Causal analysis of H1N1pdm09 influenza infection risk in a household cohort

Yohann Mansiaux,1,2 Nicolas Salez,3 Nathanael Lapidus,1,2,4 Michel Setbon,5 Laurent Andreolletti,6,7 Marianne Leruez-Villé,8 Simon Cauchemez,9 Marie-Lise Gougeon,10 Frédéric Vély,11,12,13,14 Michael Schwarzinger,15,16,17 Laurent Abel,18,19 Rosemary Markovic Delabre,1,2 Antoine Flahault,20,21 Marie-Lise Gougeon,10 Frédéric Vély,11,12,13,14 Michael Schwarzinger,15,16,17 Laurent Abel,18,19 Rosemary Markovic Delabre,1,2 Antoine Flahault,20,21 Xavier de Lamballerie,3 Fabrice Carraï1,2,4

ABSTRACT
Background Obtaining a comprehensive quantitative figure of the determinants of influenza infection will help identify priority targets for future influenza mitigation interventions. We developed an original causal model integrating highly diverse factors and their dependencies, to identify the most critical determinants of pandemic influenza infection (H1N1pdm09) during the 2010–2011 influenza season.

Methods We used data from 601 households (1450 participants) included in a dedicated cohort. Structural equations were used to model direct and indirect relationships between infection and risk perception, compliance with preventive behaviours, social contacts, indoor and outdoor environment, sociodemographic factors and pre-epidemic host susceptibility. Standardised estimates ($\beta_{std}$) were used to assess the strength of associations (ranging from –1 for a completely negative association to 1 for a completely positive association).

Results Host susceptibility to H1N1pdm09 and compliance with preventive behaviours were the only two factors directly associated with the infection risk ($\beta_{std}=0.31$ and $\beta_{std}=-0.21$). Compliance with preventive behaviours was influenced by risk perception and preventive measures perception ($\beta_{std}=0.14$ and $\beta_{std}=0.27$). The number and duration of social contacts were not associated with H1N1pdm09 infection.

Conclusions Our findings suggest that influenza vaccination in addition to public health communication campaigns focusing on personal preventive measures should be prioritised as potentially efficient interventions to mitigate influenza epidemics.

INTRODUCTION
Influenza infection results from a complex interplay between biological characteristics of the virus and host, individual and collective behaviours, social interactions and environmental factors. Numerous studies have investigated risk factors for pandemic influenza infection (H1N1pdm09): factors as diverse as individual characteristics (eg, young age, female gender,1 chronic comorbidity3 and low pre-epidemic antibody titre4), indoor and outdoor environments (eg, number of participants in the household5 and residence in an urban area6), contacts with infected individuals,1,3 and use of preventive measures such as handwashing6 have been found to be determinants of H1N1pdm09 infection. However, no study has simultaneously explored these factors to provide a comprehensive figure of the determinants of influenza infection. The transmission mechanisms of the virus are therefore unclear and a question remains of how this information can be used by policymakers to effectively plan future mitigation strategies.

Conceptually, the determinants condition two unobserved components causally linked to infection, namely exposure to the virus and host susceptibility. For example, the number of contacts or use of preventive measures is a factor related to exposure, while haemagglutination antibody (HAI) titres, acquired naturally or after influenza vaccination, are related to host susceptibility. Other factors such as risk perception, beliefs and behaviours may act on both exposure and susceptibility, for example, by increasing social distance or vaccination.

In this study, a causal model with structural equations was used to explicitly model direct and indirect relationships between H1N1pdm09 infection and perception of risk in infection, compliance with preventive behaviours, social contacts, indoor and outdoor environment, sociodemographic factors and pre-epidemic host susceptibility. Through the integration of various risk factors known to play a role in H1N1pdm09 infection, we aimed to identify the most critical determinants and thus help prioritise clear targets for future public health interventions.

METHODS
Data source
We used data from the CoPanFlu-France cohort, which was created to study H1N1pdm09 infection risk. The cohort comprises 601 households (1450 participants) randomly selected between December 2009 and July 2010 and actively monitored for influenza-like illness (ILI) during two consecutive influenza seasons (2010–2011 and 2011–2012). When ILI was reported (fever $\geq 37.8°C$ and cough and/or sore throat, with no other known cause7), nasal swabs were collected from every household member by a study nurse and screened for respiratory viruses. Serum samples were collected for pre-seasonpost-season HAI titeration. Questionnaires exploring medical history, risk perception, use of preventive measures and duration and location of social contacts were administrated at inclusion. Study protocol, data collection and representativeness of the study households’ details can be found elsewhere.8 Study protocol was approved...
by the research ethics committee ‘Comité de Protection des Personnes Ille-de-France 1’. All participants gave written informed consent. H1N1pdm09 infection was defined by either positive H1N1pdm09 RT-PCR or a positive H1N1pdm09 RespiFinder assay on a nasal swab, or seroconversion (fourfold increase in the HAI titre). Among the 1450 individuals included in the study, 132 withdrew participation before the end of the 2010–2011 season (17 for familial and/or health reasons and 64 due to the study burden; 47 participants were lost to follow-up, and 4 died). This analysis therefore focused on 1318 participants (359 households). Infection status was unavailable for 197 individuals (14.9%) because at least one serological sample was missing. These participants were nevertheless included in the analysis of other covariates.

**Hypothesised structural model**

We postulated that influenza infection results from two unobserved phenomena: virus exposure and host susceptibility (figure 1). Their simultaneous impact on H1N1pdm09 infection risk was modelled using a structural equation model (SEM, see online supplementary file). SEM allows one to estimate direct and indirect relationships between observed or latent variables. Latent variables are unobserved variables, estimated from several observed variables, called indicators.

We modelled exposure to H1N1pdm09 with seven latent variables and one observed variable: cumulative incidence of ILI in the relevant French administrative regions. Host susceptibility to H1N1pdm09 infection was modelled with a single indicator: the pre-epidemic HAI titre (see online supplementary file and Lapidus et al9 for additional information on covariates).

**Contact network**

Seven continuous indicators were used for this latent variable: daily number and duration of contacts with individuals <15 years, between 15 and 50 years, and >50 years and daily time spent in public transport. A contact was a verbal exchange (at least three words) between the participant and another individual or a physical contact. Duration and location of contacts, as well as the age of the individuals, were collected annually. For each participant, we used all the contact data collected as close to and/or during the 2010–2011 epidemic season (20 December 2010–20 February 2011 (eg, if a participant had multiple pre-epidemic or postepidemic contact data, we considered the information collected as close to the start and the end of the influenza season, respectively). The duration and number of individual contacts were averaged over the repeated contact data and summed according to the contact’s age and were log-transformed (a value of 0.01 was imputed for participants reporting a zero summed duration or number of contacts). Reported durations of a given contact were additional, that is, a 20 min contact with three individuals corresponded to a summed contact duration of 60 min.

**Perception of infection risk and of preventive measures**

Risk perception and opinions on preventive measures have been shown to impact individual preventive behaviours.10 11 Two latent variables were used to model participants’ perception and beliefs about H1N1pdm09 influenza based on responses to a dedicated questionnaire (see figure 1 for covariates used). Specifically, all household participants >15 years indicated whether they ‘totally agree’/’partly agree’/’partly disagree’/’totally disagree’ to statements such as ‘prevention highly depends on behaviours’ and ‘H1N1pdm influenza is most often fatal’. For this analysis, the answers were dichotomised (‘agree’/‘disagree’). For household participants <15 years, we imputed for each missing covariate the most frequent answer by adults in the same household.

**Compliance with preventive behaviours**

Compliance with preventive measures has been shown to reduce H1N1pdm09 infection risk.6 12 Three binary indicators were used: ‘always/often washes hands after coughing/sneezing’, ‘always/often covers mouth while coughing or sneezing’ and ‘daily frequency of handwashing (with soap or hand sanitizer) ≥5’, dichotomised around the median daily frequency of handwashing.

**Indoor characteristics**

Six categorical indicators were used to describe living room and bedroom heating (electric/gas/other), average temperatures (dichotomised around their medians) and presence of air humidifiers. The possible impact of relative humidity on influenza aerosol transmission has been reported, and high indoor heating levels, especially during winter months, may lead to continuous circulation of dry air, creating an environment favourable to infectious viral particles persistence.13

**Neighbourhood socioeconomic status**

Geocoding of participants’ addresses allowed us to assign each participant a geographic unit defined by INSEE (Institut National de la Statistique et des Etudes Economiques), IRIS zones corresponding to statistical block groups of about 2000 inhabitants.14 Three continuous indicators, provided by INSEE, were used to describe neighbourhood socioeconomic status: ‘employment rate (%) among 15–65 years’, ‘proportion (%) of inhabitants >15 years without a diploma’ and ‘mean annual income (k€)’ (log-transformed).

**Urban environment**

Three binary indicators were used for this latent variable: ‘lives in a urban area’ (defined by a 6-item variable provided by INSEE), ‘presence of agricultural land near habitation’ and ‘presence of livestock near habitation’.

**Host susceptibility to H1N1pdm09 infection**

We used a pre-epidemic HAI titre (log-transformed) to measure host susceptibility to H1N1pdm09 infection. An HAI titre is used by the European Medicines Agency (EMA) and the US Food and Drug Administration (FDA) to evaluate vaccine protection.15 16

**Additional observed covariates**

We postulated that participants’ age class (<15, 15–50, >50 years) would influence contact network, compliance with preventive behaviours, and differential susceptibility. Factors associated with high pre-epidemic HAI titres, such as influenza vaccination and ILI history during the previous season, were also investigated. Potential associations between age class, gender and vaccination were tested (relationships between observed covariates and observed/latent variables are shown in the online supplementary file).

ILI incidence in the relevant French administrative regions was used as a proxy of exposure. Using data from the French Sentinel Network, we estimated for each French administrative region the cumulative weekly incidence of ILI (defined as the abrupt onset of fever >39°C with myalgia and respiratory signs) cases per 100 000 population for the 2009–2010 (7 September 2009–27 December 2009) and 2010–2011 (20 December 2010–19 February 2011) epidemic seasons.
Figure 1 Hypothesised structural equation model. Ellipses: latent variables; boxes: observed variables. Dotted background: latent variables related to exposure to H1N1pdm09. Striped background: latent variable related to susceptibility to H1N1pdm09 infection. For clarity, relationships involving the additional observed variables are not shown.

Figure 2 Final structural equation model. All coefficients have p values <0.05. Ellipses: latent variables; boxes: observed variables. Dotted background: latent variables related to exposure to H1N1pdm09 virus. Striped background: latent variable related to susceptibility to H1N1pdm09 infection.
2010–20 February 2011) seasons. While the cumulative incidence of ILI per region in 2010–2011 was expected to be associated with the 2010–2011 infection, we assumed that the 2009–2010 cumulative ILI incidence would affect H1N1pdm09 host susceptibility.

Model estimation

Mplus 7 software was used to estimate the parameters of the structural model while accounting for within-household correlations. For easier interpretation of the coefficients, we report standardised estimates ($\beta_{std}$) with $p < 0.05$. Standardised estimates (ranging from $-1$ for a completely negative association to $1$ for a completely positive association) can be interpreted with reference to other estimates and the relative strength of associations can be compared. Goodness of fit was assessed with two criteria: the comparative fit index (CFI) and the root mean square error of approximation (RMSEA). The CFI ranges from 0 to 1, with values $>0.90$ corresponding to an acceptable fit, while an RMSEA value $<0.08$ is recommended.18 To improve goodness of fit, new relationships can be added following ‘modification indices’ (approximations of model fit improvement when new structural equations are added) or non-significant and inconsistent relationships can be removed.

RESULTS

Data

Among the 1318 participants included in the study, 46.4% were male (612 participants). Two hundred and fifty-seven participants were $<15$ years (19.5%); 24 infections; 9.3%), 545 were aged 15–50 years (41.3%; 36 infected; 6.6%) and 516 were $>50$ years (39.2%; 29 infected; 5.6%). One hundred and twenty-seven participants received the 2010–2011 seasonal influenza vaccine (9.6%), compared to 180 participants vaccinated against H1N1pdm09 for the 2009–2010 season (13.7%). Ninety-six participants (7.3%) had ILI during the 2009–2010 season. Infection status was determined for 1121 participants (89 infections, 7.9% of the participants with a known infectious status), which were used to estimate relationships with H1N1pdm09 infection.

Structural-equation model of H1N1pdm09 infection

The final SEM model is shown in figure 2 (coefficients of the relationships between latent variables and their indicators as well as a summary of the added and removed relationships are provided in the online supplementary file). H1N1pdm09 host susceptibility and compliance with preventive behaviours were the only two factors directly associated with H1N1pdm09 infection ($\beta_{std}=0.31$, $p < 0.001$ and $\beta_{std}=-0.21$, $p = 0.011$). Compliance with preventive behaviours was positively influenced by the latent variables ‘risk perception’ of H1N1pdm09 infection and ‘perception of preventive measures’ ($\beta_{std}=0.14$, $p = 0.029$ and $\beta_{std}=0.27$, $p = 0.001$). Participants $<15$ years and those aged 15–50 years were more exposed through contacts than participants $>50$ years ($\beta_{std}=0.82$ and $\beta_{std}=0.75$, $p < 0.001$). These two younger populations were less frequently vaccinated during the 2010–2011 season ($\beta_{std}=-0.26$, $p = 0.018$ and $\beta_{std}=-0.29$, $p < 0.001$) and were more likely to have had ILI during the previous season ($\beta_{std}=0.10$, $p = 0.022$ and $\beta_{std}=0.11$, $p = 0.017$). Participants $<15$ years were less compliant with preventive behaviours than those belonging to the older age groups ($\beta_{std}=-0.63$, $p < 0.001$) and, males were less compliant than females ($\beta_{std}=-0.26$, $p < 0.001$). Participants $<15$ years were the least susceptible to H1N1pdm09 ($\beta_{std}=-0.17$, $p < 0.001$), vaccination during both the 2009–2010 and 2010–2011 seasons, ILI during the 2009–2010 season, and a high cumulative ILI incidence in the participants’ administrative regions (2009–2010 season) were associated with lower host susceptibility ($\beta_{std}=-0.29$, $p = 0.007$; $\beta_{std}=-0.19$, $p < 0.001$; $\beta_{std}=-0.07$, $p = 0.033$ and $\beta_{std}=-0.13$, $p = 0.018$, respectively). Individuals vaccinated during the 2009–2010 season were more likely to be vaccinated during the following season ($\beta_{std}=0.34$, $p = 0.008$). Individuals living in an urban area with a ‘high’ neighbourhood socioeconomic status (eg, a high mean income in the IRIS zone) were more likely to be vaccinated against pandemic influenza for the 2009–2010 season ($\beta_{std}=0.32$, $p = 0.006$ and $\beta_{std}=0.23$, $p = 0.010$). Model goodness-of-fit indices were satisfactory (RMSEA=0.023, CFI=0.943).

DISCUSSION

Using structural equation modelling, we were able to obtain a coherent quantitative picture of the complex mechanisms determining H1N1pdm09 infection during the 2010–2011 influenza season. To the best of our knowledge, our study is the first to simultaneously integrate factors as diverse as contacts, risk perceptions, preventive behaviours and environmental and socioeconomic factors. Risk factor studies since 2009 have mainly been based on regression analysis, which may fail to capture the complex mechanisms underlying infection by estimating only direct, independent associations between covariates and H1N1pdm09 infection. Use of a structural-equation model allowed us to explicitly model direct and indirect relationships between highly diverse factors and their respective impacts on infection. Our results are particularly relevant to informing public health strategies; they can be used to determine which dimensions should be priority targets of public health communication campaigns.

Host susceptibility, measured here with a pre-epidemic HAI titre, was the major factor explaining infection.3 19 Influenza vaccination, ILI history during the first H1N1pdm09 season, and the level of community incidence, were the factors associated with host susceptibility.20

Our results also suggest that compliance with preventive behaviours played an important role in influenza infection.12 Compliance with preventive behaviours was related to a positive perception of the impact of such measures and to the risk perception of infection.21 Lower preventive behaviour compliance among men compared with women has previously been observed.11 Urban area of residence and neighbourhood socioeconomic status had no direct influence on infection, but both factors positively influenced the probability of 2009–2010 vaccination. Urban residence has been reported to have a positive influence on seasonal influenza vaccination.22 No relationship between vaccination and high neighbourhood socioeconomic status has previously been reported, but a high level of education has been linked to pandemic vaccine uptake,23 as has living in a household where the head of the family has a high professional status.24

The infection rate during the second season was 7.9% in the CoPanFlu cohort, a figure consistent with estimated seroconversion rates of 6.2% and 6.8% reported by Chen et al25 during the second and third H1N1pdm09 waves in Singapore. Participants $<15$ years were the least susceptible in our study (see online supplementary file). This may be explained by a higher rate of H1N1pdm09 infection during the 2009–2010 season; 10.3% of young participants in the CoPanFlu cohort reported ILI during this period. A similar relationship between a
high rate of infection among children during the 2009–2010 season and lower susceptibility in 2010–2011 has been reported.20

No relationship was found to exist between exposure through the contact network and H1N1pdm09 infection. Several studies have nevertheless linked school closures or holiday periods to a reduced infection rate,26 implying a major influence of contacts with children on the individual infection risk. However, deeper analyses (see online supplementary file) did not suggest a higher level of exposure among infected participants. There are two possible explanations for this finding. First, contact information was not necessarily collected at the time when influenza was present in the community, especially during winter months. We used a methodology similar to that of the POLYMOD study and obtained highly consistent findings regarding assortativity and intensity of contacts according to age.27 This type of measure is now often used to explore social contact influence on the spread of infectious diseases.28–31 Contacts were also stable across calendar seasons (see online supplementary file). However, we cannot formally exclude that the pattern of contacts observed might not be representative of contacts with infected participants; contact patterns/behaviours of susceptible individuals with infectious participants may be modified in this context. Second, the daily number and duration of contacts were high in all age groups. Consequently, the likelihood of contacts with one or several infectious participants over the 2-month influenza season might be so high that transmissibility would not be influenced by the number of occasions, but rather by host characteristics.

Another limitation of our study was related to the modelling of the host susceptibility, which was only based on the pre-epidemic HAI titre. Integration of alternative correlates of protection such as T-cell responses,32 or of genetic factors,33 may provide a more robust picture of host susceptibility. Among the 1450 participants included, 132 were not followed for the considered season and another 197 had a missing infectious status. Comparison of these 329 participants with the 1121 individuals with a known infectious status showed that the participants in the first group were older (46 years on average vs 38 years), were living in households with a lower number of participants (2.9 participants on average vs 3.1) and were more likely to be vaccinated during the 2009–2010 season (vaccination rate 18% vs 11%). These last two characteristics were associated with the participant’s age. No differences were observed for sex, 2010–2011 seasonal vaccination, ILI in 2009–2010, and pre-epidemic HAI titres. Thereby, considering the low differences between the two groups for most of these characteristics and that the influence of the participant’s age on several factors under study was taken into account in our analyses, we can assume that the absence of infectious status information from these 329 participants is unlikely to bias the general conclusions of our work.

Compared with previous studies, our work has two major strengths. First, households were randomly sampled through a probabilistic procedure and selected households were shown to be representative of the general population (see Lapidus et al for more details). Second, participants were prospectively followed through an active weekly surveillance system over the two influenza periods. Both limit the risk of selection and information bias.

To the best of our knowledge, our study is also the first to simultaneously integrate such a variety of influenza risk factors through a causal approach. Structural equation models, however, cannot unequivocally prove causal relationships: acceptable indices of goodness of fit merely indicate that postulated relationships are supported by the data. Other models may provide equivalent or better performance, and covariates potentially involved in the phenomenon under study might have been omitted.34

CONCLUSION

Structural equation modelling should be considered as a valuable approach to steer public health policies. While our results rely on data collected from the 2010 to 2011 influenza season, we feel the risk factors identified would be relevant in future epidemic and pandemic contexts. The nature of the influenza virus requires the production of a new annual vaccine to lower individual susceptibility. Risk perception and compliance with preventive behaviours, though modifiable, are unlikely to change over time without targeted intervention. We suggest that communication campaigns focus on infection risk and preventive measures efficacy, specifically targeting men and children. Additionally, we find that measures based on social distancing should not be a primary public health mitigation strategy, but warrant further investigation.

What is already known on this subject

- Since 2009, numerous studies have investigated risk factors for pandemic influenza infection, reporting factors as diverse as individual characteristics, indoor and outdoor environments, contacts with infected individuals and use of preventive measures.
- Questions remain about the respective roles of the reported determinants in the mechanism of influenza virus transmission.
- Knowledge on how this information can be used to guide policymakers to implement effective mitigation strategies is lacking.

What this study adds

- Using a novel methodological approach, we were able to rank the relative impacts of all the determinants of infection, thereby determining priority targets for public health communication campaigns. Pre-epidemic antibody titres, compliance with preventive behaviours, as well as risk perception, were the only factors directly or indirectly associated with the infection risk.
- Social behaviour does not seem to be a major determinant of infection.
- On the basis of our findings, we believe it is likely that public health communication campaigns stressing the efficacy of preventive measures and the risks associated with pandemic influenza infection would be effective to mitigate the influenza burden.

Author affiliations

1 INSERM, UMR_S 1136, Institut Pierre Louis d’Epidémiologie et de Santé Publique, Paris, France
2 UPMC Univ Paris 06, UMR_S 1136, Institut Pierre Louis d’Epidémiologie et de Santé Publique, Sorbonne Universités, Paris, France
Conception, Marseille, France

Astra Zeneca and Boiron and received travel support to attend scienti

Competing interests
FC reported not having any shares or paid employment with

Ethics approval
Comité de Protection des Personnes Ile-de-France 1.

from Novartis.

properly cited and the use is non-commercial. See: http://creativecommons.org/

and license their derivative works on different terms, provided the original work is

Contributors
Ho^pitaux de Paris

and analysed the data. YM and FC were involved in writing the manuscript.

Acknowledgements
The authors thank David Young for a critical review of the

manuscript.

Contributors
NS, NL, MS, LA, ML-V, SC, M-LG, FV, MSc, Lab, RMD, AF and XdL

critical revision of the manuscript. NL, MS, RMD and FC participated in

the acquisition of data. NS and XdL performed serological analyses. NL, AF and FC
drafted the protocol. NL, NS, MS, RMD, M-LG, FV, ML-V, LA, SC, LAB, MSF, AF,

XdL and FC contributed to conception of the study. YM performed the experiments

and analysed the data. YM and FC were involved in writing the manuscript.

Funding
This study was supported by the Institut de Microbiologie et Maladies

Infectieuses (IMMI-AIESAN), the French Ministry of Health and Assistance Publique

Hôpitaux de Paris—PHRRC 2010 #AOM10199, the French Ministry of Research and

the Institut de Recherche en Sante Publique (IReSP—TIGR 2009).

Competing interests
FC reported not having any shares or paid employment with

pharmaceutical companies; he received honoraria from Novartis, GlaxoSmithKline,

Astra Zeneca and Boiron and received travel support to attend scientific meetings

from Novartis.

Patient consent
Obtained.

Ethics approval
Comité de Protection des Personnes Ile-de-France 1.

Provenance and peer review
Not commissioned; externally peer reviewed.

Open Access
This is an Open Access article distributed in accordance with the

Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and

license their derivative works on different terms, provided the original work is

properly cited and the use is non-commercial. See: http://creativecommons.org/

licenses/by-nc/4.0/

REFERENCES