Title:

High-resolution influenza mapping of a city reveals socioeconomic determinants of transmission within and between urban quarters

Authors:

Adrian Egli[a,b,c,*,+], Nina Goldman[d*], Nicola F. Müller[e,f], Myrta Brunner[d], Daniel Wüthrich[a,b,f], Sarah Tschudin-Sutter[g], Emma Hodcroft[f,h], Richard Neher[f,h], Claudia Saalfrank[d], James Hadfield[i], Trevor Bedford[i], Mohammedyaseen Syedbasha[b], Thomas Vogel[d], Noémie Augustin[d], Jan Bauer[d], Nadine Sailer[d], Nadezhda Amar-Sliwa[d], Daniela Lang[a,b], Helena M.B. Seth-Smith[a,b,f], Annette Blaich[a], Yvonne Hollenstein[a,b], Olivier Dubuis[j], Michael Nägele[j], Andreas Buser[k], Christian H. Nickel[l], Nicole Ritz[m], Andreas Zeller[n], Tanja Stadler[e,f,*], Manuel Battegay[g,*], and Rita Schneider-Sliwa[d,*]

Affiliations:

[a] Clinical Microbiology, University Hospital Basel & University of Basel, Basel, Switzerland

[b] Applied Microbiology Research, University of Basel, Basel, Switzerland

[c] Department Clinical Research, University of Basel, Basel, Switzerland

[d] Human Geography, Department of Environmental Sciences, University of Basel, Basel, Switzerland

[e] Department of Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland

[f] Swiss Institute of Bioinformatics (SIB), Basel, Switzerland

[g] Infectious Diseases and Hospital Epidemiology, University Hospital Basel & University of Basel, Basel, Switzerland

[h] Biozentrum, University of Basel, Basel, Switzerland

[i] Fred Hutchinson Cancer Research Center, Seattle, USA

[j] Viollier AG, Allschwil, Switzerland

[k] Blood Donation Center, Swiss Red Cross, Basel, Switzerland

[l] Emergency Department, University Hospital Basel, Basel, Switzerland

[m] Pediatric Infectious Diseases and Vaccinology, University Children's Hospital Basel and University of Basel, Basel Switzerland

[n] Centre for Primary Health Care, University of Basel, Basel, Switzerland

* Equal contribution

+ Corresponding author: adrian.egli@usb.ch

35

36 Short title: Influenza transmission within a city

37 Keywords: Influenza, city, socioeconomic status, transmission, population density, income,
38 vaccination rate, herd immunity, public transport

Abstract. (162/150)

With two-thirds of the global population projected to be living in urban areas by 2050, understanding the transmission patterns of viral pathogens within cities is crucial for effective prevention strategies. Here, in unprecedented spatial resolution, we analysed the socioeconomic determinants of influenza transmission in a European city. We combined geographical and epidemiological data with whole genome sequencing of influenza viruses at the scale of urban quarters and statistical blocks, the smallest geographic subdivisions within a city. We observed annually re-occurring geographic clusters of influenza incidences, mainly associated with net income, and independent of population density and living space. Vaccination against influenza was also mainly associated with household income and was linked to the likelihood of influenza-like illness within an urban quarter. Transmissions patterns within and between quarters were complex. High-resolution city-level epidemiological studies combined with social science surveys such as this will be essential for understanding seasonal and pandemic transmission chains and delivering tailored public health information and vaccination programs at the municipal level.

Introduction:

Transmission of influenza is influenced by extensive but poorly understood interactions between various viral, host and environmental factors[1-3]. Influenza may serve as a model for pandemic threats including the most recent COVID-19 pandemic. Whole viral genome sequencing has enabled reconstruction of phylogenetic relatedness at high resolution. Using these approaches, the interactions and dynamics of influenza transmission events have been described across a range of scales: globally[4-6], across continents[1,7], in university campuses[8], or within households[9-12]. With two-thirds of the global population projected to be living in urban areas by 2050, understanding the transmission patterns of influenza within cities is crucial for effective prevention strategies and may help to prepare for pandemic threats. Previous work identified cities as containing critical chains of transmission outside of peak climatic conditions (Dalziel, B. D. et al Science 2019), but the resolution to look at these critical intra-city transmission chains in detail has until now been lacking. Very few studies have explored transmission events and dynamics of influenza viruses at the scale of a city[13-17]. Cities are heterogeneous with remarkably different neighbourhoods based on the socioeconomic position of the individuals living there. Consequently, a high spatial resolution of the urban space and built environment is crucial to fully understanding the impact of factors linked to health and disease[18,19].

In this study we combine epidemiologic, geographical and demographical factors at the unprecedented resolution of urban quarters (i.e. neighbourhood/area) and statistical blocks (i.e. city block/street) levels, the smallest statistical enumeration areas within a city. Basel, Switzerland, serving as our model city, we explored the local patterns of influenza distribution and transmission from 2013 to 2018. Basel has urban quarters that differ substantially in socioeconomic indicators and housing structure. By using information at the level of statistical blocks, we were able to construct a detailed picture of influenza transmission within the city. We visualized kernel density estimates of influenza cases (reflecting the clustering of cases), a fundamental data smoothing method where inferences about the population are made, together with various population-based factors on maps with high resolution. Cases were corrected for population density and socioeconomic factors such as education levels, income and available living space were then analysed. Furthermore, we complemented the data with a detailed personal survey distributed to 30,000 households during the 2015/2016 flu season, to further explore urban quarter-specific aspects of self-reported influenza-like illness and association to socioeconomic factors. Finally, we described the details of influenza transmission during the 2016/2017 season using whole genome sequencing of all collected influenza viruses, covering

more than 650 isolates, combining the data, resulting in an unprecedented spatiotemporal resolution. The details on the study design for this project have previously been published[20].

Results:

Reoccurring influenza patterns within a city

The City of Basel has 19 urban quarters and a stable mean population of 175,350 inhabitants during the 5-year study period (+/- 1,737 inhabitants) (**Figure S1A**). A total of 1,078 statistical blocks were identified within the city for the subsequent analysis. First, we collected epidemiologic and geographic information of 1,715 PCR-confirmed influenza cases over five consecutive influenza seasons from 2013/2014 to 2017/2018 and identified areas with a high burden of influenza infections (**Figure S1B-C**). Influenza viruses isolated in Switzerland were similar to viruses isolated across Europe in the study period (**Table S1**). On a weekly basis, we linked each individual case of PCR-confirmed influenza to the patient's place of residence and anonymized the data at the scale of statistical blocks. We determined both the kernel density estimates for absolute influenza cases across all urban quarters and observed regional clusters of influenza cases within the city across all five seasons (**Figure 1A**) with similar spatial patterns for each individual season (**Figure S1D-H**).

We further explored the association between influenza occurrence and socioeconomic factors as well as the built city environment. We linked socioeconomic scores to each statistical block reflecting (i) the population density (inhabitants per hectare (ha)), (ii) the living space (per capita in m^2), and (iii) the net income (median in CHF). Each of these three key factors was translated to "socioeconomic points" ranging from one to five, which were added to generate a total socioeconomic score for each housing block: a score of three reflecting the lowest and a score of fifteen the highest possible socioeconomic value[20] (**Figure S2A-D**). The median socioeconomic scores of urban quarters ranged from three to ten. We observed that city areas with a high socioeconomic score had fewer influenza cases in comparison to areas with lower socioeconomic scores (**Figure 1B**). **Figure 1C** provides a representative example of the effect of population density during the 2016/2017 influenza season comparing the urban quarters Gundeldingen (GU, high population density with high influenza burden) with Bruderholz (BR, low population density with low influenza burden). These two urban quarters are located next to each other but are notably different. The urban quarter with small detached residential buildings (single family homes) showed a mean socioeconomic score of 9.5 (+/- 1.71 standard deviation) while the other urban quarter with large (multifamily) residential buildings shows a mean socioeconomic score of 5.6 (+/- 1.87 standard deviation). The housing structures of particular urban quarters may serve

as a surrogate marker for the population density and available living space in a particular area and thereby also be an indicator of influenza burden (**Figure 1D**).

To account for this potential ecologic fallacy, we corrected the influenza incidence rates per 1,000 inhabitants for each statistical block. We still observed similar dense influenza case patterns at the level of statistical blocks across urban quarters during the examined influenza seasons (**Figure 1E-F, S3A-H, Supplementary file 1**). Besides visualizations of influenza case distributions, we further explored associations between socioeconomic factors and the influenza incidence (corrected for the population density within each statistical block). The analysis included all five seasons and data from all 1,484 statistical blocks for every year. We used a multivariable Poisson regression and observed that net income of the statistical block (per 1,000 CHF) was the strongest predictor for influenza incidence at the level of statistical blocks, independent of population density and living space in most influenza seasons (**Figure 1G, S4**).

Socioeconomic factors determinate herd immunity

To further explore the interrelationship between socioeconomic factors and rates of influenza-like illness, we conducted a detailed survey across ten urban quarters for the 2015/2016 season (**Figure S1A**). The survey included 54 questions, addressing (i) influenza-like illness and vaccination, (ii) aspects of the urban environment, (iii) access to health care information, (iv) health related data, and (v) the place of residence at the level of statistical enumeration blocks. The English version of the survey is attached as supplementary file (**Supplementary File 2**).

The share of self-reported influenza-like illness cases fulfilling all three defining criteria for each of the selected urban quarters ranged from 3.4% to 7.0% (n=358; median 4.5%) during the influenza season 2015/2016 (**Figure 2A; S5A**). The number of influenza-like illness and PCR-confirmed cases (71 influenza A and 44 influenza B) did not correlate in that particular season across the explored statistical blocks responding to the survey (r=0.0144, p=0.7411; **Figure S5B**). However, the reported frequency of influenza-like illness corresponds to previously reported attack rates of three to five percent[21-23]. Other respiratory viruses likely contribute to cases matching the non-specific influenza-like illness definition. In our analysis, 20-25% of the influenza PCRs performed on patients presenting an influenza-like illness are subsequently confirmed by PCR-based diagnostics as indicated for the 2017/2018 influenza season (**Figure S5C**).

Next, we performed regression models with self-reported influenza-like illness as a binary endpoint for each individual. The factors associated with increased relative risks for self-reported influenza-like illness in stepwise forward and backward selection multivariable analysis were: ≥3 people per household and daily use of public transport. Of note, a total of 22.8% (1859 of 8149) were from households with ≥3 people. Of which 81.3% (1511/1859) were couples or single parents with children and 8.8% (720/8149) had children under the age of 7 years. The factors associated with decreased relative risks were: vaccination against influenza, age more than 65 years, and daily physical activity (**Table 1; Table S2**). The most protective variable against influenza-like illness was vaccination. In a multilevel multivariable model, vaccination remained the most important factor even when correcting for the urban quarter an individual lived in (RR 0.4, 95% CI 0.24 – 0.67). Of note, people with and without self-reported influenza vaccination showed similar odds ratios for symptoms of common cold (respiratory symptoms not fulfilling the influenza-like illness case definition, such as running nose and sore throat; OR 0.98, 95% CI 0.78 - 1.24).

The self-reported influenza vaccination rates were heterogeneously distributed at the level of urban quarters (**Figure 2B; S5D**). The median vaccine rate was 25.8% (IQR 18.7% - 32.2%) throughout the city, which is in the range of the previously reported averages in Switzerland [24]. We performed regression models with self-reported vaccination against influenza as binary endpoint for each individual. The factors consistently associated with an increased likelihood for self-reported influenza vaccination were people belonging to a risk group of chronic disease and health care workers. Low household income (below 6000 CHF per month) was associated with a significant lower likelihood of self-reported vaccination (**Table 2; Table S3**). When adjusting the multivariable model to specific urban quarters, low income was no longer correlated with vaccination (p=n.s.), which confirms the previously mentioned profound difference of income between urban quarters (**Figure S2C**).

The survey results confirm that vaccination against influenza has a protective effect against influenza-like illness. The odds of acquiring influenza-like illness during the 2015/2016 season were 3-times less in the group of vaccinated people (0.33 OR, 95% 0.24-0.46) in comparison to non-vaccinated people. This correlates to reported vaccine effectiveness rates of 44% in children and 78% in adults against influenza for the 2015/2016 season[25,26] (**Table S4**). The median self-reported vaccination rates between urban quarters showed a direct correlation with the socioeconomic score of the respective urban quarter (R^2=0.607, p=0.0079; **Figure 2C**) – the

higher the socioeconomic score, the higher the vaccine rate. Income showed the highest correlation with vaccination rates, followed by living space and population density, respectively (R^2 0.622, p=0.0067; R^2 0.618, p=0.007; R^2 0.532, p=0.0167). Self-reported vaccination rates may serve as a surrogate marker for herd immunity[27,28]. Similarly, at the level of an urban quarter, herd immunity strongly affected the risk of acquiring influenza-like illness during the 2015/2016 season. Urban quarters with a high self-reported vaccination rate showed a significantly lower likelihood of influenza-like illness of the surveyed population compared to urban quarters with a low vaccine rate (R^2=0.61, p=0.01; **Figure 2D**). However, no protective effect against common cold was observed (**Figure 2E**, p=0.56). Therefore, unvaccinated people living in the Matthaeus quarter (MA) show an 8-times higher probability of contracting an influenza-like illness than those vaccinated while in the Bruderholz quarter (BR) unvaccinated people only show a 1.2-fold higher risk. We again observed the association of self-reported influenza-like illness and vaccination rates (in % of returned survey) at the level of individual statistical blocks (r=-0.1182, p=0.007). However, in years with a low vaccine effectiveness (**Table S4**) this association may be weaker.

Humoral immunity of healthy donors is variable across urban quarters

In order to monitor antibody titres over time in a healthy population across the city, we recruited 214 healthy blood donors living in Basel before the 2016/2017 influenza season. We determined their hemagglutination inhibition titres against the circulating virus, H3N2 (Influenza A/Hong Kong/4801/2014). We also quantified antibody titres against all other vaccine strains (Influenza A/California/7/2009 H1N1 pdm09; Influenza B/Brisbane/60/2008; and Influenza B/Phuket/3073/2013). Previous antigen exposure to other strains may affect the response rates in the general population ([29,30]**; Table S1**). Before the 2016/2017 influenza season, we observed that across all urban quarters a median of 21% (IQR 17-28.5%) had seroprotective antibody levels (defined as hemagglutination inhibition titres equal or more than 1:40[31]) (**Figure 3A**). Again, urban quarters with lower socioeconomic scores also showed low seroprotection rates (e.g. Matthaeus, Breite, Kleinhueningen and Klybeck) (**Figure S6A**). Urban quarters with higher socioeconomic scores showed a median seroprotection rate of 26.1%, whereas those with lower socioeconomic scores showed a median seroprotection rate of 14.6% (p=0.05). Blood donors with influenza vaccination showed significant higher H3N2 specific HI titers in comparison to people who were not vaccinated (p<0.0001; **Figure S6B**). Similar to the survey, in this cohort net income was associated with the vaccination status. Blood donors who were influenza vaccinated had a median higher net income per statistical block when compared to non-vaccinated blood donors (median

CHF 54,144 vs. CHF 48,898, p=0.047; **Figure S6C**), whereas population density did not differ (p=0.39).

Geographic patterns of influenza transmission within a city

Next, we studied the impact of geographic and socioeconomic factors on influenza transmission using molecular epidemiological techniques. During the 2016/2017 season, we screened patients presenting an influenza-like illness using an influenza-specific PCR. We included 663 influenza virus positive samples recruited from 12 different study sites across the city ([32]; **Figure S7A**). The largest cohort within a single city, to the best of our knowledge. First, we determined antibody titres against the circulating H3N2 virus and other vaccine viruses in the patients, who having presented within days of feeling sick still had high viral loads meaning that antibody levels could be assumed to be close to baseline. Patients with PCR-confirmed influenza showed significant lower antibody titres against influenza A H3N2 in comparison to patients without influenza, highlighting the importance of protective antibodies (**Figure S7B** and **S7C**). Of note, antibody titres against viruses not circulating (e.g. H1N1) where not significantly different between PCR positive and negative patients.

A total of 663/858 viral genomes passed our duplicate and quality control filter (methods and[33]) and were included in subsequent analyses. 427/663 (64.4%) of isolates were from people living in Basel and the remaining 236/633 (37.3%) isolates were people living outside of Basel, mainly in surrounding villages within a 20km range but occasionally (n=6/633) from further away (e.g. Columbia, Turkey or Italy).

First, we compared the hemagglutinin (HA) gene sequence from our strains to a global sample of ~1400 influenza virus HA genes from the GISAID data repository of the 2016/2017 seasons. Using the integrated visualization of a phylogenetic tree and a map implemented in Nextstrain, we investigated the geographic distribution of influenza virus diversity at the levels of cities, urban quarters, and statistical blocks (https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/ha; links to visualization of all viral segments are provided in Table S5). This visualization can be interactively explored using the narrative functions of Nextstrain. The diversity of European influenza isolates was fully represented by the isolates collected from patients living in Basel and the surrounding area. In addition, also at the level of urban quarters and statistical blocks, we

observed the same high intermixture and diversity. Although the socioeconomic scores differ between urban quarters and statistical blocks, we did not observe an association with the phylogenetic tree structure (https://nextstrain-dev.herokuapp.com/community/narratives/appliedmicrobiologyresearch/Influenza-2016-2017/baselFlu).

Next, we further explored viral transmission clusters in more detail. Based on previously published mutation rates[34,35], we assumed that influenza viruses accumulate an average of 10 single nucleotide polymorphisms during a single influenza season. Viral strains with a genetic relatedness within this range were assigned as being in the same transmission cluster. A sequence, which is more than 10 single nucleotide polymorphisms different from any other sequence was ignored, so that a transmission cluster contains at least 2 viral strains. The 54 transmission clusters identified (**Figure 3B**) incorporated 547/663 influenza strains, each comprising a minimum of two isolates, a median of three isolates, ranging from 2 to 111 within and outside of the city (**Figure S7D**). Only 116 (17.5%) cases could not be linked to another individual. Most of the transmission clusters contained samples from more than one urban quarter, and only a few clusters were predominantly located within a single urban quarter (see clusters 19, 31 and 53, **Figure 3B;** https://nextstrain-dev.herokuapp.com/community/narratives/appliedmicrobiologyresearch/Influenza-2016-2017/baselFluClusters). Removing isolates from outside the city and focusing only on transmission events within city, we observed that 368/427 (76%) influenza strains belong to 43 of the overall 54 clusters across this single influenza season.

Next, we connected all identical influenza isolates between different urban quarters. Using permutation tests, we observed a generally high exchange rate and complex transmission dynamic between the different urban quarters. Some urban quarters showed significant connections to other urban quarters. For example, isolates from Vorstaedte (VO) and Wettstein (WE) were more identical in comparison to other urban quarters ($p<0.005$, **Figure 3C**). Transmission events within the same urban quarter were explored. Interestingly, two urban quarters – Gundeldingen (GU) and Vorstaedte (VO) – showed influenza isolates that were significantly more related to other isolates from within the same urban quarter than to isolates from other quarters or outside of Basel ($p<0.001$, **Figure 3C**). These two urban quarters show a low socioeconomic score and lower pre-seasonal seroprotection rate. Phylogenetic cluster size did not correlate with any socioeconomic factors (p=n.s.).

Conclusions

Each year influenza infects millions of people around the globe[36,37]. Historically, human pandemics have had devastating outcomes, and preventive measures should have high priority. Our results reflect an unprecedented large and dense datasets of PCR-confirmed influenza cases, the largest city-wide survey on influenza-like illness performed to date, and influenza genome sequences, all mapped to the statistical block level. We observed annually re-occurring geographic clusters of influenza cases and incidences. These influenza hot spots were mainly associated with net income, and independent of population density and living space. Vaccination against influenza was heterogeneously distributed between urban quarters and dependent mainly on household income. The rate of vaccination was linked to the likelihood of contracting an influenza-like illness within a specific urban quarter. Finally, transmission events between quarters were highly complex, but for two urban quarters particularly high levels of transmission events within the same urban quarter could be observed.

Our study has certain limitations. First, in the survey, a more research-interested population may have replied to the questionnaire. In particular, 37.9% of males replied to the survey, whereas in a census 48.2% were males. However, the overall reported vaccination are in line with regular interviews conducted by the Federal Office of Public Health of Switzerland, which report vaccination rates of >30% in people with 64 years of age and older[24]. Also, the self-reported influenza-like illness rates corresponds to estimated and published attack rates[38,39]. For influenza diagnostic using PCR confirmation, we cannot exclude some recruitment bias, that particular people with a specific demographic background are more likely to be included at a tertiary healthcare centre. In order to compensate for such a bias, we also included data from other study sites including a children's hospital and family doctors and a private diagnostic laboratory which also covers family doctors and other non-University hospitals in the region (**Figure S7A**). Further, ecologic fallacy may link certain socioeconomic factors, therefore we adjusted our analysis for population density. However, when correcting for population densities, statistical associations remained stable. Although, the influenza sampling for the sequencing part of the study was reasonably thorough, certainly not all influenza cases could be captured as many patients with influenza will not seek healthcare. Thus, some transmission links will always remain unknown. For more detailed analyses, future studies would need to include hundreds of samples with high resolution epidemiological data across multiple influenza seasons to study the changing patterns of transmission over multiple years. In addition, the transmission events across urban quarters

are highly complex, driven by mobility of people living within a city and commuters, place of working, and changing antibody titres of exposed patients in each year. Although we have asked where a person works, we were not able to collect sufficient data for this question – only 24.6% provided details and people working in the same statistical block were very rare (<5%).

Our findings provide important insights demonstrating influenza transmission patterns in a city serving as a model for studying the dynamics of seasonal flu transmission and evolution within a city. These results should be repeated in cities of different sizes and complexities around the world to allow public health services to be tailored most effectively. It will be interesting to see whether further factors concerning influenza transmission can be identified in other cities as well as their role in public health measures. Importantly, since vaccination rates were strongly dependent on income and linked to influenza incidences, providing better access to vaccination for low income households would likely have a substantial impact on influenza transmission. The knowledge gained with our study can help to tailor public health measures such as urban vaccination programs to urban influenza hotspots and not just to selected target groups (e.g. high-risk populations with chronic illness). National vaccination recommendations often only include a selected population, e.g. health care workers, patients with comorbidities or particular age groups. Our results suggest that such strategies may entirely miss the populations where most transmission occurs such as elderly, chronically ill, and children. Finally, this large-scale interdisciplinary study may serve as a blueprint for investigating seasonal and pandemic viruses on the smallest scale within a given geographical context such as a city. It may form the basis to develop more effective and targeted counter-measures at the most relevant public health levels e.g. at the urban quarter, or even smaller urban subdivisions and social milieus reflected therein. This would allow us to address more and account for a greater variety of population segments and help to identify potential drivers of transmission.

Methods and Material

Data and sample collection

The overall study design has been previously published [20]. Briefly, the study had retrospective and prospective parts. The retrospective study part consisted of an analysis of the numbers of PCR-confirmed influenza cases over the course of five years from 2013 to 2018. The prospective study part consisted of (i) a household survey focusing on influenza-like illness and vaccination against influenza during the 2015/2016 season; (ii) an analysis of influenza viruses collected during the 2016/2017 season using whole genome sequencing of viral genomic sequences to determine genetic relatedness, clusters and putative transmission events; and (iii) measurement of influenza-specific antibody titres against all vaccinated and circulated strains during the 2016/2017 season from healthy individuals, allowing us to monitor herd immunity across urban quarters. Particular patient groups may more likely present at a hospital and receive PCR-based diagnostics – to minimize this potential bias, we also integrated data from a large private diagnostic laboratory, which includes mainly non-hospitalized patients. Study teams at the University Hospital Basel and the University Children Hospital collected data and recruited patients for this study. In addition, a network of 24 paediatricians and family doctors also helped in recruiting patients (see acknowledgments). Viollier, a private laboratory, providing its services to a large number of private practices within the City of Basel, also provided samples and data.

Ethics and dissemination. The study is registered (clinicaltrials.gov; NCT03010007 on 22nd December 2016) and approved by the regional ethics committee as an observational study (EKNZ project ID 2015–363 and 2016-01735).

PCR-confirmed cases. Influenza specific PCRs were performed on nasopharyngeal swabs as part of the routine diagnostic workup. For (semiquantitative) PCR detection all laboratories used either the Xpert Flu/RSV or the Xpress Flu/RSV (Axonlab, Switzerland). For each PCR-confirmed case the exact place of residency was transposed to statistical blocks in order to anonymize the data using a geoinformation system (ArcGIS by ESRI, Switzerland) and to visualize the findings on maps. A statistical block is usually bordered from all sides by roads. In individual cases, the boundary is by zone plan categories (e.g. railway areas, forest, green zone, agricultural zone, etc. [40]). For each statistical block demographic data was obtained from the Statistics Office of the Canton Basel-City: net income in CHF, population per ha, and living space in m^2 per capita. These socioeconomic factors contributed to an overall socioeconomic score for every statistical block ranging from minimum three to maximum fifteen.

383

384 *Survey on ILI and vaccination.* We selected urban quarters for the survey a priori, based on
385 identified diversity in their socioeconomic structure (**Figure S2A-D**). We translated the
386 questionnaire into the six most common spoken languages in the selected urban quarters in order
387 to reach a maximum population. The survey was distributed within one week after the end of the
388 influenza season. Of the 30,000 questionnaires distributed in 10 urban quarters, 8,149 (27.1%)
389 were returned and fulfilled the quality criteria for subsequent analysis. Returned questionnaires
390 were quality controlled and the data was entered into a database. Questionnaires were excluded
391 when no information of influenza-like illness or vaccination status was reported. Influenza-like
392 illness was defined according by the WHO as a combination of self-reported fever, coughing, and
393 illness of less than ten days [41,42]. The questionnaires allowed the option of self-identifying one's
394 place of residence within a statistical block within the urban quarter.
395 Relative risks for influenza-like illness and influenza vaccination were estimated by uni- and
396 multivariable Poisson regression with robust error variance. To deal with possible confounding,
397 all variables found to differ significantly in univariable analyses between participants with and
398 without ILI and with and without influenza vaccination, respectively, were included in the
399 multivariable generalized linear models. A Bonferroni-adjusted p-value threshold was applied to
400 select variables for inclusion into the multivariable models. Poisson regression models using
401 stepwise forward and backward selection were applied to identify variables independently
402 associated with both primary outcomes.
403 To account for socioeconomic differences related to each urban quarter and potentially
404 influencing both the risk for influenza-like illness acquisition and vaccination uptake, analyses
405 regarding individual risk factors were complemented by multivariable, multilevel mixed-effects
406 generalized linear models. The Pearson and deviance goodness-of-fit tests were performed to
407 assess the fit of the data to a Poisson distribution in the final regression models. Analyses were
408 performed using Stata statistical software, version 15.1 (Stata Corp, College Station, Texas,
409 USA).
410
411 *Sample collection.* At 12 study sites we recruited paediatric and adult patients fulfilling the
412 following criteria: cough, fever, and sudden disease onset – these patients with influenza-like
413 illness were further evaluated using the previously described influenza specific PCR.
414
415
416

Whole genome sequencing and data analysis.

Details regarding the sequencing have been previously described[33]. Briefly, samples from PCR-confirmed cases had the whole genome amplified using PCR. From the resulting amplicons we created libraries using the Nextera XT protocol and sequenced them on a MiSeq platform (Illumina, San Diego, California) with 300bp paired end reads at 48-plex.

Quality control and genome assembly. We collected all data in the 2016/17 influenza season as previously described[20]. In total, we sequenced 857 samples. All genomes are available at NCBI with accession numbers MN299375 - MN304713. A full list of all used influenza genome for this publication (including GISAID strains) is available at: https://github.com/appliedmicrobiologyresearch/Influenza-2016-2017/blob/master/data_information/accession_numbers.tsv. If a patient was sampled more than once, we only included the first isolate, resulting in 755 remaining samples. If not otherwise indicated default settings of the bioinformatic tools were used. Raw Illumina reads were trimmed with Trimmomatic 0.36[43]. Alignment of paired-end reads used bowtie 2.2.3[44] with strain A/New York/18/2014 as a reference (Fludb.org segment accession numbers: PB2, KT837257; NP, KT837224; PA, KT837236; NA, KT837237; PB1, KT837241; M1/M2, KT837260; HA, KT837206; NS1/2, KT837196). The aligned reads were sorted by using samtools 1.2[45]. Variants were called and filtered by using LoFreq 2.1.2[46]. Variant calling was done for sites with a coverage of at least 100. Sites with a coverage of less than 100 were assumed to be unknown and were denoted as N, that is any possible nucleotide. Sequences that showed a read depth of 100 for at least 80% of the positions in at least four segments were used for the analysis. Non double infections with two strains were noted. Using these parameters, we continued our analysis using 663 samples. The consensus sequences from these strains were deposited in GenBank (numbers will be available upon acceptance of the manuscript). We aligned the consensus sequences using muscle v3.8.31[47] and used the concatenated alignment of all segments to calculated a maximum likelihood tree[8,48] using RaxML ([49]; version 8.2.8, options: -x 1522 -f a -m GTRGAMMA -p 1522 -# 100) to depict the relationship between the isolates.

Identification of transmission clusters. In order to trace local influenza virus lineages, we grouped the different influenza sequences into clusters using a maximum distance of 10 SNPs using an in-house python script (https://github.com/appliedmicrobiologyresearch/Influenza-2016-2017). In a recent publication the average evolutionary rate for influenza was determined to be 2.5×10^{-3} nucleotide substitutions per site per year[50]. With a genomes size of 13,588 bp this equates to 34

mutations per year. Therefore, 10 SNPs can be estimated to corresponds to 3.5 months, the length of a typical influenza season in Basel.

Transmission within urban quarters. In order to determine direct transmission between patients, we use the same in-house python script and clustered the influenza strains that have no detected variants in-between. With this we found that 139 of 663 isolates (all full genomes) are connected in 52 clusters in which the isolates are identical (0 SNPs). To identify quartiers that show increase transmission within or to other quartiers, we performed a permutation test (10,000 repetitions) in which the connections between quarters are compared to samples to which the quartiers are randomly assigned using an in-house python script (https://github.com/appliedmicrobiologyresearch/Influenza-2016-2017). The distribution of clusters over time and in the different quarters were visualized using ggplot2[51]. The phylogenetic trees were visualized using iTOL[52]. The transmission of isolates between different quarters was visualized using Circos[53].

Visualisation in Nextstrain. The visualization at https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/ha and https://nextstrain-dev.herokuapp.com/community/narratives/appliedmicrobiologyresearch/Influenza-2016-2017/baselFlu was produced using the Nextstrain toolchain[54]. Specifically, influenza virus A/H3N2 sequences were downloaded from GISAID[55], filtered down to ~1400 sequences from 2016-07-01 to 2017-06-30, augmented with influenza reference viruses, and combined with sequence data from this study. The resulting sequences were then aligned using mafft[56], a phylogenetic tree was built using IQ-tree[57], which was subsequently turned into a time scaled tree using treetime[58] in analogy to the analysis workflow used for the weekly updated influenza phylogenies at nextstrain.org/flu. The visualization is implemented through the Nextstrain community feature.

The analysis was repeated for each segment of the influenza genome, as well as for a concatenation of all segment. While the latter is not expected to yield trees that faithfully reflect the relationship of all viruses and segments to each other, such an analysis of concatenated genomes are nevertheless useful to resolve the relationship of very similar genomes.

The WGS dataset of all sequenced strains were visualized in the Nextstrain tool with demographic and socioeconomic metadata at spatiotemporal resolution – **Table S5** provides all links for data

485 access on nextstrain. We also used the narrative function of Nextstrain for specific highlighting
486 of aspects, namely the phylogenic tree structure across cities, urban quarters and statistical
487 blocks: https://nextstrain-
488 dev.herokuapp.com/community/narratives/appliedmicrobiologyresearch/Influenza-2016-
489 2017/baselFlu
490 and for specific transmission clusters: https://nextstrain-
491 dev.herokuapp.com/community/narratives/appliedmicrobiologyresearch/Influenza-2016-
492 2017/baselFluClusters
493). All used files can be downloaded at https://github.com/appliedmicrobiologyresearch/FluBasel.
494

Conflict of interest

None of the study authors has a conflict of interest to declare.

Funding of the study

This project was funded by the Swiss National Science foundation (SNF; grant number CR32I3_166258). AE received additional grants for this project: Freiwillige Akademische Gesellschaft Basel, Blutspende Zentrum SRK, and partially salary grant from SNSF Ambizione (PZ00P3_154709/1).

Acknowledgement

We thank the family doctors and pediatricians helping to recruit the patients for this study: Praxisgemeinschaft Dornacherstrasse (Dr. Burger, Dr. Eggenschwiler, Dr. Wyss, Dr. Gessler, Dr. Nonnenmacher), Praxis Bündnerhof (Dr. Müller, Dr. Peters, Dr. Hantke), Praxisgemeinschaft Banderet-Malè (Dr. Banderet-Uglioni, Dr. Malè), Hammerpraxis (Prof. Zeller), Praxis Schneider/von Hornstein (Dr. Schneider, Dr. von Hornstein), Davidsbodenpraxis (Dr. Amacher, Dr. Hug, Dr. Voelin, Isay, Dr. Pizzagalli, Dr. Navarini), Praxis Türkoglu/Bär (Dr. Türkoglu, Dr. Bär), and Praxis Gordon / Walker (Dr. Gorden, Dr. Walker). We also thank the study nurse team of the Clinical trial unit (Silke Purschke and Karin Wild) for their excellent organization and coordination of the patient recruitment. We thank Magdalena Schneider, Rosamaria Vesco, Christine Kiessling, Elisabeth Schultheiss, and Clarisse Straub for excellent technical assistance of genome sequencing. We also thank Dr. Vladimira Hinic (University Hospital Basel) and Prof. Emma Slack (ETH Zurich) for critically reading the manuscript. We thank Andrew Jermy for excellent feedback and critical review on the manuscript (Germinate, UK).

We also acknowledge the contributing colleagues and centers to GISAID (see https://github.com/appliedmicrobiologyresearch/Influenza-2016-2017/blob/master/data_information/acknowledgement_table.tsv for a full list).

Author contributions

Study design: AE, MaB, TS, RSS

Data capture: AE, CS, NG, MB, MS, AB, YH, JB, TV, NAS

Strain collection: AE, OD, MN

Whole genome sequencing: DL, DW, HSS

Hemagglutionation inhinition assay: MS, DL

Generation of maps: MB, NA, NS

Nextstrain visualization: RN, EH

Patient recruitment: NR, CHN, AZ, AB

Bioinformatic analysis: DW, EH, RN, NFM, JH, TB, HSS

Statistical analysis: STS

Writing of first draft manuscript: AE, NG

Reviewed the manuscript: MaB, TS, RSS, MB, NFM, DW, RN, HSS, AB, STS

Data availability statement: All sequencing data (raw reads) will be made available at NCBI. As well tables with anonymized PCR-confirmed cases and anonymized survey results will be made available in a public data repository.

Code availability statement: All codes used to process the viral genomic data will be made available on github.

542 **Tables**

543

544 **Table 1. Factors associated with self-reported influenza-like illness.** Forward and backward
545 selection multivariable regression models for factors, which are associated with influenza-like
546 illness. All variables with univariable associations to influenza-like illness are shown in **Table S1**.
547 RR, relative risks.

548

549 **Table 2. Factors associated with self-reported vaccination against influenza.** Forward and
550 backward selection multivariable regression models for factors which are associated with
551 vaccination against influenza. Data on univariate associations with vaccination can be found in
552 **Table S2**. RR, relative risks. [a], Anatomical Therapeutic Chemical (ATC) according to WHO
553 definition.

554 **Figure legends**

555

556 **Figure 1: Influenza cases in Basel 2013-2018.** The graph shows the distribution of PCR
557 confirmed influenza cases in the City of Basel at the resolution of statistical enumeration blocks.
558 **(A)** Geographical clustering of influenza infections. A kernel density estimation across all five
559 influenza seasons from 2013/2014 to 2017/2018 was used to analyse the influenza case
560 distribution across the city. The 4th and 5th share of the kernels were visualized. The red clouds
561 show Influenza A cases and the blue clouds Influenza B cases. **(B)** As **A**, with underlying
562 socioeconomic score for each housing block visualized, ranging from 3 to 15. The score consists
563 of three individual factors: population density per ha, living space per person in m^2 and net income
564 in CHF. The natural break of the socioeconomic scores were used and visualized with three
565 different shades of grey. Clusters of influenza correlate with lower socioeconomic scores. **(C)**
566 Case load across representative examples of urban quarters (Bruderholz as less densely
567 populated and Gundeldingen as more densely populated). Each red dot represents an Influenza
568 A case. In order to not show the actual address of a patient, we set the dots in the middle of the
569 statistical block. **(D)** We combined Google Earth satellite images with maps showing the kernel
570 density estimates of influenza cases to indicate building density. Red clouds show Influenza A
571 cases and blue Influenza B cases, correlating with building density, reflecting population density.
572 **(E)** Influenza A virus incidence rates of season 2016/2017 per 1000 inhabitants of each statistical
573 housing block. **(F)** Influenza B virus incidence rates of season 2016/2017 per 1000 inhabitants of
574 each statistical housing block. **(G)** Multivariable Poisson regression analysis to explore the
575 association of influenza virus incidence (per 1000 inhabitants corrected for the population of each
576 housing block) and each socioeconomic factor. Average influenza incidences for the seasons
577 2013/2014 to 2017/2018 are shown. Net income relative risk per 1 CHF; populations density per
578 person per ha; living density per person in m^2.

579

580 **Figure 2: Survey on self-reported Influenza-like illness and vaccination over season**
581 **2015/2016. (A)** Distribution of self-reported influenza-like illness cases according to the WHO
582 definition[41] over the 10 selected urban quarters. The number of Influenza-like illness cases per
583 incoming answers is expressed as percentage value for each quarter. The blue colour range is
584 according to natural breaks and ranges from 3.4% to 7%. **(B)** Self-reported vaccine rates (on a
585 per person basis / all received surveys for the specific urban quarter). Natural breaks between
586 urban quarters are shown in different shades of green, ranging from 18.7% to 32.2%. **(C)**
587 Comparison of socioeconomic score with self-reported vaccine rates using logistic regression

588 shows a correlation at the levels of urban quarter in the aggregate. **(D)** Self-reported vaccine rates
589 against influenza and the likelihood for ILI, using logistic regression, shows an inverse correlation.
590 **(E)** Self-reported vaccine rates against influenza and the likelihood of common cold, using logistic
591 regression, shows no correlation.
592
593 **Figure 3: Influenza antibody titres and whole genome sequencing of influenza viruses to**
594 **explore transmission between urban quarter levels. (A)** Pre-seasonal antibody titres against
595 H3N2 from healthy blood donors in the 2016/2017 influenza season measured using
596 hemagglutination inhibition assay. Distribution of seroprotection (defined as titre >1:40) are
597 visualized as percentages using natural breaks. **(B)** Cluster analysis of each influenza case
598 genome during the 2016/2017 season. Each dot represents a sequenced influenza virus isolate.
599 Colour reflects the host address by specific urban quarter. Each horizontal line corresponds to a
600 local transmission cluster of influenza viruses. Three representative clusters 19, 31 and 53 are
601 detailed on the right using maximum likelihood trees indicating the complexity of transmission
602 clusters across time and urban quarters. **(C)** Exchange of influenza viruses throughout the city.
603 Red lines connecting two quarters indicates isolates that are identical within the same
604 transmission cluster, with the line width and colour intensity indicating the number of isolates in
605 that cluster shared by the connected quarters. Therefore, they indicate the level of transmissions
606 between two quarters. When correcting for multiple comparison, only the transmission within the
607 same urban quarters remained statistically significant ($p<0.05$), which is indicated as a coloured
608 bar at the basis of the circle. The darker the colour the more significant the association.
609
610

Supplementary material

Table S1. Influenza virus types per season in Basel and Europe.

Table S2. Univariable analysis of factors associated with self-reported influenza-like illness.

Table S3. Univariable analysis of factors associated with self-reported vaccination.

Table S4. Vaccine effectiveness from 2013 to 2018.

Table S5. List of all link to Nextstrain for visualization of phylogenetic trees.

623 **Supplementary file 1.** Animation of cumulative PCR-confirmed influenza cases during the
624 Influenza 2017/2018 seasons. Influenza A and B are shown in the same colour. Based on the
625 incidence rates of Figure S4.
626
627 **Supplementary file 2.** Questionnaire, English version
628
629 **Supplementary file 3.** Submitted paper "Characterising the epidemic spread of Influenza A/H3N2
630 within a city through phylogenetics" by Müller NF, Wüthrich D, Goldman N et al.

631 **Figure S1: Influenza cases and distribution in the City of Basel per season from 2013/2014**
632 **to 2017/2018. (A)** Map of urban quarters showing the study area. Urban quarters included into
633 the household survey are marked in red. **(B)** Absolute numbers of PCR-confirmed influenza for
634 each season. **(C)** Cumulative influenza cases for each season. **(D-H)** Influenza cases per
635 statistical block shown as kernel density estimates for each influenza season.
636
637 **Figure S2: Socioeconomic data of the City of Basel.** Combined and individual socioeconomic
638 scores are shown for each statistical block. For each map, the natural breaks of the individual
639 factors is shown. **(A)** Combination of socioeconomic factors. **(B)** Living space in m^2 per capita.
640 **(C)** Net income in CHF per statistical block. **(D)** Population density as inhabitants per hectare (ha).
641
642 **Figure S3: Influenza incidence rates in the City of Basel per season from 2013/2014 to**
643 **2017/2018. (A-H)** PCR-confirmed influenza incidence rates (per 1000 inhabitants) for each
644 season per urban quarter.
645
646 **Figure S4: Multivariable Poisson regression for influenza incidence in each influenza**
647 **season from 2013/2014 to 2017/2018.** The correlation of each socioeconomic factors – income,
648 population density, and living density – with influenza incidence is expressed as relative risk (RR).
649
650 **Figure S5: Survey for self-reported ILI and vaccination against influenza. (A)** Self-reported
651 influenza-like illness cases corrected per returned questionnaire per statistical block. **(B)** Self-
652 reported ILI and PCR-confirmed influenza infections for the 2015/2016 influenza season. ILI
653 cases according to WHO definition are shown in grey for 10 selected urban quarters and PCR-
654 confirmed Influenza A as red dots and Influenza B as blue dots. **(C)** Ratio of tested PCRs and
655 positive results (PCR-confirmed) in a single season. **(D)** Self-reported vaccination corrected per
656 returned questionnaire per statistical block.
657
658 **Figure S6: Pre-seasonal influenza antibody titres measured by hemagglutination inhibition**
659 **assay in season 2016/2017. (A)** Percentage of seroprotected healthy blood donors (HIA titre of
660 1:40 or above), prior to the influenza season, by urban quarter. As for smaller urban quarters not
661 enough measurements were available, these were combined based on geographical proximity
662 and socioeconomic similarity. **(B)** Hemagglutination inhibition antibody titres in healthy blood
663 donors before the 2016/2017 season. Influenza A H3N2 titres against the circulating virus of that
664 season is shown in vaccinated and non-vaccinated people. **(C)** Net income in CHF per statistical

665 block shown for vaccinated and non-vaccinated healthy blood donors in the 2016/2017 season.
666
667 **Figure S7: Whole genome sequencing of influenza viruses during the 2016/2017 season.**
668 **(A)** The distribution of study recruitment sites for children (dark grey dots) and adults (white dots),
669 and the kernel density estimates of overall influenza burden in the city. **(B)** Hemagglutination
670 inhibition antibody titres against the circulating Influenza A H3N2 virus and all other vaccinated
671 Influenza viruses in each patient enrolled with influenza-like illness during the 2016/2017 season.
672 Patients with PCR-confirmed influenza are compared against patients with negative PCR testing.
673 All box blots show median and interquartile ranges. Whiskers indicate 10-90% percentile with
674 outliers. **(C)** The percentage of cumulative H3N2 specific hemagglutination inhibition antibody
675 titres is shown, again highlighting the difference between PCR positive and negative titres. **(D)**
676 Number of clusters and cluster size are shown, from the whole genome sequencing data. More
677 than 50% of the clusters comprise five or fewer viruses.
678
679 **Figure S8:** Cluster analysis of each influenza case genome during the 2016/2017 season based
680 on the phylodynamic model. Each dot represents a sequenced influenza virus isolate. Colour
681 reflects the host address by specific urban quarter. Each horizontal line corresponds to a local
682 transmission cluster of influenza viruses. Three representative clusters A, B and C are detailed
683 on the right using maximum likelihood trees indicating the complexity of transmission clusters
684 across time and urban quarters.
685

686 **Supplement methods and results**

687

688 **Calculations for tables 1 and 2 and table S3.**

689 Y serves as placeholder for the outcome variable and x as placeholder for the examined variable.
690 The calculations are performed a modified Poisson regression approach to prospective studies
691 with binary data[59].

692

693 **Univariable model**: glm y x ,fam (poisson) link(log)nolog eform robust

694 glm: generalized linear model

695 fam: family (describes the distribution; here we choose a poisson distribution to estimate relative
696 risks)

697 link: link function

698 log: exponentiated coefficients are incidence-rate ratios

699 eform: report exponentiated coefficients

700 nolog: suppresses the iteration log display

701 robust: robust variance estimator

702

703 **Multivariable model**

704 Stepwise forward selection: stepwise, pe(0.05): glm y x1 x2 x3 x(…) ,fam (poisson) link(log)nolog
705 eform robust

706

707 pe: begin with empty model

708 (0.05): specification of the significance level

709 All other abbreviations see above

710

711 Stepwise backward selection: stepwise, pr(0.05): glm y x1 x2 x3 x(…) ,fam (poisson)
712 link(log)nolog eform robust

713

714 Pr: begin with full model

715 All other abbreviations see above

716

717 **Multilevel model (mixed effects generalized linear model)**

718 meglm y x1 x2 x3 x(…) , II urban quarter:, fam (poisson) link(log)nolog eform

719 Group variable : urban quarter

720 **Phylodynamic model for transmission.**

721 In addition to the previously described clustering model based on the number of single nucleotide
722 polymorphisms, we have also applied a clustering method based on a phylogenetic approach
723 (see attached supplementary publication by Mueller N, Wuethrich D et al.).

724

725 **Methods:**

726 *Initial clustering based on nucleotide differences.* We combined the Basel sequences with a global
727 sample of sequences (downloaded on 17th July 2018) from https://www.gisaid.org which had
728 been sampled between 1st January 2016 and 31st December 2017 for which at least the
729 segments HA, NA and MP were available. We then calculated the average nucleotide difference
730 between any of the sequences and the sequences from Basel. To split the dataset into
731 manageable pieces, we first grouped any two sequences from Basel together if they were within
732 an average nucleotide difference of 0.0025 per position. When the full genome for two sequences
733 was available, this would correspond to about 32 different positions on the full genome. For an
734 average clock rate of 0.005 mutations per site per year, 32 differences correspond to a pairwise
735 phylogenetic distance of about 0.5 years. Sets of sequences from Basel were only split into two
736 groups if the two closest related sequences of each group exceeded this criterium. Based on this
737 initial grouping, we added global sequences to each cluster if they were at maximum 0.0025/2
738 mutations away from any of the sequences from Basel. Factor 2 was used to reduce the number
739 of global sequences in each of these initial clusters.

740

741 *Phylogenetic trees of initial clusters.* We next estimated rates of evolution for each segment using
742 the SRD06 model[60] and a strict clock model, from 200 full genome sequences sampled in
743 California, New York and Europe between 2010 and 2015, using Beast 2.5[61]. Apart from this,
744 these sequences were not otherwise used in the analysis, and as such are an independent
745 dataset. We allowed each segment to have its own phylogeny in order to avoid the possibility of
746 a reassortment biasing the estimates of evolutionary rates. Each of the segments, as well as the
747 first two and third codon position was allowed to have its own rate scaler. We ran ten independent
748 Markov Chain Monte Carlo (MCMC) chains each for 10^8 iterations and then combined them after
749 a burn-in of 10%. These estimated evolutionary rates are long-term rates of influenza A/H3N2.

750

751 We next reconstructed the phylogenetic trees of all initial clusters by using the full genomes of all
752 samples from the initial clusters. We fixed the evolutionary rates to be equal to the mean
753 evolutionary rates as estimated using the methods above. As a population model, we used a

constant coalescent model with an estimated effective population size that was shared among all initial clusters. We then estimated a distribution of phylogenies for each initial cluster, assuming that all segments share the same phylogeny. As estimated in the previous analysis, reassortment will not bias evolutionary rates.

Local cluster identification. To identify sets of sequences from Basel that were likely to have been transmitted locally, we reconstructed the geographic origins of lineages that were introduced into Basel. Therefore, we used the phylogenetic tree distributions for each initial cluster to reconstruct the ancestral states using parsimony. We made some modifications to the standard algorithm for parsimony ancestral state reconstruction to reflect our prior assumption, that Basel is unlikely to act as a relevant source of influenza on a global scale. If any descendent was not in Basel, the node was classified as "not Basel". Since the influenza season covers only a few months, we additionally assumed that lineages are unlikely to persist in Basel for more than 5 weeks (0.1 years) without being sampled. To reflect this assumption, we classified internal nodes that are more than 0.1 years from a sample from Basel to be either outside of Basel, or to be in an unknown location. We then defined sequences to be in the same local cluster if all their ancestors are inferred to be in Basel. Local clusters are obtained for each iteration of the MCMC. The exact workflow, including BEAST2 input files can be found at https://github.com/nicfel/Flu-Basel/LocalClusters. While alternative model-based approaches exist to reconstruct locations of internal nodes[62], these approaches themselves make strict assumptions that are violated when studying the spread of diseases on a city scale. Also, it is further unknown how well they perform when migration between individual locations is very strong. From the grouping of sequences into local clusters as described above, sequences were classified into different local clusters over the course of the MCMC.

Results:

Using clusters obtained from a phylodynamic approach, we obtain a similar clustering pattern meaning the pattern is robust towards how we define clusters. In particular, using the phylodynamic method, we identified 96 clusters of at least two highly similar viral genomes, representing influenza transmission, which together incorporated 534/663 influenza strains. We determined local transmission clusters by identifying groups of local isolates that are phylogenetically more closely related to each other than to any isolate from the global collection from GISAID (www.gisaid.org). Some of the clusters included isolates from both within and outside of the city, whereas other clusters were focused mainly within the city. Most of the clusters

788 contained samples from different urban quarters, and only a few clusters were predominantly
789 located within a single urban quarter (see Cluster A to C, **Figure S8**). Within the city, we observed
790 323/427 (76%) of influenza strains belong to 69 of the overall 96 clusters across this single
791 influenza season. The clusters comprised a median of 3 isolates, ranging from 2 to 42 within and
792 outside of the city.

References:

1 Geoghegan, J. L. *et al.* Continental synchronicity of human influenza virus epidemics despite climatic variation. *PLoS Pathog.* **14**, e1006780, doi:10.1371/journal.ppat.1006780 (2018).

2 Peci, A. *et al.* Effects of Absolute Humidity, Relative Humidity, Temperature, and Wind Speed on Influenza Activity in Toronto, Ontario, Canada. *Appl. Environ. Microbiol.* **85**, doi:10.1128/AEM.02426-18 (2019).

3 Roosa, K. & Chowell, G. Assessing parameter identifiability in compartmental dynamic models using a computational approach: application to infectious disease transmission models. *Theor. Biol. Med. Model.* **16**, 1, doi:10.1186/s12976-018-0097-6 (2019).

4 Langat, P. *et al.* Genome-wide evolutionary dynamics of influenza B viruses on a global scale. *PLoS Pathog.* **13**, e1006749, doi:10.1371/journal.ppat.1006749 (2017).

5 Lewis, N. S. *et al.* The global antigenic diversity of swine influenza A viruses. *Elife* **5**, e12217, doi:10.7554/eLife.12217 (2016).

6 Neher, R. A., Bedford, T., Daniels, R. S., Russell, C. A. & Shraiman, B. I. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1701-1709, doi:10.1073/pnas.1525578113 (2016).

7 Vijaykrishna, D. *et al.* The contrasting phylodynamics of human influenza B viruses. *Elife* **4**, e05055, doi:10.7554/eLife.05055 (2015).

8 Virk, R. K. *et al.* Molecular Evidence of Transmission of Influenza A/H1N1 2009 on a University Campus. *PLoS One* **12**, e0168596, doi:10.1371/journal.pone.0168596 (2017).

9 McCrone, J. T. *et al.* Stochastic processes constrain the within and between host evolution of influenza virus. *Elife* **7**, doi:10.7554/eLife.35962 (2018).

10 Poon, L. L. *et al.* Viral genetic sequence variations in pandemic H1N1/2009 and seasonal H3N2 influenza viruses within an individual, a household and a community. *J. Clin. Virol.* **52**, 146-150, doi:10.1016/j.jcv.2011.06.022 (2011).

11 Thai, P. Q. *et al.* Pandemic H1N1 virus transmission and shedding dynamics in index case households of a prospective Vietnamese cohort. *J. Infect.* **68**, 581-590, doi:10.1016/j.jinf.2014.01.008 (2014).

12 Valley-Omar, Z. *et al.* Intra-host and intra-household diversity of influenza A viruses during household transmissions in the 2013 season in 2 peri-urban communities of South Africa. *PLoS One* **13**, e0198101, doi:10.1371/journal.pone.0198101 (2018).

13 Adeola, O. A., Olugasa, B. O., Emikpe, B. O. & Folitse, R. D. Syndromic survey and molecular analysis of influenza viruses at the human-swine interface in two West African cosmopolitan cities suggest the possibility of bidirectional interspecies transmission. *Zoonoses Public Health* **66**, 232-247, doi:10.1111/zph.12559 (2019).

14 Dalziel, B. D. *et al.* Urbanization and humidity shape the intensity of influenza epidemics in U.S. cities. *Science* **362**, 75-79, doi:10.1126/science.aat6030 (2018).

15 Todd, S. *et al.* Primary care influenza-like illness surveillance in Ho Chi Minh City, Vietnam 2013-2015. *Influenza Other Respir Viruses*, doi:10.1111/irv.12574 (2018).

16 Chen, T. M. *et al.* The transmissibility estimation of influenza with early stage data of small-scale outbreaks in Changsha, China, 2005-2013. *Epidemiol. Infect.* **145**, 424-433, doi:10.1017/S0950268816002508 (2017).

17 Uchida, M. *et al.* High vaccination coverage is associated with low epidemic level of seasonal influenza in elementary schools: an observational study in Matsumoto City, Japan. *BMC Infect. Dis.* **18**, 128, doi:10.1186/s12879-018-3025-9 (2018).

18 Goncalves, L. *et al.* Urban Planning and Health Inequities: Looking in a Small-Scale in a City of Cape Verde. *PLoS One* **10**, e0142955, doi:10.1371/journal.pone.0142955 (2015).

19 Wilking, H., Hohle, M., Velasco, E., Suckau, M. & Eckmanns, T. Ecological analysis of social risk factors for Rotavirus infections in Berlin, Germany, 2007-2009. *Int J Health Geogr* **11**, 37, doi:10.1186/1476-072X-11-37 (2012).

20 Egli, A. *et al.* Identification of influenza urban transmission patterns by geographical, epidemiological and whole genome sequencing data: protocol for an observational study. *BMJ Open* **9**, e030913, doi:10.1136/bmjopen-2019-030913 (2019).

21 Park, J. E. & Ryu, Y. Transmissibility and severity of influenza virus by subtype. *Infect. Genet. Evol.* **65**, 288-292, doi:10.1016/j.meegid.2018.08.007 (2018).

22 Huang, Q. S. *et al.* Risk Factors and Attack Rates of Seasonal Influenza Infection: Results of the Southern Hemisphere Influenza and Vaccine Effectiveness Research and Surveillance (SHIVERS) Seroepidemiologic Cohort Study. *J. Infect. Dis.* **219**, 347-357, doi:10.1093/infdis/jiy443 (2019).

23 Somes, M. P., Turner, R. M., Dwyer, L. J. & Newall, A. T. Estimating the annual attack rate of seasonal influenza among unvaccinated individuals: A systematic review and meta-analysis. *Vaccine* **36**, 3199-3207, doi:10.1016/j.vaccine.2018.04.063 (2018).

24 Zurcher, K., Zwahlen, M., Berlin, C., Egger, M. & Fenner, L. Trends in influenza vaccination uptake in Switzerland: Swiss Health Survey 2007 and 2012. *Swiss Med. Wkly.* **149**, w14705, doi:10.4414/smw.2019.14705 (2019).

25 Kimiya, T. *et al.* Effectiveness of inactivated quadrivalent influenza vaccine in the 2015/2016 season as assessed in both a test-negative case-control study design and a traditional case-control study design. *Eur. J. Pediatr.* **177**, 1009-1017, doi:10.1007/s00431-018-3145-7 (2018).

26 Puig-Barbera, J. *et al.* Low influenza vaccine effectiveness and the effect of previous vaccination in preventing admission with A(H1N1)pdm09 or B/Victoria-Lineage in patients 60 years old or older during the 2015/2016 influenza season. *Vaccine* **35**, 7331-7338, doi:10.1016/j.vaccine.2017.10.100 (2017).

27 Fine, P., Eames, K. & Heymann, D. L. "Herd immunity": a rough guide. *Clin. Infect. Dis.* **52**, 911-916, doi:10.1093/cid/cir007 (2011).

28 Kim, T. H., Johnstone, J. & Loeb, M. Vaccine herd effect. *Scand. J. Infect. Dis.* **43**, 683-689, doi:10.3109/00365548.2011.582247 (2011).

29 Plant, E. P. *et al.* Different Repeat Annual Influenza Vaccinations Improve the Antibody Response to Drifted Influenza Strains. *Sci. Rep.* **7**, 5258, doi:10.1038/s41598-017-05579-4 (2017).

30 Olafsdottir, T. A. *et al.* Age and Influenza-Specific Pre-Vaccination Antibodies Strongly Affect Influenza Vaccine Responses in the Icelandic Population whereas Disease and Medication Have Small Effects. *Front. Immunol.* **8**, 1872, doi:10.3389/fimmu.2017.01872 (2017).

31 Trombetta, C. M., Perini, D., Mather, S., Temperton, N. & Montomoli, E. Overview of Serological Techniques for Influenza Vaccine Evaluation: Past, Present and Future. *Vaccines (Basel)* **2**, 707-734, doi:10.3390/vaccines2040707 (2014).

32 Egli, A. *et al.* Identification of influenza urban transmission patterns by geographical, epidemiological and whole genome sequencing: Protocol for an observational study. *BMJopen*, accepted, doi:http://dx.doi.org/10.1136/bmjopen-2019-030913 (2019).

33 Wuthrich, D. *et al.* Evaluation of two workflows for whole genome sequencing-based typing of influenza A viruses. *J. Virol. Methods* **266**, 30-33, doi:10.1016/j.jviromet.2019.01.009 (2019).

34 Kim, J. I. *et al.* Distinct molecular evolution of influenza H3N2 strains in the 2016/17 season and its implications for vaccine effectiveness. *Mol. Phylogenet. Evol.* **131**, 29-34, doi:10.1016/j.ympev.2018.10.042 (2019).

35 Xue, K. S., Moncla, L. H., Bedford, T. & Bloom, J. D. Within-Host Evolution of Human Influenza Virus. *Trends Microbiol.* **26**, 781-793, doi:10.1016/j.tim.2018.02.007 (2018).

36 Iuliano, A. D. *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet* **391**, 1285-1300, doi:10.1016/S0140-6736(17)33293-2 (2018).

37 Organization, W. H. *Global Influenza Strategy 2019-2030*. 34 (WHO, 2019).

38 Chunara, R., Goldstein, E., Patterson-Lomba, O. & Brownstein, J. S. Estimating influenza attack rates in the United States using a participatory cohort. *Sci. Rep.* **5**, 9540, doi:10.1038/srep09540 (2015).

39 Lee, R. U., Phillips, C. J. & Faix, D. J. Seasonal Influenza Vaccine Impact on Pandemic H1N1 Vaccine Efficacy. *Clin. Infect. Dis.* **68**, 1839-1846, doi:10.1093/cid/ciy812 (2019).

40 Basel-Stadt, S. O. o. t. C. *Statistischer Block*, <https://www.statistik.bs.ch/zahlen/raumdaten/raumeinheiten/block.html> (2019).

41 Fitzner, J. *et al.* Revision of clinical case definitions: influenza-like illness and severe acute respiratory infection. *Bull. World Health Organ.* **96**, 122-128, doi:10.2471/BLT.17.194514 (2018).

42 Organization, W. H. *WHO Global Epidemiological Surveillance Standards for Influenza*. 84 (WHO, 2014).

43 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

44 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

45 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

46 Wilm, A. *et al.* LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res.* **40**, 11189-11201, doi:10.1093/nar/gks918 (2012).

47 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

48 Munir, M. Bioinformatics analysis of large-scale viral sequences: from construction of data sets to annotation of a phylogenetic tree. *Virulence* **4**, 97-106, doi:10.4161/viru.23161 (2013).

49 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313, doi:10.1093/bioinformatics/btu033 (2014).

50 Klein, E. Y., Serohijos, A. W., Choi, J. M., Shakhnovich, E. I. & Pekosz, A. Influenza A H1N1 pandemic strain evolution--divergence and the potential for antigenic drift variants. *PLoS One* **9**, e93632, doi:10.1371/journal.pone.0093632 (2014).

51 Wickham, H. *ggplot2: Elegant Graphics for Data Analysis (UseR!)*. 2nd edn, 276 (Springer, New York, 2016).

52 Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**, 127-128, doi:10.1093/bioinformatics/btl529 (2007).

53 Krzywinski, M. *et al.* Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639-1645, doi:10.1101/gr.092759.109 (2009).

54 Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121-4123, doi:10.1093/bioinformatics/bty407 (2018).

55 Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Global challenges* **1**, 33-46, doi:10.1002/gch2.1018 (2017).

56 Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059-3066, doi:10.1093/nar/gkf436 (2002).

57 Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).

58 Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus evolution* **4**, vex042, doi:10.1093/ve/vex042 (2018).

59 Zou, G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* **159**, 702-706, doi:10.1093/aje/kwh090 (2004).

60 Shapiro, B., Rambaut, A. & Drummond, A. J. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. *Mol. Biol. Evol.* **23**, 7-9, doi:10.1093/molbev/msj021 (2006).

61 Bouckaert, R. *et al.* BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **15**, e1006650, doi:10.1371/journal.pcbi.1006650 (2019).

62 Vaughan, T. G., Kuhnert, D., Popinga, A., Welch, D. & Drummond, A. J. Efficient Bayesian inference under the structured coalescent. *Bioinformatics* **30**, 2272-2279, doi:10.1093/bioinformatics/btu201 (2014).

Table 1. Factors associated with self-reported influenza-like illness. Forward and backward selection multivariable regression models for factors, which are associated with influenza-like illness. All variables with univariable associations to influenza-like illness are shown in **Table S1**. RR, relative risks.

	Univariable			**Multivariable with stepwise forward variable selection**			**Multivariable with stepwise backward variable selection**		
	RR	**95%CI**	**p-value**	**RR**	**95%CI**	**p-value**	**RR**	**95%CI**	**p-value**
Vaccination	0.33	0.24-0.46	<0.001	0.38	0.23-0.64	<0.001	0.38	0.23-0.64	<0.001
Smoking	1.33	1.16-1.51	<0.001	--	--	--	**--**	**--**	**--**
Health check-ups	0.82	0.75-0.89	<0.001	--	--	--	**--**	**--**	**--**
Physical activity	0.83	0.75-0.02	<0.001	0.86	0.75-0.99	0.046	0.86	0.75-0.99	0.046
Daily public transport use	1.17	1.07-1.29	0.001	1.27	1.11-1.44	<0.001	1.27	1.11-1.44	<0.001
Contact with > 50 people per day	1.57	1.24-1.98	<0.001	--	--	--	**--**	**--**	**--**
Age > 65 years	0.25	0.18-0.36	<0.001	0.35	0.19-0.66	0.001	0.35	0.19-0.66	0.001
Living space (≤ 30 m2/capita)	1.71	1.26-2.34	0.001	--	--	--	**--**	**--**	**--**
≥ 3 people / household	2.06	1.68-2.54	<0.001	1.53	1.14-2.06	0.005	1.53	1.14-2.06	0.005

Table 2. Factors associated with self-reported vaccination against influenza. Forward and backward selection multivariable regression models for factors which are associated with vaccination against influenza. Data on univariate associations with vaccination can be found in **Table S2**. RR, relative risks. [a], Anatomical Therapeutic Chemical (ATC) according to WHO definition.

	Univariable			**Multivariable with stepwise forward variable selection**			**Multivariable with stepwise backward variable selection**		
	RR	**95%CI**	**p-value**	**RR**	**95%CI**	**p-value**	**RR**	**95%CI**	**p-value**
Risk group based on ATC-level[a]	1.75	1.50-2.03	<0.001	1.69	1.16-2.45	0.006	1.69	1.16-2.45	0.006
Smoking	0.68	0.63-0.74	<0.001	--	--	--	--	--	--
Hand washing	1.11	1.05-1.17	<0.001	--	--	--	--	--	--
Health check-ups	1.37	1.32-1.42	<0.001	--	--	--	--	--	--
Healthy diet	0.86	0.83-0.91	<0.001	--	--	--	--	--	--
Physical activity	0.91	0.87-0.95	<0.001	--	--	--	--	--	--
Contact with more than 50 people a day	0.78	0.70-0.88	<0.001	--	--	--	--	--	--
Healthcare worker	2.40	2.13-2.70	<0.001	1.84	1.35-2.51	<0.001	1.84	1.35-2.51	<0.001
Male sex	1.13	1.05-1.22	0.001	--	--	--	--	--	--
Age > 65 years	2.72	2.53-2.93	<0.001	--	--	--	--	--	--
living space (≤ 30m^2/capita)	0.56	0.46-0.69	<0.001	--	--	--	--	--	--
≥ 3 people / household	0.69	0.63-0.77	<0.001	--	--	--	--	--	--
Lower education level	1.62	1.43-1.83	<0.001	--	--	--	--	--	--
Upper and middle management	1.37	1.20-1.55	<0.001	--	--	--	--	--	--
Low Income (≤ CHF 6000 gross household income)	0.79	0.72-0.86	<0.001	0.63	0.44-0.89	0.010	0.63	0.44-0.89	0.010
Share of green areas (≤ 20% / residential block)	1.22	1.10-1.34	<0.001	--	--	--	--	--	--

Table S1. Influenza virus types per season in Basel and Europe

	City of Basel	Switzerland[54-58]	Europe[59-63]
2013/2014	Influenza A (n=136, 97%) Influenza B (n=5, 3%)	Total samples n=580 Influenza A (98%): H1N1 41%, H3N2 56% Influenza B (2%): Yamagata 2%, Victoria 0%	Total samples n=34,210 Influenza A (94%): H1N1 41%, H3N2 47% Influenza B (6%): Yamagata 1%, Victoria 0%
2014/2015	Influenza A (n=226, 67%) Influenza B (n=110, 23%)	Total samples n=937 Influenza A (71%): H1N1 14%, H3N2 56% Influenza B (29%): Yamagata 27%, Victoria 1%	Total samples n=40,931 Influenza A (68%): H1N1 15%, H3N2 49% Influenza B (32%): Yamagata 8%, Victoria 0%
2015/2016	Influenza A (n=92, 62%) Influenza B (n=56, 38%)	Total samples n=975 Influenza A (35%): H1N1 30% H3N2 5% Influenza B (64%): Yamagata 2%, Victoria 62%	Total samples n=56,892 Influenza A (59%): H1N1 49%, H3N2 7% Influenza B (41%): Yamagata 1%, Victoria 18%
2016/2017	Influenza A (n=501, 99%) Influenza B (n=5, 1%)	Total samples n=982 Influenza A (96%): H1N1 2% H3N2 94% Influenza B (4%): Yamagata 3%, Victoria 0%	Total samples n=49,410 Influenza A (90%): H1N1 1%, H3N2 75% Influenza B (10%): Yamagata 2%, Victoria 2%
2017/2018	Influenza A (n=268, 38%) Influenza B (n=433, 62%)	Total samples n=1,292 Influenza A (29%): H1N1 23%, H3N2 5% Influenza B (71%): Yamagata 66%, Victoria 0%	Total samples n=60,658 Influenza A (37%): H1N1 20%, H3N2 11% Influenza B (63%): Yamagata 29%, Victoria 1%

Table S2. Univariable analysis of factors associated with self-reported influenza-like illness.

Variables	ILI yes		ILI no		Relative Risk	95% CI	p-Value
	n	in %	n	in %			
Vaccination (yes)	37	1.8	2036	98.2	0.33	0.24-0.46	<0.001
Risk group based on ATC-level (yes)	52	3.0	1678	97.0	0.59	0.37-0.95	0.029
Smoker							
no	250	4.0	5971	96.0	1.33	1.16-1.51	<0.001
yes, sometimes	47	5.3	837	94.7			
yes, daily	61	7.1	801	92.9			
Alcohol consumption							
no	825	4.9	1582	95.1	0.78	0.66-0.92	0.003
yes, sometimes	259	4.7	5261	95.3			
yes, daily	15	2.0	742	98.0			
Implementation of hand washing							
Very bad	0	0.0	19	0.0	1.05	0.91-1.21	0.510
Bad	4	5.8	65	94.2			
Average	33	4.1	770	95.9			
Good	126	4.6	2639	95.4			
Very good	193	4.7	3939	95.3			
Implementation of health check-ups							
Very bad	56	6.1	856	93.9	0.82	0.75-0.89	<0.001
Bad	65	6.8	886	93.2			
Average	91	4.8	1809	95.2			
Good	65	3.6	1731	96.4			
Very good	33	3.0	1061	97.0			
Implementation of healthy diet							
Very bad	1	5.6	17	94.4	0.82	0.73-0.93	0.002
Bad	10	8.3	110	91.7			
Average	63	4.8	1252	95.2			
Good	189	5.2	3475	94.8			
Very good	94	3.5	2594	96.5			
Implementation of physical activity							
Very bad	3	3.4	84	96.6	0.83	0.75-0.02	<0.001
Bad	34	7.7	409	92.3			
Average	94	5.3	1671	94.7			
Good	137	4.7	2761	95.3			
Very good	90	3.6	2395	96.4			
Public transport use					1.17	1.07-1.29	
Never	2	1.4	144	98.6			0.001
Rare	47	3.7	1214	96.3			
Multiple times a month	69	4.1	1598	95.9			
Multiple times a week	92	4.4	2007	95.6			
Daily	142	5.6	2375	94.4			
Contact with more than 50 persons a day	86	6.4	1250	93.6	1.57	1.24-1.98	<0.001
Healthcare worker	44	4.6	912	95.4	0.79	0.57-1.08	0.135
work with children	40	5.2	728	94.8	0.90	0.65-1.25	0.541
Work in an open plan office	135	6.2	2040	93.8	1.19	0.94-1.49	0.147

Gender Male Female	 129 229	 4.2 4.6	 2907 4715	 95.8 95.4	0.92	0.74-1.13	0.423
old age (≥ 65 years)	36	1.5	2349	98.5	0.25	0.18-0.36	<0.001
Foreign population	93	5.8	1508	94.2	1.41	1.12-1.77	0.004
living space (≤ 30 m^2/capita)	48	7.8	566	92.2	1.71	1.26-2.34	0.001
≥ 3 persons / household	137	7.5	1700	92.5	2.06	1.68-2.54	<0.001
Lower educational level (compulsory)	13	3.2	391	96.8	0.70	0.40-1.20	0.195
Upper and middle management	63	4.7	1270	95.3	0.80	0.61-1.05	0.108
Low Income (≤ CHF 6000 gross household income)	150	4.8	2985	95.2	1.00	0.80-1.25	0.995
Share of green areas (≤ 20% / residential block)	85	5.1	1590	94.9	0.83	0.65-1.05	0.121
Share of built up area (≥ 60% / residential block)	29	3.9	721	96.1	0.87	0.60-1.27	0.472

Table S3. Univariable analysis of factors associated with self-reported vaccination.

Variables	vaccination yes		vaccination no		Relative Risk	95% CI	p-Value
	n	in %	n	in %			
Risk group based on ATC-level	826	47.1	929	52.9	1.75	1.50-2.03	<0.001
Smoker							
no	1819	28.9	4473	71.1	0.68	0.63-0.74	<0.001
yes, sometimes	150	16.7	749	83.3			
yes, daily	130	14.9	742	85.1			
Alcohol consumption							
no	479	28.3	1215	71.7	1.03	0.95-1.100	0.477
yes, sometimes	1356	24.3	4222	75.7			
yes, daily	254	33.1	513	66.9			
Implementation of hand washing							
Very bad	4	20.0	16	80.0	1.11	1.05-1.17	<0.001
Bad	16	23.2	53	76.8			
Average	187	23.0	626	77.0			
Good	675	24.2	2113	75.8			
Very good	1160	27.2	3033	72.3			
Implementation of health check-ups							
Very bad	128	13.9	792	86.1	1.37	1.32-1.42	<0.001
Bad	163	16.9	801	83.1			
Average	439	22.8	1487	77.2			
Good	600	33.2	1209	66.8			
Very good	502	45.1	611	54.9			
Implementation of healthy diet							
Very bad	6	31.6	13	68.4	0.86	0.83-0.91	<0.001
Bad	38	31.9	81	68.1			
Average	413	31.0	920	69.0			
Good	940	25.4	2760	74.6			
Very good	614	22.5	2113	77.5			
Implementation of physical activity							
Very bad	24	27.0	65	73.0	0.91	0.87-0.95	<0.001
Bad	131	29.4	314	70.6			
Average	515	28.7	1278	71.3			
Good	712	24.4	2211	75.6			
Very good	574	22.8	1946	77.2			
Public transport use					1.05	1.01-1.08	
Never	40	27.2	107	72.8			0.005
Rare	265	20.8	1010	79.2			
Multiple times a month	423	25.0	1268	75.0			
Multiple times a week	643	30.3	1480	69.7			
Daily	643	25.3	1903	74.4			
Contact with more than 50 persons a day	281	20.8	1069	79.2	0.78	0.70-0.88	<0.001
Healthcare worker	323	33.4	645	66.6	2.40	2.13-2.70	<0.001
work with children	111	14.3	667	85.7	0.78	0.65-0.94	0.008
Work in an open plan office	396	18.0	1810	82.0	1.02	0.91-1.16	0.708
Gender					1.13	1.05-1.22	
Male	857	27.9	2217	72.1			0.001
Female	1235	24.7	3767	75.3			
old age (≥ 65 years)	1116	46.1	1303	53.9	2.72	2.53-2.93	<0.001
Foreign population	377	23.2	1249	76.8	0.87	0.79-0.96	0.005
living space (≤ 30m^2/capita)	89	14.4	527	85.6	0.56	0.46-0.69	<0.001

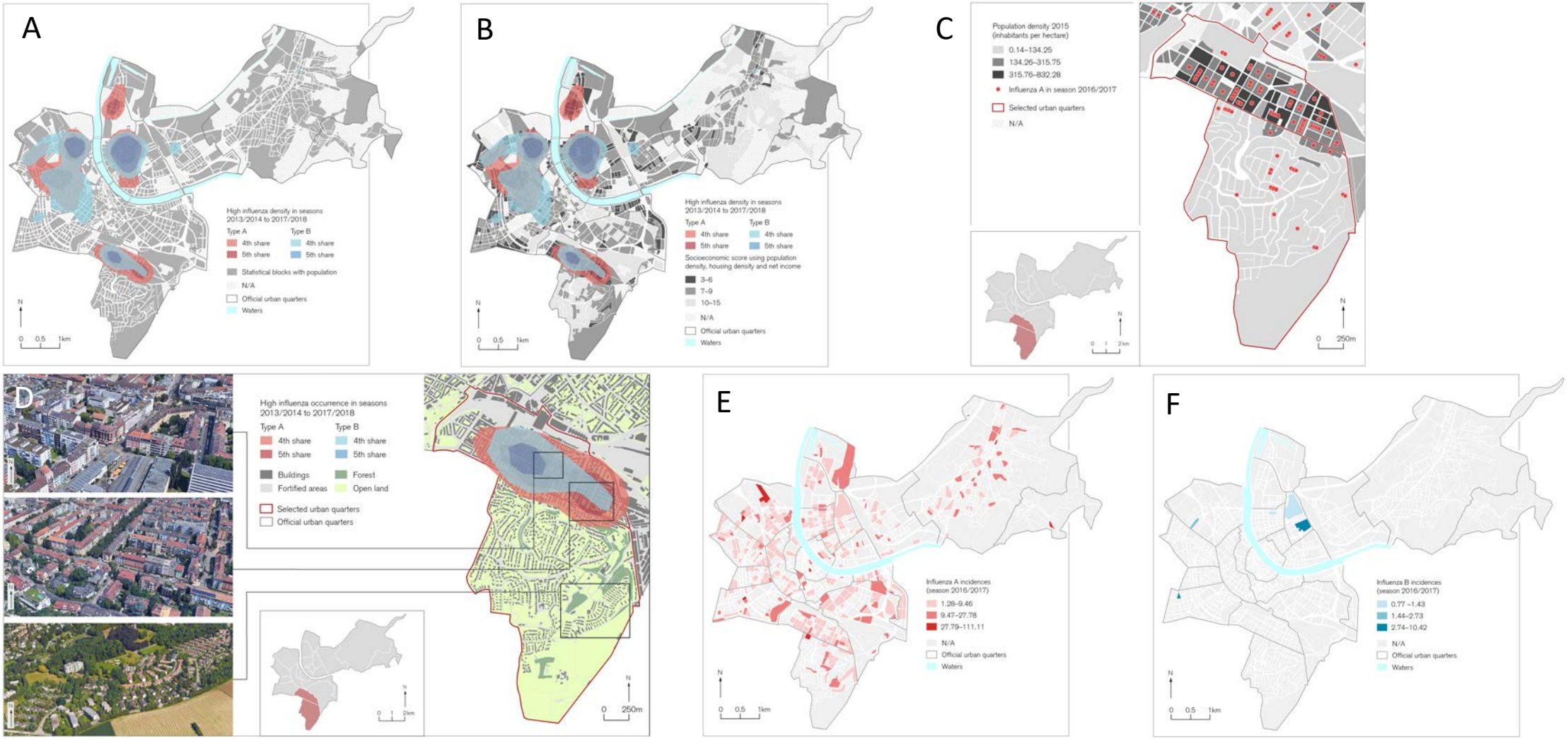
≥ 3 persons / household	359	19.3	1500	80.7	0.69	0.63-0.77	<0.001
Lower educational level (compulsory)	168	40.3	249	59.7	1.62	1.43-1.83	<0.001
Upper and middle management	297	22.0	1051	78.0	1.37	1.20-1.55	<0.001
Low Income (≤ CHF 6000.- gross household income)	673	21.3	2491	78.7	0.79	0.72-0.86	<0.001
Share of green areas (≤ 20% / residential block)	375	22.1	1324	77.9	1.22	1.10-1.34	<0.001
Share of built up area (≥ 60% / residential block)	161	21.2	597	78.8	0.81-	0.70-0.93	0.003

Table S4. Vaccine effectiveness from 2013 to 2018

	Influenza strain				Reference
	A (H1N1)	**A (H3N2)**	**B (Yamagata)**	**B (Victoria)**	
2013/2014	57.9%; 62%; 74%	-	-	-	[64-66]
2014/2015	-	-22%; -8%	-	-	[67,68]
2015/2016	20.2%, 45%; 64%; 68%	43%	49%	-33.2%; 57%	[25,69-71]
2016/2017	-	38%; 42%; 43%; 48%	-	-	[72-75]
2017/2018	55%; 68%	-42%; -27%; -16%; -1%; 7%	49%; 77%	-	[76]

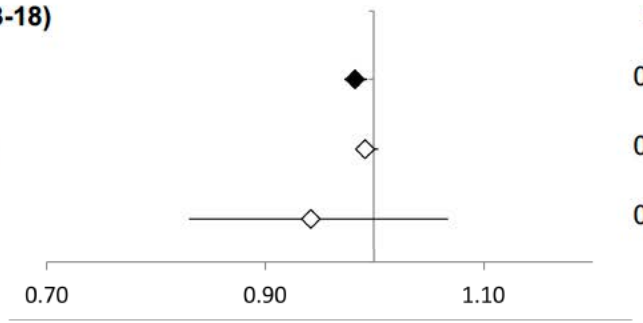
Table S5. List of all link to Nextstrain for visualization of phylogenetic trees

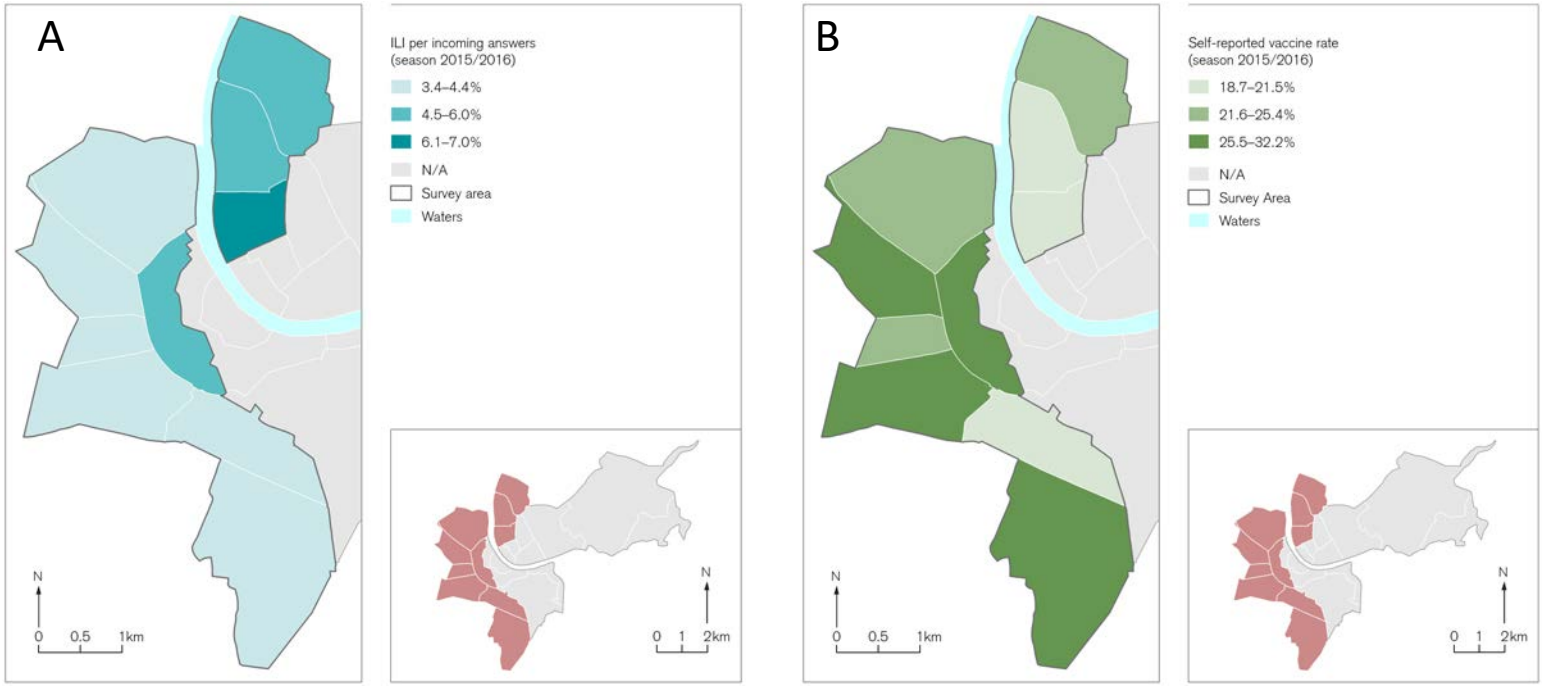
Viral segment	Links
all	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/full
HA	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/ha
NA	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/na
NP	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/np
PA	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/pa
PB1	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/pb1
PB2	https://nextstrain-dev.herokuapp.com/community/appliedmicrobiologyresearch/Influenza-2016-2017/h3n2/pb2

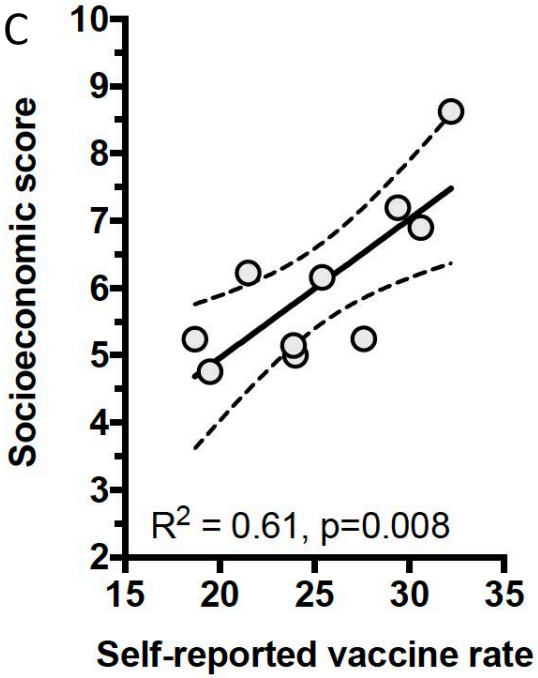


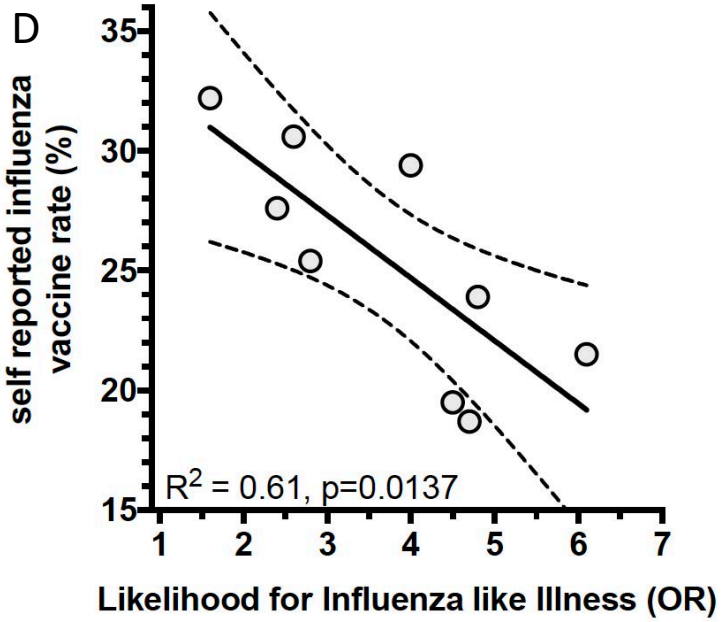
G

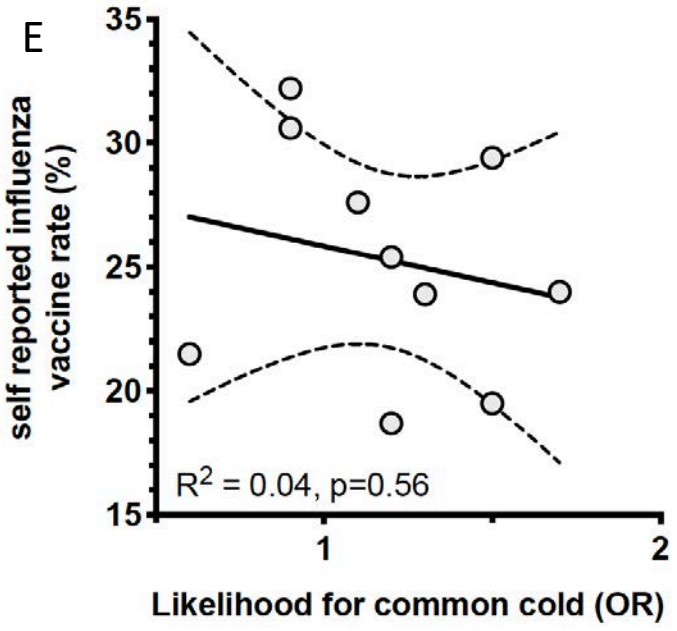
Average incidence (2013-18)	RR	95%CI	p-value
Income	0.983	0.972-0.994	0.002
Population density	0.993	0.984-1.003	0.192
Living density	0.942	0.831-1.068	0.349



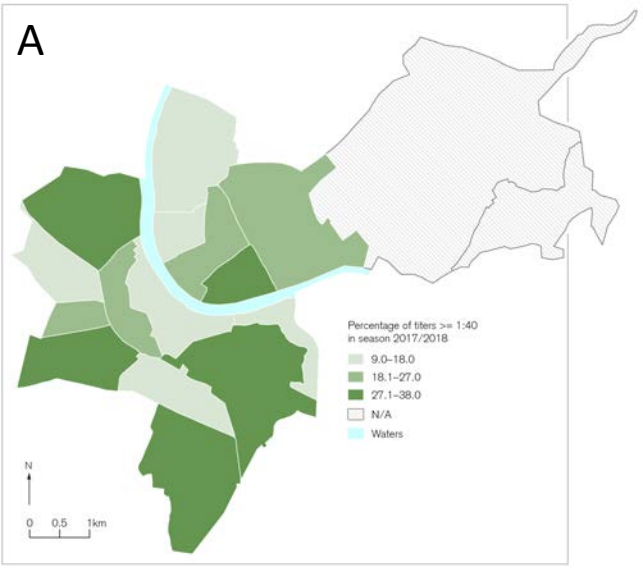

A
ILI per incoming answers
(season 2015/2016)
3.4–4.4%
4.5–6.0%
6.1–7.0%
N/A
Survey area
Waters
N
0 0.5 1km
0 1 2km
B
Self-reported vaccine rate
(season 2015/2016)
18.7–21.5%
21.6–25.4%
25.5–32.2%
N/A
Survey Area
Waters
N
0 0.5 1km
0 1 2km


C
Socioeconomic score
Self-reported vaccine rate
R2 = 0.61, p=0.008

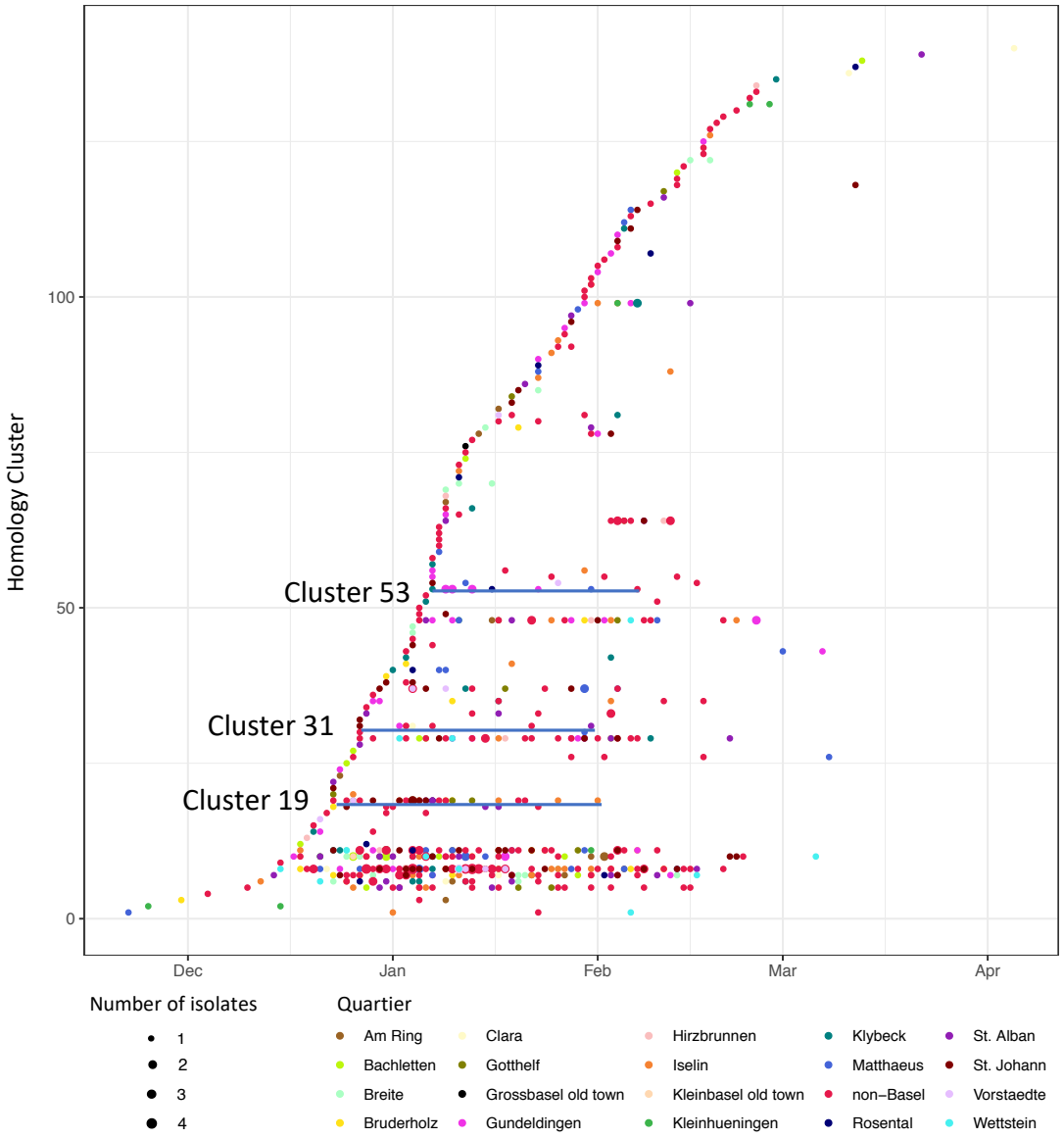

D
self reported influenza vaccine rate (%)
Likelihood for Influenza like Illness (OR)
R2 = 0.61, p=0.0137


E
self reported influenza vaccine rate (%)
Likelihood for common cold (OR)
R2 = 0.04, p=0.56

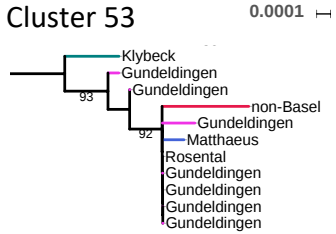
A



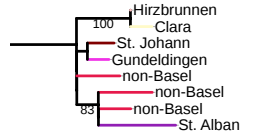
B



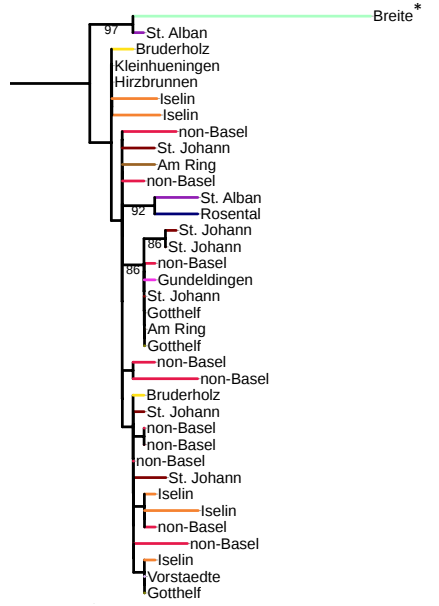
Cluster 53



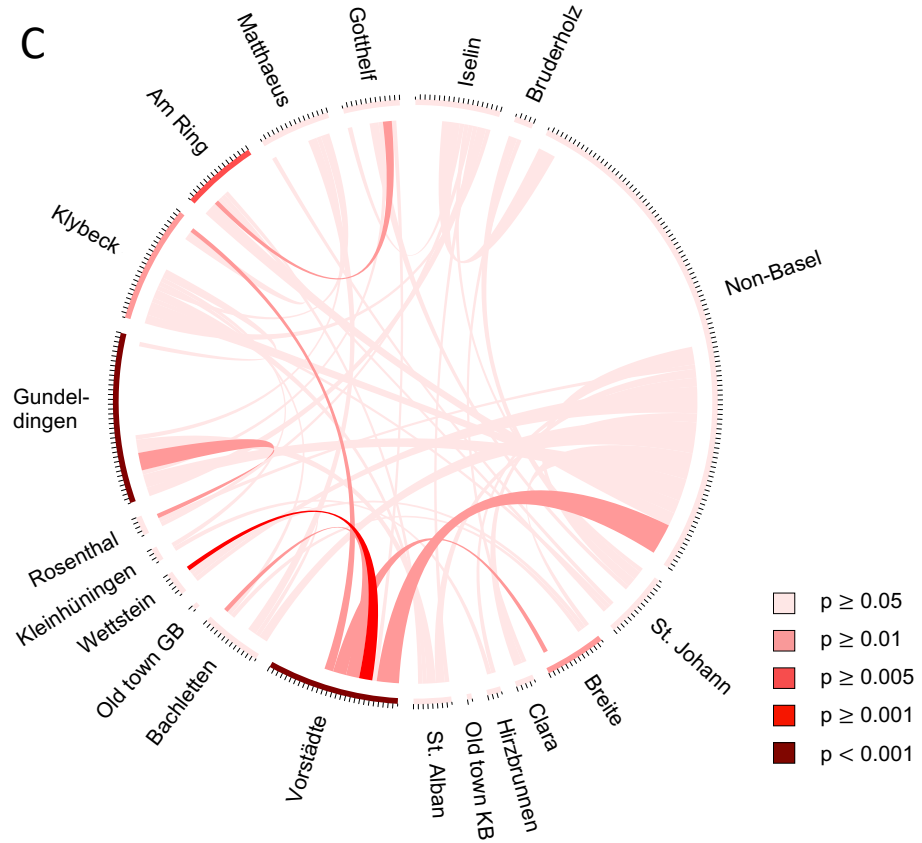
Cluster 31



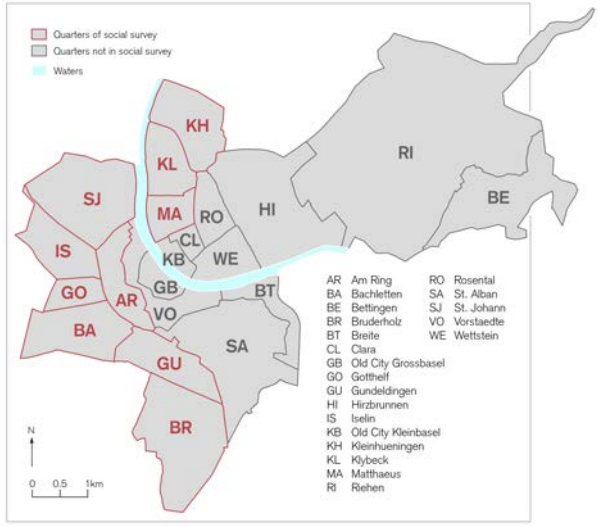
Cluster 19



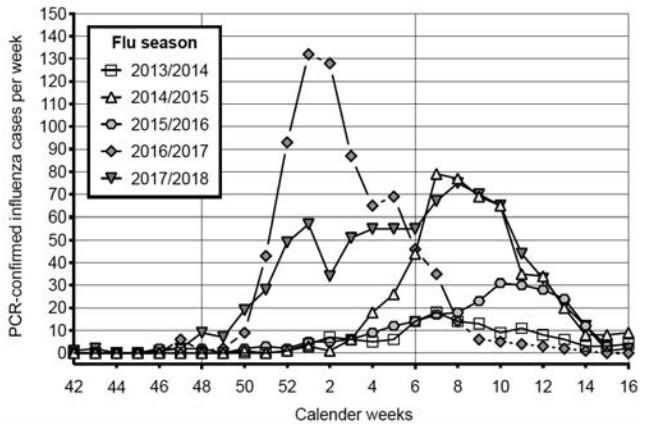
C



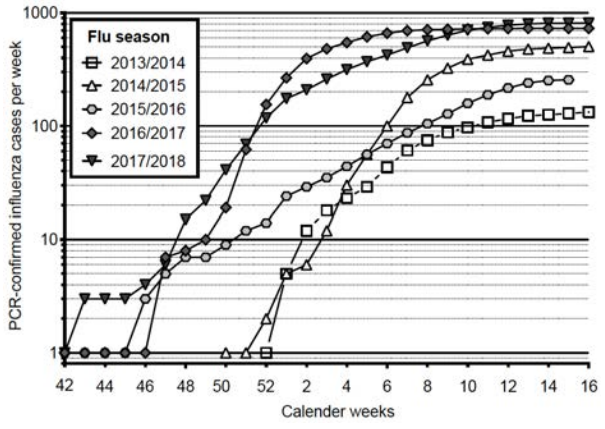
A


Quarters of social survey
Quarters not in social survey
Waters
KH
KL
SJ
MA
RO
HI
RI
BE
CL
IS
KB
WE
GB
GO
AR
BT
VO
BA
SA
GU
BR
AR Am Ring
BA Bachletten
BE Bettingen
BR Bruderholz
BT Breite
CL Clara
GB Old City Grossbasel
GO Gotthelf
GU Gundeldingen
HI Hirzbrunnen
IS Iselin
KB Old City Kleinbasel
KH Kleinhueningen
KL Klybeck
MA Matthaeus
RI Riehen
RO Rosental
SA St. Alban
SJ St. Johann
VO Vorstaedte
WE Wettstein
N
0 0.5 1km

B


Flu season
2013/2014
2014/2015
2015/2016
2016/2017
2017/2018
PCR-confirmed influenza cases per week
Calender weeks

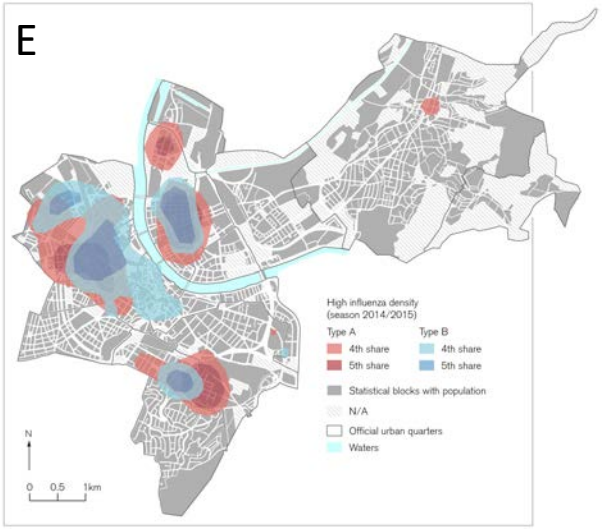
C


Flu season
2013/2014
2014/2015
2015/2016
2016/2017
2017/2018
PCR-confirmed influenza cases per week
Calender weeks

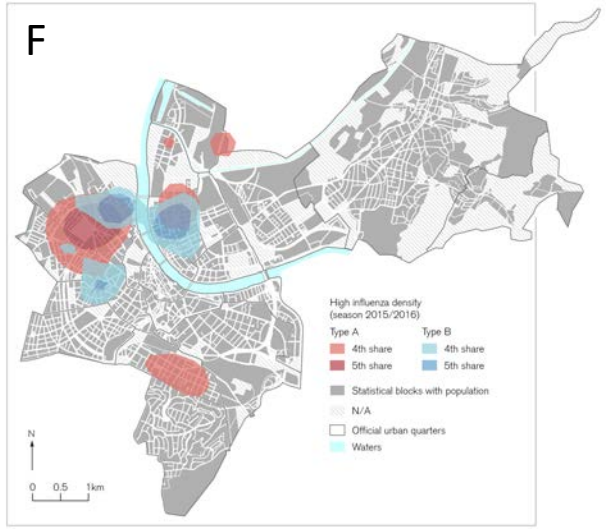
D


High influenza density
(season 2013/2014)
Type A
Type B
4th share
5th share
Statistical blocks with population
N/A
Official urban quarters
Waters
N
0 0.5 1km

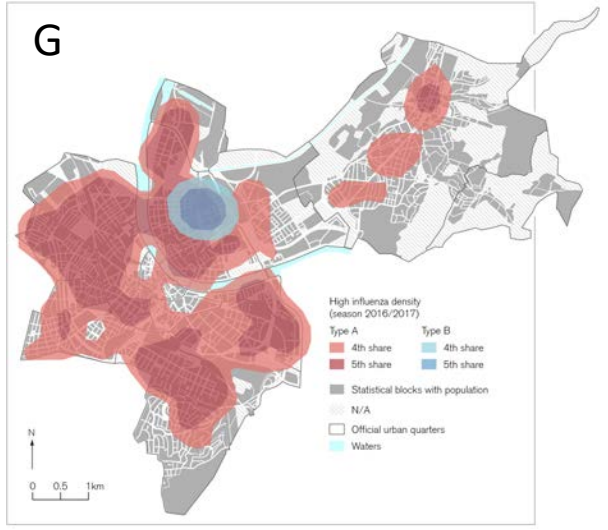
E


High influenza density
(season 2014/2015)
Type A
Type B
4th share
5th share
Statistical blocks with population
N/A
Official urban quarters
Waters
N
0 0.5 1km

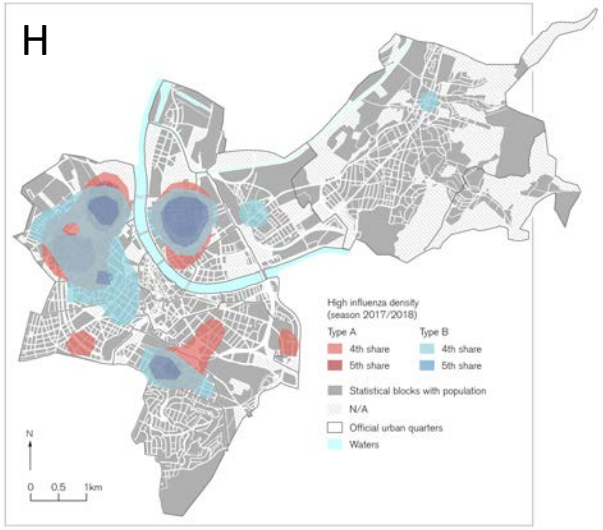
F

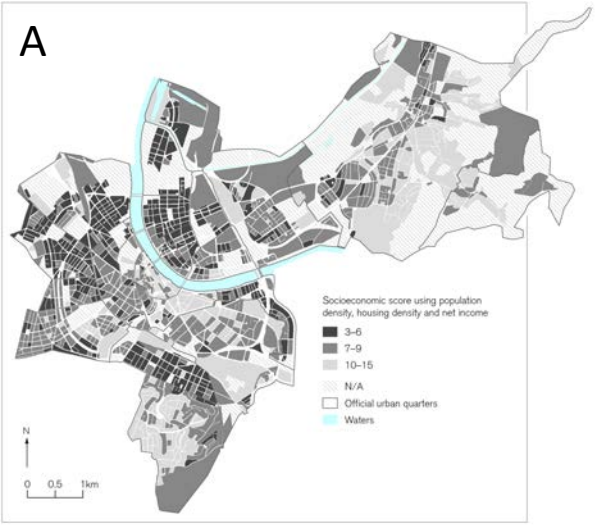

High influenza density
(season 2015/2016)
Type A
Type B
4th share
5th share
Statistical blocks with population
N/A
Official urban quarters
Waters
N
0 0.5 1km

G

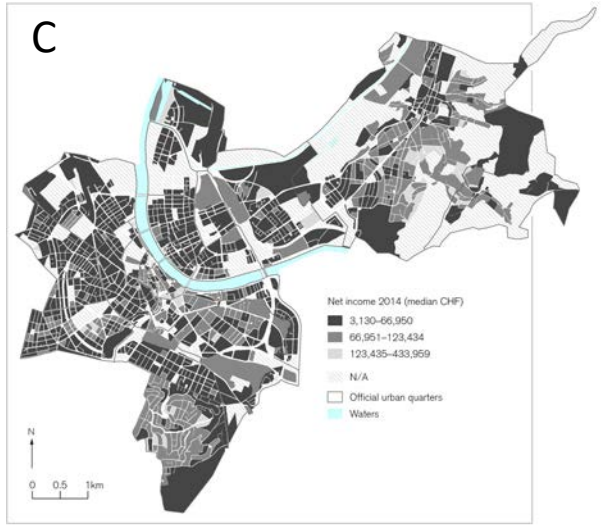

High influenza density
(season 2016/2017)
Type A
Type B
4th share
5th share
Statistical blocks with population
N/A
Official urban quarters
Waters
N
0 0.5 1km

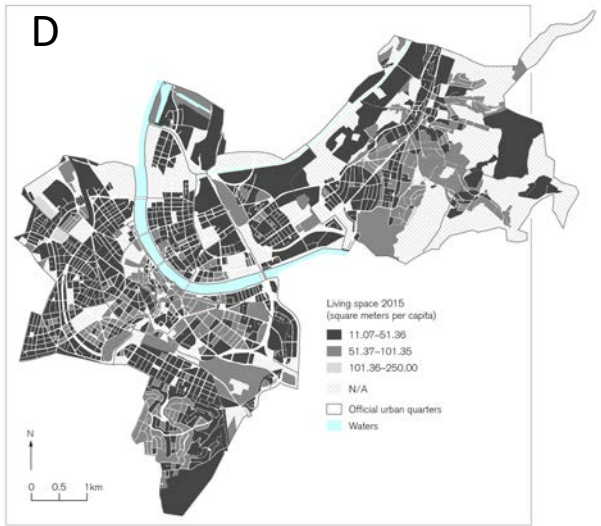
H


High influenza density
(season 2017/2018)
Type A
Type B
4th share
5th share
Statistical blocks with population
N/A
Official urban quarters
Waters
N
0 0.5 1km

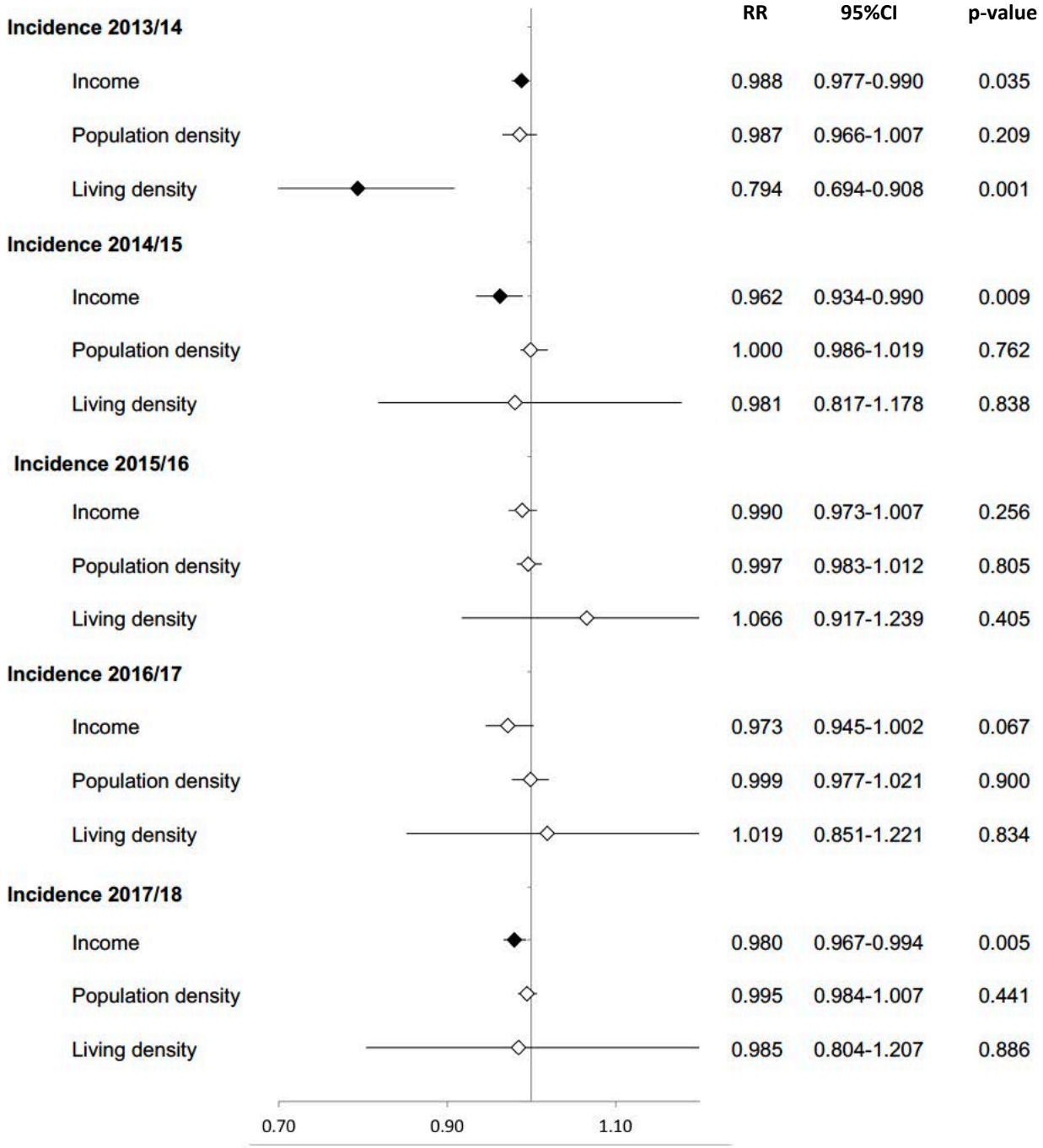

A
Socioeconomic score using population density, housing density and net income
3–6
7–9
10–15
N/A
Official urban quarters
Waters
N
0 0.5 1km

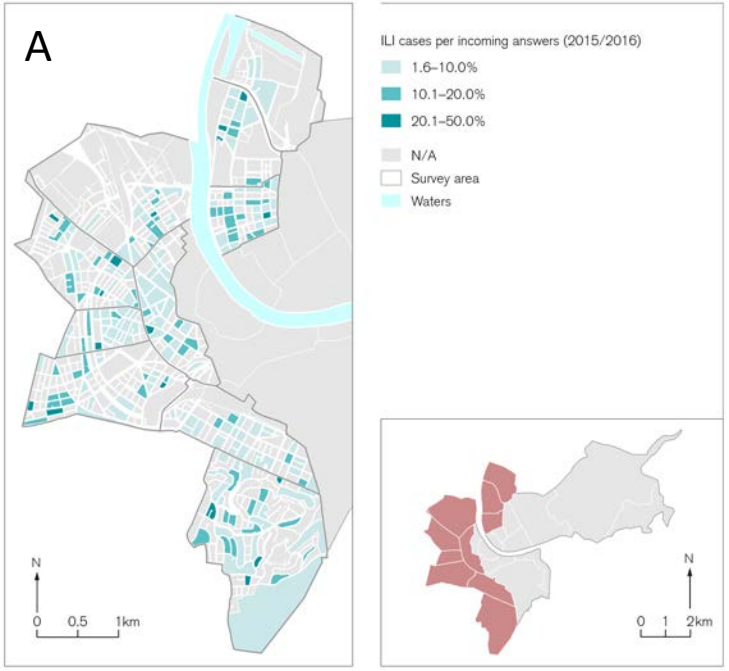

B
Population density 2015 (inhabitants per hectare)
0.14–134.25
134.26–315.75
315.76–832.28
N/A
Official urban quarters
Waters
N
0 0.5 1km

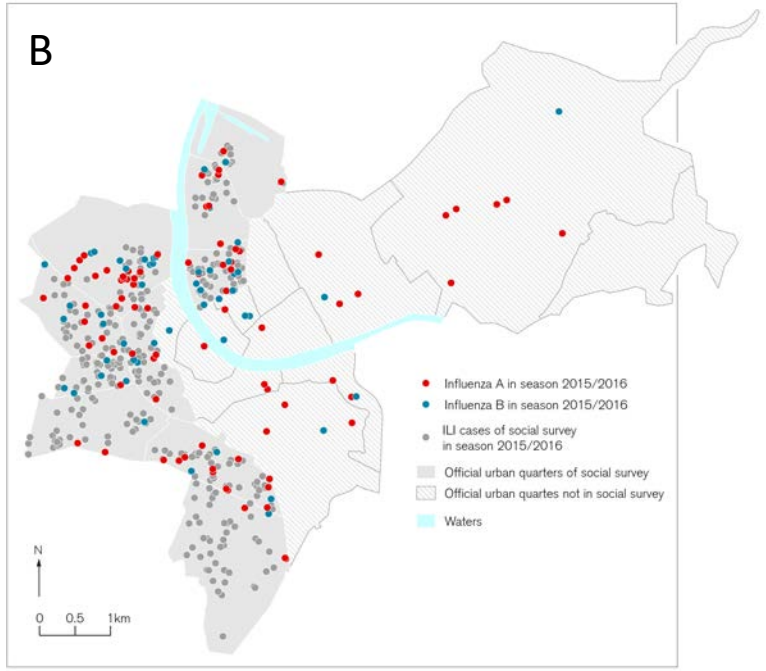

C
Net income 2014 (median CHF)
3,130–66,950
66,951–123,434
123,435–433,959
N/A
Official urban quarters
Waters
N
0 0.5 1km

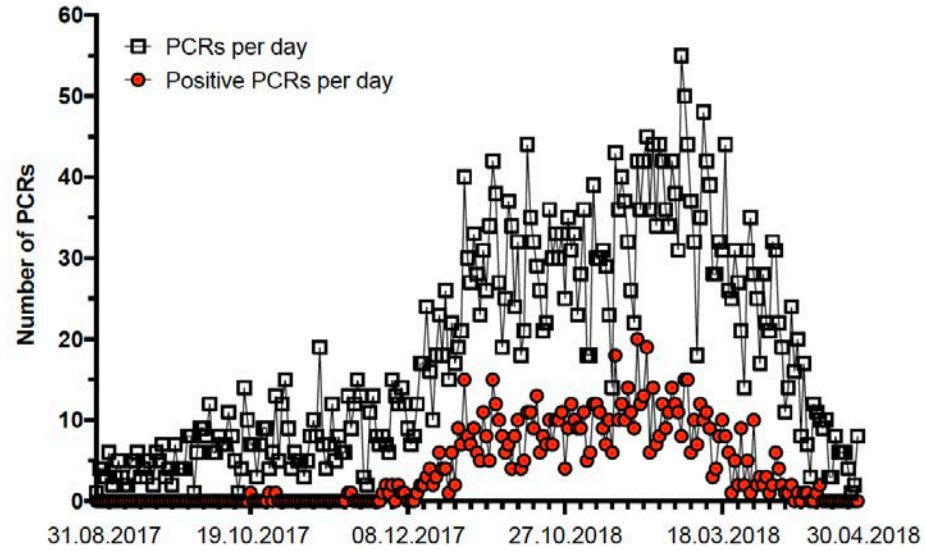

D
Living space 2015 (square meters per capita)
11.07–51.36
51.37–101.35
101.36–250.00
N/A
Official urban quarters
Waters
N
0 0.5 1km


A
Influenza A incidences
(season 2013/2014)
1.12–4.17
4.18–9.09
9.10–28.57
N/A
Official urban quarters
Waters
B
Influenza B incidences
(season 2013/2014)
1.48
1.49–4.93
4.94–9.80
N/A
Official urban quarters
Waters
C
Influenza A incidences
(season 2014/2015)
0.78–12.20
12.21–43.48
43.49–142.86
N/A
Official urban quarters
Waters
D
Influenza B incidences
(season 2014/2015)
1.12–7.30
7.31–26.32
26.33–83.33
N/A
Official urban quarters
Waters
E
Influenza A incidences
(season 2015/2016)
1.12–5.95
5.96–20.62
20.63–52.63
N/A
Official urban quarters
Waters
F
Influenza B incidences
(season 2015/2016)
0.99–3.28
3.29–8.33
8.34–19.80
N/A
Official urban quarters
Waters
G
Influenza A incidences
(season 2017/2018)
0.78–13.27
13.28–35.71
35.72–78.95
N/A
Official urban quarters
Waters
H
Influenza B incidences
(season 2017/2018)
1.12–10.07
10.08–27.78
27.79–74.07
N/A
Official urban quarters
Waters
N
0 0.5 1km

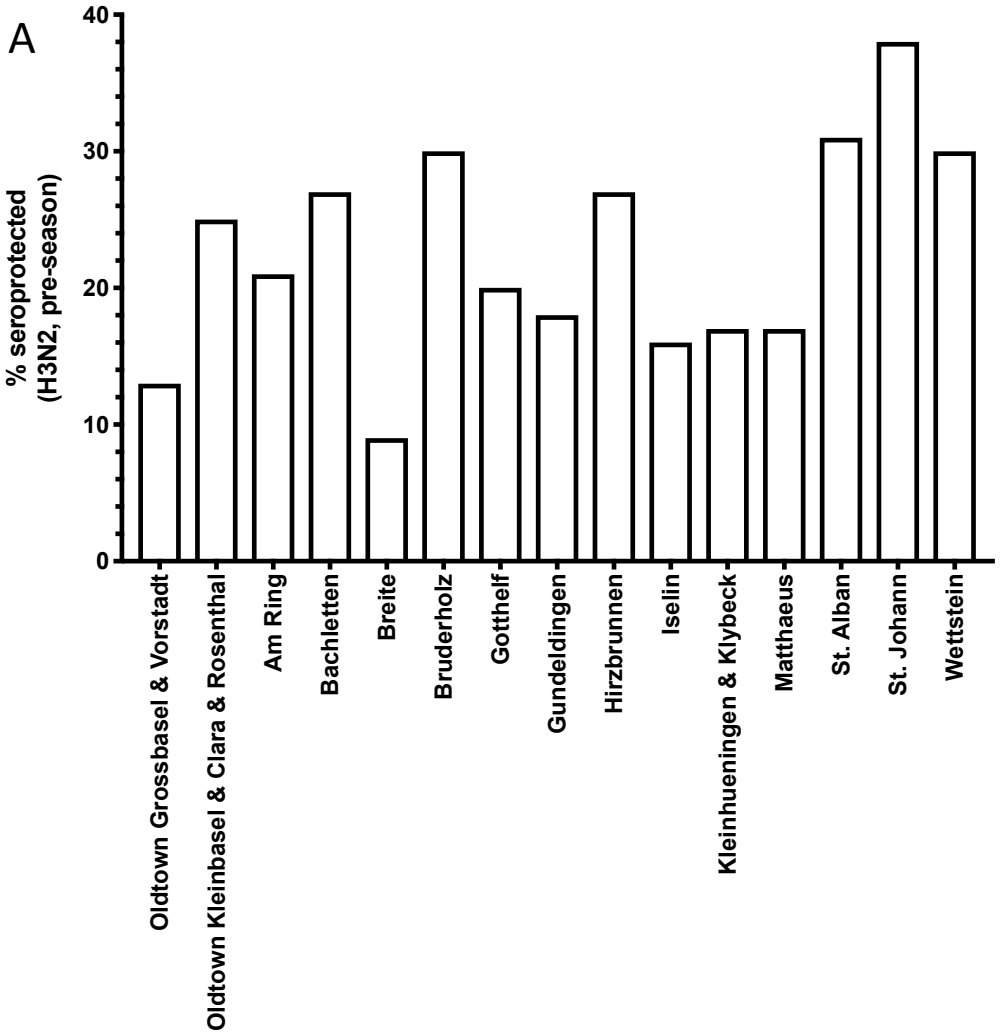

RR
95%CI
p-value
Incidence 2013/14
Income 0.988 0.977-0.990 0.035
Population density 0.987 0.966-1.007 0.209
Living density 0.794 0.694-0.908 0.001
Incidence 2014/15
Income 0.962 0.934-0.990 0.009
Population density 1.000 0.986-1.019 0.762
Living density 0.981 0.817-1.178 0.838
Incidence 2015/16
Income 0.990 0.973-1.007 0.256
Population density 0.997 0.983-1.012 0.805
Living density 1.066 0.917-1.239 0.405
Incidence 2016/17
Income 0.973 0.945-1.002 0.067
Population density 0.999 0.977-1.021 0.900
Living density 1.019 0.851-1.221 0.834
Incidence 2017/18
Income 0.980 0.967-0.994 0.005
Population density 0.995 0.984-1.007 0.441
Living density 0.985 0.804-1.207 0.886
0.70
0.90
1.10

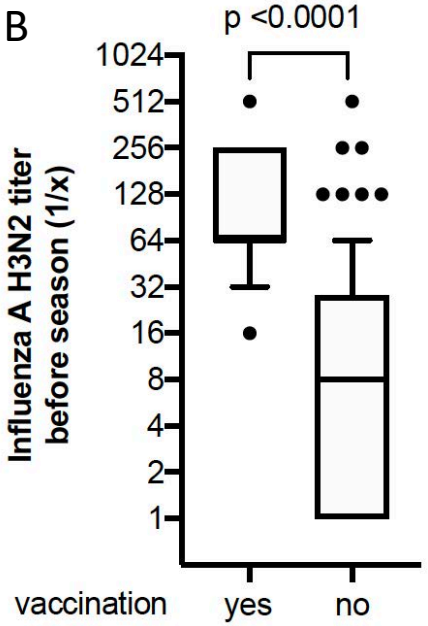

A
ILI cases per incoming answers (2015/2016)
1.6–10.0%
10.1–20.0%
20.1–50.0%
N/A
Survey area
Waters
N
0 0.5 1km
N
0 1 2km

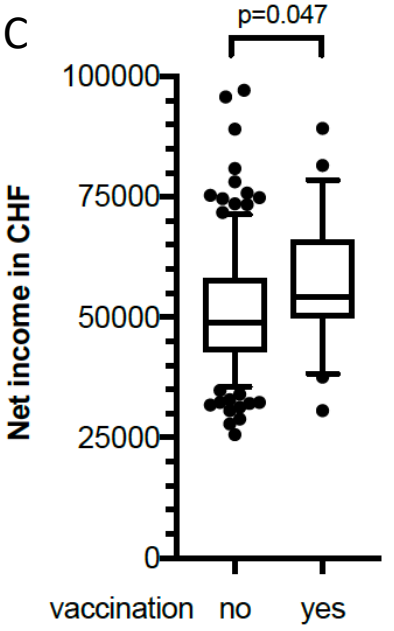

B
Influenza A in season 2015/2016
Influenza B in season 2015/2016
ILI cases of social survey in season 2015/2016
Official urban quarters of social survey
Official urban quartes not in social survey
Waters
N
0 0.5 1km

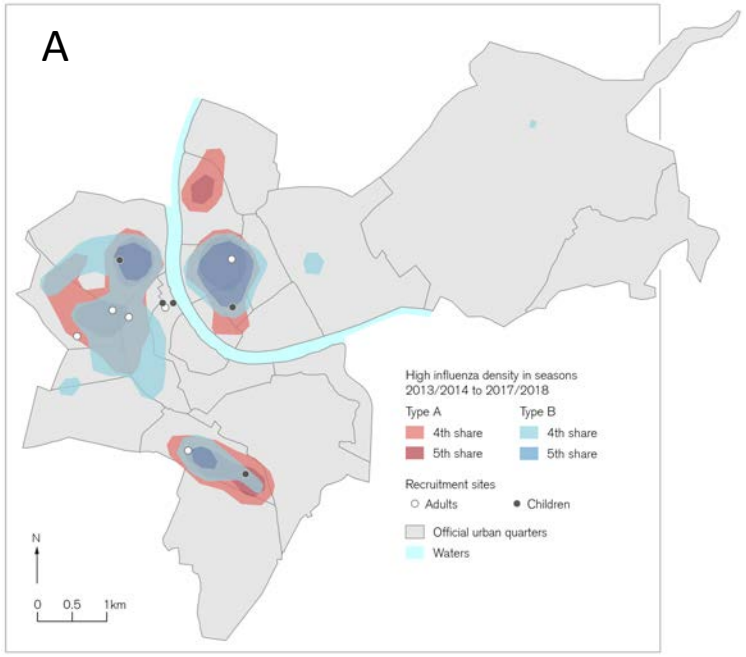

C
PCRs per day
Positive PCRs per day
Number of PCRs
60
50
40
30
20
10
0
31.08.2017
19.10.2017
08.12.2017
27.10.2018
18.03.2018
30.04.2018

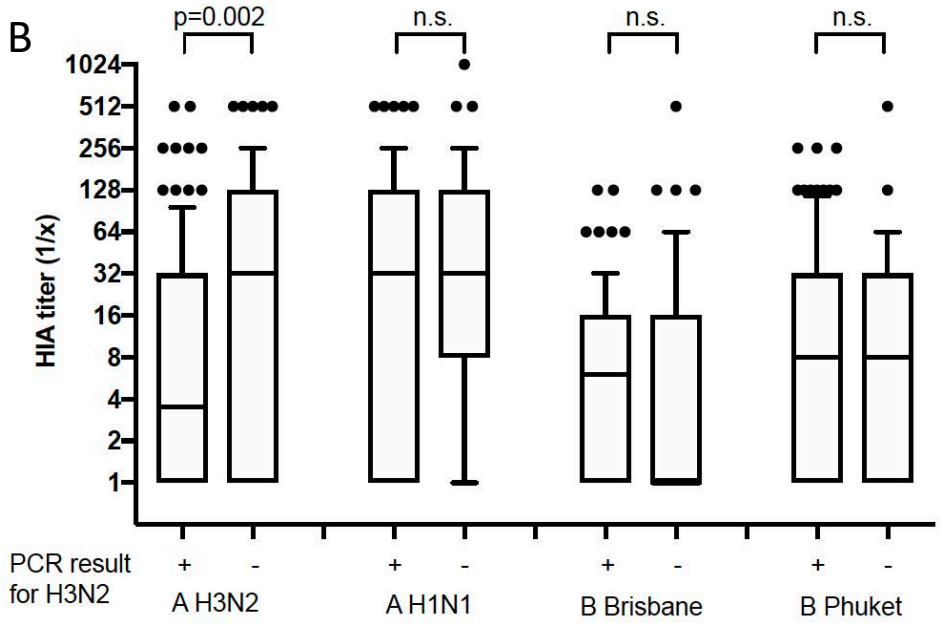

D
Self-reported vaccine rate (season 2015/2016)
5.26–30.77%
30.78–60%
60.01–100%
N/A
Official urban quarters of social survey
Waters
N
0 0.5 1km
N
0 1 2km

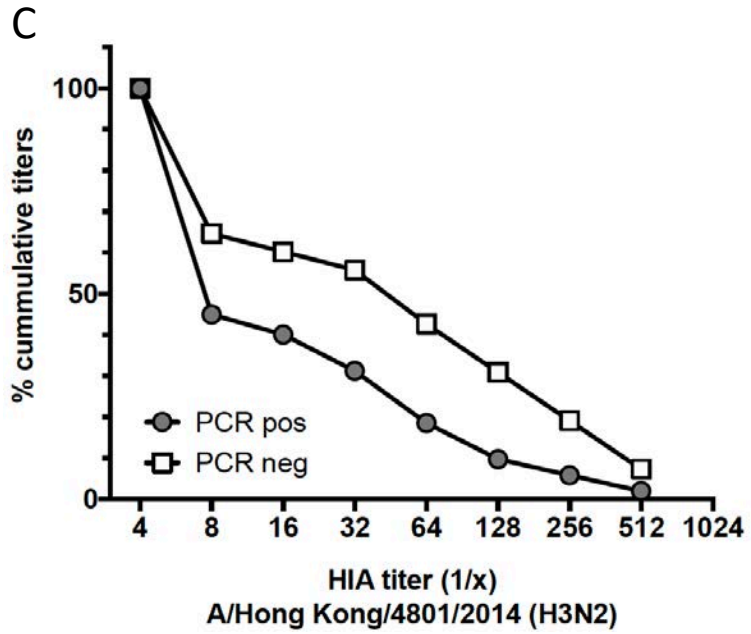

A
% seroprotected (H3N2, pre-season)
40
30
20
10
0
Oldtown Grossbasel & Vorstadt
Oldtown Kleinbasel & Clara & Rosenthal
Am Ring
Bachletten
Breite
Bruderholz
Gotthelf
Gundeldingen
Hirzbrunnen
Iselin
Kleinhueningen & Klybeck
Matthaeus
St. Alban
St. Johann
Wettstein

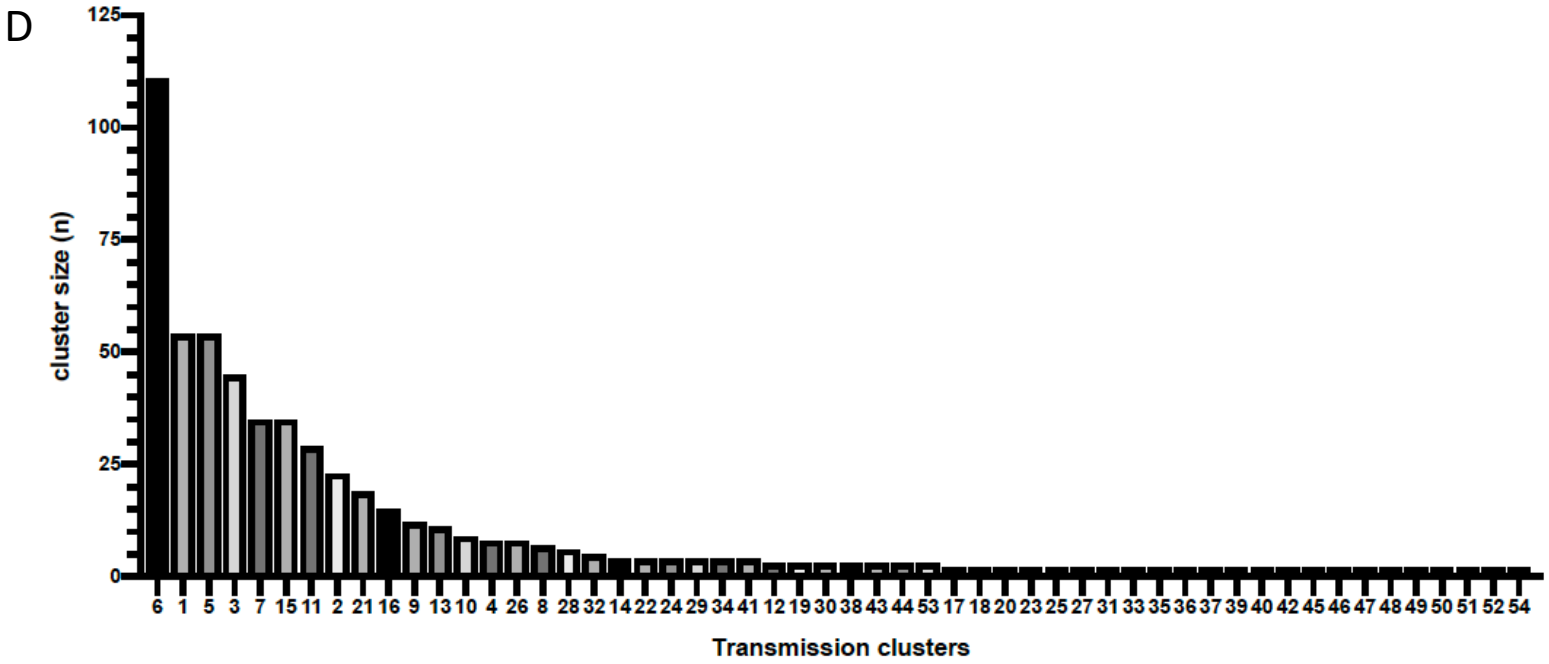

B
p <0.0001
Influenza A H3N2 titer before season (1/x)
1024
512
256
128
64
32
16
8
4
2
1
vaccination
yes
no

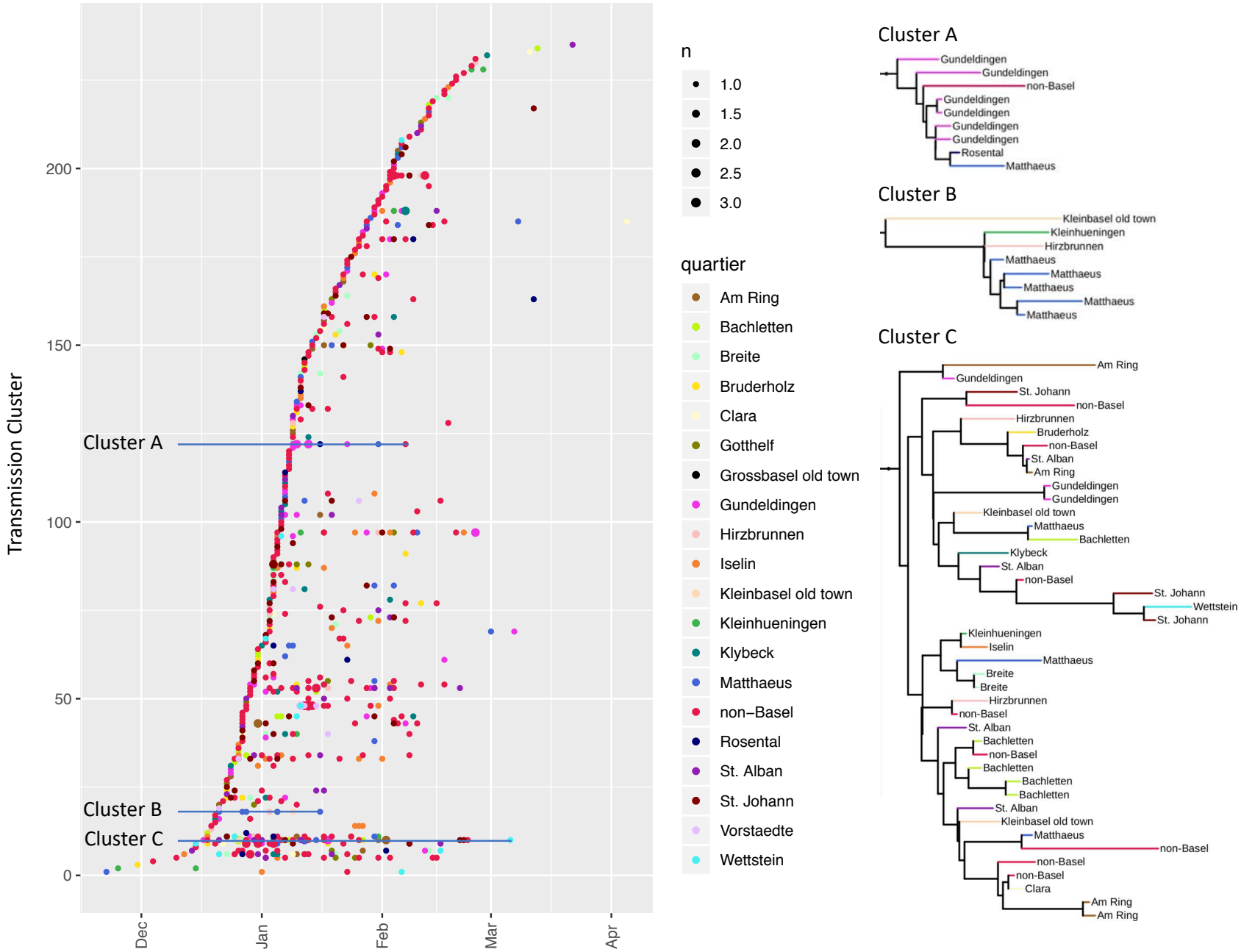

C
p=0.047
Net income in CHF
100000
75000
50000
25000
0
vaccination
no
yes


A
High influenza density in seasons 2013/2014 to 2017/2018
Type A
4th share
5th share
Type B
4th share
5th share
Recruitment sites
Adults
Children
Official urban quarters
Waters
N
0 0.5 1km


B
p=0.002
n.s.
n.s.
n.s.
HIA titer (1/x)
1024
512
256
128
64
32
16
8
4
2
1
PCR result for H3N2
+ -
A H3N2
A H1N1
B Brisbane
B Phuket


C
% cummulative titers
100
50
0
PCR pos
PCR neg
4 8 16 32 64 128 256 512 1024
HIA titer (1/x)
A/Hong Kong/4801/2014 (H3N2)


D
cluster size (n)
125
100
75
50
25
0
6 1 5 3 7 15 11 2 21 16 9 13 10 4 26 8 28 32 14 22 24 29 34 41 12 19 30 38 43 44 53 17 18 20 23 25 27 31 33 35 36 37 39 40 42 45 46 47 48 49 50 51 52 54
Transmission clusters


Transmission Cluster
200
150
100
50
0
Dec
Jan
Feb
Mar
Apr
Cluster A
Cluster B
Cluster C
n
1.0
1.5
2.0
2.5
3.0
quartier
Am Ring
Bachletten
Breite
Bruderholz
Clara
Gotthelf
Grossbasel old town
Gundeldingen
Hirzbrunnen
Iselin
Kleinbasel old town
Kleinhueningen
Klybeck
Matthaeus
non−Basel
Rosental
St. Alban
St. Johann
Vorstaedte
Wettstein
Cluster A
Gundeldingen
Gundeldingen
non-Basel
Gundeldingen
Gundeldingen
Gundeldingen
Gundeldingen
Rosental
Matthaeus
Cluster B
Kleinbasel old town
Kleinhueningen
Hirzbrunnen
Matthaeus
Matthaeus
Matthaeus
Matthaeus
Matthaeus
Cluster C
Am Ring
Gundeldingen
St. Johann
non-Basel
Hirzbrunnen
Bruderholz
non-Basel
St. Alban
Am Ring
Gundeldingen
Gundeldingen
Kleinbasel old town
Matthaeus
Bachletten
Klybeck
St. Alban
non-Basel
St. Johann
Wettstein
St. Johann
Kleinhueningen
Iselin
Matthaeus
Breite
Breite
Hirzbrunnen
non-Basel
St. Alban
Bachletten
non-Basel
Bachletten
Bachletten
Bachletten
St. Alban
Kleinbasel old town
Matthaeus
non-Basel
non-Basel
non-Basel
Clara
Am Ring
Am Ring