nature research

Corresponding author(s): Gordon, Aubree & Krammer, Florian

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	All clinical data was recorded using OpenClinica. Laboratory data was collected in Microsoft Access.
Data analysis	To infer the crude estimates of the 50% and 80% protective levels, we used a 3-parameter logistic regression model (nplr R package v0.1-7). In models 1 and 2, a smoothing spline function was used to model the effect of age on infection risk (mgcv R package v1.8-24). A mixture model was used to explore underlying non-responder sub-populations based on the observed distribution of the antibody response (mixtool R package v1.0.4). Receiver Operating Characteristic Curve (ROC) analysis (pROC R package v1.12.1) was used to estimate the sensitivity and specificity of each assay. Classification and Regression Trees analysis (rpart R package v4.1-13) were performed to identify the best combination of assays indicated by their positive and negative predictive values in identifying PCR positive individuals. False discovery rates (FDR) were calculated in GraphPad Prism (v7.04) using the two-stage step-up method of Benjamini, Krieger and Yekutieli.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data to understand and assess the conclusions of this research are available in the main text, supplementary materials, and via ImmPort. Data are available by request, as is required by the institutional review board–approved protocol for the Household Influenza Transmission Study.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All persons in the study with sufficient sample volume and type during the included years were included in the study. Based on infection probabilities, an r-squared of 0.5-0.7, an alpha of 0.05 a sample size of 46-134 is required for 80% power. Thus, the sample size of 300 used in the analyses gives >80% power.
Data exclusions	Participants were excluded from this analysis if sufficient blood samples were not available.
Replication	Replication was not performed.
Randomization	Randomization was not performed. Samples were allocated based on whether or not the individual tested positive by RT-PCR for influenza. Age was included in all analyses as a covariate.
Blinding	All assays were performed by personnel blinded to infection status.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	During the 2013 and 2015 influenza A(H1N1)pdm epidemics, 300 household members who lived with one of 88 index cases were followed for 3-5 weeks to test for infection and seroconversion. If the 300 participants, 101 were children between the ages of 0-14 years and 199 were aged 15 or more years. There were 101 females and 199 males in the study.
Recruitment	Eligible households included those that 1) had an index case that had a positive QuickVue Influenza A+B rapid test result and with acute respiratory infection (ARI) symptom onset within the previous 48 hours; and, 2) had at least one person living with the index case. Given that only 3% of houses (5/165) approached declined participation and only 1% (9/786) of potential

participants in enrolled houses declined participation, we think that it is very unlikely that it is very unlikely that self-selection bias has biased the results of the study.