Kelley *et al. BMC Public Health (2025) 25:1345*
https://doi.org/10.1186/s12889-025-22496-8

BMC Public Health

RESEARCH

Exploring influenza vaccination determinants through digital participatory surveillance

Kathleen Kelley[1*], Nicolò Gozzi[1], Mattia Mazzoli[1] and Daniela Paolotti[1]

Abstract

Background Vaccination is key for mitigating the impact of recurring seasonal influenza epidemics. Despite the efficacy and safety of influenza vaccines, achieving optimal vaccination uptake remains a challenge. This study aimed to explore the determinants of influenza vaccination uptake using data from Influweb, the Italian node of the Influenzanet participatory surveillance network.

Methods This study utilizes a longitudinal dataset of self-reported vaccination statuses from Italian participants across the 2011–2021 flu seasons. Logistic regression models were used to identify factors associated with vaccination uptake. Post-stratification weights were applied to account for demographic differences between the Influweb sample and the general population.

Results The analysis reveals that individuals using public transport and those living with minors are less likely to receive the influenza vaccination. On the other hand, university-educated individuals, and those on medication for chronic diseases are more likely to be vaccinated. Age also plays a role: individuals aged 44 and under are less likely to vaccinate compared to those aged 45–65, while those over 65 are more likely to do so. Furthermore, higher cumulative influenza-like illness incidence rates within a macro-region are associated with increased vaccination uptake. Finally, the impact of COVID-19 pandemic was associated with an increase in influenza vaccination uptake. Comparison of the Influweb data to vaccination rates reported by the Italian Health Ministry revealed higher coverage for self-reported vaccination. This could be linked to the voluntary nature of the survey, possibly attracting a more health-conscious cohort.

Conclusions Our study found that individuals living with minors and those relying on public transportation have lower odds of being vaccinated, despite having a higher documented risk of respiratory virus exposure. These findings highlight the importance of continued public health efforts targeting vulnerable groups and raising awareness about the risks of forgoing vaccination. The complex interplay of socioeconomic, demographic, and public health context significantly shapes vaccination decisions, emphasizing the need for tailored public health campaigns.

Keywords Influenza Vaccination, Participatory Surveillance, Vaccine Uptake

*Correspondence:
Kathleen Kelley
kathleen.kelley@isi.it
[1] ISI Foundation, Via Chisola 5, Turin, Italy

Introduction

Vaccination stands as a fundamental pillar in public health, playing a key role in preventing infectious diseases and mitigating the impacts of epidemics and pandemics [1]. Despite their proven efficacy and safety, vaccination programs often encounter challenges in achieving optimal uptake [2]. This issue is particularly pressing in the case of influenza, where vaccine coverage remains below targeted rates despite the availability of effective vaccines [3]. The World Health Organization (WHO) has recognized vaccine hesitancy—defined as delay in acceptance or outright refusal of vaccines despite available services—as a significant global health threat, underlining the need for targeted public health interventions to enhance vaccine acceptance [4].

It is well-established that a complex interplay of socioeconomic, demographic, and psychological factors contributes to vaccination uptake [5].

On a community level, economic factors and access to healthcare are significant determinants of vaccination rates. When vaccinations are easily available at pharmacies, workplaces, or community centers, the rate of uptake increases [6, 7]. Additionally, socioeconomic status (SES) factors, such as employment and financial stability, have been found to significantly influence influenza vaccination rates, as demonstrated in a study of high-risk groups in Italy [8].

The interaction of demographic and psychological factors also play a crucial role. Perception of risk, as a result of demographic characteristics, influences decisions. This can be related to the perceived risk of contracting the flu or potential side effects from the vaccine itself [6]. Older adults and individuals with chronic diseases are more likely to get vaccinated due to a higher perceived risk of severe influenza outcomes [7, 9, 10]. In particular, patients with chronic kidney or liver diseases are more likely to vaccinate due to a higher vulnerability to severe influenza infections [7, 9]. Better health literacy among older and more educated populations further contributes to higher vaccination rates [11]. Conversely, lower perceived risk of infection typically leads to lower vaccination rates [6]. Additionally, social encouragement from family, friends, coworkers, and especially healthcare providers, plays a crucial role in promoting vaccination [7, 9, 12, 13].

Recommendations from public health authorities also play a role. In most industrialized European countries, the population groups that tend to get vaccinated the most are those for which the vaccine is recommended, namely elderly and fragile individuals, to the point that only these are the categories for which most official data are available.

While understanding well-documented vaccine determinants is essential, capturing evolving vaccination behaviors requires timely data collection methods. In contrast to traditional surveillance methods, digital participatory surveillance systems have gained prominence for their ability to collect real-time data on public health behaviors, symptoms, and vaccination uptake directly from volunteer participants. These systems not only provide valuable insights into individual health behaviors but also offer a faster approach to monitoring disease spread compared to traditional methods.

Participatory surveillance systems have been increasingly utilized to monitor influenza-like illness (ILI) and associated health behaviors. For instance, in North America, the participatory surveillance system *Flu Near You* has been employed to assess health-seeking behaviors in individuals with likely ILI cases [14]. Similarly, Australia's Flutracking system has demonstrated adaptability by monitoring both influenza and COVID-19 incidence, highlighting the versatility of these systems in estimating illness trends [15].

In Europe, Influenzanet serves as a network of web-based platforms designed to monitor influenza and other respiratory diseases through participatory surveillance [16]. It has been used to identify key determinants associated with higher ILI risk, leveraging self-reported data to improve disease tracking across countries [17]. Studies considering data spanning over a decade, from multiple countries, have shown that the epidemiological signals generated by Influenzanet platforms are strongly aligned with those from sentinel-based surveillance systems [18–20]. Data collected through participatory surveillance has also been validated in other health-related contexts. For example, self-reported data on antibiotics and cough medication from the UK-based participatory platform exhibited good agreement with prescription data from the UK National Health System [21]. Vaccination data provided by participants in the Netherlands via the web platform has been used to study vaccine effectiveness [18].

In this study, we aim to investigate the determinants of influenza vaccination in Italy using longitudinal data over multiple influenza seasons from Influweb, the Italian node of Influenzanet. Active since 2008, Influweb was designed to complement traditional sentinel-based surveillance of influenza-like illness (ILI) by gathering information directly from the general population. The platform's primary goal has been to enhance the monitoring of ILI incidence, and its surveillance estimates have been validated over the years [16]. Our objective is to evaluate the potential of the Influenzanet platforms in identifying key factors influencing vaccination uptake and providing timely insights to guide more targeted vaccine campaigns. This approach could be particularly beneficial for improving uptake among groups that traditionally exhibit lower vaccination rates.

Through Influweb we have access to individual data on influenza vaccination decisions over multiple influenza seasons from 2011 to 2021, as well as a range of individual socio-demographic and health-related information, including age, employment status, household composition, education level, and medication use for chronic conditions. Additionally, we complement this data with reported influenza incidence rates. Logistic regression models were employed to identify determinants of vaccination uptake.

This study contributes to the existing literature by leveraging longitudinal data from a participatory surveillance platform to examine vaccination behaviors across multiple flu seasons. It shows how self-reported data from platforms like Influweb can be useful tools to collect important individual health-related behavior data. Our work offers valuable insights in support of more effective public health strategies for vaccine-preventable diseases such as influenza.

Methods

Dataset description

This study utilizes a unique dataset provided by Influweb, which operates as a participatory symptomatic surveillance survey within Italy and forms a part of Influenzanet—a survey network dedicated to monitoring influenza-like illnesses across Europe. The platform functions as a longitudinal study, leveraging voluntary participation. Individuals partake by first providing demographic and health background via an intake survey, which gathers information such as age, presence of chronic diseases, education level, vaccination status, and region of residence. The system allows participants to update their intake surveys as needed to update their information. While some participants complete intake surveys in multiple flu seasons, others may submit only one or update their intake survey in non-consecutive years.

Participants then receive weekly email reminders throughout the year to fill out a symptom survey, detailing any symptoms experienced in the past week and the health behaviors undertaken in response. The symptom surveys are routinely used to identify possible influenza-like-illnesses (ILI) cases, by referring to the ILI case definition from the European Center for Disease Prevention and Control (ECDC) [22]. Participants are asked to fill out the weekly symptom surveys year-round for continuous monitoring of ILI cases.

Recruitment for Influweb occurs annually through institutional press releases, leading to coverage in mainstream and social media. In its early years, televised appearances significantly boosted participation. Additional outreach includes science fairs, school events, and word of mouth, particularly via social media and email invitations through the Influweb system. While yearly press releases remain consistent, they often highlight platform updates to sustain public engagement.

This study utilizes data spanning from the 2011–2012 to the 2020–2021 flu seasons, focusing on the determinants of vaccination status as reported in the intake surveys. A flu season spans from November 1 to May 1 of the following year, e.g., November 1, 2011 to May 1, 2012. Each participant's data for a given flu season was represented by a single, unique intake form. In instances where multiple submissions were recorded for a participant within one season, the first one reporting vaccination was used. If no affirmative response was provided, the earliest intake form of the season was selected.

As a participatory system, Influweb relies on a self-selected sample of volunteers, which can introduce selection bias. For instance, participants in the Influweb study may have a greater interest in health-related topics, potentially leading to other characteristics or behaviors that differ from the general population. However, methodologies such as post-stratification can be employed to adjust the sample to more accurately represent the target population [23].

Despite the potential for sample biases, the utility of Influweb has been well-documented. Digital participatory surveillance systems, like Influweb, offer unique insights into disease trends that are not accessible through traditional surveillance methods [24]. Participatory surveillance systems are sensitive in detecting trends and early outbreaks because data can be collected and analyzed at a faster rate than traditional systems, which often rely on delayed reporting from primary care facilities [25]. For example, ILI forecasting is improved when using Influweb data in conjunction with the sentinel data originating from primary care facilities [26]. This is also being done with similar digital participatory systems across Europe [20, 27].

Data preparation

A complete case approach was adopted for regression modeling of seasons 2011–2012 to 2020–2021. For variables of interest with significant missing values, imputation was employed using the latest available data from participants, which significantly reduced but did not eliminate missing data points. The variable for education had 149 missing entries from 107 participants, and the macro-region variable had 78 missing entries from 56 participants.

The primary outcome, participant vaccination status, was initially categorized into three responses: yes, no, or unsure. Due to sparse data in the "unsure" category, with 58 entries from 33 participants, these responses were excluded from the final analysis.

Before applying complete case criteria, the dataset included 4636 unique participants. The exclusion of 186 participants, evenly distributed across all seasons, resulted in the final dataset of 9646 responses from 4450 unique participants, each represented by a single intake survey per flu season. Of the participants initially included in the study, 4% were excluded due to missing data, which is below commonly cited thresholds of 5% or 10% where missing data may lead to bias [28, 29]. To assess the potential impact of the complete case approach, a sensitivity analysis was conducted by comparing vaccination coverage across seasons and the demographic breakdown of the sample before and after applying the complete case criteria. This indicated that excluding cases with missing data did not significantly alter the vaccination coverage proportions across flu seasons or the demographics of the sample. These comparisons can be seen in Supplementary Figure S1 and Supplementary Table S2.

Post-stratification

For our study of the Italian population, we stratified the sample by sex, 5-year age groups and macro-region for each year included in the study. For example, a single stratum could be defined as males aged 20–24 residing in the North-East macro-region in 2015. We calculated the proportion of participants in each stratum within the final dataset ($sp_k = \frac{N_k^{sample}}{N^{sample}}$) and compared it to the proportion of the Italian population in the same strata ($rp_k = \frac{N_k^{realpop.}}{N^{realpop.}}$). Population data was sourced from the Istituto Nazionale di Statistica, using the Intercensal Register for the years 2010–2019 [30] and the Municipal Resident Population for the years 2020–2021 [31]. Weights were then calculated by taking the ratio of the proportion in stratum k of the Italian population to the proportion of Influweb sample in stratum for each year, using the formula: $W_k = \frac{rp_k}{sp_k}$. These weights were applied to the regression model and all other analyses described in later sections, unless otherwise noted.

Variable description

Variables were selected based on their anticipated influence on vaccination status, guided by a literature review. These variables, listed in Supplementary Table S1, include demographic, household, and health factors.

Public transport indicates whether public transportation is the participant's main mode of transport, in lieu of other options like bicycle or car. Household composition was assessed by determining whether participants lived with individuals in specific age groups. Participants who live with at least one person under 18 years old were classified as "living with minors," and those who live with at least one person aged 65 or older were classified as "living with elders." Both are binary variables.

Another variable was daily contacts with groups of people, informed by the study by Ibuka et al. 2016 [32]. Participants were asked if they regularly came into contact with groups of people (excluding those on public transport). If they reported daily interactions with groups such as more than 10 elderly individuals, patients, more than 10 children or teenagers (excluding their own children), or other groups of more than 10 people, the variable was coded as "True."

Educational attainment was classified into three levels. Participants who reported having a middle school or high school diploma were categorized as "High school or less," while those with a bachelor's degree or higher were grouped under "University." Those still pursuing education were classified as "Students." Employment status was also categorized into three groups: "Employed" for individuals with full-time, part-time, or self-employment; "Unemployed" for students, homemakers, and those on long-term leave from work; and "Retired" for participants who reported that they were no longer working.

Participants who reported smoking either occasionally or daily were categorized as smokers, while those who did not smoke or were uncertain of their tobacco use were classified as non-smokers. Participants were coded as "True" for the medication variable if they reported taking medication for chronic conditions such as asthma, diabetes, chronic lung disease, cardiovascular disease, renal disease, or immunosuppression due to various therapies or conditions.

To assess for the influence of the COVID-19 pandemic, a binary pandemic variable was coded as "True" for data collected during the 2020–2021 flu season and "False" for prior seasons.

Additionally, the cumulative incidence of influenza-like illness (ILI) was calculated per macro-region for each flu season. National influenza incidence data was sourced from the Respivirnet report published by the Italian National Institute of Health [33]. The regional cumulative incidence was derived by summing the weekly ILI incidence rates across the entire flu season, for each macro-region. For example, the cumulative incidence for the 2011–2012 season was calculated using the sum of weekly ILI incidence rates for all weeks in that flu season and was applied to surveys collected during the 2011–2012 period. This variable is intended to reflect trends at the macro-regional level and should be interpreted as an indicator of broader epidemic conditions. For instance, higher epidemic intensity in a macro-region may coincide with more media coverage or increased communication campaigns from local health authorities.

Preliminary analysis

Statistical analyses began with individual chi-squared tests to examine associations between vaccination status and the levels of categorical covariates. The test compares the counts observed in each category of a data set to what we would expect to see if there was no relationship at all. If the differences between these observed and expected counts are large enough, the test suggests that the variables are related [34]. To account for the repeated measures present in some participants, the Rao-Scott correction was applied to the chi-square tests. The Rao-Scott correction is a widely used method in studies involving survey data with complex designs, such as datasets with repeated observations or clustering [35–37]. This adjustment modifies the chi-square statistic by dividing it by a correction factor that accounts for both the average number of repeated observations within a cluster and the intra-cluster correlation [38]. These tests were used to identify potential factors that might influence vaccination uptake and for guiding the subsequent model selection process [39]. Statistical significance was determined using a significance level of 5%.

Regression analysis

We consider a logistic regression model designed to estimate the odds of vaccination as a function of the identified significant covariates.

Logistic regression models are used to model binary outcomes. In this case, the outcome is whether a participant was vaccinated (1) or not (0). To facilitate the model selection process, the macro-region categorical variable was transformed into dummy variables. The initial full model included all variables identified from the literature review and the chi-squared tests. The logistic regression equation can be expressed as:

$$log(\frac{p}{1-p}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_k$$

where p is the probability of vaccination, β_0 is the intercept, and $\beta_1, \beta_2, ..., \beta_k$ are the coefficients for each corresponding predictor variable$X_1, X_2, ..., X_k$. For example, β_1 might represent the effect of taking medication for a chronic disease (True/False), and β_2 could represent the age category 65+compared to the reference group 45–64.

Robust sandwich variance estimator

Our data included some participants who were observed in multiple influenza seasons. Since these repeated measures violate the assumption of independence among observations, we applied a robust sandwich estimator to adjust variance estimates and account for clustering. The estimator aggregates the score residuals for each participant. The empirical covariance matrix (B) of these residuals is then "sandwiched" between the inverse of the model-based information matrix (A^{-1}) to compute the variance of the model coefficients as: $Var(\widehat{\beta}) = A^{-1}BA^{-1}$, providing robust standard errors [40–42].

Model selection procedures

1. "Drop-one" model selection:

   Initially, a "drop one" approach using likelihood ratio tests compared the full model against models each lacking one variable. This method systematically removes each predictor from a full model and compares the reduced model to the full model using likelihood ratio tests (LRT). The LRT assesses whether the reduced model (with one less covariate) fits the data significantly worse than the full model. The test statistic is calculated as:

$$LRT = -2ln\left(\frac{L_{reduced}}{L_{full}}\right),$$

Where $L_{reduced}$ and L_{full} are the likelihoods of the reduced model and the full model. This statistic follows a chi-square distribution with degrees of freedom equal to the difference in the number of parameters between the full and reduced models (1 in this case).

A variable is considered significant if its removal resulted in a significant LRT (p-value < 0.05), indicating that the model which excludes this variable fits the data significantly worse than the full model [43]. Conversely, if the LRT yielded a p-value greater than 0.05, the variable was excluded because its removal did not significantly affect the model fit. By iteratively applying this procedure to all variables, we identified and retained only those predictors that significantly contributed to the model. This process resulted in a reduced model, referred to as the drop-one model.

2. Stepwise model selection:

   In parallel, a stepwise selection process was employed, considering both forward and backward selection. This procedure performed model selection by minimizing the Akaike Information Criterion (AIC) value. The AIC metric assesses a model's likelihood while penalizing models with many covariates, defined as:

$$AIC = -2ln\left(\widehat{L}\right) + 2k,$$

Where $\widehat{L}$ is the maximum likelihood of the model, and k is the number of parameters. The aim is to balance overall model fit with model complexity [44]. Forward selection starts with no predictors, adding them one by one based on AIC improvement, while backward elimination starts with all candidate predictors from the full model, removing the least significant ones. The combined stepwise approach iterates between adding and removing covariates to achieve the lowest AIC, resulting in the stepwise selection model.

Model comparison and final model selection
After obtaining the drop-one model and the stepwise selection model, we compared them with each other and with the full model using likelihood ratio tests to determine the preferred model [45, 46]. The comparisons were as follows:

- Drop-One Model vs. Full Model
- Stepwise Selection Model vs. Full Model
- Drop-One Model vs. Stepwise Selection Model

The LRT was used to assess whether the simpler, nested model (with fewer covariates) provided an adequate fit compared to the more complex model. A non-significant p-value (p-value > 0.05) indicates that the simpler model is preferred due to its parsimony without a significant loss in model fit.

Based on these comparisons, the drop-one model was selected as the final model because it provided the best balance between model fit and simplicity and had the lowest AIC value.

Interpretation of the final model
The coefficients from the final, drop-one logistic regression model were exponentiated as $exp(\beta)$ to obtain odds ratios (ORs), which quantify the association between each covariate and vaccination status [47]. An OR greater than 1 suggests a positive association between the outcome and the covariate (higher odds of vaccination), while an OR less than 1 suggests a negative association (lower odds of vaccination). Statistical significance was determined using a significance level of 5%.

Reasons for vaccination
As part of the survey, vaccinated Influweb participants were asked to provide their reasons for receiving the flu vaccine. Participants could select multiple reasons from a predefined list. The responses were assessed to calculate the percentage of participants indicating each reason. Percentages were computed as the number of participants selecting a specific reason divided by the total number of vaccinated participants, aggregated across all seasons.

Results
Participant characteristics
There are a total of 9,646 intake surveys from 4,450 participants that were included in the final dataset. Each participant's data for a given flu season was represented by a single, unique intake form. In instances where multiple submissions were recorded for a participant within one season, the first one reporting vaccination was used. If no affirmative response was provided, the earliest intake form of the season was selected. Table 1 presents the characteristics of the participants used in the study, for all variables considered during the features selection process. To calculate the distribution of these characteristics, the most frequently occurring value for each participant was used in cases where individuals changed categories (e.g., moved from one age group to another). Such changes were rare, occurring in only 28 participants.

The values shown represent raw counts and percentages, not accounting for post-stratification weighting. Please note that percentages may not sum to 100 due to rounding. For more details on post-stratification adjustments, their impact on the distribution of these variables, and comparisons to the national population, refer to the supplementary materials section "Population comparisons and post-stratification weighting".

Vaccination determinants
Before assessing the determinants of vaccination uptake, we compared vaccination coverage in the Influweb sample with officially reported health data to contextualize our sample within the general population and better understand its characteristics. This analysis shows that vaccination coverage in the Influweb sample is consistently higher than national estimates, likely reflecting the health-conscious nature of the Influweb population. As a result, the determinants of vaccination should be considered within this context. The full methodology and results of this comparison are detailed in the Supplementary Materials section "Vaccination coverage: comparing the Influweb sample to officially reported estimates".

Preliminary analysis
Table 2 shows the results of the chi-squared tests for categorical variables. The chi-squared tests, conducted at a significance level of 5%, indicated significant associations for all considered covariates except for Macro-Region. Due to these results, all covariates were retained for further exploration in the model selection process and the covariate "Macro-region" was broken down into dummy variables.

Table 1 Participant characteristics and their distributions in percentages

Characteristic	Categories	Influweb Sample Count (%)
Age group	0–17	504 (11)
	18–44	2019 (45)
	45–64	1486 (34)
	65 +	441 (10)
Sex	Male	2553 (57)
	Female	1897(43)
Macro-Region	Centre	560 (13)
	Islands	235 (5)
	North-East	1046 (24)
	North-West	2119 (48)
	South	490 (11)
Education Level	High school or less	1909 (43)
	Currently a student	528 (12)
	University degree	2013 (45)
Employment	Employed	2729 (61)
	Retired	477 (11)
	Unemployed	1244 (28)
Public transport primary mode of transportation	True	721 (16)
Lives with minors	True	2052 (46)
Lives with elders	True	951 (21)
Takes medication for a chronic condition	True	732 (16)
Contacts with 10 + people daily	True	3021 (68)
Smoker (occasionally, daily)	True	818 (18)

Table 2 Chi-square test comparing vaccinated and unvaccinated participants

Covariate	χ^2 (df)	*P*-value
Sex	22.729 (1)	0.006
Public transport	43.485 (1)	< 0.001
Lives with elders	288.514 (1)	< 0.001
Lives with minors	224.409 (1)	< 0.001
Contacts	26.531 (1)	0.002
Education	114.699 (2)	< 0.001
Smoker status	13.365 (1)	0.017
Employment	738.609 (2)	< 0.001
Medication	545.114 (1)	< 0.001
Age group	1071.438 (3)	< 0.001
Pandemic	287.730 (1)	< 0.001
Macro-region	4.146 (4)	0.832

Logistic regression model

The logistic regression analysis aimed to identify significant predictors of influenza vaccination status, using a significance level of 5%. The results can be seen in Table 3.

Individuals who use public transport were found to have a reduced likelihood of being vaccinated against influenza, with the odds being 28% lower compared to those who use other methods of transportation (OR = 0.718). Similarly, living with minors was associated with a decreased likelihood of vaccination. Participants living with minors had 30% lower odds of being vaccinated compared to those who do not live with minors (OR = 0.697).

Education level also played a role in explaining vaccination status. Those with a university education had 28% higher odds of being vaccinated compared to those with a high school education or less (OR = 1.276). This indicates that higher educational attainment is positively associated with vaccination uptake. Regarding employment status, there was no significant difference in vaccination likelihood between unemployed and employed individuals. Medication use emerged as a significant predictor, with individuals taking medication for chronic conditions having 142% higher odds of being vaccinated (OR = 2.422).

Age was another significant factor in vaccination decisions. Participants aged 0–17 had 70% lower odds

Table 3 Logistic regression coefficients, odds ratios, and standard errors

Variable	Odds Ratio	Coefficient	Std. Error	*P*-value
Public transport	0.718	-0.332	0.138	0.016
Lives with minors	0.697	-0.361	0.106	< 0.001
Education: Student vs high school or less	1.527	0.423	0.259	0.102
Education: University vs high school or less	1.276	0.244	0.113	0.031
Employment: Retired vs employed	1.673	0.515	0.212	0.015
Employment: Unemployed vs employed	0.997	-0.003	0.163	0.984
Medication	2.422	0.885	0.123	< 0.001
Age group: 0–17 vs 45–64	0.303	-1.195	0.291	< 0.001
Age group: 18–44 vs 45–64	0.527	-0.641	0.134	< 0.001
Age group: 65 + vs 45–64	1.684	0.521	0.198	0.009
Pandemic	8.225	2.107	0.165	< 0.001
Islands	0.675	-0.394	0.257	0.125
Macro-regional cumulative incidence	1.012	0.012	0.001	< 0.001

of being vaccinated compared to those aged 45–64 (OR = 0.303). The 18–44 age group also exhibited a lower likelihood of vaccination, with the odds being 47% lower than the 45–64 age group (OR = 0.527). Conversely, those aged 65 and older had 68% higher odds to be vaccinated compared to the 45–64 age group (OR = 1.684).

The emergence of the COVID-19 pandemic was significantly associated with vaccination behavior. The analysis revealed that individuals had 723% higher odds of being vaccinated against influenza during the pandemic season (OR = 8.225). Additionally, higher macro-regional cumulative incidence was associated with increased vaccination likelihood. For every additional reported unit case of ILI per 1,000 consulting patients over the flu season within a macro-region, the log-odds of being vaccinated increased by 1%, holding all other variables constant (OR = 1.012).

Reasons for vaccination

In addition to their demographic factors, Influweb participants who indicated they were vaccinated were asked to provide their reasoning for doing so, with the option to select multiple reasons from a given list. This can be seen in Table 4. The reasons provided by the participants align with some of the significant predictors identified in the regression model, offering additional context to understand their vaccination decisions.

Over half of the participants (53%) reported that they chose to get vaccinated to decrease their personal risk of getting influenza, while 33% of participants cited reducing the risk of spreading influenza to others as a motivation. A significant percentage (46%) also reported that they always get the vaccine, possibly due to established health routines or enhanced access through retirement health plans. Moreover, 42% of participants indicated belonging to a high-risk group as a reason for vaccination.

Table 4 Reasons for vaccination among vaccinated participants

Reason for Vaccination	Percentage (%)
Vaccination decreases my risk of getting influenza	53
I always get the vaccine	46
I belong to a risk group (e.g., pregnant, over 65, underlying health condition, etc.)	42
Vaccination decreases the risk of spreading influenza to others	33
My doctor recommended it	17
I don't want to miss work/school	16
The vaccine was free (no cost)	14
The vaccine was readily available and vaccine administration was convenient	12
It was recommended in my workplace/school	9
Other reason(s)	4

Discussion

The results of this study, conducted in Italy using data from the Influweb platform, provide a view of the factors influencing influenza vaccination uptake among the Influweb population. By examining logistic regression outcomes and qualitative reasons for vaccination, these findings can be contextualized within the broader literature on vaccination behavior.

Our study found that those living with minors were less likely to vaccinate, which is contrary to findings from France [48] and Hong Kong [49], where living with children was associated with higher vaccination rates. Within our sample, living with minors appears to be related to age. The younger age groups reported higher rates of cohabitating with children, though still 16% of those aged 65 and older report living with children. Individuals living with minors are often at a higher risk of exposure to respiratory viruses, such as influenza and COVID-19, making this an important finding that warrants further investigation [50, 51].This association may stem from individuals being inadequately informed of their own risk of infection when cohabitating with minors. Targeted public health communication could help raise risk perception among those living with minors, encouraging vaccination uptake as a protective measure against both direct and indirect exposure to respiratory viruses.

Risk perception has been widely recognized as a key motivator in vaccination decisions, as demonstrated in multiple studies [7, 9, 10, 52]. In particular, these studies found that older adults and individuals managing chronic conditions experience a heightened perceived risk of severe influenza outcomes, making them more likely to get vaccinated. This is supported by our data. The logistic regression model revealed older age and taking medication for chronic disease to be strongly associated with vaccination likelihood, aligning with research from Hong Kong [49], Italy [53], and a WHO review covering multiple countries [54]. Interestingly, while 42% of vaccinated participants cite belonging to a risk group as a motivating factor in their choice to vaccinate, only 17% note their doctor recommended vaccination to them. This, along with model results, implies the success of current public health campaigns informing those in high-risk groups of their increased risk from influenza, as well as a level of health literacy in the Influweb population.

Education level is another significant predictor of vaccination, with higher education correlating with increased vaccination rates. Participants with university-level education were more likely to vaccinate, consistent with the findings from China by Gong et. al [52], France by Vaux et. al [48], and Italy by Giacomelli et. al [55]. Similarly, a study by Wang et. al [11] suggests that better health literacy and awareness of more educated populations drives higher vaccination uptake. In our study, over half of the participants (53%) reported that they chose to get vaccinated to decrease their risk of getting influenza, further reflecting this health awareness.

Broader literature emphasizes the role of social influences in vaccination behaviors [7, 9, 12, 13, 49]. The consensus is that social encouragement from family, coworkers and, most notably, healthcare providers significantly promotes vaccination. Influweb participants indicated that they vaccinated due to recommendations from their doctors (17%) or because their school/workplace encouraged it (9%), though these were not the primary reasons given. One third (33%) of participants cited reducing the risk of spreading influenza to others as a motivation, indicating a degree of public health consciousness among the Influweb cohort. A significant percentage (45%) also reported that they always get the vaccine, possibly due to established health routines or enhanced access through retirement health plans. This routine behavior is consistent with the generally higher rate of vaccination observed in this population compared to the official average (for more details see supplementary materials "Vaccination coverage: comparing the Influweb sample to officially reported estimates").

The relationship between cumulative incidence of influenza-like illness (ILI) and vaccination rates has been explored previously, with mixed results. Research conducted across 14 European countries found inconsistent correlations between influenza vaccination coverage and ILI incidence, with significant positive correlations observed in some countries but not others [56]. In Italy, no significant correlation was found between vaccination coverage and ILI incidence at the national level over the 1999–2000 to 2013–2014 flu seasons [56]. However, our study, which focused on a smaller geographical scale at the macro-regional level, found clearer associations: higher regional cumulative ILI incidence were significantly linked to increased odds of vaccination. This association reflects trends at the macro-regional level and should not be interpreted as evidence of individual behavior. The observed relationship may reflect broader regional characteristics influencing vaccination uptake, rather than a direct link between ILI incidence and individual vaccination decisions. For example, a region with a higher incidence may respond by increasing regional health campaigns for vaccinations.

Beyond macro-regional considerations, we also tested the inclusion of survey season dummies in our logistic regression model to account for unobserved, season-specific factors. Incorporating these dummy variables left most model coefficients relatively unchanged, except for the pandemic indicator variable—which remained significant but shifted in magnitude—and the seasonal, regional ILI cumulative incidence coefficient, which lost its statistical significance. Further examination using variance inflation factors (VIF) revealed that the season dummies and cumulative incidence variable shared a considerable amount of information, indicating multicollinearity. Ultimately, for the sake of model simplicity and clearer interpretation, we retained only the cumulative incidence variable.

Our logistic regression analysis revealed that Influweb participants who rely on public transportation as their main form of travel had lower odds of being vaccinated. This finding aligns with Yang et. al [49], who highlighted transportation access as a key factor in vaccine uptake, with individuals who have private or more direct forms of transportation (e.g., a personal car or the ability to walk to a healthcare provider) being more likely to vaccinate. This result may initially seem counterintuitive given public transportation users are often at higher risk of exposure to respiratory viruses due to close contact with others in enclosed spaces [57]. In many regions worldwide, reliance on public transit may serve as a proxy for lower socioeconomic status (SES). However, the relationship between public transport use and SES in Italy is more nuanced. According to a recent national report, the percentage of high-income individuals using public transportation is actually higher than that of the lowest-income bracket [58].

In Italy, public transport access is concentrated in urban areas, where jobs with higher earning potential are more prevalent. Within these cities, central districts offer the most convenient public transportation options, but also impose restrictions on car ownership, such as limited parking, traffic restrictions, and additional costs for private parking spaces. In contrast, areas with reduced access to public transport, where private vehicles are more necessary, are often located on the outskirts of cities. This pattern extends further into rural areas, where transportation infrastructure is less developed. Another report notes that younger individuals and students—groups that tend to have lower vaccination rates—are among the most frequent users of public transportation [59].

A closer look at the Influweb data shows that the highest percentage of public transport use as the primary mode of transportation (18%) is reported by participants who identify as students or have a university-level education, whereas those with a high school education or less report the lowest usage. However, among different employment statuses, the unemployed exhibit the highest percentage of public transport use (22%). Regarding age, 14% of participants aged 0–17 report public transport as their main mode of travel, rising to 22% among those aged 18–44, then dropping to 11% for ages 45–64 and just 9% for those 65 and older. These overlapping demographic factors—age, employment status, and education—highlight the complexity of using public transport as a simple proxy for socioeconomic status (SES).

Public transportation may present barriers to vaccination access compared to more convenient options like walking or private transportation. Vaccination convenience and accessibility plays a critical role in increasing uptake, as demonstrated by research from Abbas et. al [6] and Nagata et. al [7]. These studies conclude that when vaccinations are easily accessible at pharmacies, workplaces, or community centers, the rate of uptake significantly increases. In our study, 14% of Influweb participants reported vaccinating because the vaccine was free, while 12% cited the convenience of availability and administration as key reasons for their decision. This underscores the importance of ensuring that vaccines are not only affordable but also easily accessible in locations that people frequent regularly.

Given the high vaccination coverage observed during the pandemic period (2020–2021), we further examined why vaccinated participants reported choosing to get vaccinated in this unique season by replicating Table 4 for the 2020–2021 season (see Supplementary Material, “Reasons for vaccination in pandemic season”). The results showed that the primary reasons cited during the pandemic season did not substantially differ from those reported over the entire study period.

The COVID-19 pandemic nonetheless had a significant impact on influenza vaccination behavior, with Influweb participants having higher odds to vaccinate during the pandemic. This finding aligns with global studies from Italy [60], England [61], and the United States [62], all of which reported increased flu vaccination rates during the pandemic. Many individuals may have been offered flu vaccinations alongside COVID-19 vaccinations, further boosting flu vaccine uptake. The heightened health alert during the pandemic may have contributed to the rise in vaccinations, however, this could partly reflect a pandemic-driven sample selection effect, where individuals concerned about their health or the pandemic were more likely to participate in the platform during this time.

Nevertheless, the value of participatory survey platforms should not be underestimated, as they have proven to be reliable tools for disease surveillance, particularly when combined with traditional data sources [26]. Participatory surveillance offers unique opportunities to collect health-related data that might not otherwise be available from official health authorities. For example, while the official estimates used in this study provide vaccination coverage by region and age group, they lack additional contextual data such as household composition or transportation habits, which are available from Influweb.

Additional variables that add value to participatory surveillance are weekly reports of individual symptoms (or lack thereof), lifestyle factors, healthcare seeking behaviors and reasons for choosing to vaccinate, among others. Moreover, participatory surveillance can capture populations that may be underrepresented in traditional healthcare-based surveillance, such as individuals who do not seek medical care for influenza-like illness [27]. Such data are used by public health institutions, including the Italian National Institute of Health, to complement conventional monitoring efforts. Digital surveys like Influweb's are fully anonymous and administered online, which can reduce social pressure [63].

Future research could benefit from integrating objective vaccination records to validate self-reported data and expanding studies to include more diverse participant populations to ensure broader applicability.

Limitations

Logistic regression is a standard approach for analyzing dichotomous outcomes in epidemiological research and allows for clear interpretation of findings. More complex methods, such as machine learning models, may improve predictive accuracy, but our focus was on understanding associations rather than optimizing prediction. Our covariate selection, informed by prior literature, reduced the need for more complex feature-selection techniques. More complex models can be prone to overfitting and reduced interpretability, particularly when working with moderate sample sizes [64]. We did not include interactions among covariates in our model for similar reasons, though future studies could explore interaction effects more robustly.

To better understand the composition of our study population, we compared seasonal vaccination coverage rates in Influweb to official national estimates (detailed in the supplementary materials section "Vaccination Coverage"). Initially, Influweb's vaccination coverage aligned with official estimates but later surpassed it, with the widest gap in the 2020–2021 pandemic season. Both data sources recorded their highest vaccination rates during this period. We examined changes over time in key vaccination determinants (see supplementary materials section "Trends in Influweb sample characteristics"). Age, education, medication use, and living with minors showed distributional shifts that may explain the increasing coverage in Influweb. The share of older adults grew, while younger adults declined, and the proportion of university-educated participants increased. Both of these increasing groups were associated with higher vaccination rates in the regression model.

The use of self-reported data may introduce recall bias, where participants might misremember or misreport their vaccination status.

Although we cannot verify the accuracy of the reported data, the collection of retrospective data allows us to assess consistency in participant responses across flu seasons (see supplementary materials "Retrospective data"). Only 5% of surveys from the subset of individuals that participated in consecutive seasons showed inconsistent recall from one season to the next, providing assurance that the passage of time may not greatly alter participants' reporting of vaccination events.

The voluntary nature of the Influweb platform may attract health-conscious participants who are more likely to engage in preventive behaviors like vaccination [20]. While we made efforts to improve the representativeness of the sample in terms of sex, age, and macro-region via post-stratification weights (see Supplementary Materials: Post-stratification weighting and population comparisons), they cannot fully eliminate selection bias.

Conclusions

Moving forward, leveraging participatory platforms like Influweb enhances public health efforts by providing unique insights into contextual factors such as household composition, transportation habits, and lifestyle behaviors, which are not typically recorded in official vaccination data. Integrating this type of information with traditional surveillance strengthens vaccination strategies by identifying populations at higher risk or with lower uptake.

This study confirms many established determinants of influenza vaccination, such as age, chronic disease and higher educational attainment within the Influweb context. The odds ratios for the COVID-19 pandemic and increased cumulative incidence rates of influenza-like illness in macro-regions highlight how health crises and broader epidemiological trends are associated with changes in vaccination uptake.

A key strength of this study is its transitional data spanning both pre-pandemic and pandemic periods. This allows us to assess how vaccination behaviors evolved during a major public health crisis, whether through increased public health awareness, influenza vaccine co-administration with COVID-19 vaccines, or sample selection effects. Additionally, our findings reveal that those living with minors and those who rely on public transportation have lower odds of vaccination uptake, despite their higher exposure risk for ILI.

Improving accessibility and targeting public health communication efforts at high-risk groups could increase vaccination coverage and reduce influenza transmission within these populations. For example, given the lower odds for vaccination observed among public transportation users, transit networks could serve as strategic locations for targeted vaccine messaging.

As a result, our study does not aim to produce nationally generalizable estimates but rather to assess the determinants of influenza vaccination within this engaged population. Our results reflect associations rather than causal relationships. We cannot rule out the influence of unmeasured confounding variables that may contribute to the observed trends.

We do not collect direct measures of socioeconomic status (SES) or healthcare access, for example, though education level and employment status serve as well-established proxies for SES. Additionally, macro-region indicator variables can control for organizational differences between regional healthcare services that can impact utilization and access to vaccinations, shown to vary across Italy [65]. To further address healthcare access, our model includes mode of transportation, which has been shown to be relevant in vaccination behavior [49]. Nevertheless, these proxies cannot capture all social and behavioral determinants of vaccination.

Lifestyle factors that indicate a participant's investment in their health, such as exercise and diet, may play a role in vaccination choices, but these were not directly measured. We collect data on smoking behavior, a commonly used lifestyle health indicator [66, 67], but it did not show an association with vaccination uptake in our analysis.

Vaccine hesitancy, recognized as a global health threat by the WHO [4], may be influenced by factors such as distrust in health authorities, historical injustices in healthcare, or specific religious doctrines discouraging immunizations. These beliefs and cultural norms can strongly shape an individual's decision to vaccinate, yet our study does not directly measure them.

Future efforts should broaden recruitment to improve representativeness and reduce selection bias. Enhancing the Influweb survey context to help address unmeasured confounding and provide deeper insights into vaccination behavior.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12889-025-22496-8.

Supplementary Material 1.

Acknowledgements
The authors are grateful to Matteo Delfino and Giovanni Frigione for the technical support with the data collection and extraction.

Authors' contributions
All authors contributed to the design of the study. DP designed the original survey. Data were processed and prepared for analyses by NG, MM and KK. KK performed the statistical analysis, with oversight from DP, MM and NG. The first draft of the manuscript was produced by KK and all authors reviewed, edited, and approved the final version.

Funding
The authors acknowledge support from the Lagrange Project of the Institute for Scientific Interchange Foundation (ISI Foundation) funded by Fondazione Cassa di Risparmio di Torino (Fondazione CRT). MM and DP acknowledge funding from the EU Commission in the framework of the EU Horizon Project VERDI (101045989). MM and DP acknowledge support by the ESCAPE project (101095619), funded by the European Union. DP and MM acknowledge support from the EU Commission in the framework of the EU Horizon Project SIESTA (101131957). Views and opinions expressed are however those of the authors' only and do not necessarily reflect those of the European Union or the Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Data availability
The raw data cannot be shared publicly because it contains personal individual information which could compromise users' privacy. Data requests should be addressed to the authors.

Declarations

Ethics approval and consent to participate
For influweb.it the research was conducted in accordance with Italian Data Protection Authority (Garante per la protezione dei dati personali) regulations on privacy, data collection and treatment, reported here https://www.garanteprivacy.it/web/garante-privacy-en. The Italian Data Protection Authority adheres to the norms set by the European Union's General Data Protection Regulation (GDPR). Influweb's data collection process was reviewed and approved by the institutional review board of the ISI Foundation which waived the need for ethics approval as the following applies: all use takes place in compliance with the rules contained in the GDPR; informed consent was obtained online from all participants of the platform at enrollment according to regulations, enabling the collection, storage, and treatment of data, and their publication in anonymized, processed, and aggregated forms for scientific purposes; the website has a "Privacy Statement" section in which the users who decide to enroll in the study can find all the information on who is responsible for the data acquisition and processing in each country. Informed consent was obtained online from all participants enabling the collection, storage, and treatment of data, and their publication in anonymized, processed, and aggregated for scientific purposes. The Influweb website (https://influweb.org/) has a "Privacy Policy" section in which the users who decide to enroll in the study can find all the information on who is responsible for the data acquisition and processing. To ensure privacy and data security, all collected data is pseudonymized—participants' personal identifiers (e.g., email addresses) are stored separately and are inaccessible to researchers. Survey responses are linked only via an anonymized ID, preserving full privacy. Influweb operates on Google Cloud Platform (GCP) infrastructure in Europe, maintaining GDPR compliance, with continuous backups. Additionally, security measures to protect participant data include hashed passwords, two-factor authentication (2FA) and protection of data in transit by Transport Layer Security v1.3 ("TLSv1.3").

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 8 November 2024 Accepted: 25 March 2025
Published online: 10 April 2025

References

1. Rémy V, Zöllner Y, Heckmann U. Vaccination: the cornerstone of an efficient healthcare system. J Mark Access Health Policy. 2015;3(1):27041.
2. Vallis M, Bacon S, Corace K, Joyal-Desmarais K, SheinfeldGorin S, Paduano S, Presseau J, Rash J, MengistuYohannes A, Lavoie K. Ending the pandemic: how behavioural science can help optimize global COVID-19 vaccine uptake. Vaccines. 2021;10(1):7.

3. Sheikh S, Biundo E, Courcier S, Damm O, Launay O, Maes E, et al. A report on the status of vaccination in Europe. Vaccine. 2018;36(33):4979–92.
4. World Health Organization (WHO). Ten Threats to Global Health in 2019. 2019. Available from: https://www.who.int/emergencies/ten-threats-to-global-health-in-2019
5. Roller-Wirnsberger R, Lindner S, Kolosovski L, Platzer E, Dovjak P, Flick H, et al. The role of health determinants in the influenza vaccination uptake among older adults (65+): a scope review. Aging Clin Exp Res. 2021;33:2123–32.
6. Abbas KM, Kang GJ, Chen D, Werre SR, Marathe A. Demographics, perceptions, and socioeconomic factors affecting influenza vaccination among adults in the United States. PeerJ. 2018;6:e5171.
7. Nagata JM, Hernández-Ramos I, Kurup AS, Albrecht D, Vivas-Torrealba C, Franco-Paredes C. Social determinants of health and seasonal influenza vaccination in adults≥ 65 years: a systematic review of qualitative and quantitative data. BMC Public Health. 2013;13:1–25.
8. Zanobini P, Lorini C, Caini S, Lastrucci V, Masocco M, Minardi V, Possenti V, Mereu G, Cecconi R, Bonaccorsi G. Health literacy, socioeconomic status and vaccination uptake: a study on influenza vaccination in a population-based sample. Int J Environ Res Public Health. 2022;19(11):6925.
9. Eder M, Omic H, Gorges J, Badt F, Kikic Z, Saemann MD, et al. Influenza vaccination uptake and factors influencing vaccination decision among patients with chronic kidney or liver disease. PLoS One. 2021;16(4):e0249785.
10. Caserotti M, Girardi P, Rubaltelli E, Tasso A, Lotto L, Gavaruzzi T. Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. Soc Sci Med. 2021;1(272):113688.
11. Wang L, Guo M, Wang Y, Chen R, Wei X. The relationship between influenza vaccine hesitancy and vaccine literacy among youth and adults in China. Front Immunol. 2024;15:1444393.
12. Yeung MP, Lam FL, Coker R. Factors associated with the uptake of seasonal influenza vaccination in adults: a systematic review. J Public Health. 2016;38(4):746–53.
13. Abdallah DA, Lee CM. Social norms and vaccine uptake: College students' COVID vaccination intentions, attitudes, and estimated peer norms and comparisons with influenza vaccine. Vaccine. 2021;39(15):2060–7.
14. Baltrusaitis K, Reed C, Sewalk K, Brownstein JS, Crawley AW, Biggerstaff M. Healthcare-seeking behavior for respiratory illness among Flu Near You participants in the United States during the 2015–2016 through 2018–2019 Influenza Seasons. J Infect Dis. 2022;226(2):270–7.
15. Harvey EP, Trent JA, Mackenzie F, Turnbull SM, O'Neale DJ. Calculating incidence of Influenza-like and COVID-like symptoms from Flutracking participatory survey data. MethodsX. 2022;9:101820.
16. Perrotta D, Paolotti D, Rizzo C, Bella A, Tizzoni M. Participatory online surveillance as a supplementary tool to sentinel doctors for influenza-like illness surveillance in Italy. PLoS One. 2017;12(1):e0169801.
17. van Noort SP, Codeco CT, Koppeschaar CE, van Ranst M, Paolotti D, Gomes MG. Ten-year performance of Influenzanet: ILI time series, risks, and vaccine effects in the Grote Griepmeting, Gripenet, and Influweb cohorts. Epidemics. 2015;13:128–36.
18. Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, Edmunds WJ, Kjelsø C, Mexia R, Moreno Y, Meloni S, Paolotti D. Influenzanet: Citizens among 10 countries collaborating to monitor influenza in Europe. JMIR Public Health Surveill. 2017;3(3):e7429.
19. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, Koppeschaar C, Rehn M, Smallenburg R, Turbelin C, Van Noort S. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clin Microbiol Infect. 2014;20(1):17–21.
20. Guerrisi C, Turbelin C, Blanchon T, Hanslik T, Bonmarin I, Levy-Bruhl D, et al. Participatory syndromic surveillance of influenza in Europe. J Infect Dis. 2016;214(suppl_4):S386–92.
21. Perrotta D, Delle Vedove D, Obi C, Pebody R, Schifanella R, Paolotti D. Spatio-temporal analysis of flu-related drugs uptake in an online cohort in england. InProceedings of the 9th International Conference on Digital Public Health 2019 Nov 20 (pp. 111–118).
22. Commission E. Commission Implementing Decision (EU) 2018/945 of 22 June 2018 on the communicable diseases and related special health issues to be covered by epidemiological surveillance as well as relevant case definitions. OJEU. 2018;61:170.
23. Greenacre ZA. The importance of selection bias in internet surveys. Open J Stat. 2016;6(3):397–404.
24. Leal Neto O, Paolotti D, Dalton C, Carlson S, Susumpow P, Parker M, Phetra P, Lau EH, Colizza V, Jan van Hoek A, Kjelsø C. Enabling multicentric participatory disease surveillance for global health enhancement viewpoint on global flu view. JMIR Public Health Surveill. 2023;9:e46644.
25. McNeil C, Verlander S, Divi N, Smolinski M. The landscape of participatory surveillance systems across the one health spectrum: systematic review. JMIR Public Health Surveill. 2022;8(8):e38551.
26. Perrotta D, Tizzoni M, Paolotti D. Using participatory Web-based surveillance data to improve seasonal influenza forecasting in Italy. In: Proceedings of the 26th International Conference on World Wide Web. 2017. p. 303–310.
27. Peppa M, Edmunds WJ, Funk S. Disease severity determines health-seeking behaviour amongst individuals with influenza-like illness in an internet-based cohort. BMC Infect Dis. 2017;17:1–13.
28. Bennett DA. How can I deal with missing data in my study? Aust N Z J Public Health. 2001;25(5):464–9.
29. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res. 1999;8(1):3–15.
30. Italian National Institute of Statistics (Istat). Intercensal population estimates by sex, age, and Municipality [Internet]. Data for years 2001–2019; based on the 2001 and 2011 population censuses and the 2018 Permanent Census of Population and Housing. Available from: http://dati.istat.it/Index.aspx?lang=en
31. Italian National Institute of Statistics (Istat). Resident municipal population by age, sex, and marital status [Internet]. Data at January 1st 2020 and 2021; based on the Permanent Census of Population and Housing results. Available from: http://dati.istat.it/Index.aspx?lang=en
32. Ibuka Y, Ohkusa Y, Sugawara T, Chapman GB, Yamin D, Atkins KE, et al. Social contacts, vaccination decisions and influenza in Japan. J Epidemiol Community Health. 2016;70(2):162–7.
33. Istituto Superiore di Sanità. FluNet Surveillance System [Internet]. Rome: Istituto Superiore di Sanità. Available from: https://respivirnet.iss.it/pagine/stagioni.aspx
34. Rana R, Singhal R. Chi-square test and its application in hypothesis testing. J Pract Cardiovasc Sci. 2015;1(1):69–71.
35. Grills LA, Wagner AL. The impact of the COVID-19 pandemic on parental vaccine hesitancy: a cross-sectional survey. Vaccine. 2023;41(41):6127–33.
36. Wall ER. Weighting in Survey Data. Journal of Visual Impairment & Blindness. 2022;116(1):117–8.
37. Nwanyanwu KH, Andoh J, Chen E, Xu Y, Deng Y. Social determinants of health in diabetic retinopathy in the US: evidence from the National Health and Nutrition Examination Survey (2005–2008). Invest Ophthalmol Vis Sci. 2021;62(8):1135.
38. Lavassani KM, Mohamed B, Kumar V. Developments in analysis of multiple response survey data in categorical data analysis: the case of enterprise system implementation in large North American firms. J Appl Quant Methods. 2009;4(1):36–50.
39. Nugroho WH, Handoyo S, Akri YJ, Sulistyono AD. Building multiclass classification model of logistic regression and decision tree using the chi-square test for variable selection method. J Hunan Univ Nat Sci. 2022;49(4).
40. Lumley T, Scott A. Fitting regression models to survey data. Stat Sci. 2017;1:265–78.
41. Binder DA. On the variances of asymptotically normal estimators from complex surveys. International Statistical Review/Revue Internationale de Statistique. 1983;1:279–92.
42. Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. J Hum Resour. 2015 Mar;50(2):317–72. https://doi.org/10.3368/jhr.50.2.317. Available from: https://jhr.uwpress.org/content/50/2/317.short
43. Faraway JJ. Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models. Chapman and Hall/CRC; 2016.
44. Portet S. A primer on model selection using the Akaike Information Criterion. Infect Dis Model. 2020;5:111–28.
45. Lewis F, Butler A, Gilbert L. A unified approach to model selection using the likelihood ratio test. Methods Ecol Evol. 2011;2(2):155–62.
46. Cousineau D, Allan TA. Likelihood and its use in parameter estimation and model comparison. Mesure et évaluation en éducation. 2015;37(3):63–98.
47. Tenny S, Hoffman MR. Odds Ratio [Internet]. In: StatPearls. Treasure Island (FL): StatPearls Publishing; 2025. [updated 2023 May 22; cited 2025 Apr 3]. Available from: https://www.ncbi.nlm.nih.gov/sites/books/NBK431098/.

48. Vaux S, Van Cauteren D, Guthmann JP, Le Strat Y, Vaillant V, De Valk H, Lévy-Bruhl D. Influenza vaccination coverage against seasonal and pandemic influenza and their determinants in France: a cross-sectional survey. BMC Public Health. 2011;11:30.
49. Yang L, Nan H, Liang J, Chan YH, Chan L, Sum RW, et al. Influenza vaccination in older people with diabetes and their household contacts. Vaccine. 2017;35(6):889–96.
50. Williams CJ, Schweiger B, Diner G, Gerlach F, Haaman F, Krause G, et al. Seasonal influenza risk in hospital healthcare workers is more strongly associated with household than occupational exposures: results from a prospective cohort study in Berlin, Germany, 2006/07. BMC Infect Dis. 2010;10:1–11.
51. McDonald HI, Minassian C, Brown J, Rentsch C, Mathur R, Hulme W, et al. Association between living with children and outcomes from COVID-19: an OpenSAFELY cohort study of 12 million adults in England. 2021.
52. Gong L, Zhang X, Qu Z, Francis MR, Han K, Xu C, et al. Public interest in distribution and determinants of influenza and pneumonia vaccination during the COVID-19 pandemic: an infodemiology and cross-sectional study from China. Vaccines (Basel). 2021;9(11):1329.
53. Barbadoro P, Marigliano A, Di Tondo E, Chiatti C, Di Stanislao F, D'Errico MM, et al. Determinants of influenza vaccination uptake among Italian healthcare workers. Hum Vaccin Immunother. 2013;9(4):911–6.
54. Roller-Wirnsberger R, Lindner S, Kolosovski L, Platzer E, Dovjak P, Flick H, Tziraki C, Illario M. The role of health determinants in the influenza vaccination uptake among older adults (65+): a scope review. Aging Clin Exp Res. 2021;33:2123–32.
55. Giacomelli A, Galli M, Maggi S, Noale M, Trevisan C, Pagani G, Antonelli-Incalzi R, Molinaro S, Bastiani L, Cori L, Bianchi F. Influenza vaccination uptake in the general Italian population during the 2020–2021 flu season: data from the EPICOVID-19 online web-based survey. Vaccines. 2022;10(2):293.
56. Spruijt IT, de Lange MMA, Dijkstra F, Donker GÁ, van der Hoek W. Long-term correlation between influenza vaccination coverage and incidence of influenza-like illness in 14 European countries. PLoS One. 2016;11(9)
57. Howland RE, Cowan NR, Wang SS, Moss ML, Glied S. Public transportation and transmission of viral respiratory disease: Evidence from influenza deaths in 121 cities in the United States. PLoS One. 2020;15(12).
58. Istituto Superiore di Formazione e Ricerca per i Trasporti (ISFORT). 21° Rapporto sulla mobilità degli italiani. Roma: ISFORT; 2024. Available from: https://www.isfort.it/wp-content/uploads/2024/11/21_RapportoMobilita_Sintesi.pdf
59. Istituto Nazionale di Statistica (ISTAT). Spostamenti quotidiani e nuove forme di mobilità. Rome: ISTAT; 2018. Available from: https://www.istat.it/it/files/2018/11/Report-mobilit%C3%A0-sostenibile.pdf
60. Porreca A, Di Nicola M. Flu vaccination coverage in Italy in the COVID-19 era: A fuzzy functional k-means (FFKM) approach. J Infect Public Health. 2023;16(11):1742–9.
61. Watkinson RE, Williams R, Gillibrand S, Munford L, Sutton M. Evaluating socioeconomic inequalities in influenza vaccine uptake during the COVID-19 pandemic: A cohort study in Greater Manchester, England. PLoS Med. 2023;20(9)
62. Yang Y. How influenza vaccination changed over the COVID-19 pandemic. medRxiv. 2023;2023–03.
63. Elliott R, Sawkar M, Williams M. The power of mobile platforms for data collection. Sight Life. 2019;33(1):64–9.
64. Montesinos López OA, Montesinos López A, Crossa J. Overfitting, model tuning, and evaluation of prediction performance. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction [Internet]. Cham (CH): Springer; 2022. Available from: https://www.ncbi.nlm.nih.gov/books/NBK583970/ https://doi.org/10.1007/978-3-030-89010-0_4.
65. Matranga D, Maniscalco L. Inequality in healthcare utilization in Italy: how important are barriers to access? Int J Environ Res Public Health. 2022;19(3):1697.
66. La Fauci V, Mondello S, Squeri R, Alessi V, Genovese C, Laudani N, Cattaruzza MS. Family, lifestyles and new and old type of smoking in young adults. Insights from an Italian multiple-center study. ANNALI DI IGIENE MEDICINA PREVENTIVA E DI COMUNITÀ. 2021;32(2):131–40.
67. Donfrancesco C, Buttari B, Marcozzi B, Sieri S, Di Lonardo A, Lo Noce C, Profumo E, Vespasiano F, Agnoli C, Vannucchi S, Silano M. Nutrition, physical activity and smoking habit in the Italian general adult population: CUORE project health examination survey 2018–2019. InHealthcare 2024 Feb 15 (Vol. 12, No. 4, p. 475). MDPI.

Publisher's Note