Master of Public Health

Master de Santé Publique

The H1N1/A 2009 pandemic: Important attributes that the Boruta classification method can select to better understand pandemic Vaccine Uptake

Georges Beiriger
Master of Public Health Year 2, 2020 – 2021

Location of the practicum:
Remote, EHESP

Academic supervisor:
Simon Combes, EHESP

Acknowledgment

I would like to thank my family and friends.
Without you I would not have made it!
I would also like to thank Professor Simon Combes for his help in my thesis. His support was essential to surmount this tortuous year.

Contents

Abstract

Background: The 2009 swine flu pandemic global suboptimal Vaccine Uptake demonstrates the complex nature of pandemic vaccination determinants. While there is a consensus in the literature to better understand Vaccine Hesitancy in order to implement accessible and effective vaccination campaigns, few studies have used classification methods to better determine predictors of Vaccine Uptake.

Aims:

1. Identify important predictors of H1N1/A pandemic Vaccine Uptake using the Boruta classification method.
2. Create multivariate logistic regression models using important attributes selected by the Boruta method to identify predictors' direction of association with pandemic Vaccine Uptake

Methods: We first applied the Boruta classification method to identify important predictors of Vaccine Uptake on the "Enquête sur la santé et la protection sociale 2010" survey questionnaire from IRDES. Second, we use the important features identified by the Boruta tool in nested multivariate logistic regression models to assess predictor's direction of association with pandemic Vaccine Uptake.

Results: We confirmed 17 attributes as important in classification of pandemic H1N1 Vaccine Uptake. Among the sociodemographic factor, age class achieved the highest importance score of 13.56. In regression model M3 the associated OR for the 76 to 98 age class was 4.68 with a 95% CI of [3.25, 6.11].For health access, contact with the following physician ranked high with a score of 7.32. In model M2 the attained OR was 1.66 with a 95% CI of [1.50, 1.82]. For health behaviour related variable, consumption of alcohol, reached a 10.63 importance score. However when tested in the regression models, this variable was not significant.

Conclusion: We performed an extended literature review on Vaccine Hesitancy and determinants of Vaccine Uptake in the context of pandemic A/H1N1.We have then identified several important socio-demographic, occupational health behavior and health access features using the Boruta classification method. We finally tested in nested logistic regression models these important features association with pandemic A/H1N1 Vaccine Uptake.

Keywords: Vaccine Uptake- Vaccine Hesitancy- Determinants of pandemic A/H1N1 VU- Boruta method- Nested multivariate logistic regression model

Introduction

The World Health Organization (WHO) announced on June 11 2009, the outbreak of the influenza A/H1N1 virus (WHO, 2009). In France, the first case was declared on May 1 2009 (Desenclos et al., 2010). The pandemic lasted from late October to late December 2009 (Bocquier et al., 2018; Vaux et al., 2011). Because the pandemic vaccine was not available to all, the French ministry of health proposed the vaccine for priority groups first (Inpes & ministère de la santé, 2009). Briefly, these were the 5 identified priority groups: pregnant woman, infants aged 6- 23 months at risk for influenza complications, Health Care Worker (HCW), and finally household contacts of infants aged less than 6 months (Inpes & ministère de la santé, 2009). The national vaccination campaign was carried in large vaccination center to alleviate overload on General Practitioners (GP) and pediatricians (Inpes & ministère de la santé, 2009). Unlike other national vaccination campaign that relied on Primary Care Physicians (PCP) and pharmacists among other Health Care Professionals (HCP) to deliver the vaccine, the French government opted to offer the vaccine only in vaccination centers (Vaux et al., 2011). A decision that will be retrospectively judged as detrimental (Schwarzinger et al., 2010).

First, we comprehensively identify determinants of Influenza Vaccine Coverage (IVC) among the French population with a brief overview of global A/H1N1 pandemic Vaccine Uptake (VU) determinants. Later, we define Vaccine Hesitancy (VH) and investigate successively Vaccine Acceptance (VA) and VU in relation to VH by identifying major psychological attitudes and vaccination decision making processes according to classical theories of behavior and other models of VH. Concurrently in our review, we expose methods to increase VU by better understanding VH.

A cross sectional retrospective telephone survey conducted by Vaux et al. (2011) to estimate seasonal and pandemic IVC and its determinants revealed that pandemic IVC in the French general population was estimated to be at 11.1%. Further, seasonal IVC was higher for the general population as well as for priority groups (Vaux et al., 2011). We will focus here exclusively on pandemic IVC. In this study, the strongest determinants found was a previous history of seasonal flu vaccination, followed by occupying a high managerial position. Living in a household with two children or a child less than 5 years of age, or in a household where the head has a university degree were also associated with higher pandemic VU. In addition, VH is a worldwide phenomenon, however it is more pronounced in some areas than others. In fact, a survey conducted in 2016 across 67 countries found that the perception of vaccine safety was worse in the European region, and especially in France, with 41% of French respondents considering that vaccines are unsafe (compared to a global average of 13%) (Vaux et al., 2011). In a systematic review of the determinants of the A/H1N1 pandemic VU, both Bish et al. (2011) and Brien et al. (2012) ascertain that the most cited predictors of increased pandemic VU were higher socioeconomic status , living in a suburban, less deprived area and living in larger households. Likewise, a history of previous seasonal H1N1 vaccination, as well as strong perceptions regarding susceptibility to infection, severity of illness, benefits/effectiveness of the vaccine, and not perceiving barriers to vaccination such as safety or availability of the vaccine were associated with a higher uptake of pandemic vaccination (Vaux et al., 2011) .

Like stated above for France, the A/H1N1 alarming pandemic vaccination rate was estimated at a low of 11% (Vaux et al., 2011). As a consequence, the pandemic vaccination campaign was considered a fiasco (Bocquier et al., 2018). In this study, we will classify in order of importance and assess the association of a multitude of influential factors of A/H1N1 VU in France. To grasp the complexity of pandemic vaccination, I review the literature on the topic of VH.

Literature Review

To better understand the behavioral attitudes behind the decision-making processes involved in the A/H1N1 pandemic vaccination effort, it is necessary to define VH and determine its scope. A recently established definition of VH identifies this concept first as a behavior that cause delay in acceptance or refusal of vaccination despite availability of vaccination services (Larson et al., 2011; MacDonald et al., 2015). In addition, VH is set on a continuum with vaccine hesitant individuals consisting of a wide range of heterogeneous people positioned between those who refuse with no doubts and those who accept with no doubts (Larson et al., 2011; MacDonald et al., 2015). In the middle of these two extremes are the “hesitant” individuals who may refuse some vaccines, but agree to others, delay vaccines, or accept vaccines but are unsure of doing so (Larson et al., 2014). In addition, VH is complex, community and context specific, varying across time, place and vaccines (Peretti-Watel et al., 2019). The 2011 WHO EURO Vaccine Communications Working Group first defined that VH is a confluence of three behaviours: complacency, confidence and convenience entitled the “Three Cs” model of VH (SAGE

Working Group on Vaccine Hesitancy, 2014). Confidence refers to trust in vaccine's safety and effectiveness, the system overseeing its delivery and the motivation of policy makers that guide vaccine campaign implementation (Larson et al., 2014; MacDonald et al., 2015). Complacency is the perceived belief that the risk associated with the disease is less severe than the act of vaccination itself (Larson et al., 2014; MacDonald et al., 2015). Finally, convenience is the degree of availability, appeal of immunization services and ease of accessibility of the offered vaccine (Larson et al., 2014; MacDonald et al., 2015). A more updated and exhaustive model by the WHO Strategic Advisory Group of Experts (SAGE) work group (WG), SAGE WG, on VH identified a VH Matrix of Determinants composed of three large categories displaying the multitude of factors affecting the behavioral decision to accept, delay or reject a vaccine: contextual, individual and group, and vaccine/vaccination-specific influences (MacDonald et al., 2015). This proposed model captures more extensively and effectively the extent of factors affecting the behavioral decision making process relative to vaccine hesitant individuals. The Vaccine Hesitancy determinants matrix is thus only a more detailed and categorized framework compared to the "Three C's model".

N.E. MacDonald / Vaccine 33 (2015) 4161–4164 4163

Table 1
Working Group on Vaccine Hesitancy Determinants Matrix.

Contextual influences Influences arising due to historic, socio-cultural, environmental, health system/institutional, economic or political factors	a. Communication and media environment b. Influential leaders, immunization programme gatekeepers and anti- or pro-vaccination lobbies c. Historical influences d. Religion/culture/gender/socio-economic e. Politics/policies f. Geographic barriers g. Perception of the pharmaceutical industry
Individual and group influences Influences arising from personal perception of the vaccine or influences of the social/peer environment	a. Personal, family and/or community members' experience with vaccination, including pain b. Beliefs, attitudes about health and prevention c. Knowledge/awareness d. Health system and providers – trust and personal experience e. Risk/benefit (perceived, heuristic) f. Immunization as a social norm vs. not needed/harmful
Vaccine/vaccination – specific issues Directly related to vaccine or vaccination	a. Risk/benefit (epidemiological and scientific evidence) b. Introduction of a new vaccine or new formulation or a new recommendation for an existing vaccine c. Mode of administration d. Design of vaccination programme/Mode of delivery (e.g., routine programme or mass vaccination campaign) e. Reliability and/or source of supply of vaccine and/or vaccination equipment f. Vaccination schedule g. Costs h. The strength of the recommendation and/or knowledge base and/or attitude of healthcare professionals

Figure 1: "Working Group on Vaccine Hesitancy Determinants Matrix", from (MacDonald et al., 2015)

Below, we examine some of the classical psychological theories of VH to better understand the decision making process involved with VA. We will then observe how these theories can be better adapted to capture the complexity and the wide scope of VH. However, let us first examine one of the most cited determinant of VH, namely the role of Socioeconomic Status (SES) and its consequences on VU by examining the underlying hypothesized mediators of this determinant. It is generally accepted in the literature that SES are context specific and affect VH in a poorly understood mechanism. In order to examine this association, a study by Bocquier et al. (2018) inspects the social differentiation of VH among French parents by studying the possible mediating effect of trust and commitment to health in a nationwide cross sectional study. Trust like commitment to good health behaviours are two attitudes that affect vaccination (Bocquier et al., 2018) . In fact, confidence and trust towards science in general and vaccination in particular is theorized to decrease in our society with growing disenchantment and scepticism towards science itself (Bocquier et al., 2018; Peretti-Watel et al., 2019). Commitment to making good health decisions is identified as a new societal and cultural trend encouraging the individual to seek out information about their health and exercise a heightened control on decisions affecting their health (Bocquier et al., 2018; Peretti-Watel et al., 2019). On the VH continuum, Bocquier et al. (2018) identify hesitancy as those who refuse, delay or accept with doubt. In this study, besides reporting the prevalence and the degree of VH among French parents, the authors find that more highly educated parents are delayers or refusers more often than those less educated, a finding that can be explained by the former's higher commitment to making ‘‘good” health-related decisions and lower trust toward traditional health authorities. Interestingly in this study, income as a predictor was not associated with VH, whereas educational level was, pointing to the fact that sociocognitive factors are largely more important than material one in the social differentiation of VH, when issues of convenience are removed. This study also echoes that parents who are more educated are less likely to trust official health sources and are thus more likely to pursue actively their health decision. This phenomenon has been coined as "rationalized VH" specifically among the more educated middle class people by Peretti-Watel et al. (2019). Also, in a recent study by (Weston et al. (2017) looking at both adult self-vaccination and children vaccination using the UK Flu watch pandemic cohort data, a prospective cohort study, the authors stress the importance of efficacy/safety and the perceived risk of pandemic influenza as strong predictors of both self and children vaccination. This result affirms, like the previous literature cited, the importance of both commitment to informed health decision and confidence in the vaccination as important attitudes leading to increased pandemic VU. A strong confidence in the vaccine's safety and its effectiveness accompanied by low complacency and an understanding of the risk associated with influenza disease lead to a higher pandemic VU (Weston et al., 2017).

We examine here the classical theories of health behavior influencing VH. Bish et al. (2011) proposed to adopt the Protection Motivation Theory (PMT) framework of psychological attitudes to understand pandemic VU. PMT postulates that for every health behavior there exists an inherent risk to which an individual attempts to develop coping mechanisms in order to protect one's self, commonly referred as self-response and self-efficacy (Bish et al., 2011). The major finding in this study is that, the higher the degree of threat experienced in the 2009 pandemic influenza outbreak was, along with the perception of vaccination as an effective coping strategy, the stronger the intentions and consequently the higher uptake of pandemic vaccination was. The authors define appraisal of threat as a belief to be at a greater risk of developing the disease with more concerns and worry about the severity of the disease. Coping mechanisms, i.e. self-efficacy and self-response, arises mainly from the vaccine safety and side effect profile. The study also points at evidence of social pressure and the source of vaccination information in affecting one's behavior on vaccination .In fact, the more reliable the source is the more people got vaccinated. In turn, those vaccinated exerted pressure on "hesitant" individuals. Finally, previous acceptance of the seasonal vaccine, led to a greater propensity to accept the pandemic vaccine. The latter finding is largely corroborated and replicated in other pandemic influenza VU studies (Brien et al., 2012). Hence, to increase VU, the authors encourages to highlight the risk posed by the pandemic influenza while simultaneously offering tactics to reduce this risk (e.g. vaccination).

The social and demographic factors are largely debated in the literature with mostly heterogeneous results. Factors such as access to health services or negative attitude toward vaccination act as proxies to other more precise variables and thus are not direct predictors of VU. Like we argued before and based on Bocquier et al. (2018) there are underlying factors such as confidence and commitment that mediate the role of SES on VH. However, they can be used to identify at risk groups and minority groups with less vaccine coverage. Schmid et al. (2017) include in their systematic review of pandemic A/H1N1 VU, studies conducted on the general population as well as at risk population such as HCP, pregnant women, children, the elderly or the chronically ill to better capture VU barriers at all societal levels. In this review, an extended Theory of Planned Behavior (TPB) model is proposed. TPB is a theoretical concept consisting mainly of psychological factors namely self-efficacy or locus of control, societal norm and attitudes toward vaccination that tries to explain one's behavior toward vaccination. To this model the authors adds physical, sociodemographics and contextual determinant as "modifying factors" to the classical TPB structure. The authors also include vaccine utility, past knowledge, experience and past behavior as "TPB extensions". In this model, the psychological variables entailing a low risk perception of the disease severity or high risk perception of the adverse effects of the vaccine as well as a negative attitude towards perceived pandemic vaccine utility acted as a significant barriers to VU. The social benefit variable was especially relevant for HCP since they use this argument to promote generalized vaccination. Studies showing weak HCP engagement in promoting vaccine as a social benefit acted as a barrier to VU (E. & D., 2010; Ferguson et al., 2010). In general identifying vaccination as a societal norm influenced positively VU. Additionally, we observe that a larger perceived behavioral control, a positive attitude on vaccination issues , a past behavior of receiving the seasonal flu vaccine and having experienced the disease, contribute to higher VU (Schmid et al., 2017). In fact, the major and most consistent enablers for VU on the "micro-level" were a positive attitude towards influenza vaccines, high perceived utility of vaccination, cues to action, and previous influenza vaccinations. Those finding are also corroborated in the literature as we have seen with the PMT model. However in this review, contextual and sociodemographics barriers are also examined to better capture, as explained above, a larger scope of VU determinants and were further applied on the SAGE's macro level model of vaccination. We can thus conclude with this statement on pandemic VU: for pandemic influenza vaccine, complacency was the major barrier to VU (low worry and perceived risk of the disease), followed by confidence (increased worry about the safety of the vaccine; distrust in authorities). Another systematic review of published literature specifically on childhood VH commissioned by the SAGE WG in which Larson et al. (2014) mapped on the aforementioned SAGE model of VH all relevant studies from 2007 to 2012 in order to identify key determinants of VH, assess barriers or promoters of VU and further develop SAGE's VH model, stated that the SAGE VH model attempt to classify all components of VH fails to identify VH multilevel factor's relative strengths of influence or variability, their dependency or establish their inter-relationship (Larson et al., 2014). Concurrently, the model also fail to convey that VH predictors cannot be considered in isolation as multiple factors are at play. In addition, since the majority of papers included in this systematic review were quantitative cross sectional studies, the authors call for future qualitative methodology to be adopted for a better understanding of VH's scope and expression in a given contextual level. Effectively, this strategy can strengthen understanding around decision-making processes and the ways in which explanatory factors come together to influence vaccination behavior. We can conclude from this article, as evidenced by Larson et al. (2014), that a VH metric does not exist. In fact, most VH determinants studied in the quantitative literature adopt conventional psychological health behavior constructs (e.g., Health Belief Model, Theory of Planned Behavior), as examined above, which do not adequately account for the influence of broader contextual factors. They do not apply neither universal definitions nor measures of psychological attributes. In essence, the latter described strategies fail to understand VH's scope of expression where a multidisciplinary approach, which is broad in scope but context-specific, would greatly support global understanding of vaccine VH. Peretti-Watel et al. (2019) conducted a qualitative study of two largely influential VH predictors: trust toward physician (an issue of confidence) and commitment to vaccination issues (vaccine specific issues) among parents of young children with different SES profiles. The authors look at VH as a decision making process rather than a set of attitudes, beliefs or behaviors, an approach relayed by other authors (Bocquier et al., 2018). Peretti-Watel et al. (2019) also stress on VH's notion of acceptance despite doubt and reluctance, which is effectively where a large population of vaccine hesitant people are situated. As discussed earlier, with the growing movement of healthism identified as empowerment of the individual to seek actively researched health issues, which is often accompanied with disbelief regarding health authorities, Peretti-Watel et al. (2019), seek to assess the level of distrust, commitment and resources parents engage in the vaccination of their children and potentially assess their VH profile depending on their economic status. Like we introduced beforehand in Bocquier et al. (2018) paper, this process: "rationalized VH", engage efforts and resources in the decision of vaccination. What Peretti-Watel et al. (2019) acknowledge too, is the social embeddedness of individuals in their process of vaccination decision making. This is particularly true for women, who were largely the study's respondents and those who more actively participate in making decision about their children vaccinations' decision. For parents who have lost confidence in health authorities, they will need to actively engage resources to gain trust with one figure of the health system, namely the GP or the pediatrician, within the French context. As stated in the study most of the participants, regardless of their economic status, seek vaccination information on the internet but do not particularly trust its information. They are more inclined to build trust with a face to face interaction. Because actively researching health information is a time and effort consuming task, parents need to engage with the “right” figure of the health system, the GP or the pediatrician, for which they have engaged much efforts to identify and can thus delegate confidently more aspects of their decision making process related to their health at broad or vaccine related issues . As shown above, to increase VU it is necessary to understand vaccine behavior’s determinants as dynamic factors anchored within the larger context of the community, which possess its own institutions mandated by different policies. This is an approach often neglected by classical theories and models of health behavior. It is especially important to adapt this approach when dealing with VH and VU. A study of determinants of A/H1N1 pandemic VU in the United States by Kumar et al. (2012) adopted the Social Ecological Model framework (SEM) in attempt to understand the complexity of the various factors that can affect VU and further assess their relative influence on VA behavior. Unlike classical models of health behaviors, the SEM is an overarching more comprehensive framework of determinants. The latter model not only takes into account intrapersonal variables (the micro level layer in Schmid et al. (2017) paper) but other levels of VA predictors to explain VU. In fact, the SEM is formed of intrapersonal, interpersonal, institutional, community and finally policy level. They are interdependent levels with interactions among each other that evolve with time. The strength of this model is that it recognizes that individuals are embedded within social networks, which are in turn part of institutions and communities impacted by policies, in order to assess their mutual and interdependent influence on VA (Kumar et al., 2012). In this study the policy level is measured by health insurance and being part of an immunization priority group. The institutional level is the amount of information provided by HCW on the swine flu disease. The community level is the perceived risk of H1N1 pandemic in the social network of participants. Intrapersonal and interpersonal level refers respectively to all perceived risks, attitudes toward the vaccine and the disease itself, and the social network: specifically the number friends who had received the pandemic vaccine. A strong argument is made to look at VU through the lens of these multiple SEM levels to guide future public health campaign aimed to increase VU. In fact, while it has been showed in Kumar et al. (2012) that VU can be explained mainly through the intrapersonal level approach, a finding corroborated by most other studies in the literature, other levels must be incorporated to better explain variability in VA. Importantly, this study shows that determinants of the SEM also accounts for variance in the intent of getting vaccinated as opposed to having got the pandemic vaccine, effectively short-circuiting a major limitation of cross sectional studies in identifying the direction of the association of attitudes and behaviors leading to VU. In fact in this study, it is indeed the aforementioned attitudes at the multiple level of SEM that affects behaviors and thus VU, not the other way around (Kumar et al., 2012). As it is illustrated by the authors, though the intrapersonal level explained a larger proportion of the variance in VU compared to other levels, explaining 47% of the variance in vaccine behavior, interventions targeting this level have been found to be most effective in conjunction with the interpersonal, institutional, or policy levels (Kumar et al., 2012). As also demonstrated in this paper, a larger proportion of VU was accounted when explained with the intrapersonal level in conjunction with the interpersonal level. In effect this means that an observed variance in VU can be better understood when intrapersonal and interpersonal levels - or any other relevant level affecting VU- are employed together to explain VU variability. Also, when vaccine access is promoted at the policy level, interventions targeted to increase VU at the intrapersonal level can be more effective. In essence, the authors demonstrate that when different levels are considered together, they can capture more VU variance than a single level taken alone. This has implications for future PH campaigns which have to increase the range of their interventions, with effective integration of the SEM's multiple levels, especially through a targeted communication strategy at the interpersonal channel to increase VU (Kumar et al., 2012) .

We can thus conclude that PH interventions can increase VU if they concurrently use a methodological approach based on a universally accepted theoretical model that frames VH as not only a behavioral attitude that fits in a larger model of non-vaccination but also recognizes the influences of other, arguably influential, ecological factors.

Objective

Thus, this study seeks to classify in order of importance and then assess sociodemographic, occupational, health access and health related behavior variables' impact on the A/H1N1 pandemic 2009/2010 VU in France.

Methods

Study design

First, under R version 4.0.5 the Boruta method is used to create a classification model on H1N1/A vaccination status using the 2010 French "Enquete sur La Santé et la Protection Sociale" (ESPS). Second, we create nested multinomial logistic regression models to assess the Boruta selected important predictor's association with H1N1/A vaccination status. ESPS is a CNIL (Commission nationale de l'informatique et des libertés) approved survey study carried in two data collection waves one in March to June 2010 and one in October to December 2010 from IRDES (Institut de Recherche et Documentation en Économie de la Santé). The ESPS aims to guide public policy and evaluate the French health care system equity status by studying association of individual health condition, access to health services, access to public and private insurance and their socio-economic status (Dourgnon et al., 2012). The ESPS is a nationally representative sample of the French Population who are covered under the three principal schemes of French insurance plans "Assurance Maladie". Other regimes are not included in the sample (Dourgnon et al., 2012). This implies that transborder workers, some immigrant workers are not part of the population studied. In total, the population represents 85% of the people residing in France. Study participants who were included were asked to complete a CAPI (Computer-Assisted Personal Interviews) or CATI (Computer Assisted Telephone Interviewing) investigator led telephone survey or a face to face interview survey containing questionnaires items on: (a) "Médecin traitant": Referring physician, (b) "Revenue du foyer": Household income, (c) "Transmissions intérgenérationnelles des inegalités de santé": Intergenerational transmission of health inequalities, (d) "Etat de santé", Health condition, (e) "Loi Evin": Evin Law and (f) "Origines familiales et culturelles": Family and cultural origin. To note we did not have acces to item (f) until it was too late to include it in the final data. For more information on the sampling technique of ESPS please refer to the IRDES study methodology section. The full second wave questionnaire is also available on line. For our research purposes we merged data obtained from both waves.

Measures

Guided by the scientific literature, we identified several health behavior, socio-demographic and occupational factors in our data to estimate important attributes associated with being vaccinated against pandemic A/H1N1.

Health access and health behavior variables:

We defined a **"Health service options' use when faced with a health issue"** variable consisting of use of alternative medicine, go to the GP quickly, engage other services and use self-medication. In addition, we created a **"Contact with a health professional"** variable with 2 levels consisting of: having seen the GP and one other health professional, or not having seen any health professional. **"Smoking habits"** consists of the smokers only. **"Alcohol consumption habits"** consists of: non consumers, risky alcohol consumption, and moderate alcohol consumption. **"BMI status"** included obese, overweight and underweight groups. A **"Perceived dangerous health behavior"** variable was formed and included: individuals engaged or not in a perceived risky health behavior. We also constructed variables relative to **"Self-perceived health"** and **"Self-perceived health of teeth"** and were constituted of two levels each. **"Having no chronic disease"** variable was also created. Furthermore, **"Health coverage"**, a factor variable with three levels: universal health coverage, private complementary health coverage and no private complementary health coverage was additionally produced. A **"Referring physician"** variable was created to include having or not a referring physician. Finally, a **"Give up healthcare"** variable was created and included individuals renouncing to health services due to financial limitations and those who did not.

Socio-demographic and occupational variables:

We have included in our model the listed variables below. We start with the **"Age"** and **"Gender"** variables and **"Area of residence"** which included the north, south, east and west, southeast and southwest geographical region of France with the Paris agglomeration and its region. Concerning living conditions we have constructed a variable relative to **"Housing Contract"** which included being a tenant or not. Also, a **"Housing type"** factor was comprised of: living in a farm, in an independent house, an apartment, living in precariousness or other housing type. In addition, a feature relative to the **"Living situation in the household"** was created to assess whether one lives alone or with someone in the household and encompassed person living alone, single parent family, couple with no children or with children and other unspecified living situation in the household. Concerning education we have listed an **"Educational level"** variable with 3 levels consisting of primary, secondary and tertiary (university) education. Moreover, **"Mother educational level"** and **"Father Educational level"** are factor variables with four levels in years of education: primary, secondary, lower tertiary education equivalent to Bac +2 and higher tertiary education equivalent to above Bac +2.

For the *occupation* variables we created: **“Professional occupation”** which was structured to include 5 groups: farming, artisanal and commerce jobs in the first group, the 3 other groups included managerial, intermediate and blue collar workers with the final group including other unspecified types of jobs. **“Mother professional occupation**” and **“Father professional occupation”** were also included and structured like the “**Professional occupation** “variable above. Finally a **“Type of working contract with current professional activity”** was added and was comprised of: active but in maternal leave, active but in paternal leave, active with a CDI (Contrat a Durée Indeterminée) open ended employment contract and active without a CDI (fixed term contract). Unemployed, students, inactive because invalid, and pre-retirement and retirement were also included as groups for this variable. Lastly, **“Income”** variable is a factor variable with three levels: minimum income, no minimum income and minimum income received by one member in the household.

Classification tool: The Boruta method

The Boruta classification technique was implemented to determine the importance of the above selected socio-demographic and health behavior variables as predictors of A/H1N1 pandemic vaccine status. We used this method since it allows for a faster and increased model accuracy compared to other classification method (Speiser et al., 2019). Boruta is an all relevant feature selection wrapper algorithm built around the classical Random Forest classification method (Kursa & Rudnicki, 2010). It is described as a feature selection model for finding all predictor variables by iteratively removing features that are less relevant than random probes called “shadow variables” by a statistical test (Kursa & Rudnicki, 2010).These shadow variables are obtained by random shuffling of the original attributes and thus are created to constitute a reference level that permits to decide if the importance of an attribute is significant or not using a Z score (Kursa & Rudnicki, 2010). According to the original authors, Boruta is based on the same idea which forms the foundation of the random forest classifier, namely, that by adding randomness to the system and collecting results from the ensemble of randomized samples one can reduce the misleading impact of random fluctuations and correlations (Kursa & Rudnicki, 2010).We describe here, as the authors did, the task performed by the Boruta algorithm to assess the importance of a given attribute:

1. Create “shadow variables” by adding copies of all variables.

2. Shuffling of the added attributes to remove their correlations with the response.

3. A random forest classifier is wrapped around the “shadow variables” and the original attribute with Z scores computed and gathered.

4. A maximum Z score of shadow attribute is computed (MZSA), a hit is attributed to every predictor that scored better than MZSA.

5. A two-sided test of equality with the MZSA is performed for attribute who have undetermined importance score (Importance score close to the highest MZSA).

6. Using MZSA, all attributes which have importance significantly lower than MZSA are rejected.

7. Attributes which have importance significantly higher than MZSA are deemed “important” and are thus accepted.

8. All shadow attributes are removed.

9. Repeat the procedure until the importance score is assigned to each attribute or when random forest runs limit is reached.

We have treated in our classification model tentative attributes with “TentativeRoughFix”, an inbuilt package of the Boruta classifier which remove tentative attributes, so that we achieved only confirmed and rejected attributes.

Nested multivariate regression models

For the multivariate regression, we created six nested models using only the variables selected by the Boruta classification method. We used this approach to assess the direction of association of the selected important features with pandemic vaccination. This strategy allows for testing whether a given association changes with addition of further variables. From model 1 to model 4 we introduce the socio-demographic, occupational and health services predictors. M4 to M6 include the aforementioned variables in addition to the health behavior related predictors. Model M1 consists only of the “Contact with a professional” variable. M2 is formed of M1 in addition to “Age”, “Gender” and “Living situation in the household” variables. M3 is model M2 plus the professional occupation variables: “Professional occupation”, “Type of working contract with current professional activity”, “Father professional occupation” and “Mother professional occupation”. M4 is M3 in addition to the “Health coverage” variable. M5 consists of M4 and the “Self-perceived health”, “Self-perceived health of teeth” and “Having a chronic disease” variables. Finally, model M6 consists of model M5 plus the health behavior related variables, namely the “Smoking habits”, “Alcohol consumption habits”, “BMI status” and “Perceived dangerous health behavior”.

Model 1	• “Contact with a health professional”
Model 2	• M1 • “Age ” • “Female” • “Living situation in the household”
Model 3	• M2 • “Professional occupation” • “Type of working contract with current professional activity” • “Mother professional occupation” • “Father professional occupation ”
Model 4	• M3 • “Health coverage”
Model 5	• M4 • “Self-perceived health” • “Self-perceived health of teeth” • “Having no chronic disease”
Model 6	• M5 • “Smoking habits”, • “Alcohol consumption habits” • “BMI status” • “Perceived dangerous health behavior”

Table 1: Nested logistic regression models after selection of important features by the Boruta method

Results

Descriptive statistics:

Descriptive analysis revealed different socio-demographic, occupational and health behavior characteristics between those who were vaccinated against pandemic H1N1 and those who were not. For instance, in individuals who were vaccinated against H1N1/A the percentage of those who had received a tertiary education was 36.7% compared to 21.9% for those not vaccinated. Looking at health coverage, a variable that includes universal health coverage and private complementary insurance, the percentage among those who were vaccinated was approximately similar to those who were not vaccinated. For gender also there was no difference to note among those who were and were not vaccinated. Globally, for older individuals the percentage of being vaccinated was higher than not being vaccinated, individuals aged 65 to 76 were 13.1% vaccinated vs. 10.5 % not vaccinated. For the contact with a health professional variable, individuals who tended to see only their following physician were more frequently unvaccinated at 25.1% compared to 18.6% for vaccinated people. However, for individuals who resorted more often to use additional health services as seeing their following physician in addition to one other health professional the vaccinated percentage was 63.4% compared to 51.7% for the unvaccinated. Concerning the living situation variable those who are in a couple were more vaccinated than those who were not, 74.8% vs. 65.8%. For those who received a revenue, meaning they were not receiving the minimum social income, percentage of vaccination was 94.9% compared to 91% for the non-vaccinated. For health behavior related variables, vaccination was higher among those individuals who perceived their health as bad (36.9% vs. 29%). In addition, individuals who had a chronic disease were more vaccinated than those who did not have a chronic disease (48.4% vs. 36.4). For those who had a higher BMI, vaccination was higher for the vaccinated 26.9% as compared to the non-vaccinated 24.9%. For those who identified their alcohol consumption behavior as risky the percentage of vaccination was lower for the vaccinated 5.1% as compared to the non-vaccinated 5.4%. Those who smoked were less vaccinated, 24.4% for the not vaccinated vs. 19.8 % for the vaccinated. For those who engage in a perceived risky health behavior vaccinated and unvaccinated percentages were the same at 6%. Also, for those who had a following physician the percentage of vaccinated and unvaccinated were similar at 0.3%. Health service option's use variable showed that for those who engaged immediately the help of health professional, these individuals were more vaccinated at 17.2% than for the not vaccinated12.8%.

Boruta classification of important predictors:

Based on the Boruta classification method we have identified several socio-demographic and health behavior predictors of pandemic H1N1 vaccine status with different importance levels. We will start by identifying the significant attributes with the highest level of importance. As the first figure shows below, for health behavior related variables the highest scoring attributes are the “Alcohol consumption levels” with a 10.6 mean importance score followed by “Perceived dangerous health behavior” variable with a score of 10.0. The “Self-perceived health” and “Self-perceived health of the teeth” are also significant predictors and score high 5.9 and 8.9 respectively on the mean importance level. The “Contact with a health professional” was deemed important with a score of 7.3. Also, with an intermediate importance score is the “Universal health coverage” variable with a score of 3.9 and “Smoking habits” with a score 4.2. The importance score of 3.4 was assigned to the “Health service options’ use” variable. “Chronic disease” and “Smoking habits” scored 2.8 and 2.6 respectively. After running “TentativeRoughFix”, other health related variables deemed tentative were removed, namely the “Give up health services” variable and the “Referring physician” variable was also eliminated. For the socio-demographic and occupational variables “Age” was attributed the highest importance score among all other variables with 13.5 mean importance score. The second highest scoring attribute with a 9.0 importance level was “Professional occupation”. The “Area of Residence” followed with an importance score of 5.4. “Mother professional occupation” was selected to be an important feature with a score of 4.5 whereas “Father professional occupation” scored 4.9. “Housing type” and “Housing Contract” features were selected by the method with scores of 5.0 and 4.3 respectively. Finally “Living situation in the household” was a confirmed attribute with the lowest importance score 3.4 among the socio-demographic and occupational variable. As shown by figure 2, and after running “TentativeRoughFix”, Educational level, Parental education and gender were rejected.

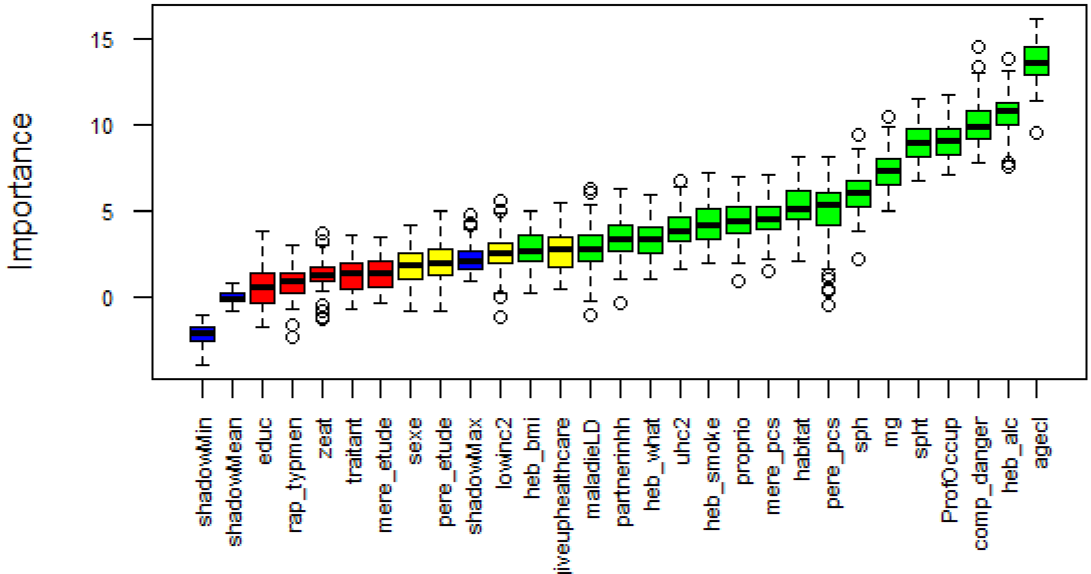

Figure 1: Predictors importance score with tentative attributes retained

To note Blue boxplots correspond to minimal, average and maximum Z score of a shadow attribute. The red, yellow and green boxplots represent Z scores of rejected, tentative and confirmed attributes respectively. Refer to the Annex for a full description of the variables and the regression results.

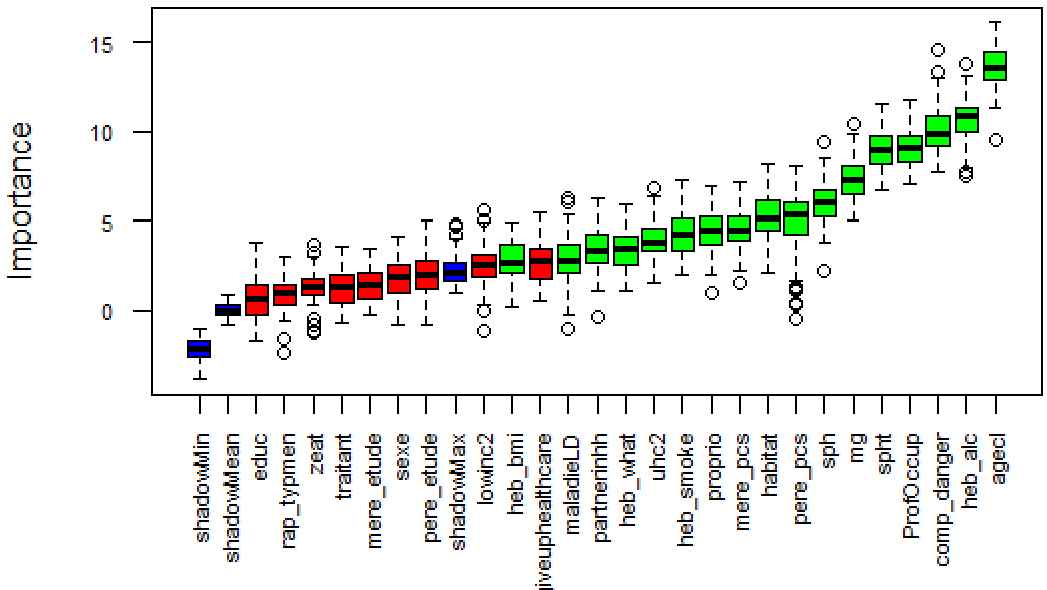

Figure 2: Final classification result, accepted and rejected attribute's importance score with tentative variables removed

Nested multivariate regression models results:

The mean Z score obtained for the "Contact with a health professional" attribute was 7.3. This variable was the first to be included in our model to predict pandemic vaccination status. It included two levels, the first being "Having contact with the GP and one other health care professional (Seen 2 HCP)" and was significant in all models of the nested multivariate logistic regression models with the highest OR observed in model 2, 1.66 with a 95% CI of 0.18. Compared to having no contact with a HCP, having a contact with a GP and another health professional increases the odds to be vaccinated by 66%. The OR observed for this level of the variable across all other models was approximately similar. The second level of the "Contact with a professional" consisted of having contact only with the GP (Seen 1 HP) and was also significant but, only in M5 and M6. These 2 models included the health behavior related variables in addition to the socio-demographic and occupational ones. The OR for the two characteristics of this predictor (Seen 1 HP or Seen 2 HP) are similar at around 1.6. Seeing at least one HP increases the odds of vaccination compared to seeing no HP. The gender attribute was deemed not significant by the Boruta method. Age classes showed a U shaped correlation with the youngest and elderly more likely to vaccinate and the 35 to 55group, as likely as the youngest age group to get the pandemic vaccine. The 28 to 39 years of age and the 54 to 98 years of age intervals results were significant and positively associated with being vaccinated across all models. The strongest association observed for an age class, was obtained for the 76 to 98 age class in model 3 with an OR of 4.68. So the odds of getting vaccinated was 3.68 times higher for this age class compared to the 18-24 years old. People living alone were less likely to be vaccinated than those living with a partner with OR at 0.73 in model M5. Though this result was significant in models M2, M3 and M5 it was insignificant in M4 and M6. “Mother professional occupation” attribute obtained a mean Z score of 4.54 in the Boruta classification. However, this predictor including all of its levels, was deemed insignificant in models M1 to M6 even when the reference group was changed to blue collar workers. In contrast, if the father occupied a high managerial position the observed ORs were significant across model 3 to 6 with the observed OR at around 2.1 showing that this category is approximately 2 times more likely to be vaccinated compared to blue collar workers. For other father professional occupations, results obtained were not significant. Results for “Housing contract type” and “Housing type” were insignificant across all models. For other housing contracts and housing types, results obtained were insignificant compared to a reference of having a house and being the owner of the house. For universal health coverage as well as self-perceived health and self-perceived health of the teeth attributes deemed important in the classification method, results obtained in the models were not significant. Having a chronic disease was associated with an OR of 1.37 and a CI of 0.14 in model M5 and M6. For health behavior related variables namely, the “Health service options’ use when faced with a health issue”, “Alcohol consumption habits”, “Perceived dangerous health behavior” and “BMI status” were not significant at all levels in our models even when compared to reference levels. To note, in the Boruta classification method, the “Alcohol consumption habits” and “Perceived dangerous health behavior” scored high with importance score of 10.63 and 3.30 respectively. In contrast, for “Smoking habits” being a smoker was associated with an OR of 0.78 and a CI of 0.09.

Discussion

The Boruta classification method followed by nested multivariate regression models

VU is a key variable to consider in vaccination strategies. We opted, in this work, for a method that finds the best predictors of VU. We have thus selected the Boruta classification method whose algorithm can select the minimal number of relevant features and produce the best possible classification model (Speiser et al., 2019). In addition, Boruta solves the problem of identifying all relevant features as opposed to finding only non-redundant attributes, a process that is commonly performed by other classification models. Furthermore, Boruta does not require as much computational power and time to produce results as compared to other methods. In effect, Boruta confirmed after just approximately 1 minute on a standard computer, 7 important attributes in our data set. In total, this method confirmed 17 of our predictors, rejected 5 and deemed 4 attributes as tentative in roughly 7 minutes only. Since Boruta cannot identify the direction of association of important predictors with pandemic vaccination, we proceeded to use the nested multivariate regression models. To note reference levels were set in our logistic models

Health access variables: the Role of "Health services options' use", "Health coverage", "Contact with a health professional" "Give up health care" and "Following physician" variables

Our approach identified both "Health services options' use when faced with a health issue", "Health coverage" and "Contact with a health professional" as important determinants of pandemic H1N1 VU and were thus included in the regression models. The "Give up health care" was deemed unimportant and was excluded. However, we found that the "Health coverage" with its associated levels in addition to "Health services options' use when faced with a health issue" and its levels not statistically significant in the nested multivariate regression models. In contrast, "Contact with a health professional" was significant. Indeed, in the French context, on the one hand, there was an effort from the sanitary authorities to convey the risk severity of the disease with alarming media campaigns on hospital admissions, ICU occupancy and increasing fatality rates. Yet, and on the other hand, there was little communication aimed at reassuring about vaccine safety, which constituted the major barrier to vaccination for the general public. Accompanied with little personal experience, which did not confirm the threat of the pandemic, as reported in Schwarzinger et.al (2010) study, VA was at its lowest among other EU countries at 17 % for the general population with an achieved coverage of 7.1 % in the population aged 18 to 69. In fact, and according to the author, getting a positive advice on the safety of the vaccine from the PCP was a major determinant for the acceptability of A/H1N1 pandemic vaccination. This finding is paralleled in our result that shows a 66 % increase in the odds of pandemic vaccination in model 6 if one has seen 2 HCP. Nevertheless, a small proportion of the population was advised to take the vaccine by their PCP. Furthermore, among HCP who constituted the first priority group to access pandemic vaccines, VU was low at 10.9%. Since the PCP were responsible for administering all scheduled vaccines and especially the seasonal flu vaccine and were excluded from the mass vaccination campaign, we can argue that this decision by the French sanitary authorities as well as the low acceptability of the pandemic vaccine by the PCP might have added to public refusal of the pandemic vaccine. PCP would have conveyed best information on pandemic vaccine safety profile and must be delegated further responsibilities in future public health campaigns to increase pandemic VU.

Health behavior variables

Chronic disease attribute was deemed important in our model and was significant in regression models M5 and M6. As evidenced in the literature, having a chronic disease meant that an individual was included in a priority group to receive pandemic vaccine and thus, in individuals who presented chronic diseases VU was higher. Self-perceived health in relation to pandemic vaccination if defined in terms of the classical theories of health behavior, as the Extended Theory of Planned Behavior and the Health Belief Model, encompasses many variables that we did not examine in our method. In fact, our variable did not include the classical characteristics of perceived control and self-efficacy on one's own health, neither did it incorporate variables related to perceived susceptibility, barriers, severity or benefits related to pandemic vaccine. Indeed, our variable was a subjective answer on one's perceived health and hence presents limitation in its interpretation. In short, our classification method deemed both self-perceived health and self-perceived health of the teeth and perceived risky health behavior as strong predictors of pandemic VU classification. However, these variables were not significant in the regression method because the variables we constructed in our model were most likely too general to answer the many layers and contrasts relative to self-perceived health in relation pandemic VU. For other health related variables such as smoking and alcohol consumption and BMI, our classification method deemed these factors as important. However, when tested in the regression models we did not obtain any statistically significant results, except for the smokers group in which the percentage of VU was 12% less relative to non-smokers. We hypothesize that for these variables, a better approach to test their relative association with pandemic VU, would have been to adopt a strategy in which these health behavior were defined in relation to theories of health behaviors.

The Social and Demographic factors role

Age:

The strongest predictor identified in our method of classification of pandemic H1N1 VU predictors was age in class. Nevertheless, there are conflicting results in the scientific literature on the role of age and its association with pandemic VU. Furthermore, attitudinal factors proper to an age class and their relative association with being vaccinated against pandemic H1N1 have also inconsistent results in the literature. In fact, there are studies reporting that the general public in the UK, France and the USA have found that older people are more likely to intend to be vaccinated (Bish et al., 2011). It is also assumed that for the younger age groups, who were not initially part of priority groups, there was a reported attitudinal difference in regards to pandemic vaccination, with younger people being more likely to report attitudes "expressing degrees of 'passivity' such as 'no one told me to be vaccinated'; 'I didn't get round to it', which may have affected their behavior However, a study by Vaux et.al (2011) on H1N1 pandemic vaccination found no association between H1N1 pandemic vaccination and age.

Gender:

Gender as a feature was not selected by our method. Like age, sex as a predictor of pandemic H1N1 VU has differing results. Bish et al. (2011) states that amongst both the general population and health professionals, men were more likely to intend to be vaccinated and to be vaccinated than women. However, as demonstrated by Vaux et al. (2011), depending on the population studied and its socio-cultural context, different studies can reveal varying results on the association of gender with pandemic VU.

Education, occupation and income level:

For our classification method we included three factors relative to education in order to better capture its role as a potential predictor of pandemic VU. Both father's and mother's education and the general education factor were thus tested to assess the role of education. However, all three factors were rejected as important features of pandemic A/H1N1 VU and were thus not included in the regression models. In the same logic, we created three variables related to occupation. The type of working contract in addition to the professional occupation of the father and of the mother were selected as strong features with high importance scores in our classification method. Interestingly, when tested in the regression models, we obtained statistically significant results only for the high managerial position of the father profession. Indeed, there is mixed evidence in the literature on the role of education and employment status. A study by Schwarzinger et.al (2010) with the general population of France showed that those with the lowest and the highest education level who work in a non-office job were most likely vaccinated or had intended to be compared to other employment and educational level group. In Vaux et. al (2011) study, those who lived in a household where the head of the family is a university degree holder and occupied a managerial position, were more likely to have been vaccinated as compared to other groups. To note, low income was rejected from our model. This is consistent with the fact that when issues of access to health services, namely geographical access and with free vaccination offered, low income can be eliminated as a barrier of vaccination. This might hold true, even with existing disparities of access to health services within one national territory, for developed countries where the vaccine was offered for free. However, income level remains a major barrier to vaccination in developing and underdeveloped countries.

Household related characteristics variables:

The housing type, housing contract, and the size of the household attributes were deemed important by our classification method. In model M5 living alone was associated with a 17% decrease in pandemic VU. However, there is no clear path to statistical significance in regards to household characteristics. Vaux et al. (2011), Sypsa et al. (2012), Rubin et al. (2012), and Schwarzinger et al. (2010) found a positive association when other studies by Horney et al. (2011) and Zijtregtop et al. (2011) did not demonstrate any significance for these variables and pandemic VU.

Area of residence: Urban vs. rural

Across the scientific literature there is heterogeneity of association between geographical region of residence and pandemic A/H1N1VU. Most of the studies only reports coverage difference in diverse regions without further investigating the association or the difference seen. A study by Ozer et al. (2012) did not find any statistical significance between region of residence and pandemic vaccination. In our model however, this variable has not been retained as an important variable and hence was not tested in the regression models.

Strengths and Limitations

Our data set was composed of a large number of observations and features. We have thus selected the Boruta method which is a convenient algorithm that can select all relevant variables needed to build a classification model. It consequently allowed us to define the variables that we want to observe. One major strength of the Boruta algorithm is that is a wrapper algorithm built around a random forest. It permits to detect the importance of each feature by providing its associated Z score. Though this score is not directly a statistical measure of the significance of the feature, the Boruta algorithm compare it to a random permutations of variables or a selection of variables and test if it is higher than the scores achieved from the random shadow variables. By adopting this approach, the misleading impact of randomness in the original sample is reduced. A limitation of the Boruta classification is that its algorithm does not treat collinearity while selecting important variables. However, we have taken precautions to limit in our models highly correlated variables. In addition, no imputation method for missing variables was done. Nevertheless, in our data set we did not have a lot of missing variables. Yet, this could have been fixed if we had used the “party” package instead of the classical “RandomForest” package. Finally, this classification method provides only ranking for features and does not allow for an estimate of the direction of association of a given feature. We have thus completed our method with nested multinomial logistic regression models to answer this problematic. Our study method is based on face to face interviewers led, or computer based questionnaires. It thus shares the general limitations of results based on this type of methodologies. Hence our study’s design might have yielded higher social desirability bias. Because our results are based on a cross sectional observational study design no causal inference can be concluded on VU.

Conclusion

We have described a classification method using the performing Boruta method that can conveniently rank important features related to pandemic VU. We have completed our approach by nested multivariate regression models to test the direction of association of important attributes selected by our classifier. This approach can be implemented in future pandemic influenza waves in order to increase VA and VU among the population.

References

Bish, A., Yardley, L., Nicoll, A., & Michie, S. (2011). Factors associated with uptake of vaccination against pandemic influenza: A systematic review. *Vaccine*, *29*(38), 6472–6484. https://doi.org/10.1016/j.vaccine.2011.06.107

Black, S., Eskola, J., Siegrist, C., Halsey, N., Macdonald, N., Miller, E., Andrews, N., Stowe, J., Salmon, D., Vannice, K., Izurieta, S., Akhtar, A., Gold, M., Oselka, G., Zuber, P., & Pfeifer, D. (2010). *influenza vaccines*. *374*(9707), 2115–2122. https://doi.org/10.1016/S0140-6736(09)61877-8.Importance

Bocquier, A., Fressard, L., Cortaredona, S., Zaytseva, A., Ward, J., Gautier, A., Peretti-Watel, P., & Verger, P. (2018). Social differentiation of vaccine hesitancy among French parents and the mediating role of trust and commitment to health: A nationwide cross-sectional study. *Vaccine*, *36*(50), 7666–7673. https://doi.org/10.1016/j.vaccine.2018.10.085

Brien, S., Kwong, J. C., & Buckeridge, D. L. (2012). The determinants of 2009 pandemic A/H1N1 influenza vaccination: A systematic review. *Vaccine*, *30*(7), 1255–1264. https://doi.org/10.1016/j.vaccine.2011.12.089

Desenclos, J., Gallay, A., & Benrekassa, J. (2010). Bulletin épidémiologique hebdomadaire-Numéro thématique - Epidémie de grippe A(H1N1)2009": premiers éléments de bilan en France. BEH. *InVS*. http://invs.santepubliquefrance.fr//beh/2010/24_25_26/beh_24_25_26_2010.pdf

Dourgnon, P., Guillaume, S., & Rochereau, T. (2012). Enquête sur la santé et la protection sociale 2010, Les Rapports de l'IRDES, n° 1886, juillet. In *Irdes* (Vol. 553). http://www.cmh.ens.fr/greco/enquetes/XML/lil-0454.xml

Dubé, E., Laberge, C., Guay, M., Bramadat, P., Roy, R., & Bettinger, J. (2013). Vaccine hesitancy: An overview. *Human Vaccines and Immunotherapeutics*, *9*(8), 1763–1773. https://doi.org/10.4161/hv.24657

E., S., & D., T. (2010). Knowledge, attitudes and anxiety towards influenza A/H1N1 vaccination of healthcare workers in Turkey. *BMC Infectious Diseases*, *10*, 281.

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed9&NEWS=N&AN=20863386

Fabry, P., Gagneur, A., & Pasquier, J. C. (2011). Determinants of A (H1N1) vaccination: Cross-sectional study in a population of pregnant women in Quebec. *Vaccine*, *29*(9), 1824–1829. https://doi.org/10.1016/j.vaccine.2010.12.109

Ferguson, C. D., Ferguson, T. E., Golledge, J., & McBride, W. J. H. (2010). Pandemic influenza vaccination: Will the health care system take its own medicine? *Australian Journal of Rural Health*, *18*(4), 137–142. https://doi.org/10.1111/j.1440-1584.2010.01148.x

HCSP. (2009). Haut Conseil de la santé publique 157. *Haut*, *6409472916*, 1–4.

Henrich, N., & Holmes, B. (2008). The public's acceptance of novel vaccines during a pandemic: A focus group study and its application to influenza H1N1. *Emerging Health Threats Journal*, *2*. https://doi.org/10.3134/ehtj.09.008

Inpes, & ministère de la santé. (2009). *Lancement de la campagne vaccinale contre la grippe A(H1N1) dans les centres de vaccination*. 1–20. http://inpes.santepubliquefrance.fr/70000/dp/09/dp091109.pdf

Kumar, S., Quinn, S. C., Kim, K. H., Musa, D., Hilyard, K. M., & Freimuth, V. S. (2012). The social ecological model as a framework for determinants of 2009 H1N1 influenza vaccine uptake in the United States. *Health Education and Behavior*, *39*(2), 229–243. https://doi.org/10.1177/1090198111415105

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, *36*(11), 1–13. https://doi.org/10.18637/jss.v036.i11

Larson, H. J., Cooper, L. Z., Eskola, J., Katz, S. L., & Ratzan, S. (2011). Addressing the vaccine confidence gap. *The Lancet*, *378*(9790), 526–535. https://doi.org/10.1016/S0140-6736(11)60678-8

Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M. D., & Paterson, P. (2014). Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: A systematic review of published literature, 2007-2012. *Vaccine*, *32*(19), 2150–2159. https://doi.org/10.1016/j.vaccine.2014.01.081

MacDonald, N. E., Eskola, J., Liang, X., Chaudhuri, M., Dube, E., Gellin, B., Goldstein, S., Larson, H., Manzo, M. L., Reingold, A., Tshering, K., Zhou, Y., Duclos, P., Guirguis, S., Hickler, B., & Schuster, M. (2015). Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, *33*(34), 4161–4164. https://doi.org/10.1016/j.vaccine.2015.04.036

Peretti-Watel, P., Ward, J. K., Vergelys, C., Bocquier, A., Raude, J., & Verger, P. (2019). 'I Think I Made The Right Decision … I Hope I'm Not Wrong'. Vaccine hesitancy, commitment and trust among parents of young children. *Sociology of Health and Illness*, *41*(6), 1192–1206. https://doi.org/10.1111/1467-9566.12902

Prematunge, C., Corace, K., McCarthy, A., Nair, R. C., Pugsley, R., & Garber, G. (2012). Factors influencing pandemic influenza vaccination of healthcare workers-A systematic review. *Vaccine*, *30*(32), 4733–4743. https://doi.org/10.1016/j.vaccine.2012.05.018

SAGE Working Group on Vaccine Hesitancy. (2014). *Report of the Sage Working Group on*. *October*, 64. https://www.who.int/immunization/sage/meetings/2014/october/1_Report_WORKING_GROUP_vaccine_hesitancy_final.pdf

Schmid, P., Rauber, D., Betsch, C., Lidolt, G., & Denker, M. L. (2017). Barriers of influenza vaccination intention and behavior - A systematic review of influenza vaccine hesitancy, 2005-2016. In *PLoS ONE* (Vol. 12, Issue 1). https://doi.org/10.1371/journal.pone.0170550

Schwarzinger, M., Flicoteaux, R., Cortarenoda, S., Obadia, Y., & Moatti, J. P. (2010). Low acceptability of A/H1N1 pandemic vaccination in french adult population: Did public health policy fuel public dissonance? *PLoS ONE*, *5*(4). https://doi.org/10.1371/journal.pone.0010199

Shaham, A., Chodick, G., Shalev, V., & Yamin, D. (2020). Personal and social patterns predict influenza vaccination decision. *BMC Public Health*, *20*(1), 1–12. https://doi.org/10.1186/s12889-020-8327-3

Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, *134*, 93–101. https://doi.org/10.1016/j.eswa.2019.05.028

States, U., & Republic, D. (2009). New influenza A(H1N1) virus infections in France, April - May 2009. *Euro Surveillance : Bulletin Européen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *14*(21), 20–22. https://doi.org/10.2807/ese.14.21.19221-en

Tout, V. (2010). *Suivi de Pharmacovigilance des vaccins grippaux A (H1N1)*. 1–6.

Vaux, S., Van Cauteren, D., Guthmann, J. P., Le Strat, Y., Vaillant, V., De Valk, H., & Lévy-Bruhl, D. (2011). Influenza vaccination coverage against seasonal and pandemic influenza and their determinants in France: A cross-sectional survey. *BMC Public Health*, *11*, 1–9. https://doi.org/10.1186/1471-2458-11-30

Weston, D., Blackburn, R., Potts, H. W. W., & Hayward, A. C. (2017). Predictors of self and parental vaccination decisions in England during the 2009 H1N1 pandemic: Analysis of the Flu Watch pandemic cohort data. *Vaccine*, *35*(31), 3875–3882. https://doi.org/10.1016/j.vaccine.2017.05.061

WHO. (2009). *Transcript of statement by Margaret Chan, Director-General of the World Health Organization*. *11 June*, 1–8. https://www.who.int/mediacentre/influenzaAH1N1_presstranscript_20090611.pdf?ua=1

WHO. (2012). *Report of the WHO Pandemic Influenza A(H1N1) Vaccine Deployment Initiative*. 1–52.

Wong, L. P., & Sam, I. C. (2010). Factors influencing the uptake of 2009 H1N1 influenza vaccine in a multiethnic Asian population. *Vaccine*, *28*(28), 4499–4505. https://doi.org/10.1016/j.vaccine.2010.04.043

Wu, J. T., Ma, E. S. K., Lee, C. K., Chu, D. K. W., Ho, P., Shen, A. L., Ho, A., Hung, I. F. N., Riley, S., Ho, L. M., Lin, C. K., Tsang, T., Lo, S., Lau, Y., Leung, G. M., & Cowling, B. J. (2011). *NIH Public Access*. *51*(10), 1184–1191. https://doi.org/10.1086/656740.The

Young, M. E., Norman, G. R., & Humphreys, K. R. (2008). Medicine in the popular press: The influence of the media on perceptions of disease. *PLoS ONE*, *3*(10), 1–7. https://doi.org/10.1371/journal.pone.0003552

ANNEX:

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Contact with a health professional: Seen the GP and one other health professional	1.65***	1.66***	1.57***	1.58***	1.49***	1.47***
	(0.18)	(0.18)	(0.17)	(0.18)	(0.17)	(0.17)
Contact with a health professional: Seen the GP only	1.04	1.14	1.11	1.11	1.53*	1.56**
	(0.14)	(0.16)	(0.15)	(0.15)	(0.28)	(0.32)
Gender: Female		1.00	0.97	0.97	0.99	1.01
		(0.09)	(0.09)	(0.09)	(0.09)	(0.10)
Age [28,34]		2.20***	2.02**	2.03**	2.01**	2.03**
		(0.46)	(0.44)	(0.45)	(0.44)	(0.45)
Age [34,39]		2.28***	2.07**	2.09***	2.04**	2.01**

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
		(0.47)	(0.46)	(0.46)	(0.45)	(0.45)
Age [39,44]		1.38	1.29	1.30	1.26	1.25
		(0.31)	(0.31)	(0.31)	(0.30)	(0.30)
Age [44,49]		1.17	1.20	1.20	1.14	1.16
		(0.27)	(0.30)	(0.30)	(0.29)	(0.29)
Age [49,54]		1.39	1.51	1.52	1.41	1.39
		(0.31)	(0.37)	(0.37)	(0.34)	(0.34)
Age [54,59]		1.99**	2.24***	2.26***	2.05**	2.00**
		(0.43)	(0.54)	(0.54)	(0.50)	(0.49)
Age [59,65]		1.84**	2.67***	2.67***	2.41**	2.31**
		(0.40)	(0.75)	(0.75)	(0.69)	(0.67)
Age [65,76]		2.15***	3.56***	3.54***	3.10***	2.91***
		(0.45)	(1.09)	(1.09)	(0.97)	(0.92)
Age [76,98]		2.80***	4.68***	4.63***	3.90***	3.63***
		(0.60)	(1.43)	(1.43)	(1.22)	(1.16)
Living situation in the household: Alone		0.66**	0.73*	0.74	0.73*	0.74

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
		(0.10)	(0.12)	(0.12)	(0.12)	(0.12)
Mother professional occupation: Farming, artisanal and commerce jobs			1.45	1.44	1.47	1.49
			(0.41)	(0.40)	(0.41)	(0.42)
Mother professional occupation: Managerial			1.76	1.75	1.79	1.80
			(0.53)	(0.53)	(0.54)	(0.55)
Mother professional occupation: Intermediate employee			1.18	1.18	1.19	1.18
			(0.33)	(0.33)	(0.33)	(0.33)
Mother professional occupation: Blue collar worker			1.03	1.03	1.02	1.03
			(0.25)	(0.25)	(0.25)	(0.25)
Father professional occupation: Farmer			0.94	0.94	0.94	0.93
			(0.17)	(0.17)	(0.18)	(0.17)
Father professional occupation: Artisanal and commerce jobs			1.13	1.13	1.12	1.10
			(0.19)	(0.19)	(0.19)	(0.18)

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Father professional occupation Managerial			2.09***	2.09***	2.10***	2.09***
			(0.27)	(0.27)	(0.27)	(0.27)
Father professional occupation: Intermediate employee			1.40*	1.40*	1.38*	1.36
			(0.22)	(0.22)	(0.22)	(0.21)
Type of working contract with current professional activity: Active but in maternal leave			1.60	1.56	1.63	1.62
			(0.74)	(0.73)	(0.76)	(0.76)
Type of working contract with current professional activity: Active but in paternal leave			3.55***	3.47***	3.61***	3.64***
			(1.17)	(1.16)	(1.21)	(1.23)
Type of working contract with current professional activity: Active with a CDI			1.22	1.19	1.23	1.19
			(0.24)	(0.25)	(0.26)	(0.25)
Type of working contract with current professional activity: Active without a CDI			1.08	1.06	1.10	1.08

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
			(0.24)	(0.24)	(0.25)	(0.25)
Type of working contract with current professional activity: Student			1.22	1.17	1.23	1.16
			(0.59)	(0.57)	(0.60)	(0.57)
Type of working contract with current professional activity: Inactive because invalid			2.53**	2.48**	2.04*	2.05*
			(0.79)	(0.78)	(0.65)	(0.66)
Type of working contract with current professional activity: Pre-retirement			0.70	0.69	0.66	0.68
			(0.45)	(0.45)	(0.43)	(0.44)
Type of working contract with current professional activity: Retired			0.88	0.87	0.85	0.83
			(0.24)	(0.24)	(0.23)	(0.23)
Housing type: Farm house/apartment/ independent house			1.06	1.05	1.04	1.05
			(0.13)	(0.13)	(0.13)	(0.13)
Housing type: Town house			1.02	1.02	1.00	1.01

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
			(0.14)	(0.14)	(0.14)	(0.14)
Housing type: Precarious housing			1.75	1.64	1.51	1.56
			(1.91)	(1.82)	(1.69)	(1.76)
Housing type: other			1.03	1.03	1.05	1.06
			(0.41)	(0.41)	(0.42)	(0.42)
Housing Contract: Owner			1.29	1.28	1.30	1.25
			(0.19)	(0.19)	(0.19)	(0.19)
Housing contract: Tenant			1.13	1.14	1.14	1.13
			(0.19)	(0.19)	(0.19)	(0.19)
Health coverage: No private complementary health coverage				0.59	0.60	0.51
				(0.40)	(0.41)	(0.35)
Health coverage: Private complementary health coverage				0.60	0.64	0.53
				(0.39)	(0.41)	(0.34)
Health coverage: Universal health coverage				0.55	0.57	0.50
				(0.37)	(0.38)	(0.33)

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Self-perceived health: Good					1.21	0.98
					(0.42)	(0.39)
Self-perceived health: Bad					1.44	1.22
					(0.51)	(0.49)
Having a chronic disease					1.37**	1.37**
					(0.14)	(0.14)
Self-perceived health teeth: Good					1.24	1.09
					(0.42)	(0.38)
Self-perceived health teeth: Bad					1.11	1.00
					(0.37)	(0.35)
BMI: Obese						0.84
						(0.12)
BMI: Overweight						0.93
						(0.10)
BMI: Underweight						0.84
						(0.26)

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Alcohol consumption habits: Non consumer						0.90
						(0.20)
Alcohol consumption habits: Risky consumption						0.91
						(0.25)
Alcohol consumption habits: Moderate consumption						1.13
						(0.26)
Smoking habits: Smoker						0.78*
						(0.09)
Health service options' use when faced with a health issue: Use of alternative medicine						0.65
						(0.18)
Health service options' use when faced with a health issue: Go to the GP quickly						1.28
						(0.17)

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Health service options' use when faced with a health issue: Engage other services						0.88
						(0.13)
Health service options' use when faced with a health issue: Use of self-medication						1.01
						(0.12)
Perceived dangerous health behavior: No						0.85
						(0.22)
Perceived dangerous health behavior: Yes						0.74
						(0.22)
Num. obs.	6541	6541	6541	6541	6541	6541
Log Likelihood	-2103.18	-2066.79	-2015.95	-2015.56	-2004.89	-1993.87
Deviance	4206.36	4133.57	4031.91	4031.13	4009.78	3987.75
AIC	4212.36	4163.57	4127.91	4133.13	4121.78	4129.75
BIC	4232.72	4265.36	4453.63	4479.21	4501.78	4611.54

Nested Statistical models Results

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$