td>&lt;0.005</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Postdoc/staff/none</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

\( \text{AIC} = 0.9 \) Some categories were aggregated due to low numbers. For type of accommodation, “Other accommodation” is the aggregation of “Rented house”, “Own home” and “Other”. All non-student options were aggregated into “Postdoc/staff/none” for year of course and subject. Midpoints (or endpoints for half-open intervals) were used for the lectures and supervisions variables (for use as categorical variables, see Table S4).
of symptom onset compared to other undergraduates (Fig. 8(b)). Other estimates from the response rates of the only available survey suggest higher risk of symptom onset in undergraduates compared to postgraduates (over 1.5 times as high; Fig. 8(c)), and in undergraduates living
in halls compared to those in other accommodation (over 2 times as high; Fig. 8(d)).

5. Discussion

In this paper, we have shared and described the data from a UK university survey on self-reported coughs, colds and influenza-like symptoms and lifestyle factors from 2007–2008. The ongoing COVID-19 pandemic has motivated our retrospective analysis and dissemination of findings to help understand transmission of respiratory infection amongst university populations.

Overall, the data from this study shows some basis for the term ‘freshers’ flu’ as risk ratio estimates suggest first-year undergraduates had an overall higher risk than other undergraduate years (Table 3). These findings could have been confounded by reporting rate if perhaps freshers were more likely to sign up for a survey on ‘freshers’ flu’, but the first-year bias was also seen in other studies (Aggarwal et al., 2021; Warne et al., 2021; Ala’Aldeen et al., 2011; Bruce et al., 2001). However, our multivariate analysis found first-year undergraduates were only more likely to report infection earlier than third-year or fourth-year undergraduates and were not significantly different from second-years. The raw data and analysis also suggested undergraduates were more at risk of symptom onset and reporting illness earlier than postgraduates (Figs. 3(b), 3(f), Table 2, Table 3), agreeing with the higher undergraduate incidences seen in Aggarwal et al. (2021), Pons et al. (1980), Layde et al. (1980). Further, our study saw evidence that this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for this effect was more extreme for undergraduates studying medicine or any science, but this could instead reflect different reporting rates for
response rates and incidences to be calculated more easily, but other reporting biases would still remain. In addition, given improvements in cheap and rapid diagnostics, we would recommend future similar surveys to consider pairing survey responses with direct testing, such as sentinel testing, to relate reported illness with actual infectious disease levels in the university and identify the pathogens involved in self-reported ‘freshers’ flu’. Participants could also be swabbed after completing the survey, but this could affect response rates. While more complex studies could be planned, there is huge value in running similar studies to this one that are relatively easy, quick and inexpensive to arrange and execute. Repeating this in multiple universities in the future can give a basis of comparison without needing negative responses, and can help to gather the data needed to answer the pertinent questions of infectious disease transmission and control within universities.

Funding

KTDE was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) and Emmanuel College, University of Cambridge. MLT was supported by EPSRC [grant number EP/N509620/1]. MJT, MJK and EMH were supported by the Medical Research Council through the COVID-19 Rapid Response Rolling Call [grant number MR/V009761/1]. MJT, JMR, MJK and JRG were supported by UK Research and Innovation through the JUNIPER modelling consortium [grant number MR/V038613/1].

CRediT authorship contribution statement

Ken T.D. Eames: Conceptualisation, Methodology, Formal analysis, Investigation, Data curation, Validation, Writing – review & editing.
Maria L. Tang: Methodology, Software, Formal analysis, Investigation, Data curation, Validation, Visualisation, Writing – original draft, Writing – review & editing.
Edward M. Hill: Validation, Writing – review & editing.
Michael J. Tildesley: Validation, Writing – review & editing.
Jonathan M. Read: Validation, Writing – review & editing.
Matt J. Keeling: Validation, Writing – review & editing.
Julia R. Gog: Supervision, Validation, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The anonymous and non-identifiable dataset from this study is available as supplementary data on the online version of this article.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epidem.2022.100659.

References


