UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN

Statistica Applicata

CICLO XXVII

COORDINATORE Prof. Fabio Corradi

BEHAVIOUR TOWARD IMMUNIZATION

Settore Scientifico Disciplinare SECS/S-04

Dottorando
Dorgali Maria Veronica

M. Veronica Dorgali
(firma)

Tutore
Prof. Pietro Angelo Manfredi

(firma)

Co-Tutore
Prof.ssa Petrucci Alessandra

(firma)

Coordinatore
Prof. Corradi Fabio

(firma)

Anni 2012/2016 *(di inizio e fine corso)*

"For a minute, there
I lost myself, I lost myself
Phew, for a minute there
I lost myself, I lost myself"
Radiohead-Karma Police

"...pagina senza testo e punteggiatura
tu la chiami bianca io la chiamo paura,
e l'ho provata uscendone accecato non l'ho cercata
è lei che mi ha trovato..."
Fabri Fibra-Panico

"All statistics have outliers."
Nenia Campbell, Terrorscape

Table of contents

Introduction

The focus of this thesis in on individual behaviour in regard of immunization, and on the use of psychological-cognitive models, and the related statistical implementation, in order to measure behaviour. The thesis is subdivided into two parts.

The *first part (*in collaboration with a group from the university of Antwerp) dealt with social orientation and vaccination attitude. A representative survey of the population in Flanders, Belgium (N= 1050), was collected in order to investigate whether respondents' attitude to vaccination was associated with their basic disposition toward other community members or the society as a whole, as measured by the Triandis and Gelfand social orientation scale. Singelis et al. and Triandis et al. developed an influential framework to study social orientation. This framework distinguishes two basic relational dimensions: a collectivism/individualism axis that reflects someone's sense of social cohesion and his/her willingness to prioritize common goals over personal ones, and a horizontal/vertical axis that indicates to what degree an individual expects equality or accepts inequality in social relationships. This generates the 4-way typology horizontal and vertical individualism and collectivism (HI, VI, HC and VC). In the survey, the Triandis and Gelfand scale to determine an individual's position on the four dimensions of social orientation was presented. It consists of sixteen value judgments on which every respondent is asked to express his/her agreement on a five-point likert scale ranging from "strongly disagree" to "strongly agree". All 4 types of social orientation are characterized by 4 statements a priori designated to reflect one specific type.

To determine the respondents' general attitude toward vaccination, we asked them to indicate their agreement with the following statement: '*If a vaccine exists for a certain disease, then vaccination is usually a good way to protect someone against this disease'*.

Through a logistic regression model, it was found that sceptics individuals have a different social orientation compared to non-sceptics, i.e. sceptics where those that declared a negative attitude towards vaccination, even after controlling for available covariates such as age, gender, educational attainment, professional group, personal experience with severe illness, experience with severe illness in the family, origin of mother and father, experience with travel vaccination, experience as health care worker, family size, vertical individualism and vertical collectivism. More specifically, vaccine sceptics scored significantly lower on both horizontal individualism and horizontal collectivism, indicating a lower disposition to see other people as equals.

This study identified social orientation as an important determinant of attitude toward vaccination. The results show that the values of horizontal individualism and horizontal collectivism are associated with whether people take a positive stand toward vaccines. These findings are notable for two reasons. First, given that social orientation was identified as a more important determinant than other more commonly reported socio-demographic variables (such as e.g. education or professional group), our findings underscore the value of investigating the psychological determinants of vaccine scepticism. Second, vaccination is a solidaristic intervention that benefits both the vaccinated individual and the wider community, whereas the risks associated with vaccination remain strictly private. Therefore, one might intuitively expect that people's attitude toward vaccination will be determined by the individualistic vs. collectivistic social orientation of individuals. However, our findings demonstrate that it is not so much individualism or collectivism that is of importance, but particularly the degree to which people value equality in their social relations. The more people see others as equals, the more positive they stand toward vaccination. Findings like these may help to improve the design of effective communication strategies for vaccines. In several countries, marketing research has shown that advertisements are more persuasive when their appeal is matched to the social orientation of their targeted audience.

The *second part*, which is the core of this thesis, has dealt with the complete development of a field study, departing from the formulation– with my supervisor - of a basic scientific question, to the design of the questionnaire, the collection of field data, up to the analysis of the collected data.

The basic scientific question dealt with preventive behaviour in relation to cervical cancer, namely why vaccine coverage against Human Papilloma Virus (HPV) is so low in many countries including Italy, particularly which behavioural factors are at work to cause the basic rational model of immunization to fail so markedly. By "basic rational model" we mean a simple decision model of agents comparing the direct perceived benefits and costs of immunization. As perceived benefits - protection from the most deadly disease, namely cancer - should be extremely high, and costs – mainly risks of possible side effects of immunization - low due to the age at immunization (i.e., parents' worries about potential side effects should be strongly mitigated compared to infants' immunization) one would expect high coverage easy to reach.

Vaccination against HPV was included in the Italian national immunization programme in 2007. The vaccine is offered free of charge to 11-year-old girls who represent the primary target group. Additionally, a catch-up campaign has been promoted in some regions to extend the active offer to older female age groups (secondary target group). In Tuscany, the vaccine is

offered free of charge to girls aged 12-18 yr. The vaccination is actively offered -with a letter sent by the local health unit- only to 12 and 16-year-old girls while a parent's application is requested for all other young women. Currently, 10 years after the initiation of the programme, HPV vaccination coverage in the two target populations is substantially lower than WHO targets (95%).

To determine which further factors, beyond the rational ones might be important for decision makers, a different theoretical approach was considered. In the last twenty years, behavioural theories or social cognition models as they are sometimes called, have become increasingly important and popular in the field of health. Research guided by psychosocial theories have added to our understanding how cognitive and social determinants might affect healthy or risky practices. From this perspective, a deepen analysis of the social, cultural and cognitive aspects of health behaviours is fundamental. Prior to the questionnaire definition a large methodological review of the existing literature was conducted. The concepts of social norms. perceived susceptibility to illness and attitudes/beliefs towards a certain health behaviour were examined in depth. These are core constructs of the most important health behaviour theories often used to explain and predict individual habits. Social norms are the customary rules that govern behaviour in groups and societies. Many researches demonstrated the important role of social norms as "social proof as information toward right living" (Cialdini &Trost, 1998) and how humans tend to rely on "social reality" when the appropriate behaviour is unclear (Festinger,1954). For this reason, social norms may have a key role in explaining vaccination choices. Perceived susceptibility is an individual's belief about the likelihood of a health threat's occurrence or the likelihood of developing a health problem. Existing research suggest that disease risk perceptions are a critical determinant of health behaviour but it is not clear how parents perceive the susceptibility of their young daughter (11 or 12 years old) for a sexual transmitted disease. Attitudes are the way a person expresses or applies is an internal feeling (beliefs) about a certain behaviour and may explain parents' decisions about HPV vaccination. For this purpose, constructs belonging to three of the most well-known theories, i.e. the *health belief model* (Becker, 1974; Rosenstock, 1974), the *theory of planned behaviour* and *the social influence theory* (Cialdini & Trost, 1998), were implemented in a newly research instrument. Published measured and original question as well as cultural and social aspects of respondents' environment were balanced to create a proper survey questionnaire. It was self-administered among young adult women aged 19-22 attending their bachelor at the universities of Pisa and Florence with the purpose of measuring the determinants of their vaccination status as well as their attitudes, beliefs and knowledge towards HPV, HPV vaccination and cervical cancer screening practices (Pap-test). After the age of 18 girls become more responsible in managing different aspects of their life, including

their health status. The rate at which adolescents' capacity for medical decision-making evolves depends on the complexity of their medical condition, their understanding or experience of it, and other contextual factors such as culture, family and belief systems. Evidence suggests that already at 14-15-year olds girls have similar capacities to adults (Weithorn and Campbell,1982). Autonomous decision-making processes should imply a greater lever of consciousness and knowledge, especially for choices related to the individual well-being. Concerning this, the research aims to investigate how much consciously respondents are involved in their health protection.

The study is based on a convenience sample collected between December and June 2016. A mixed-mode approach, i.e. online and paper-based questionnaire, was considered in order to reach the largest possible number of respondents. A total of 491 students participated in the study. Fifty-nine questionnaires were incomplete (more than 50% of the items were missed) and excluded from the final analysis. The majority of the respondents (69%) are in the age range of 19-22 years old and reside in Tuscany (72%). The 6% (25) of interviewed girls do not know the HPV and the 9% (35) have never heard about the associated vaccine. 71% (262) of the students were vaccinated against HPV, 23%(85) were unvaccinated and 6%(21) did not remember their vaccination status. The mean vaccination age was 14 years [sd: 1.9313; 95% C.I(14.05544;14.55401)]. Among vaccinated, 48% got the vaccine because their parents decided it. Non-vaccinated seemed more independently in their vaccination decision, in fact 30% (24) has taken the decision about the vaccination autonomously and 42%(33) with their parents. Among the latter, 34% and 39% respectively, declare that they were "very" or "extremely" likely to take the vaccine if it would be offered again from the national health system. HPV and Pap-smear knowledge were assessed by 14 items (yes/no/don't know) and 4 items (yes/no/don't know), respectively. Knowledge questions highlight a slightly better performance for unvaccinated girls while the information level considerably changes across faculties. Responses quality varies according to the question level, the more detailed is the information requested, the lower is the level of accuracy.

Standard and penalized logistic regression models were considered to identify the most significant factors associated with the realized immunization choice. Since the number of independent variables in the original model was high, different selection procedures were performed in order to choose the approximate best subset of covariates, one that was as simple as possible while still providing good predictive performance. In fact, traditional stepwise selection methods, such as forward and backward, suffer from high variability and low prediction accuracy, especially when the number of possible covariates is large and/or the level of multicollinearity is notable (Hastie, Tibshirani, and Friedman 2001). In such situation, penalized regression methods, such as LASSO, Elastic Net and penalized maximum likelihood

estimation (PMLE), combine the advantages of selection procedure preserving the model prediction accuracy. Logistic and penalized regression models underlined the existence of several determinants associated with declining the HPV vaccination. These included citizenship, father's age, mother's education, HPV knowledge score (number of correct answers), sources of HPV information, duration of HPV vaccine protection and attitudes towards HPV vaccination (safety and side effects). To evaluate the significance of interrelationships between psychological latent constructs (i.e. Theory of Planned Behaviour, Health Beliefs Model dimensions) partial least square structural equation models were employed. The structural equation results showed that psychological dimensions are significantly interrelated, underlining the influence of social norms and attitudes in respondents' behavioural intentions, i.e. future vaccination intentions (for unvaccinated girls) and future intentions to undergo a pap-test.

The studies presented in this thesis revealed that health behaviour, especially those related to vaccination choices, are shaped through a complex interplay of social-cognition determinants. Disposition towards other community members, social norms, beliefs and attitudes played an important role in predicting immunization decisions and should be taken into account when effective interventions are implemented. At the same time, the obtained results showed young girls still unable to make conscious and informed decision choices to protect their own health. Except for some unvaccinated respondents, girls appear unaware of the reason behind their vaccination and ignored most of the basic knowledge concerning HPV, HPV vaccination and pap-test screening.

Part 1.

Kicking against the pricks: vaccine sceptics have a different social orientation

Abstract

Background: In any country, part of the population is sceptical about the utility of vaccination. To develop successful vaccination programmes, it is important to study and understand the defining characteristics of vaccine sceptics. Research till now mainly focused either on the underlying motives of vaccine refusal, or on socio-demographic differences between vaccine sceptics and non-sceptics. It remained till now unexplored whether both groups differ in terms of basic psychological dispositions.

Methods: We held a population survey in a representative sample of the population in Flanders, Belgium (N= 1050), in which we investigated whether respondents' attitude to vaccination was associated with their basic disposition toward other community members or society in general, as measured by the Triandis and Gelfand social orientation scale.

Results: We found that sceptics and non-sceptics have a different social orientation, even when several variables are controlled for. More specifically, vaccine sceptics scored significantly lower on both horizontal individualism and horizontal collectivism, indicating a lower disposition to see others as equals.

Conclusion: These findings need confirmation in the context of different countries. Such insights can be valuable to optimize the design of effective communication strategies on vaccination programmes[1].

[1] This chapter is an extended version of the following publication: Luyten J., Desmet P., Dorgali V., Hens N. & Beutels P., 'Kicking against the pricks: vaccine skeptics have a different social orientation', European Journal of Public Health, 2014, 24 (2), p. 310-314.

1. Introduction

A central mission of public health policy is to ensure sufficiently high vaccination coverage in the population [1, 2]. Two goals motivate this objective [3]. First, to maximally prevent infections from occurring: either directly in those being vaccinated or indirectly in unprotected individuals through herd immunity (which is a consequence of reduced circulation of pathogens in a largely vaccinated population [4]). Second, to eradicate pathogens, as was the case for smallpox, and is the intention for polio through continued high coverage polio vaccination around the world [5]. Several countries, however, experience difficulties in reaching optimal participation in vaccination programs because increasing numbers of individuals hesitate or refuse to become vaccinated [6-9], in part precipitated by misinformation spread by anti-vaccination lobbies. Therefore, an essential policy challenge consists of establishing an effective dialogue between scientists, policy makers and the public at large, with the aim to sustain public trust in public health policy and to convey the need for continued vaccination efforts [10-15]. Such a strategy requires a thoroughgoing understanding of who refuses vaccination, and for which reasons [6, 13]. In the next sections a brief description of the main reasons of vaccination refusal will be provided. In addition, the extent literature underlined that some socio-demographic determinates are significant associated with sceptical groups. As it will be explained further below a new survey to explore whether vaccine sceptics are different from non-sceptics in terms of social orientation.

1.1 Motivations and demographic determinants of vaccine refusal

Research so far has placed opposition to vaccination in four general motivational categories [15–21]:

i. *general distrust of vaccination*: those who are not convinced by the ratio of risks vs. benefits of vaccines because of a lack of confidence in science, the pharmaceutical industry or public policy.

ii. *Free rider motives*: those who trust vaccination but who consider it an unnecessary intervention for themselves as long as enough others choose to become vaccinated.

iii. *Cognitive biases*: omission bias (the tendency to consider a similar risk worse when it results from action rather than from omission) or hyperbolic discounting (undervaluing the benefits of future disease protection as compared with the present in which adverse effects may occur).

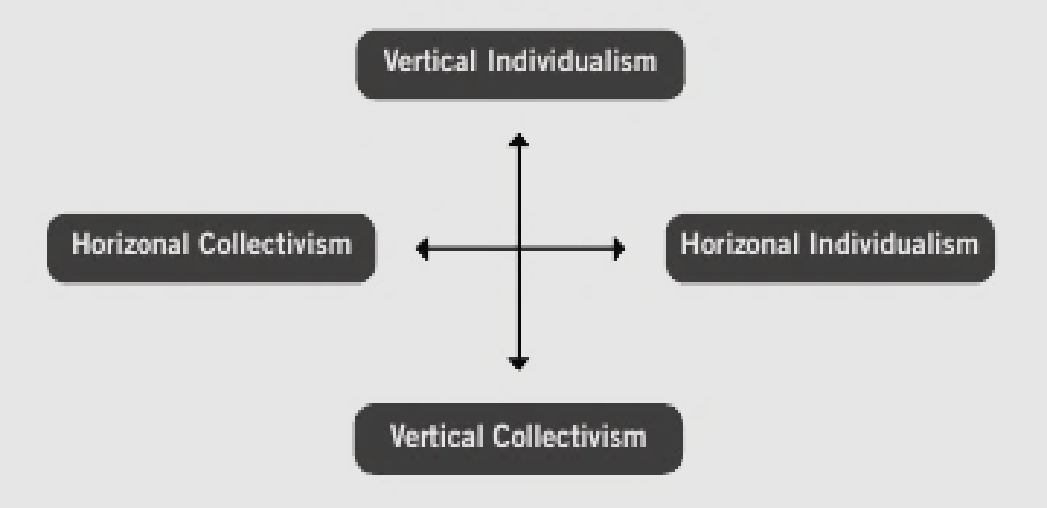
iv. *Fundamental objections*: those who hold religious or philosophical worldviews that are irreconcilable with vaccination.

Socio-demographic factors have also been shown to differ between vaccine refusers and non-refusers, including household income [22–24], level of education [23–26] marital status [6], race [15] and family size [24] with refusal often being more prevalent in relatively higher educated, wealthier groups [6,26]. To our knowledge, little is known about the more general cognitive profile of vaccine sceptics. Nonetheless, such information could be essential for the design of social marketing strategies for vaccines [12]. As vaccination is—at least partly—a matter of being solidaristic with others or not, we hypothesized that differences in social orientation, i.e. in one's basic attitude toward other community members and society in general, may translate into a different valuation of vaccines.

Singelis et al. [27] and Triandis et al. [28] developed an influential framework to study social orientation. This framework distinguishes two basic relational dimensions: a collectivism/individualism axis that reflects someone's sense of social cohesion and his/her willingness to prioritize common goals over personal ones, and a horizontal/vertical axis that indicates to what degree an individual expects equality or accepts inequality in social relationships [28,29]. This generates the 4-way typology horizontal and vertical individualism and collectivism (HI, VI, HC and VC) represented in Figure 1.

More specifically, in HI, people want to be unique and distinct from groups, i.e. they are likely to say "I want to do my own thing" and are highly self-reliant, but they are not especially interested in becoming distinguished or in having high status. In VI, people often want to become distinguished and acquire status, and they do this in individual competitions with others. They are likely to say "I want to be the best." An HC orientation emphasizes interdependency, cooperation with others and communal sharing, on a basis of equality. Finally, VC is an orientation that also promotes subordination of personal under group goals while differences in social status and hierarchy are acknowledged.

Figure 1- The 4-way typology of social orientation



2. Methods

2.1 Sample selection

Between March and July 2011, 3740 people were contacted by professional telephone operators using random digit dialing of fixed and mobile telephone numbers (during weekdays from 10 am to 9 pm, on Saturdays from 11 am to 7 pm). Every number was called once, except when respondents asked to call back at a more convenient time. Contacted persons were greeted and asked if they were willing to participate in a scientific survey concerning health policy (without
knowing that the subject was vaccination). The 1540 respondents willing to cooperate (41% of those contacted) were consequently asked for their age, gender, educational attainment and location over the five Flemish provinces and were selected when they fulfilled the predetermined quota for these criteria. Consequently, participants were asked whether they wanted to receive the survey either on paper or through an Internet-link, and whether they wanted to return their responses by post (using a prepaid envelope) or online. The sample size of this survey was determined in function of the initial purpose of the survey (a study published elsewhere [30]).

2.2 The Survey

In our survey, we presented the Triandis and Gelfand scale to determine an individual's position on the four dimensions of social orientation [34]. It consists of sixteen value judgments on which every respondent is asked to express his/her agreement on a five-point likert scale ranging from "strongly disagree" to "strongly agree" (see Figure 2). All 4 types of social orientation are characterized by 4 statements a priori designated to reflect one specific type. E.g. "Winning is everything" is held to be a typical VI statement, whereas "I feel good when I cooperate with others" is held to be typically HC. For every respondent, a score is obtained for all 4 dimensions (HI, VI, HC, VC) through calculating the average score attributed to the statements belonging to that dimension. The scale has proven to be a valid instrument to indicate how individuals see themselves in relation to other individuals and society as a whole and has been used and validated in numerous studies (for reviews see [36, 37]). Since the original scale was in English and the respondents are native Dutch speakers, translation and back-translation was conducted to ensure that all items contained equivalent meaning to the original.

To determine the respondents' general attitude toward vaccination, we asked them to indicate their agreement with the following statement on a 5-point Likert scale ranging from 'strongly disagree' to 'strongly agree' (Table 2):

'*If a vaccine exists for a certain disease, then vaccination is usually a good way to protect someone against this disease'*.

In addition, we asked for respondents' age, sex, level of education, profession, ethnic origin of mother and father, family size, age of family members, experience as health care worker, height, weight, smoking status (smoker/non-smoker), experience with travel vaccination, experience with severe illness (personal or within the family) and province, and subjected them to the EQ-5D-3L health survey, including the Visual Analogue Scale. The EQ-5D (3L) is a summary index that indicates a person's health state. It is computed by applying a formula that essentially attaches weights to different levels in the dimensions that constitute health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The EQ-5D-score is thus a weighted summary of these five dimensions. Each dimension is composed respectively by three levels, no problems (1), some problems (2), extreme problems (3). The VAS (Visual Analogue Scale) records the respondent's self-rated health on a 20cm vertical, visual analogue scale with endpoints labelled 'the best health you can imagine' and 'the worst

health you can imagine'. While the EQ-5D is an 'objective' scale (i.e. it calculates someone's health with an algorithm), the VAS represents a patient's 'subjective' perception of his/her health state [33]. In addition, combining the information about the length and the weight of the respondent the Body Mass Index (BMI) was also computed. The BMI is a measure computed by dividing the weight of an individual by the square of his/her length. A frequent use of the BMI is to assess how much an individual body weight departs from what is normal or desirable for a person of his/her height.

Figure 2 Triandis' and Gelfand's 16-item scale for horizontal and vertical individualism/collectivism.

Horizontal Individualism
1. I'd rather depend on myself than others.
2. I rely on myself most of the time; I rarely rely on others.
3. I often do "my own thing."
4. My personal identity, independent of others, is very important to me.

Vertical Individualism
5. It is important that I do my job better than others.
6. Winning is everything.
7. Competition is the law of nature.
8. When another person does better than I do, I get tense and aroused.

Horizontal Collectivism
9. If a coworker gets a prize, I would feel proud.
10. The well-being of my coworkers is important to me.
11. To me, pleasure is spending time with others.
12. I feel good when I cooperate with others.

Vertical Collectivism
13. Parents and children must stay together as much as possible.*
14. It is my duty to take care of my family, even when I have to sacrifice what I want.
15. Family members should stick together, no matter what sacrifices are required.
16. It is important to me that I respect the decisions made by my group*

**Because of unsatisfactory factor loadings, these two items were excluded from the main analysis and were only used in sensitivity analysis.*

2.3 Analysis

For the statistical analysis, we used SAS and SPSS. We conducted a confirmatory factor analysis of the 16 items of the scale and found the original 4-factor structure replicated. In our sample, all items, except two (VC13 and VC16), highly loaded on the intended dimension. Therefore, in line with the methodology followed in other studies [34,35], we excluded these two items from the main analysis. The response measuring respondents' attitude toward vaccination was dichotomized, giving 0 to those who indicated 'Agree' and 'strongly Agree' (a positive attitude toward vaccination) and 1 otherwise (a non-positive attitude). First, a

univariate regression was performed between every variable in the data set and attitude to vaccination. Then, a logistic regression model was built through forward and backward stepwise selection of all covariates to determine for the 1050 respondents, which variables were significantly associated with attitude to vaccination. The sensitivity of the results was explored by repeating the analyses using different categorizations of the dependent variable, and inclusion of the two previously excluded statements from the scale. Additionally, a mean response model was used to evaluate the impact on the results when the response is treated as continuous (i.e. without dichotomizing the dependent variable).

3. Results

We reached a sample of 1050 respondents (Table 1), considered representative for the population in Flanders (6 208 877 inhabitants, about 60% of Belgium36). Forty-one percent of the 3740 contacted persons (1540 individuals) consented to participate. Sixty-eight percent of these 1540 candidates were effectively recruited, i.e. 28% of those who were initially contacted. Fifty-one percent of the sample was female, 50% received higher education (university or non-university degree) and the mean age was 43 years.

Table 1 Characteristics of the surveyed population (N= 1050)

Variable	Classes	Percentage(N)
Age	18-25	19.24% (202)
	26-35	21.14% (222)
	36-50	22.95% (241)
	51-60	18.48% (194)
	61-76	18.19% (191)
Sex	Male	49.24% (517)
	Female	50.76% (533)
Education	NONE	1.05% (11)
	BASIC	1.62% (17)
	PROFESSIONAL	7.24% (76)
	LOWER TECHNICAL	4.48% (47)
	LOWER SECONDARY	6.1% (64)
	HIGHER TECHNICAL	11.33% (119)
	HIGHER SECUNDARY	18.67% (196)
	HIGHER NON_UNIVERSITY	35.62% (374)
	UNIVERSITY OR POST UNIV	13.90% (146)
Profession	Employees	44.57% (468)
	Independent	29.33% (308)
	Worker	14.95% (116)
	Other	11.05% (157)

Five percent of the respondents stated explicitly to be against vaccination, 15% stated to be neutral and 80% considered vaccination a good way to prevent disease (Table 2).

Table 2 Overview of the sample's attitude to vaccination

"Vaccination is a good way to prevent disease..."	**Not Agree at all**	**Not Agree**	**Neutral**	**Agree**	**Totally Agree**
Percentage	2 %	3 %	15 %	49 %	31 %
Count	16	35	159	514	326

The multiple regression results indicated that individuals who indicated a sceptical (i.e. a non-positive) attitude scored significantly lower on the dimensions HC and HI and were less likely to smoke. Per unit increase in their score on HI and HC, respondents had a 28 and 25% lower odds of being vaccine-sceptic, respectively. The odds of nonsmokers being vaccine-sceptic were 40% bigger than those of smokers. None of the other variables, including educational attainment, age, current health state, profession or experience with severe illness, had a significant predictive value in the model. On a univariate level, 'vertical collectivism' was the only additional variable with a significant (negative) association. This was no longer significant in the multiple regression model, in which we adjusted for the influence of other variables. The results using the mean response model confirmed a positive association between a more equally oriented profile and a more positive attitude toward vaccination. Furthermore, a second logistic regression model was fitted after reclassifying the neutral responses. However, moving neutral responders from the negative to the positive attitude category resulted in a very small number being classified as having a negative attitude (n = 51, or 5%), such that this approach could not produce interpretable estimates. Inclusion of the hitherto omitted Triandis and Gelfand statements (items 13 and 16 in Figure 2) did not alter these results.

Table 3 Overview logistic regression results of the significant covariates of a sceptical attitude toward vaccines

Response: Vaccination attitude (REF: Agree)	**Crude odds ratio**			**Adjusted odds ratio [a]**		
Effect	Point Estimate	95% LR CI	P-value	Point Estimate	95% LR CI	P-value
Smoking (Yes vs. No)	0.640	0.467-0.882	**0.0059**	**0.597**	**0.424-0.841**	**0.0031**
Horizontal Collectivism	0.711	0.611-0.825	**<0.001**	**0.759**	**0.642-0.897**	**0.0012**
Horizontal Individualism	0.791	0.681-0.917	**0.0018**	**0.724**	**0.617-0.849**	**<0.001**

[a] *List of non-significant parameters: age, gender, educational attainment, professional group, EQ-5D score, Visual Analogue Scale score, personal experience with severe illness, experience with severe illness in the family, origin mother, origin father, experience with travel vaccination, experience as health care worker, family size, age family members, province, height, weight, vertical individualism and vertical collectivism.*

4. Discussion

This study identified social orientation as an important determinant of attitude toward vaccination. Our results show that the values of horizontal individualism and horizontal collectivism are associated with whether people take a positive stand toward vaccines. These findings are notable for two reasons. First, given that social orientation was identified as a more important determinant than other more commonly reported socio-demographic variables (such as e.g. education or professional group), our findings underscore the value of investigating the psychological determinants of vaccine scepticism. Second, vaccination is a solidaristic intervention that benefits both the vaccinated individual and the wider community, whereas the risks associated with vaccination remain strictly private. Therefore, one might intuitively expect that people's attitude toward vaccination will be determined by the individualistic vs. collectivistic social orientation of individuals. However, our findings demonstrate that it is not so much individualism or collectivism that is of importance, but particularly the degree to which people value equality in their social relations. The more people see others as equals, the more positive they stand toward vaccination. Findings like ours may help to improve the design of effective communication strategies for vaccines. In several countries, marketing research has shown that advertisements are more persuasive when their appeal is matched to the social orientation of their targeted audience [37,38].

Individuals who scored high on HI, VI, HC or VC were found to be more susceptible for messages that emphasized self-direction, power, universalism and tradition, respectively [32]. For instance, the higher the HC orientation of participants, the more they were in favour of a brand selling a shopping bag with which 'you're doing your part to save the environment'. The higher the HI orientation, the more participants liked a brand selling T-shirts for which you could 'pick your colour, pick your message, and pick your style' [32]. Vaccination could be promoted by appealing to HC values like solidarity and interdependence (e.g. 'vaccination is a matter of taking care of each other') or HI values such as individual freedom and self-expression (e.g. 'vaccines enable you to safely explore the world by traveling' or 'vaccines enable your children to play safely with other children'). Our findings suggest that such marketing strategies are less likely to be persuasive in vaccine-sceptical groups because these groups have a significantly lower HC and HI orientation. However, as we also found a positive association between valuing equality and having a more pro-vaccination attitude, our findings also project that vaccine scepticism may diminish through more structural strategies aimed at stimulating HI and HC orientations. One could speculate that large events that speak to HI or HC values such as music festivals or election days present an opportunity to embed pro-vaccine messages.

Few studies were conducted so far in order to measure the effect of collectivism and individualism on attitudes towards vaccination. Velan (2012) [43] evaluated the attitude of the Israeli population. asking respondents to express their standpoints regarding five different vaccination programs. These included: pandemic influenza vaccination, seasonal influenza vaccination, travel vaccines, Human Papilloma Virus vaccine and childhood vaccinations. Analysis of the responses revealed that the individualism profile was characterized by the opinion that vaccination should be left to personal choice.

Betch and colleagues (2017) [41] compared vaccination decisions of 2,000 participants from South Korea, India, Vietnam, Hong Kong, the United States, Germany and the Netherlands. Countries were clustered in "eastern" and "wester" to allow for cultural comparison based on collectivistic versus individualistic orientation. The analysis revealed that participants from eastern countries perceived greater disease risk and had higher levels of collectivism. Both these elements have a mediation effects and increased vaccination intentions. In addition, these findings support the idea that vaccination choices of eastern participants are influenced by their collectivist orientation more than the western counterpart. On the other hand, Betsch and colleagues (2017) showed that the communications based on the concept of heard immunity increase vaccinations especially where the collectivistic perception is lacking, i.e. in culture focused on individual than collective benefits. Then, vaccination intentions may be positively influenced by an increase in collectivism inclinations due to the conveyed herd

immunity concept. Finally, Boas et al. (2016) [42] described that an important element of Israel's vaccination response was to promote immunization as being of benefit to the family. In this context, the family is considered as metaphor for social solidarity.

These studies underlined that the promotion of a collectivist orientation, may influence individuals' attitudes to vaccination., increasing subjects' disposition to safe immunization practice.

Some limitations of our study should be mentioned. First, our study may be influenced by the Flemish context (culture, education, experience with vaccination, etc.). It is noteworthy that most (but not all) childhood vaccines are given free of charge in Belgium. Although only polio vaccination is compulsory [39], uptake is high for most vaccines [40], and has remained largely unscathed by general and specific anti-vaccine lobby campaigns (e.g. false claims of causal links between measles-mumps-rubella vaccine and autism [41] and hepatitis B vaccine and multiple sclerosis42). Second, non-response bias is always a potential concern in survey-based research. We believe that the overall response (41% of contacted persons consented to participate) was acceptable, given the design and nature of the study. Third, this was an exploratory study. Further research in this area could expand the methodology we used with alternative and/or more specific instruments to capture attitude to vaccination and social orientation. Certainly, it would be interesting to validate our study in other countries and to investigate the influence of other psychological attributes, perhaps also on other forms of health care refusal. As the success of infectious disease prevention largely depends on collective cooperation, all knowledge about what differentiates sceptics from non-sceptics allows more successful anticipation, communication and education. The identification of social orientation as an explanatory factor for vaccine scepticism presents an opportunity to conduct further research in this direction.

References

1. Omer SB, Salmon DA, Orenstein WA, et al. Vaccine refusal, mandatory immunization, and the risks of vaccine-preventable diseases. N Engl J Med 2009;360:1981–8.
2. Dawson A. The moral case for the routine vaccination of children in developed and developing countries. Health Aff (Millwood) 2011;30:1029–33.
3. Diekema D, Marcuse E. Ethical issues in the vaccination of children. In: Bayer R, Gostin L, Jennings B, Steinbock B, editors. Public Health Ethics: Theory, Policy and Practice. New York: Oxford University Press, 2007: 279–88.
4. Fine P, Eames K, Heymann DL. ‘‘Herd immunity’’: a rough guide. Clin Infect Dis 2011;52:911-6.
5. Dowdle WR, Cochi SL. The principles and feasibility of disease eradication. Vaccine 2011;29(Suppl 4):D70–3.
6. Gust DA, Darling N, Kennedy A, et al. Parents with doubts about vaccines: which vaccines and reasons why. Pediatrics 2008;122:718–25.
7. Gill DG. Vaccine refusal and the risks of vaccine-preventable diseases. N Engl J Med 2009;361:723; author reply 4.
8. Poland GA, Jacobson RM. Understanding those who do not understand: a brief review of the anti-vaccine movement. Vaccine 2001;19:2440–5.
9. Cooper LZ, Larson HJ, Katz SL. Protecting public trust in immunization. Pediatrics 2008;122:149–53.
10. Benin AL, Wisler-Scher DJ, Colson E, et al. Qualitative analysis of mothers’ decision-making about vaccines for infants: the importance of trust. Pediatrics 2006; 117:1532–41.
11. Leask J, Chapman S, Hawe P, et al. What maintains parental support for vaccination when challenged by anti-vaccination messages? A qualitative study. Vaccine 2006;24: 7238–45.
12. Opel DJ, Diekema DS, Lee NR, et al. Social marketing as a strategy to increase immunization rates. Arch Pediatr Adolesc Med 2009;163:432–7.
13. Hammer LD, Curry ES, Harlor AD, et al. Increasing immunization coverage. Pediatrics 2010;125:1295–304.
14. Kata A. A postmodern Pandora’s box: anti-vaccination misinformation on the Internet. Vaccine 2010;28:1709–16.
15. Asveld L. Mass-vaccination programmes and the value of respect for autonomy. Bioethics 2008;22:245–57.
16. Blume S. Anti-vaccination movements and their interpretations. Soc Sci Med 2006; 62:628–42.

17. Luyten J, Vandevelde A, Van Damme P, et al. Vaccination policy and ethical challenges posed by herd immunity, suboptimal uptake and subgroup targeting. Public Health Ethics 2011;4:280–91.
18. Lee M, Male M. Against medical advice: the anti-consumption of vaccines. J Consum Mark 2011;28:484–90.
19. Kaplan M, Y K. Anti-consumption of public services: vacci(not)ion for swine flu. Manag Res Rev 2011;34:353–63.
20. Jacobson RM, Targonski PV, Poland GA. A taxonomy of reasoning flaws in the antivaccine movement. Vaccine 2007;25:3146–52.
21. Kaplan M, Kaplan Y. Anti-consumption of public services: vacci(not)ion for Swine Flu. Manag Res Rev 2011;34:353–63.
22. Kennedy AM, Brown CJ, Gust DA. Vaccine beliefs of parents who oppose compulsory vaccination. Public Health Rep 2005;120:252–8.
23. Mollema L, Wijers N, Hahne SJ, et al. Participation in and attitude towards the national immunization program in the Netherlands: data from population-based questionnaires. BMC Public Health 2012;12:57.
24. Smith PJ, Chu SY, Barker LE. Children who have received no vaccines: who are they and where do they live? Pediatrics 2004;114:187–95.
25. Wei F, Mullooly JP, Goodman M, et al. Identification and characteristics of vaccine refusers. BMC Pediatr 2009;9:18.
26. Hak E, Schonbeck Y, De Melker H, et al. Negative attitude of highly educated parents and health care workers towards future vaccinations in the Dutch childhood vaccination program. Vaccine 2005;23:3103–7.
27. Singelis T, Triandis H, Bhawuk D, et al. Horizontal and vertical dimensions of individualism and collectivism: a theoretical and measurement refinement. Cross Cult Res 1995;29:240–75.
28. Triandis HC, Gelfand MJ. Converging measurement of horizontal and vertical individualism and collectivism. J Pers Soc Psychol 1998;74:118–28.
29. Triandis HC. Individualism-collectivism and personality. J Pers 2001;69:907–24.
30. Luyten J, Dorgali V, Hens N, et al. Public preferences over efficiency, equity and autonomy in vaccination policy: An empirical study. Soc Sci Med 2013;77:84–9.
31. Oyserman D, Coon HM, Kemmelmeier M. Rethinking individualism and collectivism: evaluation of theoretical assumptions and meta-analyses. Psychol Bull 2002; 128:3–72.
32. Shavitt S, Torelli C, Riemer H. Horizontal and Vertical individualism and collectivism:implications for understanding psychological processes. In: Gelfand M, Chi-yue C, Hong Y, editors. Advances in Culture and Psychology. New York: Oxford

University Press, 2010: 384. 33 Brazier J, Ratcliffe J, Salomon J, et al. Measuring and Valuing Health Benefits for Economic Evaluation. Oxford: Oxford University Press, 2007.

33. Li F, Aksoy L. Dimensionality of individualism-collectivism and measurement equivalence of Triandis and Gelfand's scale. J Bus Psychol 2007;21:313–29. 35 Ng KY, Van Dyne L. Individualism-collectivism as a boundary condition for effectiveness of minority influence in decision making. Organ Behav Hum Decis Process 2001;84:198–225.
34. Statbel. Population. Brussels. http://statbel.fgov.be/nl/statistieken/cijfers/bevolking/ (15 June 2013, date last accessed) (2011).
35. Shavitt S, Lalwani AK, Zhang J, et al. The horizontal/vertical distinction in cross-cultural consumer research. J Consum Psychol 2006;16:325–42.
36. Torelli CJ, O¨ zsomer A, Carvalho SW, et al. Brand concepts as representations of human values: do cultural congruity and compatibility between values matter? J Mark 2012;76:92–108.
37. Haverkate M, D'Ancona F, Giambi C, et al. Mandatory and recommended vaccination in the EU, Iceland and Norway: results of the VENICE 2010 survey on the ways of implementing national vaccination programmes. Euro Surveill 2012;17:article 4.
38. Theeten H, Vandermeulen C, Roelants M, et al. Coverage of recommended vaccines in children at 7-8 years of age in Flanders, Belgium. Acta Paediatr 2009;98:1307–12.
39. Farrington CP, Miller E, Taylor B. MMR and autism: further evidence against a causal association. Vaccine 2001;19:3632–5.
40. Ascherio A, Zhang SM, Hernan MA, et al. Hepatitis B vaccination and the risk of multiple sclerosis. N Engl J Med 2001;344:327–32.
41. Betch C, Böhm R, Korn L, and Holtmann C. On the benefits of explaining herd immunity in vaccine advocacy. N Hum Beh 2017,1:6-1.
42. Boas H, Rosenthal A, Davidovitch N. Between individualism and social solidarity in vaccination policy: the case of the 2013 OPV campaign in Israel. Israel J Health Policy Res. 2016; 5:64.
43. Velan B. Vaccine hesitancy as self-determination: an Israeli perspective. Israel J Health Policy Res. 2016; 5:13.

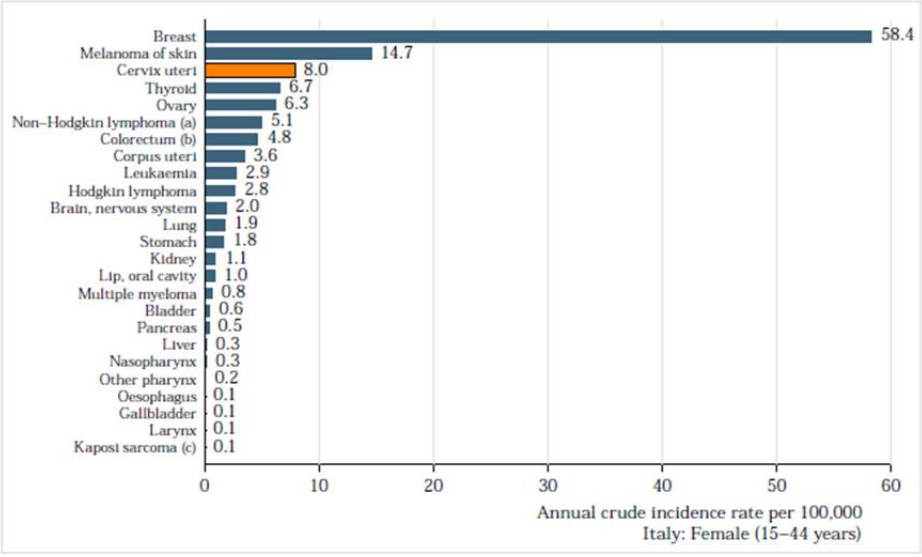
Part II

Attitudes, knowledge and intentions towards HPV immunization and cervical cancer prevention among female university students in Tuscany

1 Introduction

Worldwide, cervical cancer is the fourth most frequent cancer in women with an estimated 530,000 new cases in 2012 representing 7.5% of all female cancer deaths. There were an estimated 266,000 deaths from cervical cancer worldwide in 2012, accounting for 7.5% of all female cancer deaths. Almost nine out of ten (87%) cervical cancer deaths occur in the less developed regions. Mortality varies 18-fold between the different regions of the world, with rates ranging from less than 2 per 100,000 in Western Asia, Western Europe and Australia/New Zealand to more than 20 per 100,000 in Melanesia (20.6), Middle (22.2) and Eastern (27.6) Africa (WHO, 2016: sito internet; GLOBOCAN, 2012: http://globocan.iarc.fr; Jin et al., 1999).

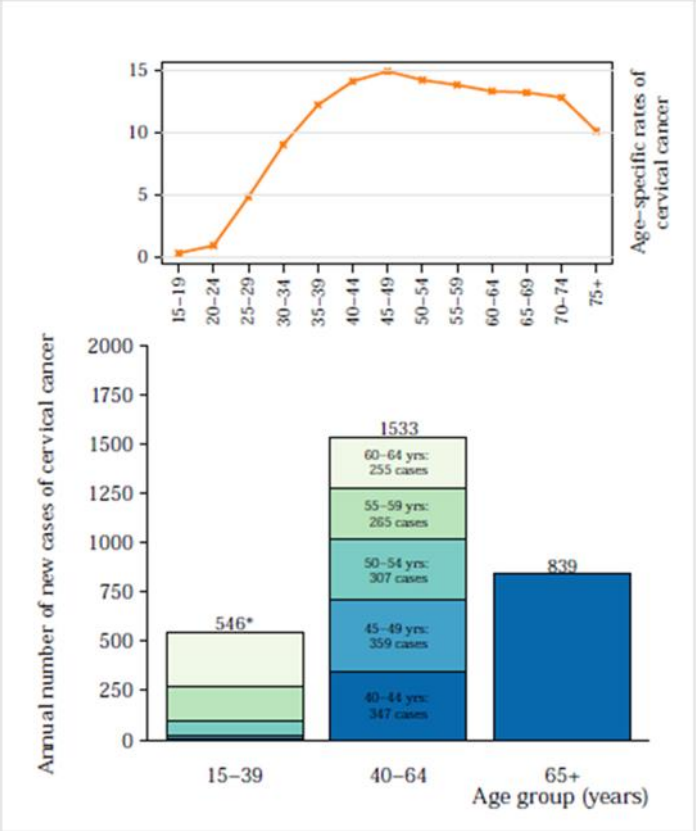
Figure 3 Comparison of age-specific cervical cancer to age-specific incidence of other cancer among women 15-44 years of age in Italy (estimates for 2012)



(*Source: Ferlay et al., 2012, available at: http://globocan.iarc.fr*)

Cervical cancer is the second most common cause of cancer death among young women (between the ages of 15–44) in Europe (Bonanni et al., 2011) and is a leading cause of cancer death in women worldwide (WHO, 2016; GLOBOCAN, 2012: http://globocan.iarc.fr; Jin et al., 1999). In Italy cervical cancer is the 3rd (Figure 3) most common female cancer and the 7th leading cause of cancer deaths in women aged 15 to 44 years, with about 2,918 new cases diagnosed and 1,016 deaths in 2012 (GLOBOCAN, 2016: http://globocan.iarc.fr).

Figure 4 Annual number of cases and age-specific incidence rates of cervical cancer in Italy estimates for 2012

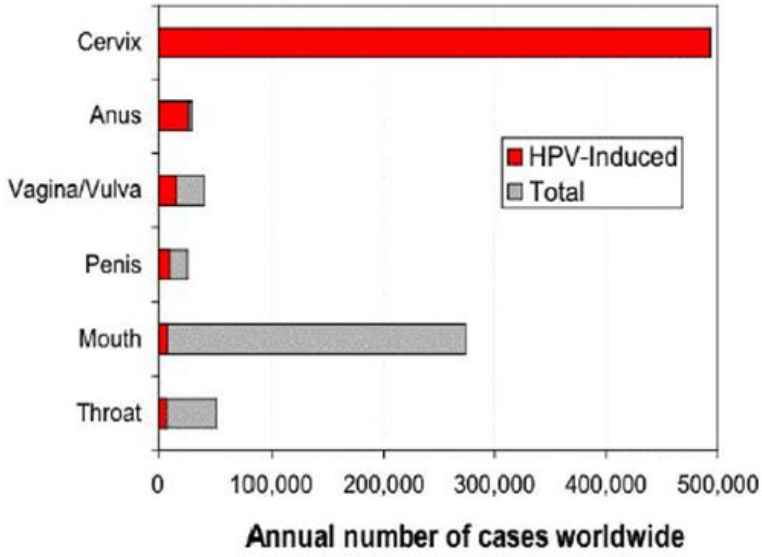


(*Source: Ferlay et al., 2012, available at: http://globocan.iarc.fr*)

In 1976, the German virologist Harald Zur Haunsen hypothesized that human papillomavirus played an important role in the cause of cervical cancer. In 1983/84, he discovered that two HPV genotypes (16 and 18) were associated with cervix uteri cancer, establishing the oncogenic risk related to some HPV serotypes (Walboomers et al 1999). Zur Haunses's studies pave the way for a vaccine development, which was introduced in most national immunization programmes in 2006. Since that time, knowledge about HPV-associated lesions has increased, revealing the great complexity of the ecosystem of this pathogen (Insiga et al.,

2007; Orlando et al., 2012; Panatto et al, 2013). The majority of HPV-related cancers derive from the uterine cervix, although the virus is also associated with five other cancers (Figure 5): penis, vulva, vagina, anus and oropharynx (including the base of the tongue and tonsils) (Parking, 2005; CCDR, 2006; Winters et al., 2006).

Figure 5 Worldwide number of cervical cancer in 2002.The fraction of cancers estimated to be induced by HPV types is shown in red.



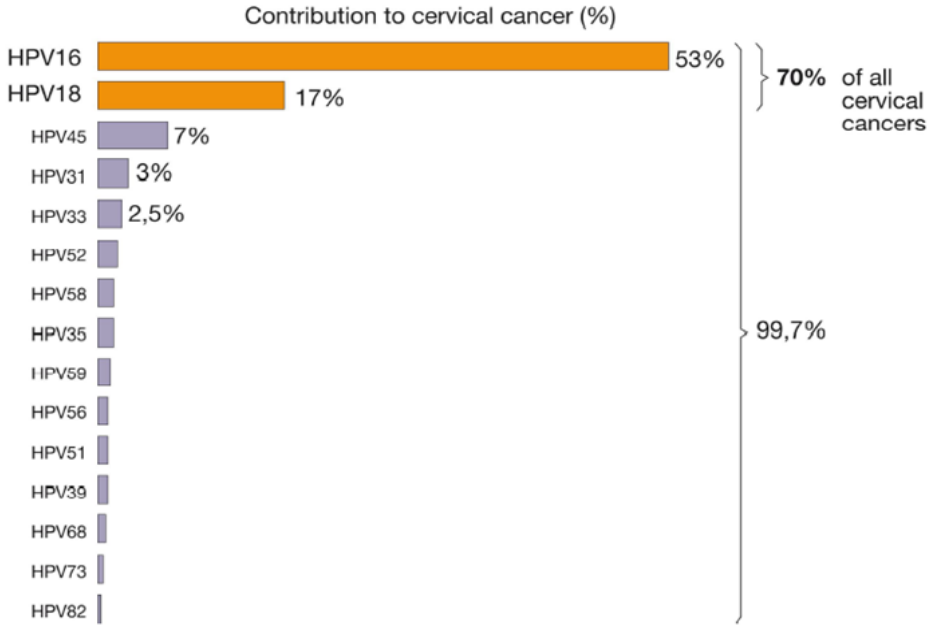
(Sources: Parking, 2006; Karolinska Institutet & The Noble Assembly at Karolinska Institutet, 2008)

Nowadays many types of human papillomavirus (HPV) are known. Of these, about 50 have a high tropism for the anogenital mucosa and are sexually transmitted. According to the most recent classification of the International Agency for Research on Cancer (IARC) 12 genotypes are defined as high-risk (HR) oncogenic (HPV-16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59). Moreover, a probable oncogenic risk has been assigned to genotype 68, and a possible oncogenic risk to other 12 types (HPV-26, 53, 66, 67, 70, 73, 82, 30, 34, 69, 85, and 97). Many low-risk (LR) genotypes that cause benign lesions and warts are also known (HPV 6, 11, 28, 32, 40, 42, 43, 44, 54, 55, 57, 61, 62, 71, 72, 74, 81, 83, 84, 86, 87, and 89) [IARC, 2016].

The two most prevalent virus genotypes are HPV16 and HPV 18 which have been identified as the most common in invasive cervical cancer during the period 1940–2007 (Alemany et al, 2014). The risk of developing squamous cell carcinoma of the cervix is about 400 times higher following infection with HPV-16 and about 250 times higher following infection with HPV-18 compared to the risk in uninfected women (De Sanjose et al., 2010). HPV16 and HPV 18 are

responsible of approximately 70% of invasive cervical cancers worldwide (Figure 6) (Roden, 2006).

Figure 6 The cumulative frequency of HPV genotypes present in cervical cancer.



(Sources: Roden, 2006; Karolinska Institutet & The Noble Assembly at Karolinska Institutet, 2008)

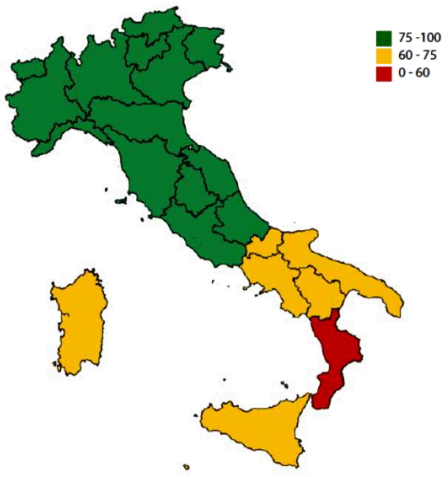
In general, the incidence of infection with high-risk HPV types tends to be higher than that with low-risk ones (Moscicki et al., 2001; Richardson et al., 2003; Muñoz et al., 2004) and to be greater among younger than older women (Franco et al., 1999; Muñoz et al., 2004), although median duration of infection appears to be comparable by age (Muñoz et al., 2004). Women infected with a given HPV type may be co-infected or subsequently infected with several types which may cause cervical lesions (Insiga et al., 2007).

While infection with a high-risk oncogenic HPV type is the underlying cause of almost all cases of cervical cancer, these infections do not always cause cancer. Most women infected with high-risk HPV do not develop cancer because most infections are short-lived and spontaneous clearance of the virus usually occurs within 2 years. Infection with high-risk HPV persists only in a small percentage of women; only a small fraction of these chronic infections progress to precancer and of these, even fewer will progress to invasive cancer (WHO, 2014; Burd, 2003). The widespread diffusion of HPV infection puts all female sexually active population at risk of cancer. For this reason, in the past decades, and before the vaccine

development, many countries worldwide have initiated cervical screening programs, aimed at the timely detection and treatment of precancerous lesions (Vaccarella et al., 2013).

Three different types of screening tests are currently available, conventional (Pap) test and liquid-based cytology (LBC), visual inspection with Acetic Acid (VIA) and HPV testing for high-risk HPV types (WHO, 2016: http://www.who.int/). The Papanicolaou test, commonly referred to as the "Pap test" or "Pap smear," is one of the most reliable and widely used cervical cancer screening tests available, and it is also relatively low-cost (Hildesheim et al.,2007). The test is conducted during a pelvic examination in which cells are collected from the cervix and analysed under a microscope for evidence of precancers. Through such screening programs, cell changes on the cervix can be detected and treated appropriately before potentially developing into cervical cancer. Screening guidelines vary by country, but in general, screening is recommended to start at about the age of 20 or 25, continue until about the age of 50 or 60, and occur every three to five years (Strander, 2009; Dara Lee et al., 2015). Widespread screening and subsequent diagnosis of pre-cancerous lesions and early-stage cancer has led to a dramatic drop in cervical cancer rates and deaths in developed countries (Schiffman et al, 2007; Arbyn et al., 2010; Dara Lee, 2015). In Italy, a screening program targeting women between 25 and 70 years has been in place since 1996, allowing a sharp reduction in cancer incidence (National Observatory on Screening, 2015). Between 2011 and 2013 three in four Italian women (25-64) undergo a pap-smear (77%) [Osservatorio Nazionale di Screening, 2015)]. The proportion of women who had pap-smear between 2011-2013 was higher in North/Central Italian regions (Figure 7).

Figure 7 National coverage by pap-test (2011-2013) [Source: Osservatorio Nazionale di Screening, 2015]



(

The discovery that the HPV was the necessary cause of cervical cancer opened a new era in oncology prevention (Bonanni et al., 2015). A new vaccine, able to prevent infections caused

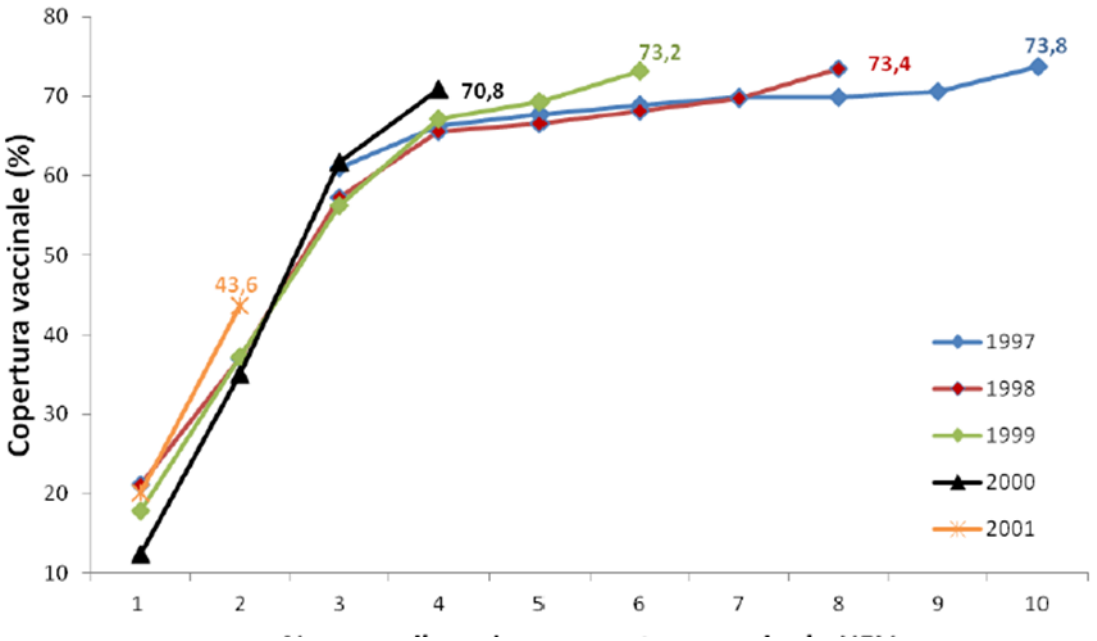
by high-risk HPV serotypes, was developed and included in the national immunization programmes of most developed countries. So far, two HPV vaccines have been licensed: a bivalent vaccine (Cervarix® GlaxoSmithKline Biologicals S.A.) and a quadrivalent vaccine (Gardasil® Merck and Co). Both vaccines are prophylactic (prevent HPV infections by HPV-16 and HPV-18.) and are most effective when administered prior to infection with HPV, which is acquired by most individuals shortly after sexual debut. For this reason, routine vaccination is highly recommended to girls of 9-13 years of age, while females 15–26 years old (those most at risk of exposure), represent the catch-up population (Koutsky, 1997; Bonanni et al, 2011). The vaccines are not therapeutic (cannot be used to treat existing HPV and HPV-related disease), nor do they have any effect on progression to disease (precancer and cancer) in persons who have HPV infection at the time of vaccination (WHO,2013). Both vaccines have been shown to elicit cross-protection against other high-risk HPV types (Malagón,2012; Schiller et al., 2012; Panatto et al., 2015). Gardasil® also offers protection against genital warts and low-risk lesions caused by HPV-6 and HPV-11 (Panatto et al., 2015). The vaccines have shown an efficacy close to 100% in preventing precancerous lesions in HPV-naïve individuals (Guzzetta et al., 2013; Schiller et al., 2012; Harper, 2009), without significant evidence for waning immunity throughout the duration of the trials (Guzzetta et al., 2013; Rowhani-Rahbar et al., 2012; Villa et al., 2006). In Europe, universal HPV vaccination of female adolescents was first introduced in 2007 in Belgium, France and Germany (Bonanni et al, 2011). Since 2007, a total of 19 out of 29 countries in the EU/EEA (Norway and Iceland but not Liechtenstein) have implemented a routine HPV vaccination programme and 10 countries have introduced catch-up programmes. As the European Centre for Disease Prevention and Control (ECDC) attested in 2012, a high heterogeneity in the strategies for implementation of HPV vaccination has been registered among EU/EEA countries. Recommendations for the vaccination age are diverse, ranging from 9 to 18 years, as they are for catch-up rounds, where they range from 12 to 40 years. The adopted vaccination policy targets girls/women in all the countries where HPV vaccine has been introduced, except for Austria, where boys/men are also included. In most cases the vaccination programmes are financed by the national health systems. However, in Austria the vaccination is entirely covered by the recipient, and Belgium and France have adopted a co-financed system, where the vaccine recipient contributes to the payment (25% and 35%, respectively). (Bonanni, 2011; ECDC, 2012). Coverage rates, where data are available, range from 17–84% and are generally lower than expected (ECDC,2012).

Since 2007-2008 universal vaccination against Human Papilloma Virus (HPV) was included in the Italian national immunization programme. The vaccine is offered free of charge to 11-year-old girls who represent the primary target group. Additionally, a catch-up campaign has been promoted in some regions to extend the active offer to older female age groups

(secondary target group)[2]. However, the transfer of responsibility for health to regional authorities has resulted in a diverse situation within the country: only seven out of the 21 Italian regions and autonomous provinces have introduced catch-up programmes, covering around 1/4 of the national population. The target age groups for catch-up immunisation are also variable from region to region (ECDC, 2012).

In Tuscany, the vaccine is offered free of charge to girls aged 12-18 yr. The vaccination is actively offered -with a letter sent by the local health unit- only to 12 and 16-year-old girls while a parent's application is requested for all other young women. Currently, 9 years after the initiation of the programme, HPV vaccination coverage in the two target populations is substantially lower than WHO targets (95%).

Figure 8 HPV vaccination coverage birth cohorts 1997-2001



(Source: Istituto Superiore di Sanità 30/06/2013)

[2] The target age groups for catch-up immunisation vary from region to region according to local objectives and legislations.

2 The failure of rational model in explaining HPV vaccination intake

Standard economic theory represents the starting point for modelling many aspects of individuals' behaviours. If a certain behaviour involves a choice between different options, with clearly perceived costs and benefits for the individual, economic theory represents the basic starting point to analyse the decision process (Darnton, 2008). Elster (1989) stated the essence of rational choice theory when he said that "when faced with several courses of action, people usually do what they believe is likely to have the best overall outcome" (Ogu, 2013). The 'rationality' defined by the rational choice theory adopts a more specific and narrowed definition, which simply means that "an individual acts as to balance costs against benefits to arrive at action that maximizes personal advantage." (Friedman, 1953).

Rational choice theory traditionally assumes that individuals make behavioural decisions based on a calculation of the expected costs and benefits of a certain behaviour. Strictly speaking, rational choice theory requires only 'well-ordered' and consistent preference mappings over the relevant period; it does not attach any welfare attributes to these preferences.

According to the basic economic approach, individuals make a rational decision by evaluating the consequences of vaccination in terms of net utility (or benefit), i.e. the balance between benefits and costs of vaccination decision. After a careful analysis of the information collected, individual's behaviour will be driven by the above-mentioned balance (net benefit perceived). If benefits prevail over costs, the choice will be in favour of vaccination, otherwise immunization will be not considered. This can be formalized as follows:

$$\Delta \mathrm{E}(\mathrm{t}) = \delta(t) m\lambda(t) - \alpha(t) \qquad (1)$$

where $\Delta E\ (t)$ indicates the net benefit perceived, or payoff perceived, resulting by the implemented behaviour in a certain instant of time t. In fact, it is assumed that variation of model's variables over time. The terms $\delta\ (t) m\lambda\ (t)$ and $\alpha\ (t)$ in the equations 1, represent, respectively, the benefit and the cost of the behaviour (get or not the HPV vaccination).

The cost-benefits model assumes that rational and informed individuals should be aware of the causal role of human papillomavirus infections in cervical cancer. Knowing that, the net benefit of HPV vaccination represents the possibility to avoid the future cost of illness (i.e. the risk of developing or dying for cervical cancer). In the models (1) the latter is determined by the product of the perceived risk of becoming infected with HPV (given by $m\lambda(t)$, where $\lambda(t)$ is the risk of getting infected with HPV during life) by the risk $\delta\ (t)$ of developing cervical cancer as consequence of the infection.

Considering the health consequences of HPV infection, the benefit offered by the HPV vaccination (i.e. the protection against the risk of contracting one of the most dangerous disease, as the cervical cancer is), should be perceived as extremely high if compared with its cost. In fact, by excluding the economic component covered by the Italian National Health System, the only cost component is represented by the fear of vaccine side effects.
According to the available scientific literature, the proportion of subjects reporting serious adverse effects (AE) or discontinuing due to an AE is low and similar for the two vaccines (Gonçalves et al., 2014). In consideration of the above, the costs associated with HPV vaccination should be essentially zero. Therefore, if the decision about vaccination was taken balancing costs and benefits, adherence should have been higher than the current coverage. The simple cost-benefits approach failed in explain the reasons behind HPV vaccination choices.

3 Research objectives

The basic scientific question dealt with preventive behaviour in relation to cervical cancer, namely why vaccine coverage against Human Papilloma Virus (HPV) is so low in many countries including Italy, particularly which behavioural factors are at work to cause the basic rational model of immunization to fail so markedly. As stated above, by "basic rational model" we mean a simple decision model of agents comparing the direct perceived benefits and costs of immunization. As perceived benefits - protection from the most deadly disease, namely cancer - should be extremely high, and costs – mainly risks of possible side effects of immunization - low due to the age at immunization (i.e., parents' worries about potential side effects should be strongly mitigated compared to infants' immunization) one would expect high coverage easy to reach.
To determine which further factors, beyond the binomial cost-benefit, might be important for decision makers, a different theoretical approach was considered. In the last twenty years, behavioural theories or social cognition models as they are sometimes called, have become increasingly important and popular in the field of health (Glanz et al., 2008). Research guided by psychosocial theories have added to our understanding how cognitive and social determinants might affect healthy or risky practices. From this perspective, a deepen analysis of the social, cultural and cognitive aspects of health behaviours is fundamental. Prior to the questionnaire definition a large methodological review of the existing literature was conducted. Among the existing approaches, more attention was given to some of the most important health models, such as the *health belief model* of Rosenstock, the *social cognitive*

theory of Bandura, the Ajzen & Fishbein's *theories of reasoned action and planned behaviour*, the Roger's *protection motivation theory*, the *social influence theory* of Cialdini & Trost and the *transtheoretical model* of Prochaska & Di Clemente. The concepts of social norms, perceived susceptibility to illness and attitudes/beliefs towards a certain health behaviour were examined in depth. These are core constructs of the above-mentioned models and of most important health behaviour theories used to explain and predict individual habits. Social norms are the customary rules that govern behaviour in groups and societies. Many researches demonstrated the important role of social norms as "social proof as information toward right living" (Cialdini &Trost, 1998) and how humans tend to rely on "social reality" when the appropriate behaviour is unclear (Festinger,1954). For this reason, social norms may have a key role in explaining vaccination choices and predict screening practices. Perceived susceptibility is an individual's belief about the likelihood of a health threat's occurrence or the likelihood of developing a health problem. Existing research suggest that disease risk perception is a critical determinant of health behaviour, but it is not clear how parents perceive the susceptibility of their young daughter (11 or 12 years old) for a sexual transmitted disease. On the contrary, the perceived susceptibility of cervical cancer should guide future decisions about pap-test screening. Attitudes are the way a person expresses or applies is an internal feeling (belief) about a certain behaviour and may explain parents' decisions about HPV vaccination as well as the respondents' future cervical cancer screening habits. Whereas respondents' vaccination status was consequences of a past behaviour, future cervical screening practices were assessed by intentions. According to Ajzen and Fishbein (1975) these are direct antecedent of a behaviour and can be used as a proximal measure of it.

For this purpose, constructs belonging to three of the already mentioned theories, i.e. the health belief model (Becker, 1974; Rosenstock, 1974), the theory of planned behaviour and the social influence theory (Cialdini & Trost, 1998), were implemented in a newly research instrument. Published measured and original question as well as cultural and social aspects of respondents' environment were balanced to create a proper survey questionnaire. The amount of paper and theories dealing with health and psychology is large and often difficult to interpret. For this reason, the largest portion of the research time was spent on the theory review and implementation of the questionnaire. It was self-administered among young adult women aged 19-22 attending their bachelor at the universities of Pisa and Florence with the purpose of measuring the determinants of their vaccination status as well as their attitudes, beliefs and knowledge towards HPV, HPV vaccination and cervical cancer screening practices (Pap-test). After the age of 18 girls become more responsible in managing different aspects of their life, including their health status. The rate at which adolescents' capacity for medical decision-making evolves depends on the complexity of their medical condition, their

understanding or experience of it, and other contextual factors such as culture, family and belief systems. Evidence suggests that already at 14-15-year olds girls have similar capacities to adults (Weithorn and Campbell,1982). Autonomous decision-making processes should imply a greater lever of consciousness and knowledge, especially for choices related to the individual well-being. Concerning this, the research aims as well to investigate how much consciously respondents are involved in their health protection.

Thus, the objectives of the current study are:

i. to understand why the simple rational model (net perceived pay off= perceived benefits–perceived costs of immunization) failed;
ii. to identify which further factors, beyond cost-benefits, might be important for decision makers (social norms, conformism, attitudes);
iii. to investigate factors related with realized immunization choice (3-fold strategy: Yes, No, I Don't remember);
iv. to assess young women attitudes, knowledges and intentions toward HPV and HPV vaccination;
v. to investigate how young women perceive their health status and how they plan to act to prevent cervical cancer in the future (Pap-test intentions).

A summary of the implemented theoretical approaches will follow.

4 Methods

4.1. Behavioural theories

4.1.1. Social Influence: social norms and conformity

The influence process is a social instrument through which human being may generate and manage changes in the social world. As many things, this process can be handled well or poorly. It can be employed to change individuals' negative behaviours and to encourage growth, thereby creating the condition for new challenging opportunities. There exists an extended literature of social science that explains how, why and when the influence process works most effectively. Social norms and conformity are two of the major components of that

body of information. In the following paragraphs, the relevant literatures on social norms and conformity will be presented.

4.1.1.1. Social Norms

A norm is, like other psychological phenomena, a construct that has widespread usage because it helps describe and explain human behaviour. Social norms are rules and standards accepted by the members of a group, that guide and/or limit the social behaviour of agents without the force of law. These standards emerge from the interaction with others; they may or may not be explicitly declared, and any penalties that arise for deviations of these originate from social relations, not from the legal system (Cialdini & Trost, 1998). Social norms can include general, societal expectations for our behaviour (Blake & Davis, 1964; Pepitone, 1976); the expectations of valued others for our behaviour (Fishbein and Ajzen's, 1975); our own expectations for our behaviour (Schwartz,1977); and standards that develop out of our observations of others' behaviour (Cialdini, Reno, and Kallgren, 1990; Cialdini &Trost, 1998). The substance of a social norm is neither inherently good nor inherently evaluable; its effectiveness and its power is granted by its acceptance within the culture and/or the communities within which it comes (Berger & Luckmann, 1966; Solomon, Greenberg, & Pyszczynski, 1991).

In short, social norms represent the community expectations with respect the group members should behave (Levine e Moreland, 1990). In summary, social norms are informal understandings that govern the behaviour of members of a society (Jackson, 1965). These rules have regarded to exist as collective representations of acceptable group conduct as well as individual perceptions of particular group conduct (Lapinsky & Rimal, 2005). They can be viewed as cultural products (including values, customs, and traditions) (Sherif, 1993) which represent individuals' basic knowledge of what others do and think that they should do (Cialdini, 2003).

There are three main types of social norms, descriptive norms, injunctive norms and subjective norms. Each of these has a specific role in affecting behaviour and in facilitating the accomplishment of basic social influence goals (Cialdini, Kallgren &Reno, 1991; Cialdini, Reno & Kallgren, 1990; Cialdini& Trost, 1998). A brif descrition of these three norms categories in presented in the following paragraphs.

4.1.1.2. Descriptive norms

In general, human being are motivated to act in ways that are effective in achieving their goals; they want to make accurate decisions (Cialdini & Trost, 1998). White in 1959, indicated that interest in accurately perceiving and dealing with our environment is an adaptive strategy that is present from birth (in the form of focal attention and object perception), and that this "reflectance" motivation goes beyond object manipulation to exploration of the social environment, in order to understand and interact effectively with others (Cialdini & Trost, 1998:155). One source of evidence that people look to when trying to maximize the effectiveness of their social behaviour is the descriptive norm operating in the situation (Cialdini, Kallgren, & Reno, 1991; Cialdini, Reno, & Kallgren, 1990; Schaffer, 1983). Descriptive norms refer to the individual's perception of what other people do in a given situation (Cialdini& Trost, 1998). Watching others allows to obtain information on what is considered normal in some novel or ambiguous contexts (Gilbert, 1995; Stiff, 1994). When the appropriate behaviour is unclear, we tend to rely on "social reality" as displayed by others (Festinger,1954). In addition, other's behaviour provides information about what is "socially approved": the greater the number of people who respond to the same situation in the same way, the greater the perception that a certain behaviour is correct (Thibaut & Keley, 1959).

People are most likely to imitate other's behaviour when the situation is novel, ambiguous or uncertain (Sherif, 1936; Deutsch & Gerard, 1955; Tesser, Campbell, & Mickler, 1983), and especially when the source of reference is similar to them (Festinger, 1954). It is possible to maximize the effectiveness of this tendency modelling the behaviours after others and following those who are not only similar to us, but successful as well. Rather, those who successfully model effective behaviour will have an advantage over those who do not. As noted by Allison (1992: 284), "Imitation may be ubiquitous but it is not indiscriminate" and people should be more likely to imitate those who have visible signs of success, such as wealth, power, or status (Cialdini & Trost, 1998:255).

A series of studies were carried out by Cialdini, Reno and Kallgren (1990) in order to examine the effect of descriptive norms in littering behaviour. To influence the individuals' littering behaviour descriptive norms were manipulated by controlling the amount of littering in a variety of location (e.g., parking garage, parks, etc.). In fact, the experimental environment was either clean (antilittering descriptive norm) or littered (pro-littering descriptive norm). In general, the studies supported the importance of descriptive norms in eliciting norm-consistent behaviour, regardless of whether the norm was either pro- or anti-littering: people tended to litter significantly more into a littered environment than into a clean environment (Cialdini & Trost, 1998).

Other researches demonstrated the important role of social norms as "social proof as information toward right living" (Cialdini &Trost, 1998). Latane & Darley (1968a; 1968b) and Latane & Nida (1981) illustrated how social rules influence bystander intervention during emergencies. Their studies showed that, when alone, bystander always help during emergency; but in the presence of others who are not moving to help, most people withhold assistance (Batson, 1998, in this Handbook; Latane & Nida, 1981). Phillips (1974) identified the power of social proof to legitimize suicide as a method for dealing with the troubles of life (Cialdini & Trost, 1998). The author dubbed the phenomenon the "Werther effect", to underline the effect of Goethe's 1774 novels on suicide behaviour in Europe. In fact, *The Sorrows of Young Werther* spurred a rash of imitative suicides across Europe (Phillips, 1974; Cialdini & Trost, 1998). Similar social phenomena have been noted after the suicide deaths of Marilyn Monroe in 1962 and Yukiko Okada, a popular Japanese singer, in 1986 (UPI, April 23, 1986). Stack (1990) found significant increases after both celebrity and noncelebrity suicides, but larger increases after celebrity suicides. Self-murder is often driven by the desire of imitation (Phillips, 1989) and the social proof of similarity (Cialdini, 1993). The first cause was identified by Phillips (1989), who argued that people imitate famous suicides because the ensuing media coverage demonstrates that, rather than being punished, the deceased is accorded attention and status not conferred in life. On the other hand, Cialdini (1993) underlined that mechanisms of similarity operate during self-murder. In fact, people are most likely to commit a copycat suicide when the precipitating suicide is committed by someone who is similar in age or sex (Schmidtke & Hafner, 1988; Cialdini & Trost, 1998). Another area in which descriptive norms present a potent and potentially damaging is drug and alcohol abuse.

4.1.2. The goal of building and maintaining social relationship

According to Cialdini & Trost (1998), social norms have the power to influence individuals' conducts clarifying the behaviours that are expected of them by those in their social world. In this section, we will examine types of norms that go beyond simply describing appropriate behaviour to *prescribing* it, as well as *proscribing* inappropriate behaviour.

4.1.2.1. Injunctive norms

The "norm" construct is most popularly used to refer to behaviours that are accompanied by social acceptance or approval by others (Allison, 1992; Opp, 1982; Cialdini & Trost, 1998). According to the definition of Cialdini, Kalllgren and Reno (1991:203) injunctive norms

"*characterize the perception of what most people approve or disapprove*" (Cialdini & Trost, 1998). They specify what should be done and represent the moral rules of a certain group or community. Injunctive norms motivate behaviour by promising social rewards or punishments for it. They stimulate or inhibit the implementation of a certain behaviour by promising respectively, remuneration or social punishments depending on the course of action chosen. (Allison, 1992,1982; Cialdini & Trost, 1998). As for the descriptive norms, Cialdini, Kallgren and Reno in 1990, investigated the effect of priming either an antilittering norm or other social rules that varied in their conceptual closeness to littering (e.g. recycling, energy conservation, support local museum). The norm was emphasized by distributing a handbill with a corresponding message (e.g. "April is Keep Arizona Beautiful Month. Please Do Not Litter.") under the windshield wipers of library patron cars. The greater the message's conceptual distance from littering (i.e. support local museums), the highest was the percent of handbills thrown to the ground. An elaborate follow-up study strengthened the priming effect through perceptual narrowing by manipulating physical arousal and found a similar trend in the arousal condition (Kallgren, Cialdini, & Reno, 1989). In general, injunctive norms have increased the capacity to produce a positive behaviour than descriptive ones in many empirical situations.

Injunctive norms are powerful and represent an important tool for changing human behaviours, but to encourage positive behaviours they should attract individuals' attention and induce a conscious process of thinking (Cialdini & Trost, 1998). In addition to injunctive rules, social responsible behaviours are influences by other's approval. For example, Berkowitz and Lutterman (1968) measured individuals' knowledge of norms and found that those who were high on social responsibility tended to do whatever they perceived their social group thought was right, rather than being motivated by a sense of obligation to others in general. Most recent research indicates that injunctive norms need not to be expressed in order to direct behaviour, and that the reward of popularity is sufficiently powerful to elicit even health-threatening behaviour (Cialdini & Trost, 1998; Crandal, 1988).

4.1.2.2. Subjective norms

Knowing how we perceive others' expectations for us is important in understanding how those perceptions influence our behaviour, and Fishbein and Ajzen (1975) argued that in order to predict behaviour, we must also express our willingness to go along with those expectations. As explained above, the concept of subjective norms was developed by Fishbein and Ajzen (1975; Ajzen & Fishbein, 1980) in the TRA model (page 40). They argued that part of the difficult in predicting behaviour from attitudes was that an important determinant of

behaviours, and behavioural intensions, combined with attitude valence: subjective norms. A subjective norm is "the person's perception that most people who are important to him think he should or should not perform the behaviour in question" (Fishbein & Ajzen, 1975: 302).

A supplementary research material on the theory of conformity is provided in appendix 4.

4.2. Socio cognition models

During the early 1950s, academic social psychologists were developing an approach to understanding behaviour that grew from learning theories derived from two major sources: Stimulus Response (S-R) Theory (Watson, 1925) and Cognitive Theory (Lewin, 1951; Tolman,1932). S-R theorists believed that learning results from events (termed reinforcements) reduce physiological drives that activate behaviour. Skinner (1938) formulated the widely-accepted hypothesis that the frequency of a behaviour is determined by its consequences or reinforcement. For Skinner, the mere temporal association between a behaviour and an immediately following reward was regarded as sufficient to increase the probability that the behaviour would be repeated. Cognitive theorists, however, emphasize the role of subjective hypotheses and expectations held by individuals, believing that behaviour is a function of the subjective value of an outcome and of the subjective probability, or expectation, that a particular action will achieve that outcome. Mental processes such as thinking, reasoning, hypothesizing, or expecting are critical components of all cognitive theories. Cognitive theorists believe that reinforcements operate by influencing expectations about the situation rather than by influencing behaviour directly (Glanz et al., 2008).

A brief description of the cognitive models considered during the study will follow.

4.2.1. The Theory of Planned Behaviour

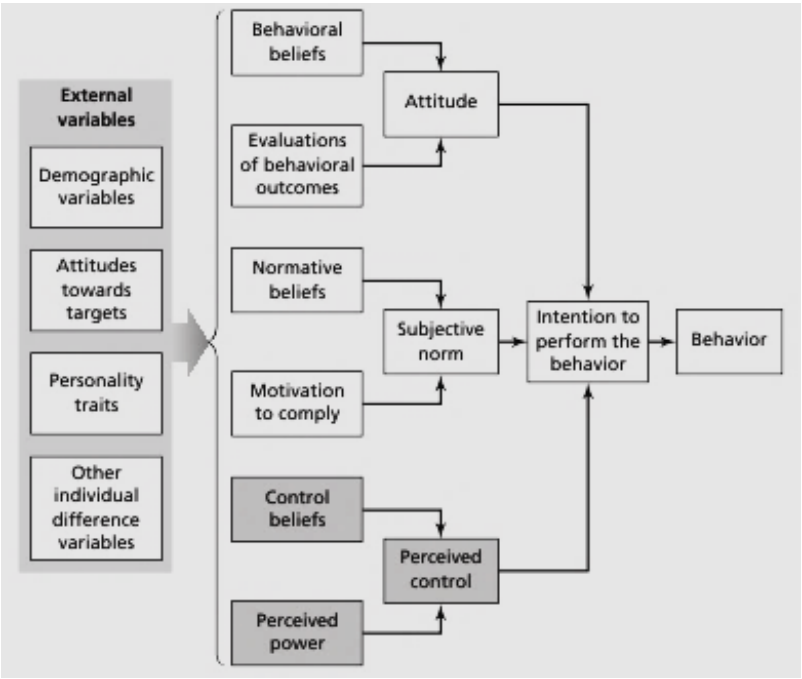
The theory of Planned behaviour (TPB) is an extension of the Theory of Reasoned Action (TRA), proposed in 1980 by Ajzen and Fishebein, and defined the individual motivation factors related to the likelihood of performing a specific behaviour (Glanz et al, 2008). TRA and TPB assumed that the strongest predictor of any given behaviour is the behavioural intention, which in turn is determined by attitudes toward the behaviour and social normative perception regarding it (Ajzen, 1991). The TRA was intended to be applied to the prediction of purely volitional behaviours but, as Ajzen (1988) later argued, many behaviours are not under

complete volitional control. He therefore expanded the TRA by adding the concept of perceived behavioural control, which refers to people's appraisals of their ability to perform the behaviour and originates from self-efficacy theory (SET) (Ajzen, 1991; Ajzen and Driver, 1991; Ajzen and Madden, 1986; Glanz et al., 2008). The self-efficacy concept came from social cognitive theory and was proposed for the first time by Bandura in 1977. According to Ajzen (1988), perceived behavioural control should predict behavioural intention and, when people's perceptions of control accurately reflect their control over behaviour, it should predict actual performance of the behaviour too (Rutter & Quine, 2002; Glanz et al., 2008). In recent years, Fishbein and colleagues have further expanded TRA and TPB to include components from other major behavioural theories and have proposed use of an Integrated Behavioural Model (IBM) (Glanz et al, 2008).

4.2.1.1. The Theory of Reasoned Action and the Theory of Planned Behaviour

TRA and TPB state that intentions are direct antecedent of behaviour. In turn, intentions depend on the attitude toward the behaviour and on the subjective norms associated with it. TPB adds perceived control over the behaviour, taking into account situations where one may not have complete volitional control over a behaviour (Glanz at al. 2008). To be thorough, the main model constructs follow and Figure 9 illustrates the relationships among dimensions.

Figure 9- Theory of Reasoned Action (light area) and Theory of Planned Behaviour (entire figure)



(Source: Glanz et al, 2008)

i. *Attitudes:* is the product of a set of salient beliefs about the consequences of performing the, each weighted by an evaluation of the importance of each of the consequences (Rutter & Quine, 2002; Glanz et al., 2008). Thus, a person shows a positive attitude toward a certain behaviour, if he/she holds strong beliefs that positively valued outcomes will result from performing that particular action. Conversely, we said that a person has a negative attitude if he/she thinks that the behaviour will result in negative outcomes (Glanz et al., 2008).

ii. *Social norms*: Fishbein and Ajzen (1975, Ajzen & Fishbein, 1980), argued that part of the difficulty in predicting behaviours from attitudes was that, an important determinant of behaviours and behavioural intentions, had to be combined with attitude valence: subjective norms (Cialdini & Trost, 1998). A subjective norm is *the person's perception that most people who are important to him think he should or should not perform the behaviour in question* (Fishbein &Ajzen, 1982:302). As regard to how explained previously, the norm dimension is related with two important dimensions, the individual perception of social normative pressure which describes how others' beliefs (*normative beliefs*) incentive or not the behaviour in question and how the individual is *motivated to comply* with those referents (Rutter & Quine, 2002; Glanz et al.2008). Referents or significant others are individuals whose preferences about a person's behaviour in this domain are important to him or her (Corner and Armitage, 1998). A person, who believes that certain referents think she/he should perform a behaviour and he/she is motivated to meet the expectations of those referents, will hold a positive subjective norm. Conversely, a person, who believes that these referents think she should not perform the behaviour and he/she is not motivated to meet the expectations of those referents, will have a negative subjective norm. Finally, a person that is not interested in complying with any of the important references, it is said to have a neutral subjective norm. (Rutter & Quine, 2002; Glanz et al., 2008).

iii. *Perceived behavioural control*: the perceived control was added to the original TRA model in order to account for factors, outside individual control, that may affect intentions and behaviours (Glanz et al, 2008). Perceived behavioural control refers to the person's perceptions of their ability to execute the behaviour. It is similar to Bandura's (1986) construct of self-efficacy (see 'Perceived control' and 'Self-efficacy and health') and it is determined by the total set of accessible control beliefs that indicate the presence or absence of factors that may facilitate or impede the behaviour achievement. The Ajzen's idea (1991) to include the perceived control on the original TRA model was based on the conviction that the behavioural performance was jointly affected by two principal factors, motivation (intention) and capacity

(control over behavioural performance). TPB also postulates that perceived control is an independent determinant of behavioural intention, along with attitude toward the behaviour and subjective norm. Holding attitude and subjective norm constant, a person's perception of the ease or difficulty of behavioural performance will affect his behavioural intention (Glanz et al, 2008). Ajzen's (1985) early presentations of the TPB suggested that PBC and intentions would interact in their predictions of behaviour such that intentions would become stronger predictors of behaviour as PBC increased (Corner and Armitage, 1998).

iv. *Intention*: According to Ajzen (2002b), intention is *the cognitive representation of a person's readiness to perform a given behaviour*, and is considered to be *the immediate antecedent of behaviour* (Ajzen, 1985, 1987,1991; Glanz et al., 2008). Intentions capture the motivational factors that influence a behaviour and are indications of how hard people are willing to try in order to perform the behaviour (Ajzen, 1985, 1987,1991). Ajzen (1991) stated that "*the stronger the intention to perform the behaviour, the more likely should be its engagement*" (Ajzen, 1991:181). Intentions and behaviour are held to be strongly related when measured at the same level of specificity in relation to the action, target, context, and time frame (Fishbein & Ajzen, 1975; principle of compatibility; Corner and Armitage, 1998) and when the time interval is short enough to ensure that intentions have not changed (Randall & Wolff, 1994; Corner and Armitage, 1998). In summary, intentions are direct subsequent of attitudes, subjective norms and perceived behavioural control and the influence of these predictors on this dimension change in respect to the behaviour in question and the population under study.

4.2.2. The Health Belief Model

Since the early 1950s, the Health Belief Model (HBM) has been one of the most widely used conceptual frameworks in health behaviour research, both to explain change and maintenance of health-related behaviours and as a guiding framework for health behaviour interventions. The Health Belief Model (Rosenstock 1966, 1974a, 1974b) proposes that people will be motivated to carry out preventive health behaviours in response to a perceived threat to their health. It was developed in the 1950s as a way to explain why medical screening programs offered by the US Public Health Service, particularly for tuberculosis, were not successful (Hochbaum, 1958; Rosenstock, 1960, 1974; Glanz et al.,2006). Later, the model evolved gradually in response to very practical public health concerns, such as the study of people's responses to symptoms (Kirscht, 1974; Glanz et al., 2006) and adherence to medical regimens (Becker, 1974; Glanz et al., 2006).

In this model, two classes of variable are important, the psychological *state of readiness* to take a specific action and the extent to which a particular course of action is believed to be beneficial in reducing the threat (Rosenstock, 1966:98, Rutter & Quine, 2002). According to Rosenstock, both variables are characterized by two main dimensions. The *state of readiness to act* depends on individuals' perception of personal *susceptibility* or *vulnerability* to a particular health threat, and on the perceptions of the *severity* with which that threat might affect their life (Rutter & Quine, 2002). On the other hand, the extent to which a course of action is believed to be beneficial is the results of beliefs about the *benefits* to be gained by a particular action weighted against the *costs* or *barriers* to action (Rutter & Quine, 2002:10). Rosenstock argued that the level of readiness is related with the motivation to act and that, the perception of barriers and benefits determine which path of action would be chosen. Therefore, individuals are most likely to follow a particular health action if they believe themselves to be susceptible to a particular condition which they also consider to be serious and believe that the benefits outweigh the costs of the action taken to counteract the health threat (Steptoe et al., 2010)

However, the combination of these could reach considerable levels of intensity without resulting in overt action unless some instigating event occurred to set the process in motion or trigger action in an individual psychologically ready to act (Rosenstock 1966: 102; Rutter & Quine, 2002:10). The consideration above led Rosenstock to introduce a new dimension to the original model, the *cues to action*. Cues to action are assumed to include a diverse range of triggers to the individual taking action which may be internal (e.g., physical symptom) or external (e.g., mass media campaign, advice from others) to the individual (Janz and Becker, 1984; Steptoe et al., 2010). More recently, other constructs are added to the HBM; thus, the model has been expanded to include also the self-efficacy dimension and some motivating factors. To be thorough, the main model constructs follow and Figure 10 illustrates the relationships among constructs.

4.2.2.1. HBM constructs

i. *Perceived Susceptibility* ("How likely am I to get ill?") refers to beliefs about the likelihood of getting a disease or condition.

ii. *Perceived Severity* ("How serious would the illness be?") refers to the feelings about the seriousness of contracting an illness or of leaving it untreated, include evaluations of both medical and clinical consequences (for example, death, disability, and pain) and possible social

consequences (such as effects of the conditions on work, family life, and social relations). The combination of susceptibility and severity has been labelled as *perceived threat*.

iii. *Perceived Benefits*: refer to an individual's assessment of the efficacy of a certain behaviour in reducing the risk of a possible disease (Janz et al., 1984). If an individual believes that a particular action will reduce his/her susceptibility to a certain health disease or decrease its seriousness, he/she is likely to engage the behaviour (Rosenstock,1974). Other non-health-related perceptions, such as the financial savings related to quitting smoking or pleasing a family member by having a mammogram, may also influence behavioural decisions. Thus, individuals exhibiting optimal beliefs in susceptibility and severity are not expected to accept any recommended health action unless they perceived the behaviour as potentially beneficial in reducing the threat (Glanz et al., 2006).

iv. *Perceived Barriers*: refer to the potential negative aspects of a particular health action that may act as impediments to undertaking recommended behaviours. A kind of nonconscious, cost-benefit analysis occurs wherein individuals weigh the action's expected benefits with perceived barriers. Thus, combined levels of susceptibility and severity provide the energy or force to act and the perception of benefits (minus barriers) provide a preferred path of action(Rosenstock,1974). Even if an individual perceives a health condition as threatening and believes that a particular action will effectively reduce the threat, barriers may prevent engagement in the health-promoting behaviour. In other words, the perceived benefits must outweigh the perceived barriers for behaviour change to occur (Janz et al., 1984; Glanz et al., 2008).

v. *Cues to Action*: include a diverse range of triggers including individual perceptions of symptoms, social influence (doctor advices) and health education campaigns. Various early formulations of the HBM included the concept of cues that can trigger actions. Hochbaum (1958), for example, thought that readiness to take action (perceived susceptibility and perceived benefits) could only be potentiated by other factors, particularly by cues to instigate action, such as bodily events, or by environmental events, such as media publicity. He did not, however, study the role of cues empirically. Although the concept of cues as triggering mechanisms is appealing, cues to action are difficult to study in explanatory surveys.

Figure 10- The Health Belief models

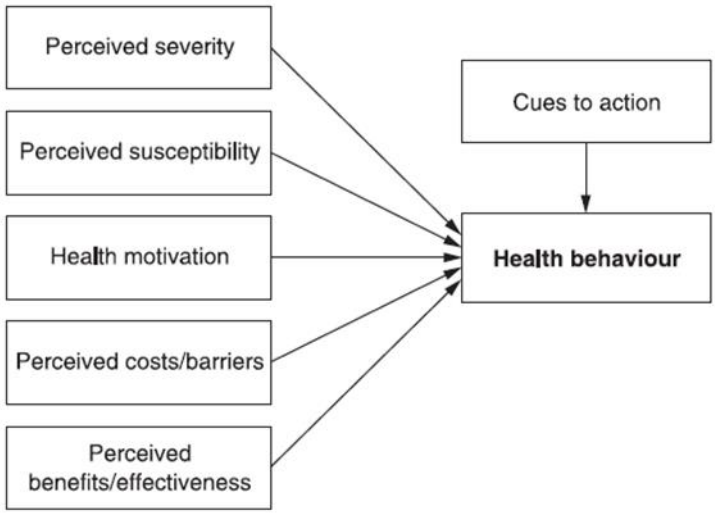


Fig 1 The Health Belief Model.

4.3. Psychological theories and health behaviours

The health belief model (HBM) and the theory of planned behaviour are two of the most widely used theories to explain and predict health behaviours. The HBM and TPB were compared in predicting college women's HPV vaccine uptake or intention in two separate studies; both studies reported that the TPB consistently outperformed the HBM based on amount of variance explained (Bennett, Buchanan & Adams, 2012; Gerend & Shepherd, 2012). Bennett et al. (2012) determined that the TPB explained 52% of the variance in intentions to vaccinate while Gerend et al. (2012) reported that the TPB explained 39% of the variance in vaccine uptake. When HBM and TPB constructs were combined into one model, the HBM constructs only accounted for an additional 4% beyond that explained by the TPB constructs (Gerend et al., 2012). Fisher et al. (2013) tested the TPB in predicting HPV vaccination intentions of college-age students, and found that attitudes and social norms accounted for 53% and 44% of the variance in women's and men's vaccination intentions, respectively. Based on these findings, the combination of TPB and HBM appears to be the most useful framework for predicting HPV vaccination intentions among university students.

As Ayers et al (2008) indicates, the TPB presents a number of advantages over the traditional 'social cognition model'. First, the model is a general theory, and according to Stroebe (2000) general theories should be preferred to health or behaviour specific model for reason of

parsimony. Secondly, unlike the HBM, the constructs and their causal relationship are clearly specified and there exist an exhaustive literature, provided by the author, describing how the dimensions should be operationalized. Finally, many studies (meta-analysis and observational) show that the model explains a good amount of intentions and behaviour variance. Evidence regarding the models dimensions efficacy in predict behaviour comes from hundreds of studies that have been summarized in several meta-analyses and reviews (Rutter & Quine, 2002;Armitage and Conner, 2001; Albarracin, Johnson, Fishbein, and Muellerleile, 2001; Albarracin and others, 2003; Albarracin, Kumkale, and Johnson, 2004; Albarracin and others, 2005; Downs and Hausenblas, 2005; Durantini and others, 2006; Hardeman and others, 2002; Sheeran and Taylor, 1999; Webb and Sheeran, 2006). TRA and TPB have been used successfully to predict and explain a wide range of health behaviours and intentions, including smoking, drinking, health services utilization, exercise, sun protection, breastfeeding, substance use, HIV/STD-prevention behaviors and use of contraceptives, mammography, safety helmets, and seatbelts (Rutter & Quine, 2002;Albarracin, Fishbein, and Goldestein de Muchinik, 1997; Albarracin, Johnson, Fishbein, and Muellerleile, 2001; Bandawe and Foster, 1996; Bosompra, 2001; Bogart, Cecil, and Pinkerton, 2000; Fishbein, 1993; Montaño and Taplin, 1991; Morrison, Spencer, and Gillmore, 1998; Steen, Peay, and Owen, 1998; Trafimow, 1996). Findings have been used to develop many effective behaviour change interventions (Fishbein, 1990; Fisher, Fisher, and Rye, 1995; Gastil, 2000; Hardeman and others, 2005; Jemmott, Jemmott, and Fong, 1992; Jemmott and Jemmott, 2000; Rutter & Quine, 2002). Meta-analytic reviews of the TPB provide strong support for the predictive validity of the TPB in terms of the percentage of variance explained in behaviour and intentions by the components of the TPB (Godin & Kok, 1996; Sutton, 1998).

There are two main quantitative reviews of the HBM model, the one realized by Janz and Becker in 1984 and the second proposed by Harrison et. al in 1992. Different strategies in quantifying findings were adopted by the authors. Janz and Beckers's study provided a summary of 49 HBM related investigators, 18 prospective and 28 retrospectives (Janz and Becker, 1984). Among the included, twenty-four examined preventive health behaviours (PHB) while nineteen considered sick role behaviours (SRB). Summary results have shown how the different constructs differs in predicting health behaviours of different nature. Perceived barriers were the most powerful single predictor across all studies and behaviours. Although both perceived susceptibility and perceived benefits were important overall, the first one was a stronger predictor of preventive health behaviour than sick-role behaviour. The reverse was true for perceived benefits. Overall, perceived severity was the least powerful predictor; however, this dimension was strongly related to sick-role behaviour (Glanz et al.,

2008). Harrison and colleagues originally identified 234 published empirical tests of the HBM. Of these, only 16 studies (i.e. 6.8%) measured all four major components and included reliability checks. Considering a Perason's r index, the correlation between the HBM components and the health behaviour was analyzed for each study. Although all the correlations tested were statistically significant, their effect size was very small. Across studies, individual constructs accounted for a small percentage (from 0.5% to 4%) of the health behaviour variance. Unlike Janz and Becker (1984), Harrison et al. found that HBM components had different associations in cross-sectional versus longitudinal designs. Both benefits and barriers had significantly larger effect sizes in prospective than in retrospective research, whereas in the case of severity, the effect size was significantly larger in retrospective studies (Cambridge book). Orbell et al (1995) reported perceived susceptibility and barriers to entirely mediate the effects of social class upon uptake of cervical screening. The HBM has also inspired a range of successful behavior change interventions (e.g., Jones et al, 1987). The main strength of the HBM is the common-sense operationalization it uses including key beliefs related to decisions about health behaviors. However, further research has identified other cognitions that are stronger predictors of health behaviour than those identified by the HBM, suggesting that the model is incomplete. This prompted a proposal to add self-efficacy and intention to the model to produce an "extended health belief model" (Rosenstock et al, 1988) which has generally improved the predictive power of the model (e.g., Hay et al, 2003).

4.4.The questionnaire

In this study, the instrument used was newly developed, as no gold standard questionnaire currently exists on the topic of HPV vaccination and cervical cancer prevention among Italian university women. The most important study on HPV vaccination in Italy was founded by the Italian Ministry of Health (ProgettoValore), coordinated by the National Institute of Public Health (Istituto Superiore di Sanità-ISS) and carried out between November 2011 and July 2012. The survey addressed families of unvaccinated girls born in 1997 or 1998 who were offered vaccination in 2008-2010. In the study, a 23-item questionnaire was developed in order to investigate demographic information, vaccination status, barriers/reasons for non-vaccination, HPV knowledge, source of information about HPV and future intention about HPV vaccination (Giambi et al., 2014). Although the ISS research brought important information on HPV vaccination refusal, the instrument used was developed without the support of any behavioural theory. There are many Italian, European and international studies which aim to identify HPV vaccination uptake and cervical cancer determinants. However, by considering that the HPV vaccine is recommended to 11 and 12-year old girls before they engage in any skin-to-skin sexual contact, most studies have focused on parents' attitudes and beliefs for vaccinating their daughters (Feiring et al., 2015; Ogivile et al., 2010; Rondy et al., 2010; Gefenaite et al., 2013). In addition, some of them are not supported by any theoretical background (Zimet et al., 2010; Ogilvie et al., 2010) and others do not investigate HPV vaccination and cervical cancer at the same time (Zimet et al., 2010; Zhuang et al., 2016). Those characteristics have made the existing questionnaires unsuitable to achieve our research objectives. In addition, social influence and HBM/TPB dimensions are often strictly connected with the culture in which they originate and operate (Cialdini & Trost, 1998). For this reason, a new questionnaire structured according to the theory of planned behaviour, the health belief model and the social influence theory was specifically designed for this project. As explained above, to obtain valid questions, behavioural items formulation was supported by an accurate and extensive literature review. Published measured and original question as well as cultural and social aspects of respondents' environment were balanced to create a proper survey instrument.

First, the questionnaire was developed by filters and questions available only for group of respondents, i.e. answer(s)/section(s) based on previous question(s). For example, section e (Appendix 5) was specifically created to investigate unvaccinated respondents. These sections allowed us to obtain specific information for some population subgroups (i.e. sexually active, vaccinated and unvaccinated) that may be found as important for policy interventions.

Research objectives were measured by different questions. Past vaccination choices were assessed by the question "Are you vaccinated?" (Question 38-Appendix 5) which was characterized by three possible answers, "Yes", "No" and "I don't remember". Future preventive behaviours instead, were assessed by questions concerning intentions, intentions of future HPV vaccination for unvaccinated respondents (Section e-Appendix 5) and future intentions of undergo a pap-test for the rest of the sample (Section h-Appendix 5).

Intentions is a latent dimension belonging to the TPB. According to Fishbein and Ajzen's (1975) TRA/TPB, intentions represent the direct antecedent of human behaviour. Authors' guidelines underlying that intentions and behaviour should be investigated separately, i.e. two surveys should be realized in order to identify the amount of behaviour variance predicted by previous intentions. In general, intentions appear before then behaviour and this time gap should be respect during data collection. For this purpose, this survey investigated intentions and its determinants. The realized behaviour will be inquired in a future survey. Further behavioural questions were formulated on the basis of two of the six domains of the Health Belief Model [the perceived susceptibility (to develop cervical cancer in futture) and the perceived severity of the disease (cervical cancer)] and all the remaining constructs of the Theory of Planned Behaviour (Attitudes, Subjective Norms, Self-efficacy). Descriptive norms question was based on the Social Influence theory. For the TPB domains direct measurement were preferred.

All the items used to measure psychological dimensions are summarized in tables 4 and 5. The tables presented as well as the subgroup of interest and the type of measurement scales/questions considered.

Table 4 HBM general constructs and vaccination TPB constructs

Psychological theory	Item	Subgroup	Dimension measured	Scale/question
Health belief model	I believe HPV can be extremely harmful (Credo che il tumore al collo dell'utero sia una malattia molto grave)	Respondents that know HPV	Perceived severity of disease 1	5-point Likert scale
	The thought of get cervical cancer in future scares me (Mi spaventa pensare che potrei ammalarmi di cancro al collo dell'utero in futuro)	Respondents that know HPV	Perceived severity of disease 2	5-point Likert scale
	Chance of a girl to get cervical cancer in the future without vaccination and pap-test (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che non si vaccinate e non si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	Respondents that know HPV	Perceived susceptibility 1	Thermometer scale (0%-100%)
	Chance of a girl to get cervical cancer in the future if vaccinated but not undergoing a pap-test (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che sia vaccinata ma non si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	Respondents that know HPV	Perceived susceptibility 2	Thermometer scale (0%-100%)
	Chance of a girl to get cervical cancer in the future if unvaccinated but undergoing a pap-test (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che non si vaccinata ma si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	Respondents that know HPV	Perceived susceptibility 3	Thermometer scale (0%-100%)
	Chance of a girl to get cervical cancer in the future if vaccinated and undergoing a pap-test (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che sia vaccinata e si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	Respondents that know HPV	Perceived susceptibility 4	Thermometer scale (0%-100%)
Theory of planned behaviour	I believe that the vaccine is effective in preventing HPV infection (Credo che il vaccino sia efficace nel proteggere contro l'infezione da HPV)	Respondents that know HPV	Vaccination Attitudes/Beliefs 1	5-point Likert scale
	I believe that the vaccine distribution is strongly influenced by pharmaceutical companies. (Credo che la distribuzione del vaccino da parte dello stato sia influenzata dalle case farmaceutiche)	Respondents that know HPV	Vaccination Attitudes/Beliefs 2	5-point Likert scale
	I am very worried about the long-term	Respondents	Vaccination	5-point Likert

	side effects of the HPV vaccination (Ho paura che il vaccino possa causare problemi di salute anche molto tempo dopo la vaccinazione)	that know HPV	Attitudes/Beliefs 3	scale
	I believe that the vaccine induces only partial-protection against the risk of getting infected by HPV (Credo che la vaccinazione fornisca una protezione parziale dal rischio di contrarre l'HOV in eventuali rapporti non protetti)	Respondents that know HPV	Vaccination Attitudes/Beliefs 4	5-point Likert scale
	I believe that the protection induced by the vaccine is more important than its side effects (I credo che la protezione offerta dal vaccino sia più importante dei suoi eventuali effetti collaterali)	Respondents that know HPV	Vaccination Attitudes/Beliefs 5	5-point Likert scale
	I believe that the vaccine is safe. (Io credo che il vaccino sia sicuro)	Respondents that know HPV	Vaccination Attitudes/Beliefs 6	5-point Likert scale
	Girls do not need Pap tests after having the HPV vaccine (Credo che chi si vaccina contro l'HPV non abbia più bisogno di sottoporsi all'esame del pap-test in futuro)	Respondents that know HPV	Vaccination Attitudes/Beliefs 7	5-point Likert scale
	I think I can deal alone (take a decision without the help of my parents) my health-related choices. (Io penso di poter gestire da sola (senza l'ausilio dei miei genitori) le scelte che riguardano la mia salute)	Respondents that know HPV	Behavioural Control	5-point Likert scale
	Premarital sex is not allowed in my religion and I do not need HPV vaccination. (La mia religione non consente rapport prematrimoniali quindi non ho bisogno del vaccino)	Unvaccinated respondents	Vaccination Attitudes/Beliefs 8	5-point Likert scale
	I think I should get more information before deciding about vaccination. (Credo che dovrei informarmi meglio prima di prendere una decisione sulla vaccinazione anti-HPV)	Unvaccinated respondents	Vaccination Attitudes/Beliefs 9	5-point Likert scale
	I am against vaccination in general. (Sono contraria a tutte le vaccinazioni)	Unvaccinated respondents	Vaccination Attitudes/Beliefs 10	5-point Likert scale
	I am worried about post-vaccination side effects. (Mi spaventano i possibili effetti collaterali post-vaccinazione)	Unvaccinated respondents	Vaccination Attitudes/Beliefs 11	5-point Likert scale
	Which person's opinion may influence your future choice about HPV vaccination? (Se domani ti venisse riproposta la vaccinazione, il parere di quali persone potrebbe influenzare la tua scelta	Unvaccinated respondents	Subjective norm (normative beliefs)	Multiple choice

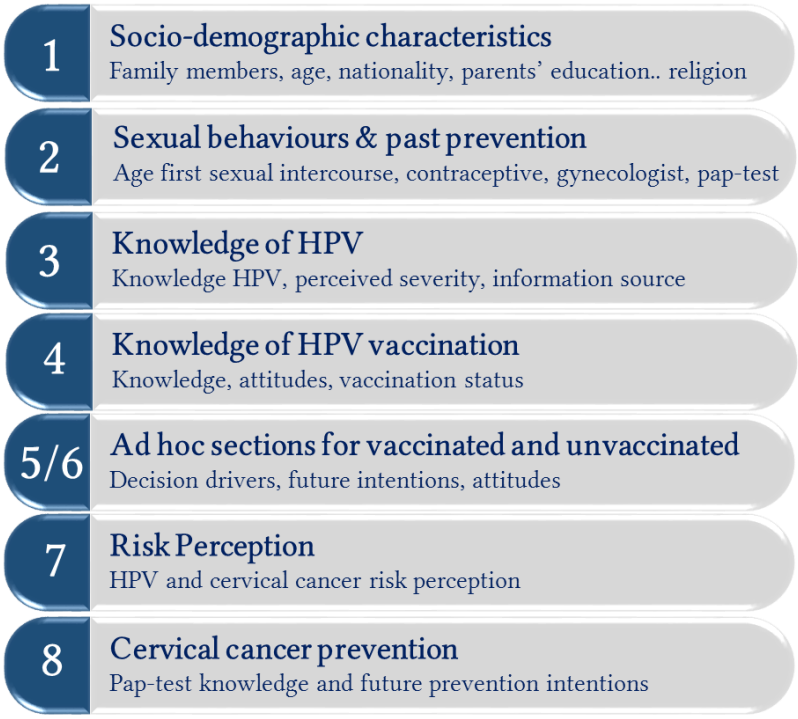
	futura?)			
	Which person would support your decision about HPV vaccination? (Quali di queste persone sarebbe a favore di una rua eventuale vaccinazione?)	Unvaccinated respondents	Subjective norm (motivation to comply)	Multiple choice

Table 5 Pap-test TPB constructs

Psychological theory	Item	Subgroup	Dimension measured	Scale
Theory of planned behaviour (Pap-test items)	Having a regular pap screening is helpful to avoid cervical cancer in future. (Effettuare il pap-test regolarmente fa diminuire la possibilità che mi ammali di cancro cervicale in future)	Respondents that know HPV and Pap-test	Pap-test Attitudes/Beliefs 1	5-point Likert scale
	I am scared to have a pap-test because I might have an abnormal result. (Ho paura di effettuare il pap-test perché mi spaventerebbe un eventuale risultato negativo)	Respondents that know HPV and Pap-test	Pap-test Attitudes/Beliefs 2	5-point Likert scale
	I am too young to undergo a pap-test. (Credo di essere troppo giovane per dover effettuare l'esame del pap-test)	Respondents that know HPV and Pap-test	Pap-test Attitudes/Beliefs 3	5-point Likert scale
	I am embarrassed when I have a pap test. (L'esame del pap-test mi mette a disagio)	Respondents that know HPV and Pap-test	Pap-test Attitudes/Beliefs 4	5-point Likert scale
	Pap testing is helpful for early detection of cervical cancer. (Credo che grazie al pap-test eventuali forme pre-cancerogene possano essere trattate precocemente)	Respondents that know HPV and Pap-test	Pap-test Attitudes/Beliefs 5	5-point Likert scale
	Which person's opinion may influence your future choice about pap-test? [Se domani ti venisse riproposta la vaccinazione, il parere di quali persone potrebbe influenzare le tue scelte di prevenzione futura (pap-test) del tumore al collo dell'utero?]	Respondents that know HPV and Pap-test	Subjective norm (normative beliefs)	Multiple choice
	Which person would support your decision to undergo a pap-test? (Quali di queste persone sarebbe sarebbero d'accordo con te se tu effettuassi il pap-test in futuro?)	Respondents that know HPV and Pap-test	Subjective norm (motivation to comply)	Multiple choice
	How many of your friends undertook a pap-test examination? (Quante delle tue amiche hanno già effettuato il pap-test?)	Respondents that know HPV and Pap-test	Descriptive norm	Multiple choice
	How likely do you undergo a pap-test in future? (Quanto è probabile che ti sottoponga all'esame del pap-test in futuro?)	Respondents that know HPV and Pap-test	Pap-test intentions	Multiple choice
	How many times do you undergo a pap-test in future?	Respondents that know	Pap-test Intentions	Multiple choice

A cover letter, containing a brief description of research objectives and the privacy protection norms, was added in the first page of the instrument. The regional etic commission authority consensus was obtained for the privacy protection. The final questionnaire was composed by 63 questions and eight sections: Socio-Demographic characteristics, Sexual Behaviour and past prevention, Knowledge of HPV, Knowledge of HPV vaccination, Ad hoc Section for vaccinated and unvaccinated (Determinants and Attitudes), Risk Perception, Cervical Cancer Prevention (Figure 11). A description of each section will follow whereas the whole questionnaire is shown in appendix 5.

Figure 11 Questionnaire sections



Socio-Demographic characteristics. The demographic section consists of 13 questions investigating, respectively, the academic year of respondents' enrolment, the family components and their age, the citizenship, the province of residence when the respondent was 16, the parents' level of education and religion, the respondent's religion, the religious practice of parents and respondent, the smoke behaviour and the daily number of cigarettes smoked. It was asked for the province of residence when the interviewed was 16 because the most part of respondents fell in the regional catch-up vaccination programme. The academic year of enrolment and the age were measured in years. Citizenship, residence, level of education and

religious categories were sourced from the Italian national institute of statistics (ISTAT) "Multiscopo" questionnaire.

Sexual behaviour and past prevention. The sexual behaviour and past prevention section consists of 15 questions investigating, respectively, sexual debut, age at first sexual intercourse, number of sexual partners in the last 12 months, proportion of occasional sexual partners in the last 12 months, number of sexual partners in the entire life, type of contraceptives used, knowledge of contraceptives that protect against sexual transmitted diseases (STD), sources of sexual information, past gynaecologist visit and pap-test, reasons of past visit and pap-test, childhood vaccinations. The age at first intercourse was measured in years. All other variables were characterized by multiple choice questions.

Knowledge of HPV. The Knowledge of HPV section consists of 4 questions investigating, respectively, knowledge of HPV, perceived severity of the diseases, and HPV information sources. A new HPV knowledge scale was defined, composed by 14 statements (Question 31-Appendix 5) regarding established facts about HPV and HPV risk factors. Each statement was characterized by three answer options, 'Yes', 'No', or 'I don't know'. The 'I don't know' option was included to reduce or prevent participants from guessing; participants were explicitly instructed to choose this option if they would consider their answer to be a guess. Perceived severity, as explained above, was one of the *health belief model* dimensions considered in the questionnaire and was assessed with two questions measured by a 5-point Likert scale. Sources of HPV information were measured by a multiple-choice question.

Knowledge of HPV vaccination. The Knowledge of HPV Vaccination section consists of 6 questions investigating, respectively, knowledge of HPV vaccinations, attitudes towards HPV vaccinations, self-efficacy, vaccination status and vaccination age. HPV vaccine knowledge was assessed through 3 multiple choice questions investigating, respectively, if the respondents had ever heard about HPV vaccination, the efficacy of the vaccination if administered to sexual active woman and the protection conferred by the vaccine. Attitude and self-efficacy items were newly developed according to the Theory of Planned Behaviour, as explained above. An ad hoc scale was created in order to adapt TPB dimensions to the studied population. Attitudes were assessed by 7 items and self-efficacy by a single item (Tables 4 and 5). All the items were measured by a 5-point Likert scale. The sources of information were measured by a multiple-choice question. The vaccination status was measured by a 3-fold questions investigating if the respondent was vaccinated, unvaccinated or did not remember her status. Age of vaccination was measured in years.

Unvaccinated ad hoc section (Decisions and attitudes towards HPV vaccination). The Unvaccinated ad hoc section consists of 7 questions investigating, respectively, the determinants of non-vaccination, who made the decision about vaccination, the intentions towards a possible future HPV vaccination, social norms, and attitudes towards vaccination. Determinants of non-vaccination, vaccination decision-maker and intentions were assessed considering multiple-choice questions. Both, normative beliefs and motivation to comply (subjective norms) were measured by two multiple-choice questions. A special scale to measure attitude to HPV vaccination was developed for non-vaccinated women. The scale is characterized by 4 items measured by a 5-point Likert scale. All the items investigated were written according to the TPB principles.

Vaccinated ad hoc section (Decisions and attitudes towards HPV vaccination). The Vaccinated ad hoc section consists of three questions investigating, respectively, who made the decision about vaccination, social norms and vaccination collateral effects. All the variables were measured by multiple-choice questions.

Risk Perception. The Risk Perception section consists of 4 questions investigating the perceived risk to develop cervical cancer during life. Each question presents a specific scenario characterized by different combinations of HPV vaccination and screening through pap-test. Specifically, the first question asks what the possibility to develop cervical cancer without vaccination and pap-test is, the second without vaccination but by undergoing a pap-test, the third without pap-test but by being vaccinated, and the fourth considering both prevention methods, pap-test and vaccination. For each scenario respondents were asked to indicate in a thermometer scale (ranging from 0% to 100%) the probability of develop cervical cancer in future. These measures are related with the perceived susceptibility dimensions of the health beliefs model.

Cervical Cancer Prevention. The Cervical Cancer Prevention section consists of 9 questions investigating respectively, descriptive norms about pap-tests, knowledge about pap-test, attitudes towards pap-test, future intentions of undergoing a pap-test, additional prevention methods (i.e. gynaecologist visits and use of contraceptive to avoid STD), subjective norms. Descriptive norms, social norms, intentions and additional prevention methods were assessed by multiple choice questions. A new knowledge scale was developed to investigate the level of pap-test awareness. The question consists of 4 items measured by a Yes/No/Don't know scale. To measure attitude scale, as above, a new item scale was defined. The question consists of 5

items measured by a 5-point Likert scale. Subjective norms and attitude constructs were formulated according to the TPB principles. Descriptive norms question has been created following the social influence theory.

4.5.The pilot test

A new questionnaire should be validated. According to that, a pilot study was carried out in spring 2015 at the University of Pisa and Florence. During the test eight courses were involved: the four bachelor courses of statistics of the Department of Economic and Management (University of Pisa) and the four bachelor courses of statistics of the School of Economics and Management (University of Florence). A total of 207 questionnaires were collected during the test. The pilot study showed some issues related to the instrument. Some questions or answer modes were perceived as difficult to interpret from respondents. The number of missing values was quite high for these variables. In particular, the pre-test underlined that the semantic scales used to measure subjective norms were inappropriate. Even if semantic scale represents the standard choice for TPB dimensions, a simpler alternative was chosen. In the final questionnaire, subjective norms were investigated through multiple choice questions. In addition, the use of semantic scales force to increase the number of questions presented. Because the questionnaire length was already high, a simpler and shortest option was preferred. Poor respondent understanding characterized some attitude items as well. For this reason, some items were changed and other eliminated from the final version. Questions investigating sexual behaviours were expanded to collect more precise information in relation to the number of occasional partners.

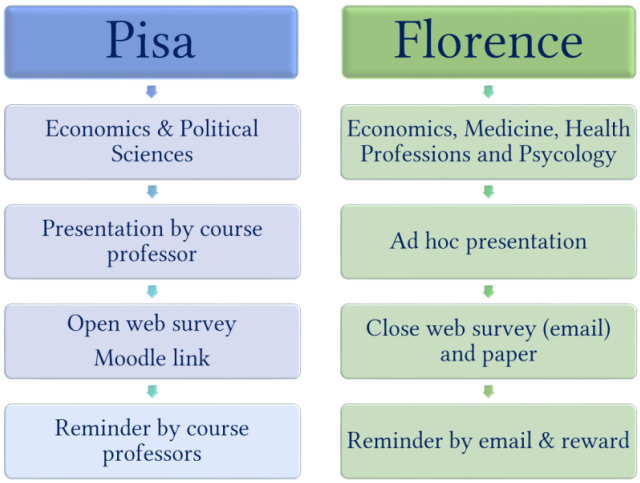
4.6.Questionnaire administration

The sample was collected between December and June 2016 and was selected using non-probability technique (convenience sample). Convenience sampling (also known as Haphazard Sampling or Accidental Sampling) is a type of nonprobability or non-random sampling where members of the target population that meet certain practical criteria, such as easy accessibility, geographical proximity, availability at a given time, or the willingness to participate are included for the purpose of the study (Dörnyei, 2007; Etikan et al., 2016;). It is also referred to the researching subjects of the population that are easily accessible to the researcher (Given Lisa, 2008; Etikan et al., 2016;). Convenience samples are sometimes regarded as ‘accidental samples’ because elements may be selected in the sample simply as

they just happen to be situated, spatially or administratively, near to where the researcher is conducting the data collection. A mixed-mode approach, i.e. online and paper-based questionnaire, was considered in order to reach the largest possible number of respondents. Two university (Pisa and Florence) and five faculties (Economics, Psychology, Medicine, Health Professions and Political Science) were involved in the study (Figure 12). The online administration was done with the support of the open source survey software *Lime Survey*. Different strategies of survey announcement were adopted according to the faculty and the administration method (Figure 12). More specifically, in the Economics and Political Sciences courses (University of Pisa) the survey announcement was done by the course professor, who described the main objective of the research and indicated the compilation method. For these faculties, in fact, the survey link was advertised in the course page of the online Moodle platform.

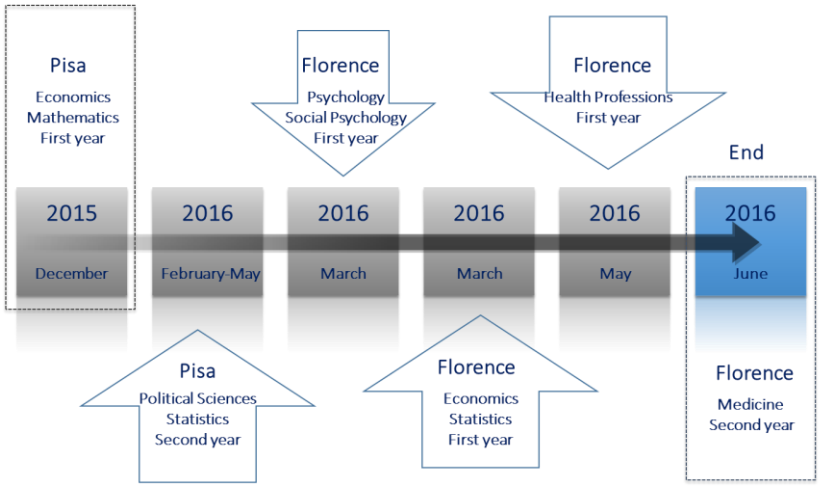
Among the University of Florence courses, a different strategy was adopted. After obtaining a permission from class instructors, verbal announcements were made at the end of the lectures. These announcements, supported by visual slides, consisted of a short description of the study objectives and its relevance. Two different administration methods were adopted for these courses. The students of Psychology, Medicine and Health Professions were reached by email (collected in class after the research announcement) whereas Economics respondents complete the paper questions after the research announcement.

Figure 12 Administration strategies and research announcement



Data were collected between December 2015 and June 2016, according to the timeline showed in Figure 13.

Figure 13 Survey timeline



4.7.Statistical Methods

During the study, different statistical methods were considered to find and answer specific research objectives. The objectives, methods and software considered are summarized in table 6. A brief description of the most important statistical methods considered is provided in appendix 1 and appendix 3.

Table 6- Statistical methods

Objective	Statistical method	Subdataset considered	Software and package
Missing data filling	Multiple imputation by chained equations (Van Buuren, 2012; Azur et al., 2011; Enders, 2010)	Complete dataset (N =432)	MICE R package(Van Buuren & Groothuis-Oudshoorn, 2011)
Determinants of realized immunization choices	Appendix 1. Logistic regression (Agresti,2002; Hosmer and Lemeshow, 2000) Penalized maximum likelihood (Moons et al., 2014) Lasso and elastic-net penalized regressions (Hastie et al., 2001)	Vaccinated and Unvaccinated (N=342)	R packages Stats, rms (Harrell et al., 2017), Glmnet (Friedman et al., 2017)

	ROC curves and Brier's score (Hastie et al., 2001)		
Psychological scale	Factorial analysis	Vaccinated and Unvaccinated that knew pap-test (N=351) Unvaccinated respondents (N=85)	R package Psych (Revelle,2017)
Interrelationship between psychological constructs	Appendix 3. Partial least square structural equation models (Hair et al.,2014). Handbook of Partial Least Squares (Vinzi et al., 2010). PLS Path Modelling with R (Sanchez, 2013).	Vaccinated and Unvaccinated that knew pap-test (N=351) Unvaccinated respondents (N=82)	R package plspm (Sanchez et al., 2017)

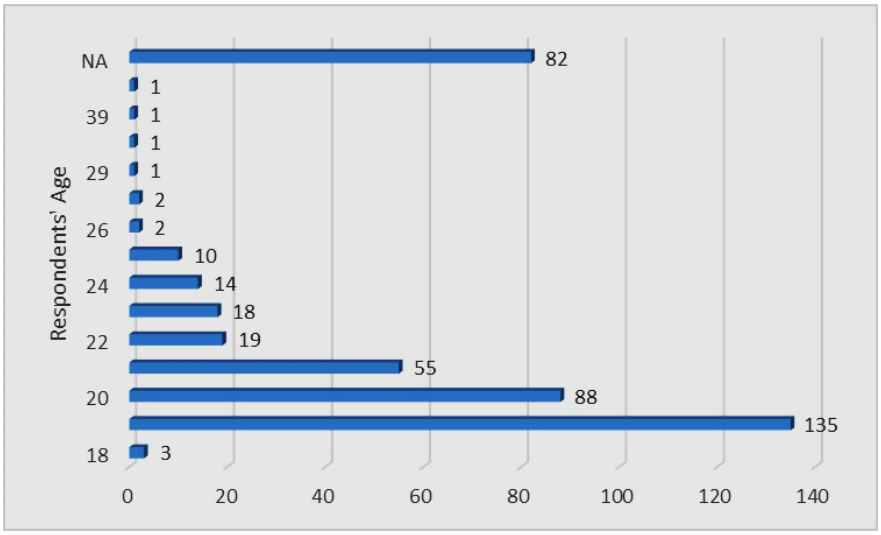
5 Results

5.1. Exploratory data analysis

A total of 491 students participated in the study. Fifty-nine questionnaires were highly incomplete and were excluded from the final analysis. Thus, the following results are based on the responses of 432 participants.

Respondents socio-demographic profile. The majority of the respondents (69%) are in the age range of 19-22 (Figure 14) years old and reside in Tuscany (72%).

Figure 14 Age of respondents' distribution (Absolute values)

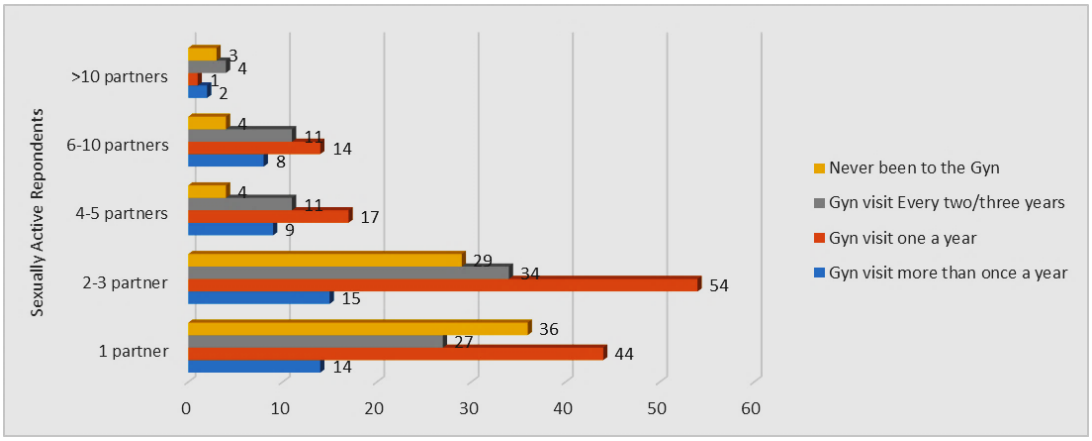


52% (223) of respondents' mothers have high school diploma, 25% (109) junior-high school and 16% (69) university degree. Similar percentage were obtained for the father. In fact, 45% (192) of respondents' fathers have high school diploma, 31% (137) junior-high school and 16% (68) university degree. According to respondent's answer, the 78% of parents are Catholics. Mothers' mean age was 50 years [95%C.I.:(49.91;50.87)] and fathers' 53 years [95%C.I.:(52.99;54.12)].

Sexuality behaviours and past cervical screening. The 20% (87) of respondents are not sexually active and among them 67% declared to be Catholic and 87% are used to smoke. Among sexually active girls (80%; N=344), those who had more than 3 partners during their life had at least one gynaecologist visit per year and underwent a Pap-test.

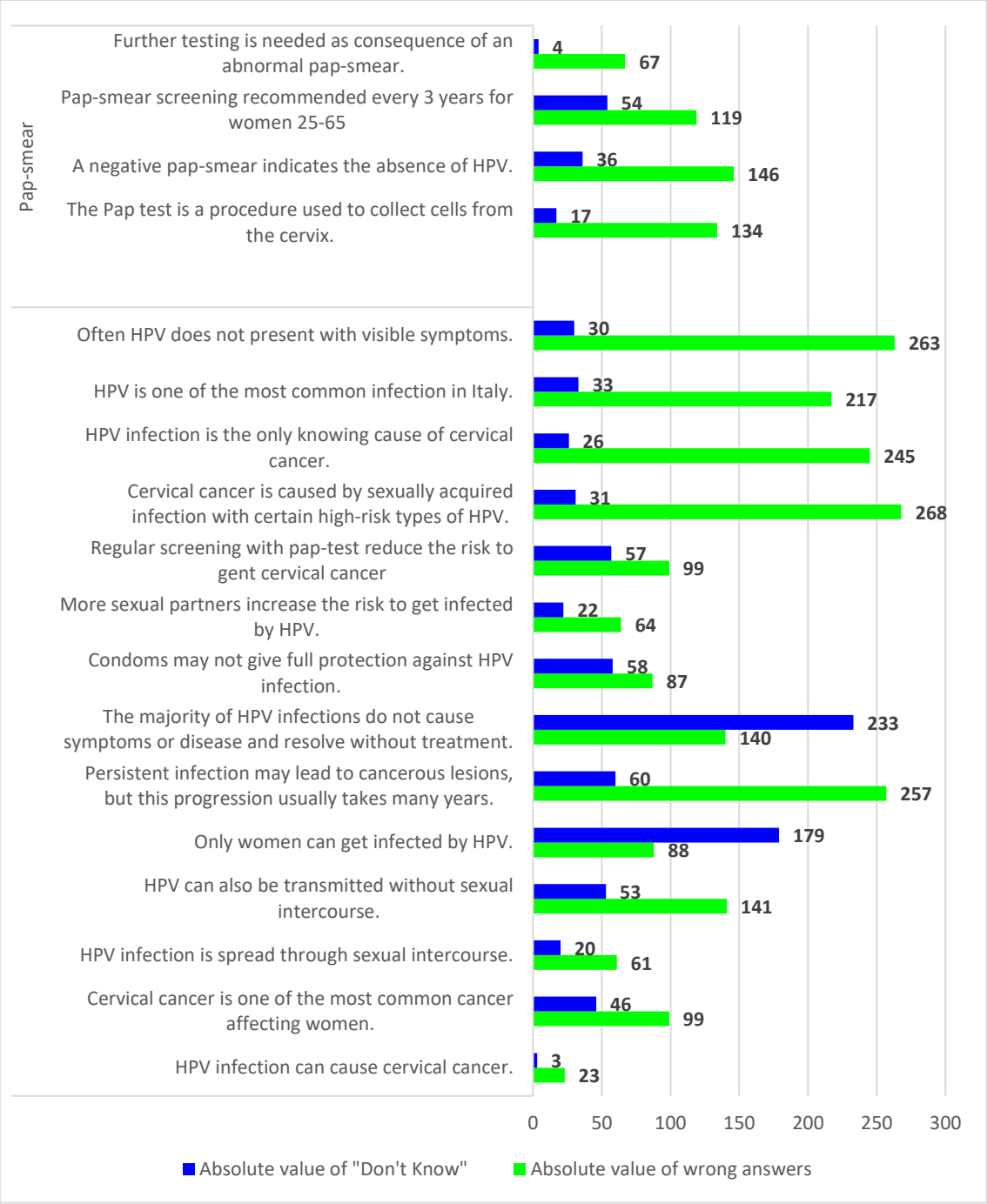
The 72% (214) of respondents that had a gynaecologist visit decided alone to undertake it. 111 (25%) students declared to have undertook at least one pap-test in the past years and the 57% (64) of them followed the medical doctor's advice. 78% (267) of sexually active girls who practice contraception declared to use condom and the 16% the pill. The main sources of sexual information among respondents are friends (27%-117) and health care workers (ASL, family counselling) (53%-229).

Figure 15 Sexually active respondents, life number of partners and past gynaecologist visits.



HPV knowledge. The 6% (25) of interviewed girls do not know the HPV and the 9% (35) have never heard about the associated vaccine. HPV and Pap-smear knowledge were assessed by 14 items (yes/no/don't know) and 4 items (yes/no/don't know), respectively. Knowledge questions highlight a slightly better performance for unvaccinated girls while the information level considerably changes across faculties. Responses quality varies according to the question level, the more detailed is the information requested, the lower is the level of accuracy (Figure 16).

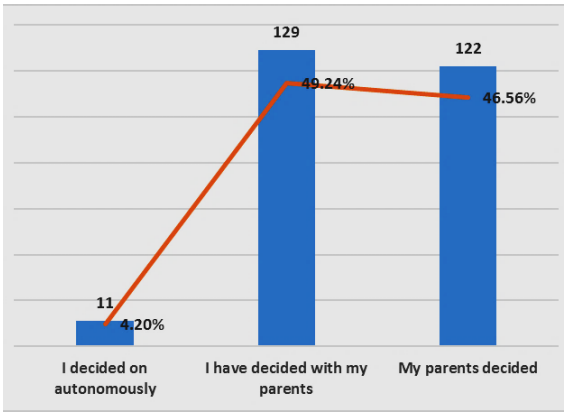
Figure 16 Knowledge questions absolute values of "Don't know" and wrong answers



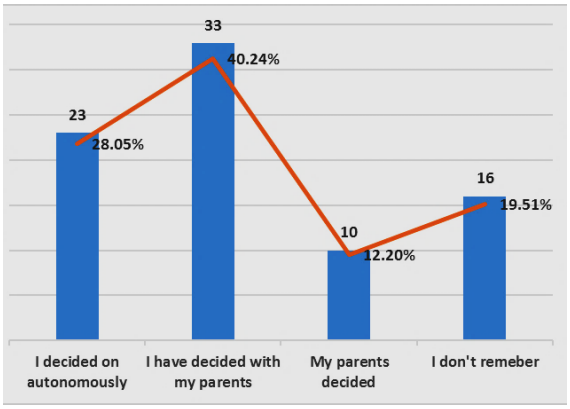
71% (262) of the students were vaccinated against HPV, 23%(85) were unvaccinated and 6%(21) did not remember their vaccination status. The mean vaccination age was 14 years [sd: 1.9313; 95% C.I (14.05544;14.55401)]. Among vaccinated, 48% got the vaccine because their parents decided it. Non-vaccinated seemed more independently in their vaccination decision, in fact 30% (24) has taken the decision about the vaccination autonomously and 42%(33) with their

parents (Figure 17). Among the latter, 34% and 39% respectively, declared that they were "very" or "extremely" likely to take the vaccine if it would be offered again from the national health system. The 52% (224) ignore the duration of protection provided by HPV vaccination and the types of virus covered by the vaccines. The 26% (113) think that is valid for 10 years and the 22%(94) for the whole life. The 38% know that the vaccine is active against genotypes 16 and 18.

Figure 17 Vaccination decision among vaccinated and unvaccinated women.



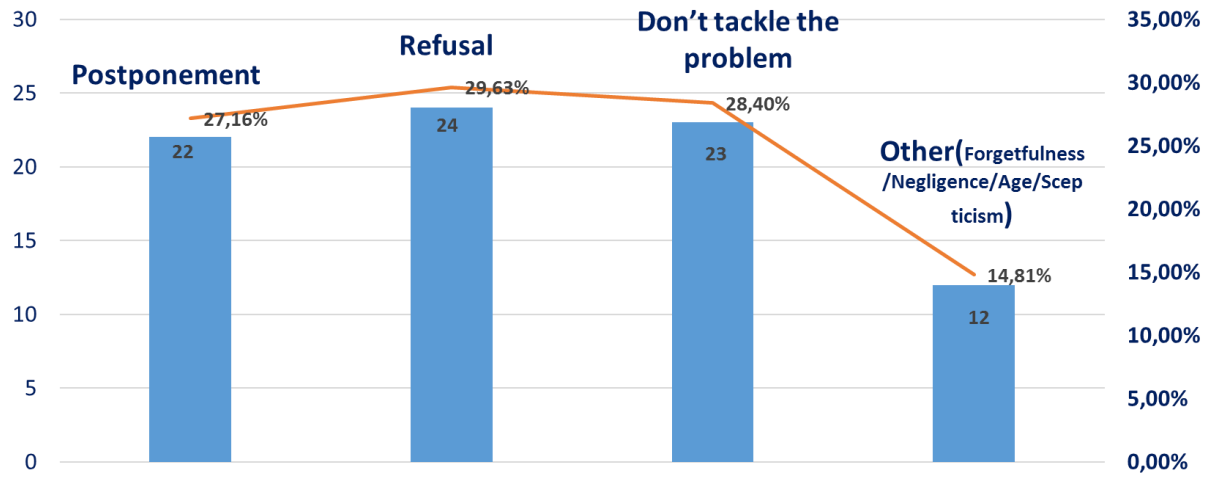
a. Vaccinated girls (N=262)
Mean age: 20.14
[sd: 1.4035; 95% C.I(19.9587;20.3309)]



b. Unvaccinated girls(N=85)
Mean age: 22.28
[sd: 4.3769; 95% C.I(21.1879;23.3746)]

According to the survey results, the main reasons of vaccination refusal (Figure 18) were postponement (27%), refusal (30%) and lack of awareness (28%).

Figure 18 Reasons of vaccination refusal



The most common sources of information for HPV vaccination were parents (41%), gynaecologist/ family physician (36%) and school/university (26%). Finally, the 70 % of respondents (171) declared that will definitely undergo a pap-test examination in the future.

5.1.1. Vaccinated and unvaccinated respondents

Psychological item answers between vaccinated and unvaccinated respondents were analysed more in depth. According to figures 19 and 20, that represent the perceived risk of cervical cancer investigated by questions 51-54 (appendix 5), both vaccinated and unvaccinated respondents are aware that vaccination and pap-test screening will reduce consistently the future risk of cervical cancer.

Figure 19 Perceived risk of cervical cancer among unvaccinated

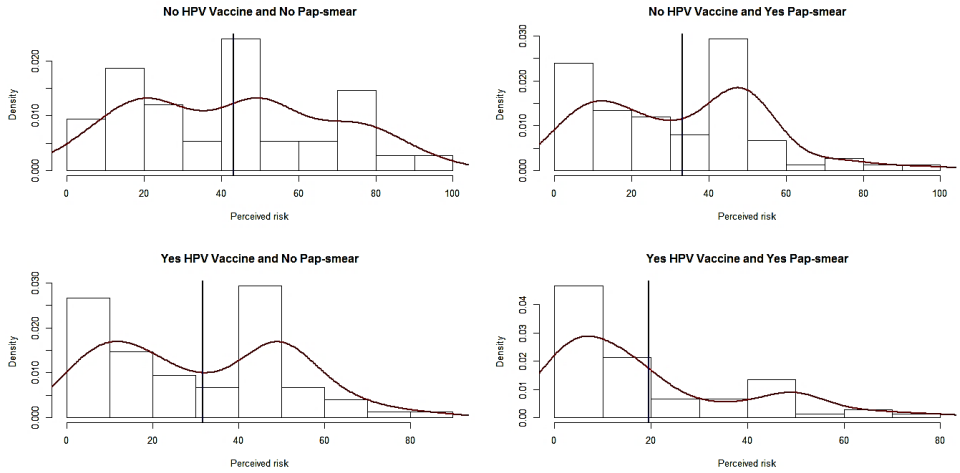


Figure 20 Perceived risk of cervical cancer among vaccinated

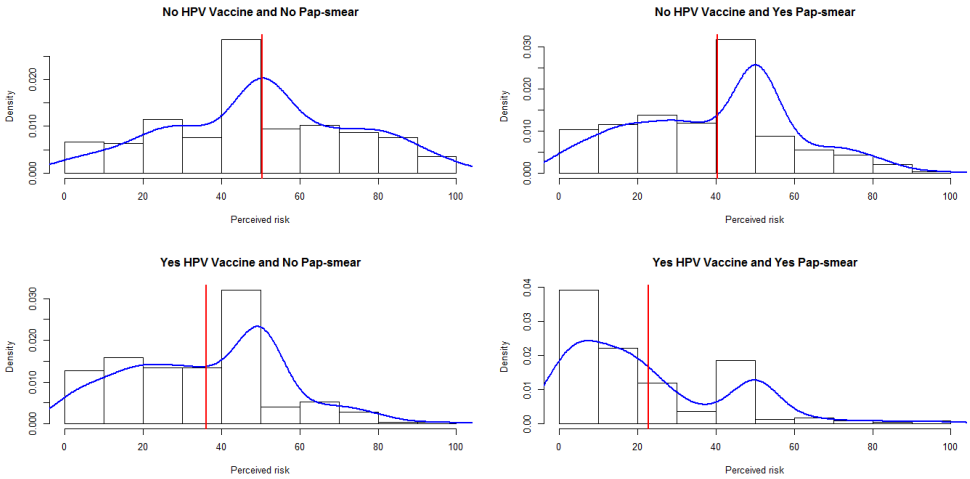


Figure 21 TPB attitudes of vaccinated and unvaccinated

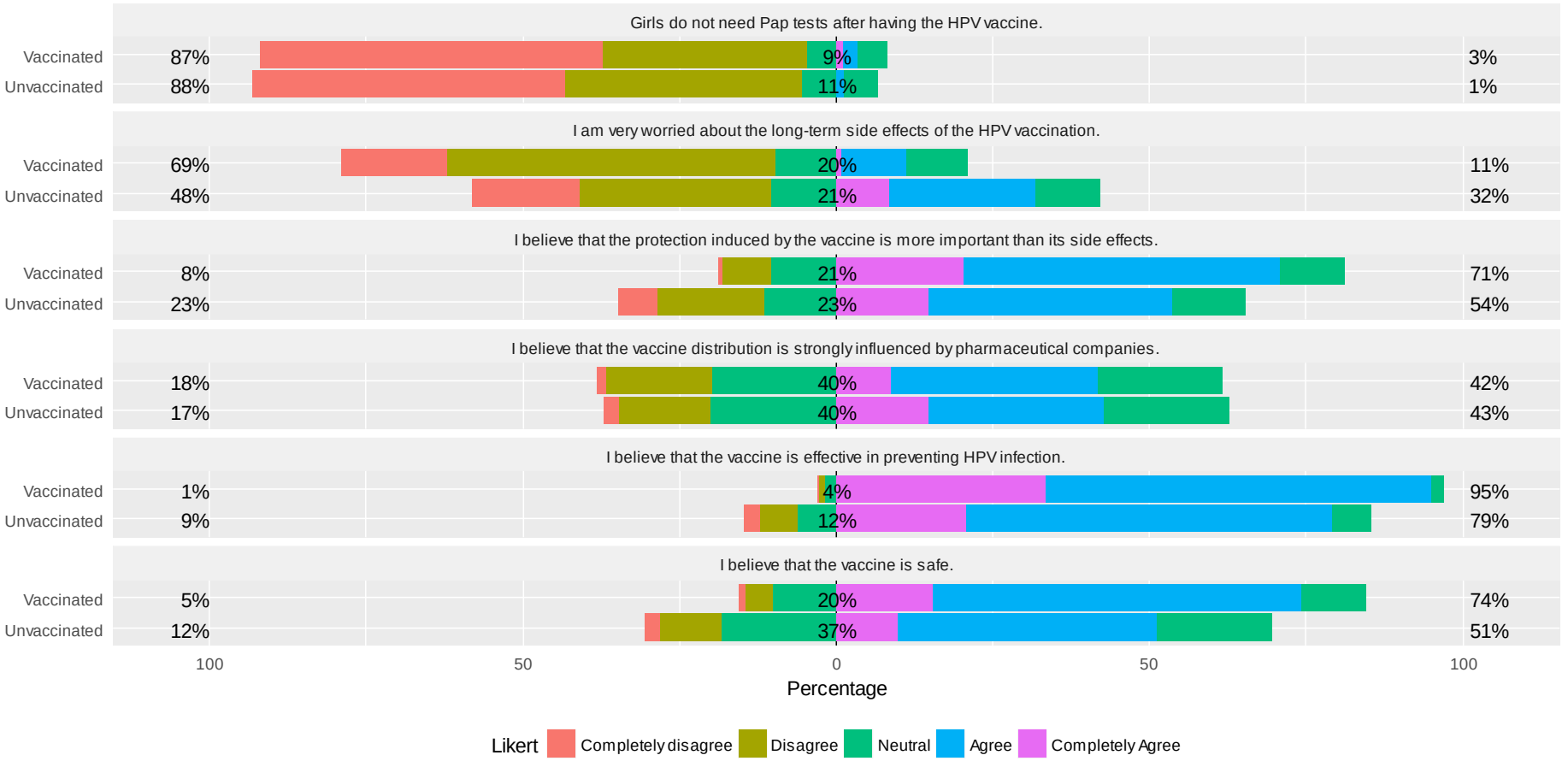


Figure 21 shows the different answers of vaccinated and unvaccinated girls to common attitudes items. The constructs showed in figure 21 measure the respondents vaccination believes towards HPV vaccination. Inspecting the results, it is possible to see that vaccinated students are likely to "agree" or "completely agree" items referred to vaccination safety and efficacy. In respect to the other psychological dimensions, the analysis did not show important difference between the two subgroups.

5.2. Missing data imputation

Missing data occur in survey research for a variety of reasons and are a cause of concerns for data analysis. In general, three main sources of missingness may arise in survey research: *noncoverage* (exclusion of an element from the target population), *total nonresponse* (a sampled element refuse to participate) and *item nonresponse* (a respondent may miss one or more questions on a survey) (Brick and Kalton, 1996). The impact of missing data on quantitative research can be serious, leading to biased estimates of parameters, loss of information, increased standard error and weakened generalizability of findings (Dong & Peng, 2013). Given the sampling nature of the data, a special attention was paid for handling missing data of partially observed respondents. The exploratory analysis underlined the presence of several variables affected by missing values. The analysis done in this thesis focused on the problem of *item non-response* because this was the reason of missigness in the collected data.

In general, the missing data problem at item level needs to be tackled from three aspects, the proportion of missing data, the missing data mechanisms and the patterns of missing data.

Proportion of missing data. The proportion of missing data is directly related to the quality of statistical inferences. In literature, there are not clear guidelines about the acceptable percentage of missing data for valid statistical inferences (Dong & Peng, 2013). For example, Schafer (1999) asserted that a missing rate of 5% or less is insignificant whereas Bennet (2001) fixed the cutoff point at 10% of the total sample. Furthermore, the amount of incompleteness is not the sole criterion by which researcher assesses the missing data problems. In fact, missing data mechanisms and patterns have greater impact on results than does the proportion (Tabachnick and Fidell, 2012; Dong & Peng, 2013). Table 7 listed the number of missing values, as well as the proportion of missing for all the variables presenting a percentage of missing greater than 2. The highest proportion of missing was found in the age of respondent whereas

the most part of variables had a percentage lower than 5%. The amount of missingness for the whole dataset was around 5%.

Table 7 Missing values: counts and proportions

Variable	*Count*	*Proportion*	*Variable*	*Count*	*Proportion*
Respondent's age	72	16.7053	Attitudes toward pap-test 3	23	5.3364
Mother age	18	4.1763	Attitudes toward pap-test 4	23	5.3364
Father age	15	3.4803	Attitudes toward pap-test 5	23	5.3364
Age first intercourse	12	2.7842	Future pap-test intentions	10	2.3202
Proportion of partner at risk in the last 12 months	25	5.8005	Future pap-test frequency	16	3.7123
Reason of last gynaecologist examination	11	2.5522	Future gynaecologist visit frequency	20	4.6404
Unvaccinated section:			Social Norms: gynaecologist	19	4.4084
Future vaccination cost	20	4.6404	Social Norms: parents	19	4.4084
Social Norms: gynaecologist	10	2.3202	Social Norms: sisters/brothers	19	4.4084
Social Norms: parents	10	2.3202	Social Norms: friends	19	4.4084
Social Norms: sisters/brothers	10	2.3202	Social Norms: healthcare professionals	19	4.4084
Social Norms: friends	10	2.3202	Social Norms: (Others)	17	3.9443
Social Norms: healthcare professionals	10	2.3202	Social Norms: gynaecologist	17	3.9443
Social Norms: (Others)	10	2.3202	Social Norms: parents		
Pap-test section:			Social Norms: sisters/brothers	17	3.9443
Pap-test knowledge 1	18	4.1763	Social Norms: friends	17	3.9443
Pap-test knowledge 2	18	4.1763	Social Norms: healthcare professionals	17	3.9443
Pap-test knowledge 3	18	4.1763	Social Norms: gynaecologist	17	3.9443
Pap-test knowledge 4	17	3.9443	Parent's religion	13	3.0162
Attitudes toward pap-test 1	22	5.1044	Religious practice of parents	12	2.7842
Attitudes toward pap-test 2	22	5.1044			

Missing data mechanisms: According to Rubin (1976) there are three mechanisms under which missing data can occur: missing at random (MAR), missing completely at random (MCAR) and missing not at random (MNAR). If the probability of being missing is the same for all cases, i.e. it is by chance, then the data are said to be MCAR. This effectively implies that causes of the missingness are unrelated to the data. If the probability of being missing is the same only

within groups defined by the observed data, then the data are missing at random (MAR). If the missing data cannot be assumed to be neither MCAR nor MAR and the probability of being missing varies for reason that are unknown to the researcher, the data is said to be MNAR. In this case, the missing data mechanism cannot be ignored and must be considered in the imputation model (Little and Rubin, 2002; Van Buuren, 2012; Dong & Peng, 2013).

Given the nature of the questionnaire -characterized by filter answer(s)/section(s) based on previous question(s)- the imputation procedure was done in two-stages. In the first stage filters free variables, i.e. questions answered by all respondents, were imputed whereas filter related covariates were filled in the second stage. A MAR mechanism was assumed for both stages. The tenability of MCAR can be examines using Little's multivariate test (Little and Schenker, 1995), However, it is impossible to test whether the MAR condition holds, given only the observed data (Carpenter & Goldestein, 2004; Horton and Kleinman, 2007; White et al, 2011). Diggle et al. (1995) and Tabachnick and Fidell (2012) suggested a simple t-test of mean differences between the group with complete data and that with missing data. On the contrary, Schafer and Graham (2002) criticized the practice of dummy coding missing values, because this technique redefines the parameters of the population. To this purpose, the existing literature often suggests including additional variables into a statistical model in order to make the missing data mechanisms to be ignored (Collins et al., 2001; Graham, 2003; Rubin, 1996). This suggests that the researcher should be proactive about satisfying the MAR assumption by measuring variables that might explain missingness. In addition, an accurate knowledge of the data collection mechanisms should help the researcher to assess with precision the type of process leading to the formation of the missing (Allison, 2001).

In this study, an accurate control of the data collection processes allowed to identify some auxiliary variables able to make the missing data mechanisms to be ignored. Considering the small proportion of missigness, especially for variables considered in further analysis, and that- as shown by Collins et al. (2001)- the violation of MAR assumption does not seriously distort parameter estimates, auxiliary indicators were used to make the missing mechanism ignored. In fact, the number of missing data depended on the questionnaire administration methods and on the faculties attended. Two categorical variables indicating the type of administration methods and the university faculties were added to the imputation methods as possible predictors. These indicators were highly associated with variables presenting high proportion of missing data.

Pattern of missing. The missing data pattern was arbitrary.

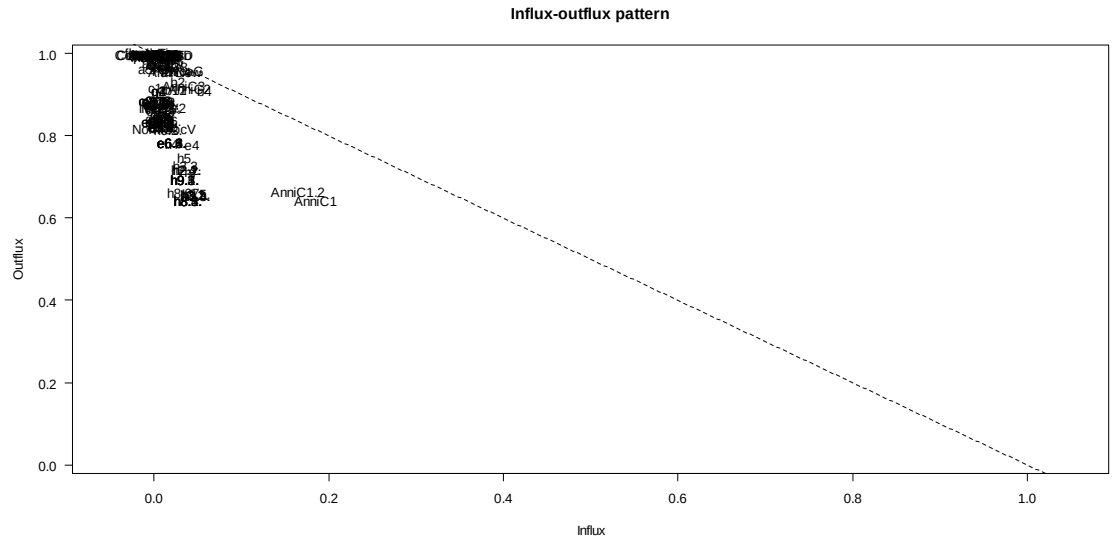
Multiple imputation. When missing data are MAR, listwise deletion may lead to biased results (Rubin 1976). Various procedures have been suggested in the literature over the last several decades to deal with missing data (i.e., Anderson 1957; Rubin 1972; Rubin 1987). The technique of multiple imputation (MI), which is a flexible, simulation-based statistical technique for handling missing data, has gained popularity increasingly over the years (Rubin 1976; Rubin 1987). Specifically, MI acknowledges the uncertainty by generating a set of m plausible values for each unobserved data point, resulting in m complete data sets, each with one unique estimate of the missing value. Since, MI is considered by many statisticians to be the most appropriate technique for addressing missing data in many circumstances, this procedure was preferred for the analysis. The chosen multivariate imputation method was multivariate imputation using chained equations (MICE), one of the most popular choices used in practice, also known as imputation using fully conditional specifications (van Buuren, Boshuizen and Knook 1999) and as sequential regression multivariate imputation (Raghunathan et al. 2001). The MICE method used a Gibbs-like algorithm to impute multiple variables sequentially using univariate fully conditional specifications. Creating multiple imputations, as opposed to single imputations, accounts for the statistical uncertainty in the imputations. In addition, the chained equations approach is very flexible and can handle variables of varying types (e.g., continuous or binary) as well as complexities such as bounds or survey skip patterns. In fact, the algorithm requires a specification of univariate imputation method separately for each incomplete variable. Thus, to impute categorical variables, logistic, ordered logistic, or multinomial logistic regressions were used. To impute continuous variable, predictive mean matching was considered. The imputation procedure was performed using R Statistical Software and the package MICE (Van Buuren et al., 2011).

Imputation model and number of imputation. The imputation model should include useful variables and be general enough to capture assumed structure of the data (Dong & Peng, 2013; Azur et al., 2011). According to Schafer (1997) and Van Burren et al. (1999) variables of theoretical interests, variables associated with the missing mechanisms and variables correlated with those incompletes were included in the procedure. Given the two-step imputation procedure, different group of predictors were chosen in each of them. Before the imputation process, the dataset did not contain derived variables, i.e. sum scores, ratios, interactions, etc. Covariates transformation was done entirely after MI. The number of possible predictors was chosen according to the influx and outflux indexes proposed by Van Buuren (2012). The influx of a variable quantifies how well its missing data connect to the observed data on other variables. Influx depends on the proportion of missing data of the variable. Influx of a completely observed variable is equal to 0, whereas for completely

missing variables is 1. The outflux of a variable quantifies how well its observed data connect to the missing data on other variables. It is an indicator of the potential usefulness of a certain variable for imputing other covariates. Outflux depends on the proportion of missing data of the variable. Outflux of a completely observed variable is equal to 1, whereas outflux of a completely missing variable is equal to 0.

The influx-outflux pattern is showed in figure 22 By inspecting the plot below can be noted that all variables are characterized by high outflux and low influx values. The result is related to the proportion of missingnes which is relatively low for most of them. according to Van Buuren (2011) variables that are located higher up in the display are more complete and thus potentially more useful for imputation.

Figure 22 Influx-Outflux pattern

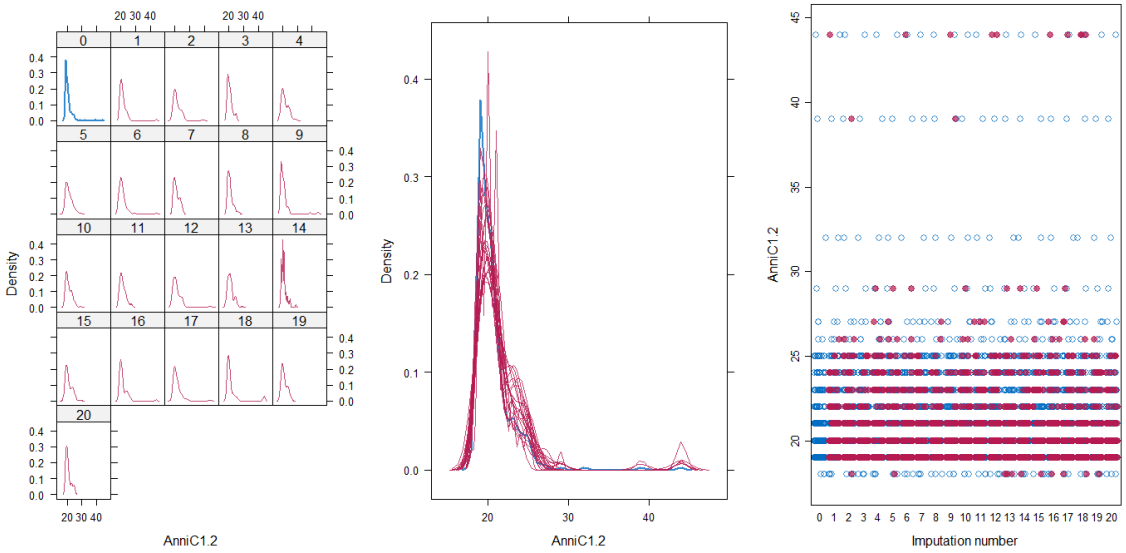


Given the influx-outflux results, almost all the variables included in the dataset were used in the imputation process. In fact, Schafer (1997) recommended that is often beneficial to choose as large a number of predictors as possible to make the MAR assumption more plausible. To simplify the predictors selection and make the imputation faster, the MICE package contains several tools that aid an automatic predictor selection. In addition, the imputation function detects multicollinearity, and solves the problem by removing one or more predictors from the matrix. The number of imputations needed in MI is a function of the rate of missing information in a dataset. Although Rubin (1987) proposed a formula to compute the relative efficiency of imputing m times, methodologists have not agreed with the optimal number of imputations (Dong & Peng, 2013). Schafer and Olsen (1998) suggested that just 3-5 imputations are sufficient to obtain valid results while Schafer and Graham (2002) indicated

that 20 imputation ore enough in many practical applications to avoid uncertainty (Dong & Peng, 2013). As suggested by Schafer and Graham (2002) 20 imputations were performed.

Assessing the Imputation Procedure. According to Van Buuren (2012) graphical comparisons are one of the best tools to assess imputations plausibility. In general, good imputations should have a distribution similar to the observed data and plots allow to easily inspect the discrepancy between observed and imputed data. In figure 23 kernel densities and stripplots of imputed and observed data for three of the most affected variables (with highest percentage of missing) are shown. Blue represents the observed data and red shows the imputed ones.
For the three variables, the distributions match up well. Similar results were obtained for the remaining variables.

Figure 23 Respondents' age [missing percentage=16.7% (N=72/431)]



As shown by figure 23, the most part of imputed data belong to the class 19-25, which is the most frequent class of the original variable. The same trend characterized the other imputations, as shown in figure 24 and 25. In fact the most frequent categories for the variables "Proportion of partner at risk" and "Reasons of past gynaecologist visit" were, respectively, 1 ("None") and 3 ("Parents'advice").

Figure 24 Proportion of partners at risk in the last 12 months [Percentage of missing =5.8% (N=25/431)]

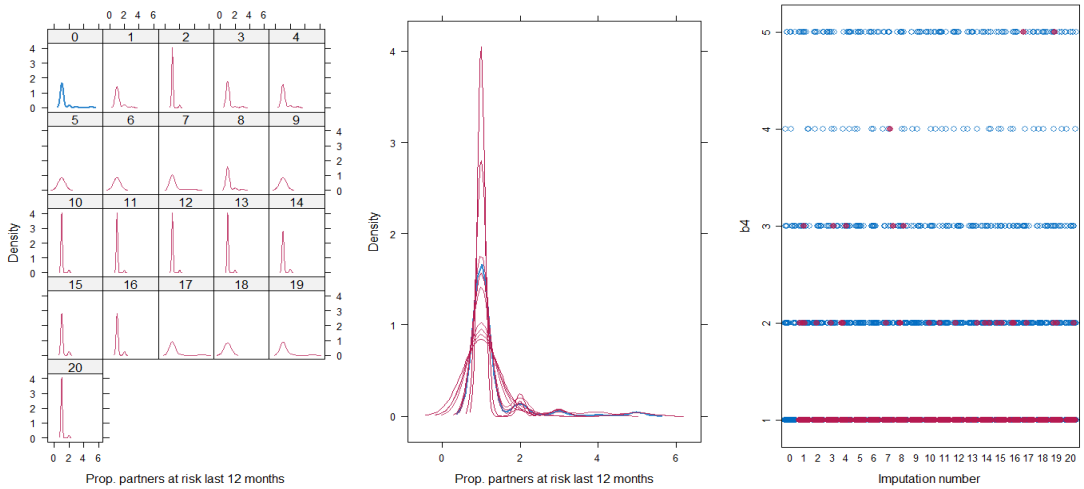
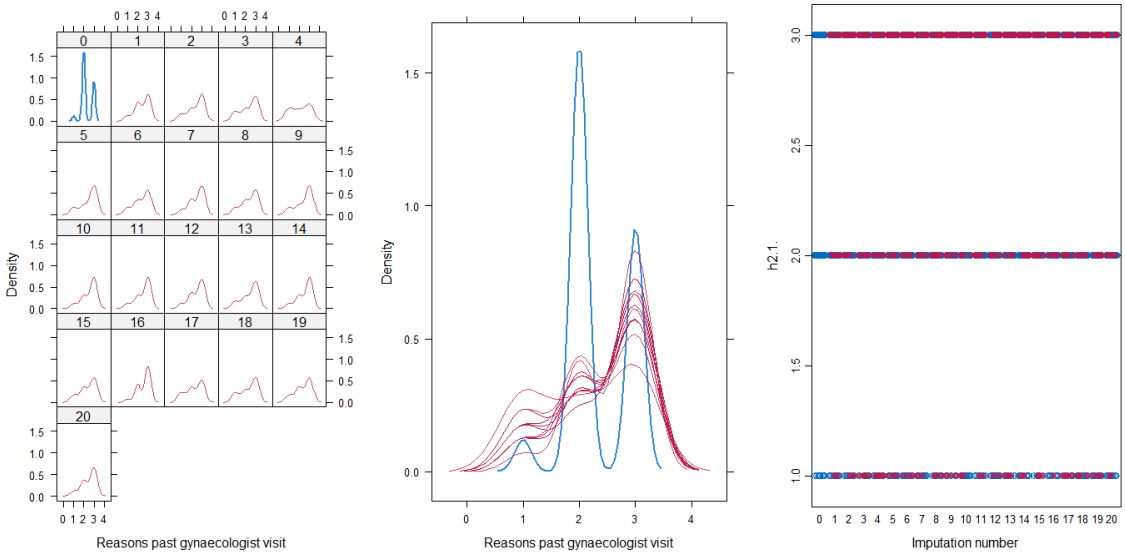
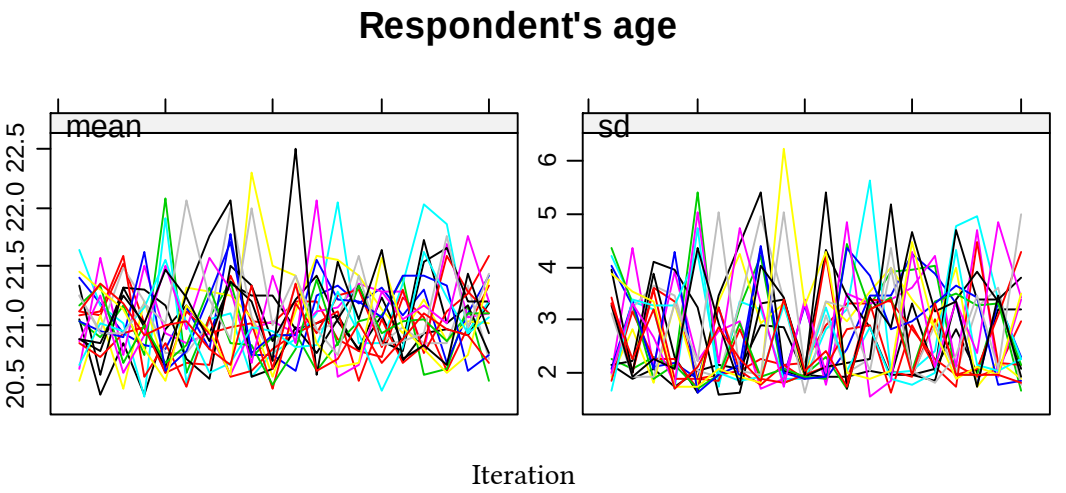


Figure 25 Reasons of past gynaecologist visit [Percentage of missing =4.2% (N=18/431)]



There is not a clear-cut method to assess the MICE algorithm convergence. Van Buuren (2012) suggested to plot one or more parameter against the iteration number and to inspect convergence graphically. According to the author, on convergence, the different streams should be freely intermingled with one another, without showing any definite trends. Convergence is diagnosed when the variance between different sequences is no larger than the variance within each individual sequence. Observing plots in figure 26, inspection of the stream does not reveal particular problem of the imputation model. Same results were obtained for the remaining variables.

Figure 26 Convergence plots



The appropriate way to analyse multiply imputed data is to perform the complete data analysis on each imputed dataset and then combining the estimates from each dataset to obtain the final results (pooling). The variance estimates calculated in the pooling procedure involve both the "within" variance calculated for each dataset individually, as well as the "between" variance that reflects the uncertainty in the imputations—how variable the results are across the imputed datasets. (Van Buuren, 2012; Azur et al., 2011). MI procedure was tested fitting the logistic model considered in section 9.2 (to identify determinants of realized choices) in three different datasets. To obtain final results estimates were combined (pooled) for all the 20 datasets. Coefficients and standard error estimates are shown in table 8. Inspecting the table can be seen that estimates are the same for each dataset considered. The low proportion of missingness, characterizing the selected covariates and the dependent variable, is the cause of this equality.

Table 8 Coefficients and standard errors estimated for each selected dataset.

Variable	***Used dataset***		***Dataset 5***		***Dataset 15***		***Pooled estimates***	
	Estimate	*SD*	*Estimate*	*SD*	*Estimate*	*SD*	*Estimate*	*SD*
(Intercept)	1.0946	1.0020	1.0946	1.0020	1.0946	1.0020	1.0946	1.0020
Father's age (Ref:"51-60") **								
Father's age: 41-50	-0.2727	0.3733	-0.2727	0.3733	-0.2727	0.3733	-0.2727	0.3733
Father's age: 61+	-1.6429	0.5892**	-1.6429	0.5892**	-1.6429	0.5892**	-1.6429	0.5892**
Father's age Don't live with father	-0.1473	0.5044	-0.1473	0.5044	-0.1473	0.5044	-0.1473	0.5044
Citizenship (Ref: Abroad) *								
Citizenship: Italy	1.8493	0.7019**	1.8493	0.7019**	1.8493	0.7019**	1.8493	0.7019**
Mother's education (Ref:" High school") .								
Mother's education: None, primary, secondary	-0.7637	0.4509	-0.7637	0.4509	-0.7637	0.4509	-0.7637	0.4509
Mother's education: Graduate or professional degree	-0.0967	0.5638	-0.0967	0.5638	-0.0967	0.5638	-0.0967	0.5637
Attitude 1 (Ref: "Agree") *** [1]								
Att1: Disagree	-1.2681	0.5393*	-1.2681	0.5393*	-1.2681	0.5393*	-1.2681	0.5393*
TPB Attitude 2 (Ref: "Agree") **[2]								
Att2: Disagree	-0.6839	0.4909	-0.6839	0.4909	-0.6839	0.4909	-0.6839	0.4909
TPB Attitude 3 (Ref: "Agree") **[3]								
Att3: Disagree	0.8815	0.4275*	0.8815	0.4275*	0.8815	0.4275*	0.8815	0.4275*
Score HPV knowledge	-0.3425	0.0698***	-0.3425	0.0698***	-0.3425	0.0698***	-0.3425	0.0698***
Sources of HPV information (Ref:" School/University") **								
InfoHPV: Others (Health workers, internet, radio and tv, books, etc.)	-0.0466	0.6625	-0.0466	0.6625	-0.0466	0.6625	-0.0466	0.6624
InfoHPV: Parents	1.2188	0.4518**	1.2188	0.4518**	1.2188	0.4518**	1.2188	0.4518**
InfoHPV: Medical doctor/Gynaecologist	0.5622	0.5111	0.5622	0.5111	0.5622	0.5111	0.5622	0.5111
Vaccine duration of protection (Ref:" I don't know") *								
Protection: 10 years	0.3271	0.3647	0.3271	0.3647	0.3271	0.3699	0.3270	0.3647

Protection: Lifetime	1.1997	0.4388**	1.1997	0.4388**	1.1997	0.4388**	1.1997	0.4388**
Interactions: Mother's education: Attitude 2 (Ref =" High school:Agree")* MotherEd*Att2: None_primary_secondary: Disagree	0.1624	0.7441	0.1624	0.7441	0.1624	0.7441	0.1624	0.7441
MotherED*Att2: Graduate or professional degree: Disagree	-2.1946	0.9148*	-2.1946	0.9148*	-2.1946	0.9148*	-2.1946	0.9148*

1. Attitude 1: I believe that the vaccine is effective in preventing HPV infection
2. Attitude2: I believe that the vaccine is safe.
3. Attitude3: I am very worried about the long-term side effects of the HPV vaccination.
*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

MICE limitations. Although the MICE approach is a principled method of addressing missing data, it is important to acknowledge certain complexities and limitations of the approach. While MICE offers great advantage over other missing data techniques in terms of its flexibility, a primary disadvantage is that MICE does not have the same theoretical justification as other imputation approaches. In particular, fitting a series of conditional distributions, as is done using the series of regression models, may not be consistent with a proper joint distribution. Initial research suggests that this may not be a large issue in applied settings (Brand, 1999; Schafer & Graham, 2002); however, further research is needed into the implications for practice. When there are relatively few variables needing imputation and a multivariate normal model would be appropriate (i.e., the variables are continuous and approximately normally distributed), a joint model such as a multivariate normal may be preferable (Azur et al., 2011).

5.3. Determinants of realized immunization choices

One of the main objective of this research was to investigate factors related with realized immunization choice. The logistic regression model was considered in order to find out the covariates highly related with the realized vaccination behaviour. As explained above, the dataset is characterized by different subset of respondents. Each subset includes a certain number of common variables and few dimensions that are specific for the subgroup considered. For the analysis, the subset (N=342) of vaccinated and unvaccinated respondents that knew HPV and its associated vaccine was considered. It contains the most crucial variables collected during the survey.

Three options were considered to investigate the respondents' "vaccination status", i.e."*I am not vaccinated*", "*I am vaccinated*", "*I do not remember my vaccination status*". Since only few girls (4.87%) did not remember their vaccination status, the category was dropped from the analysis to avoid possible sparseness bias and complete-separation issues. The frequency distribution of the variable is presented in table 9

Table 9 Vaccination status distribution

Unvaccinted	Vaccinated
82	260
23.97%	76.03%

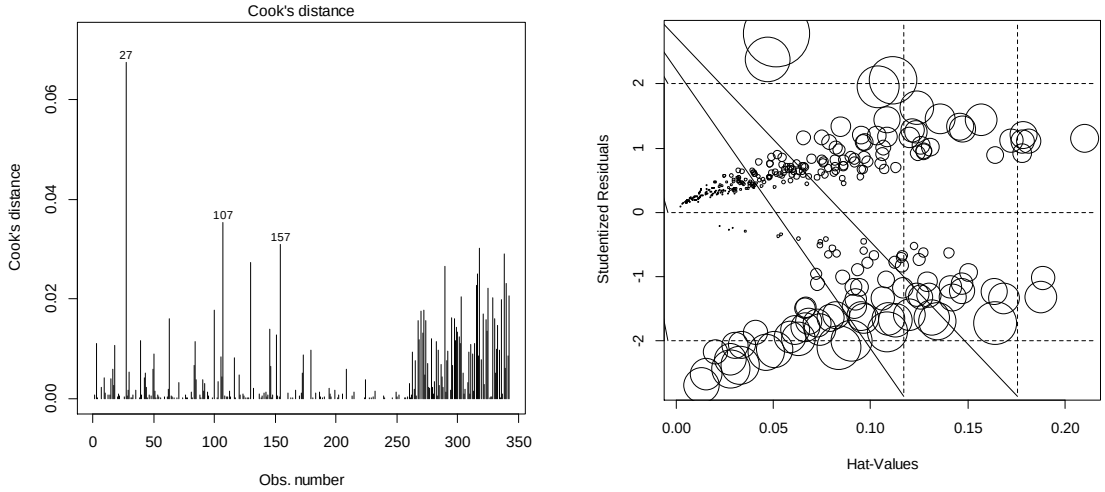
To avoid sparseness, most of the variables were reclassified. Subjective norms were measured by two multiple choice questions investigating respectively, the *social pressure* and the *motivation to comply*. Both dimensions were reclassified according to the frequency distribution of each answer option. A new variable was created for each dimension considering a decreasing number of values, starting from the most chosen (highest value) and reaching the less chosen referent (smallest value). For example, the answer options for subjective norms were, gynaecologist, parents, related, brother/sister, friends and health care workers. The respondent was asked to select the most influents among the five listed. If, for example, the gynaecologist was the first most chosen, parents were the second and brother/sisters the third, the new variable was created imputing the number 3 for the gynaecologist, the 2 for parents and the 1 for brothers/sisters. The same procedure was adopted to reclassify the source of information for sexuality and HPV vaccination. Both variables were assessed by multiple choice questions, as done for subjective norms.

Being religious (yes/ no), religious practice (yes/no) and citizenship (Italy/other) were dichotomized. As indicated in the questionnaire section, HPV knowledge was investigated by a 14-statement scale. Each statement was characterized by three answer options, 'Yes', 'No', or 'I don't know' (Question 31-Appendix 5). An IRT model for nominal response was first considered to analyse knowledge items. IRT revealed that the "don't know" option was the one associated with the highest level of knowledge, i.e. girls unaware about the answer were those with the highest level of HPV knowledge. Since IRT results were quite difficult to interpret and indicator variable, summing the number of correct answers for each respondent, was computed. From question 35 (appendix 5) investigating the vaccine protection knowledge, three different variables were obtained, one indicating if the girl was aware of the answer, one accounting for the type of viruses included in the vaccine and one explaining the knowledge of the duration of vaccine protection. Since the three indicators were highly correlated only the last was kept in the model. Psychological items were dichotomized and analysed separately, i.e. each item was considered as a covariate. Father and mother age were categorized to account for respondents that did not live in the nuclear family.

Standard and penalized logistic regression models were considered to identify the most significant factors associated with the realized immunization choice. Since the number of independent variables in the original model was high, different selection procedures were performed in order to choose the approximate best subset of covariates, one that was as simple as possible while still providing good predictive performance.

Logistic regression. First a multiple logistic regression model was fitted. The analysis was performed in R Statistical Software and the function *glm* from package STATS was considered. A univariate analysis was performed to select a smallest group of the original covariates. Only those with a p-value lower or equal to 0.25 were kept in the final model (Hosmer and Lemeshow, 2000). Then, a multiple logistic regression model was built through forward and backward stepwise selection to determine which predictors was significantly associated (p-value smaller than 0.05) with the vaccination status. Since the number of variable was already large, only one interaction term was added to the original set of predictors. Due to sparsity, the resulting estimates and confidence intervals were huge and wider. To assess the plausibility of the results residual deviance, likelihood ratio test, Hosmer and Lemeshow goodness-of-fit test, pseudo R-square, Cook's distance and standardized Pearson residuals plots were considered (Figure 27).

Figure 27 Cook's distance and studentized residuals



Standardized residual plot underlined the existence of some possible influential observations outside the range of [-2;2]. To evaluate the effect on the model fit, 14 observations were excluded from the original dataset. A new logistic regression model was fitted and an improvement of all goodness of fit index was obtained, as shown in table 10 However, reducing the number of observations sparseness issues increased (table 30 appendix 2).

Table 10 Goodness of fit indexes

Index	***Reduced dataset [-2;2]***	***Original dataset***
Residual deviance	166.81	261.84
AIC	206.81	301.84
Likelihood Ratio Test (p-value)	0.1414	0.0203
Pseudo R2		
McFadden	0.5088	0.3049
Maximum likelihood	0.4105	0.2854
Cragg and Uhler's	0.6354	0.4274
Hosmer and Lemeshow (p-value)	0.5755	0.8154

The sensitivity of the results was explored evaluating different methodological approaches.
In fact, traditional stepwise selection methods, such as forward and backward, suffer from high variability and low prediction accuracy, especially when the number of possible covariates is large and/or the level of multicollinearity is notable (Hastie, Tibshirani, and

Friedman 2001). In such situation, penalized regression methods, such as LASSO, Elastic Net and penalized maximum likelihood estimation (PMLE), combine the advantages of selection procedure preserving the model prediction accuracy. Both datasets, the complete and the reduced ones (without influential observations), were considered in the further analysis.

*Penalized maximum likelihood (*appendix 1*).* Starting from the original set of covariates, an alternative selection procedure was considered. The analysis was performed in R Statistical Software and the package *rms* (Harrell, 2017). Internal validation was considered to assess overfitting and model predictive accuracy (Harrel et al.,). Overestimation occurs when many parameters and smaller sample sizes characterized the original model (Harrel, 2001). The key elements in assessing the predictive performance of a fitted logistic regression model involve the evaluation of model discrimination, i.e. how the model discriminates between those who were and were not vaccinated, and model calibration, i.e. the agreement between observed outcomes and predictions (Harrells, 2001). Indexes such as Brier-score, Harrell's C index, Somer's D, ROC-area and calibration plot were used to measure model discrimination and calibration. (appendix 1).

A resampling validation of the complete model, with backward variables deletion was performed. Bootstrap techniques were applied to quantify the amount of overfitting of the final model and to adjust for the model's predictive accuracy. 500 hundred bootstrap samples of size 342 were drawn with replacement from the original sample. From each bootstrap sample, a final prediction model was fitted deleting predictors with p-value bigger than 0.05. The predictive accuracy of each bootstrap model was estimated in the bootstrap and in the original sample to quantify the difference. In fact, the validate function in Harrell's rms package (Table 11), provides original discrimination indexes, the training data, the original test data, the optimism value, and the corrected estimate for both models. The optimism value is calculated subtracting the test data value from the training data value, and the corrected index value is calculated subtracting the optimism value from the original index value. At the end of the procedure the 500 differences were averaged to a single estimate.

Table 11- Validation indexes

	Index	Original	Training	Test	Optimism	Index Corrected
Dataset with outliers	Somer's D	0.72	0.7733	0.6561	0.1172	0.6028
	R^2	0.4444	0.5179	0.3602	0.1577	0.2867
	C-index	0.3491	0.4223	0.2728	0.1494	0.1997
	Brier Score	0.1173	0.1036	0.131	-0.0274	0.1447

Dataset without outliers	Somer's D	0.866	0.8837	0.8052	0.0785	0.7874
	R^2	0.6282	0.6663	0.5423	0.1239	0.5042
	C-index	0.5175	0.5652	0.4298	0.1355	0.3821
	Brier Score	0.083	0.0737	0.0987	-0.0250	0.1080

The group of covariates selected was equal to the previous one (multiple logistic regression). To prevent future overfitting and to stabilize estimates, still huge due to sparseness, a penalized maximum likelihood estimation was performed on the final model. This technique is a generalization of the ridge regression method and introduce a penalty factor into the estimation process that discourages large value for the coefficients. In this way, estimates are less vulnerable to overfitting and more accurate for prediction (Baayen, 2008; Moons et al., 2004). The best penalty factor was selected according to the modified Akaike's information criterion (AIC). The optimal factor was the one that maximized the AIC (Moons et al., 2004; Ambler et al., 2002; Harrell, 2001; Gray et al., 1992)

Lasso and Elastic Net regression. To validate the set of predictors selected by the previous statistical procedures, penalized regressions (Tibshirani, 1996) were considered. Penalized regression are powerful techniques generally used for creating parsimonious models in presence of large numbers of candidate predictor variables. Ridge, LASSO and Elastic-Net (appendix 1) algorithms work on same principle, they penalize the magnitude of coefficients to select the most significant group of variables. Each method is characterized by two tuning parameters, the alpha, which defines the type of regularization (penalization) considered, and lambda, that depends on alpha and controls the amount of shrinkage on coefficients. As shrinkage penalty λ increases, the coefficient estimates will tend to approach zero. When $\lambda = 0$, no shrinkage is performed (logistic regression), and as λ increases, the coefficients are shrunk ever more strongly. When α is set to zero, the L2 regularization is considered, i.e. the Ridge regression, while when α is set to one, a L1 regularization is preferred, i.e. Lasso regression. The Elastic-Net ($0 < \alpha < 1$) is a combination of the L1 and L2 penalties. Since Lasso and EN perform model selection, they were preferred over the ridge algorithm. Cross-validation (Friedman et al., 2017; Hastie et al., 2001; Tibshirani, 1996; Cessie & Van Houwelingen, 1990), Brier's score (Brier, 1950) and ROC curves were considered to choose the best combination of α and λ. The best combination was the one presenting the lowest cross-validation error, the lowest Brier's score and the highest area under ROC curve (appendix 2). As showed in tables 31 and 32 in appendix 2 the selected best subset of variables was equal to

previous models. The analysis was performed in R Statistical Software and the package *Glmnet* (Friedman et al., 2017).

The different approach considered identified the same subset of covariates significantly related with the "vaccination status". The best model was chosen evaluating the predictive performance of each of them. Brier's score and ROC curves were considered as measures of model predictive accuracy. As showed in table 12 the PMLE model fitted on the reduced dataset (without influential observations) is the best one.

Table 12- ROC areas and Brier's scores

Model	ROC Area	Brier's score
Elastic Net Complete	0.75	0.15
Elastic Net Reduced	0.80	0.10
PMLE Complete	0.80	0.12
PMLE Reduced	0.85	0.09

The PMLE model underlined the existence of several determinants associated with declining the HPV vaccination (table 13). These included citizenship, father's age, mother's education, HPV knowledge score (number of correct answers), sources of HPV information, duration of HPV vaccine protection and attitudes towards HPV vaccination (safety and side effects). The interaction between mothers' highest educational level and the safety vaccination attitudes was founded significant as well. The reference category for the dependent variable (vaccination status) were the vaccinated respondents. Plots of significant covariates are shown in figure 28.

Table 13- Determinants of HPV realized immunization choices (PMLE regression model)

Variable	***Estimates***	***S.E.***	***P-value***	***Penalty***
Intercept	-0.043	0.6836	0.9498	0
Father's age (Ref:"51-60") **				
Father's age: 41-50	-0.2500	0.3142	0.4262	2.2079
Father's age: 61+	-0.9412	0.3901	0.0158	2.2079
Father's age Don't live with father	-0.3667	0.3622	0.3113	2.2079
Citizenship (Ref: Abroad) *				
Citizenship: Italy	0.8868	0.435	0.0415	1.8028
Mother's education (Ref:" High school") .				
Mother's education: None, primary, secondary	-0.6814	0.3392	0.0445	2.0817
Mother's education: Graduate or professional degree	-0.4341	0.3803	0.2537	2.0817
Attitude 1 (Ref: "Agree") *** [1]				
Att1: Disagree	-0.8945	0.3846	0.02	1.8028
TPB Attitude 2 (Ref: "Agree") **[2]				
Att2: Disagree	-0.6517	0.3488	0.0617	1.8028
TPB Attitude 3 (Ref: "Agree") **[3]				

Att3: Disagree	0.8046	0.3336	0.0159	1.8028
Score HPV knowledge	-0.2935	0.0608	<0.0001	7.1423
Sources of HPV information (Ref:" School/University")**				
Source: Others (Health workers, internet, radio and tv, books, etc.)	0.4164	0.4302	0.3331	2.2079
Source: Parents	1.2057	0.3478	0.0005	2.2079
Source: Medical doctor/Gynaecologist	0.4168	0.3736	0.2646	2.2079
Vaccine duration of protection (Ref:" I don't know") *				
Protection: 10 years	0.2209	0.3051	0.469	2.0817
Protection: Lifetime	0.939	0.3466	0.0067	2.0817
Interactions: Mother's education: Attitude 2 (Ref =" High school:Agree")*				
MotherEd*Att2: None_primary_secondary: Disagree	-0.1826	0.5794	0.7526	0.7259
MotherED*Att2: Graduate or professional degree: Disagree	-2.0773	0.7164	0.0037	0.5669

1. Attitude 1: I believe that the vaccine is effective in preventing HPV infection
2. Attitude2: I believe that the vaccine is safe.
3. Attitude3: I am very worried about the long-term side effects of the HPV vaccination.
*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

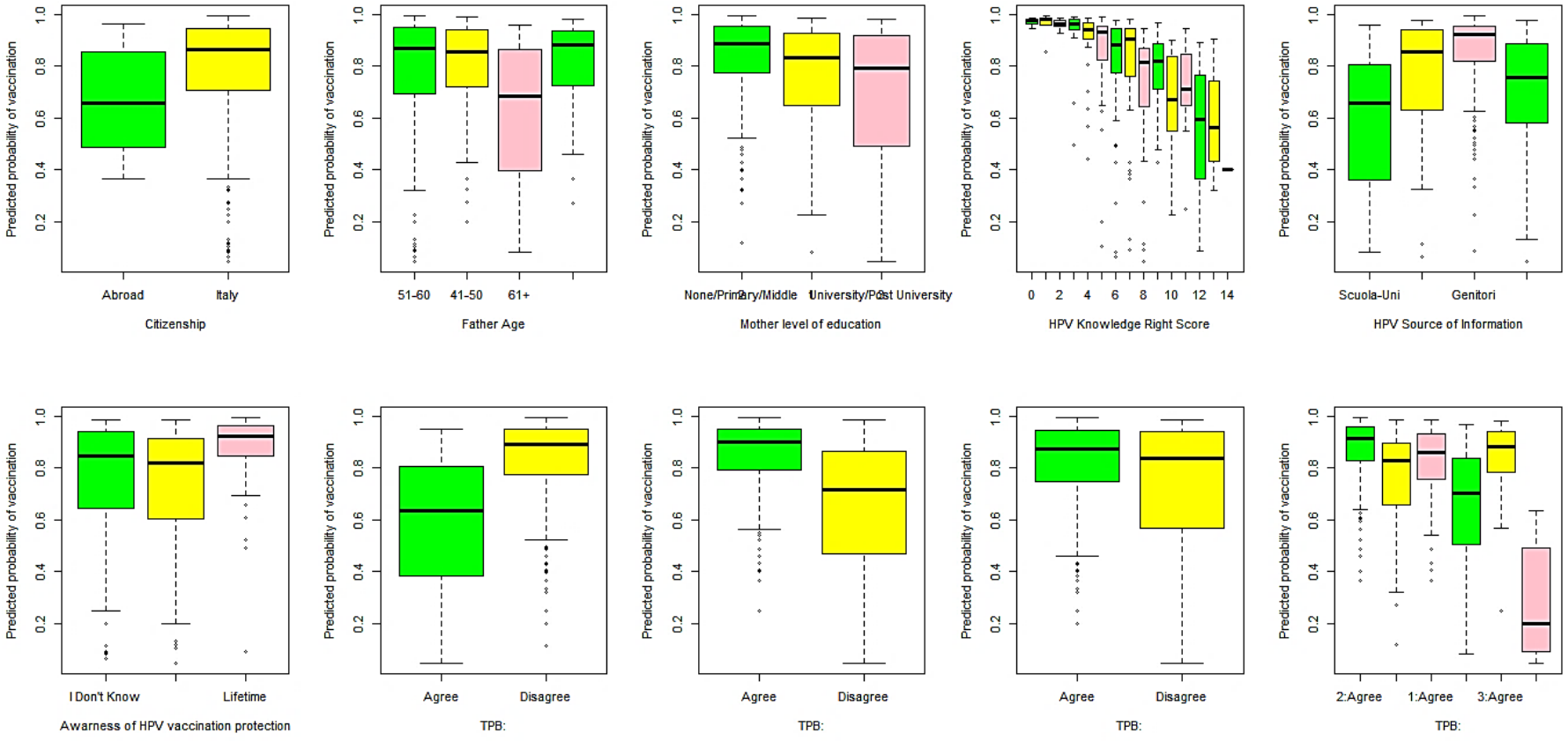


Figure 28 Significant covariates plot

5.4.Determinants of intentions

This research applied the theory of planned behaviour (TPB) and part of the health belief modes (HBM) and the social influence theory (SIT) to investigate behaviours related with prevention of HPV infections and cervical cancer. Psychological theories were considered in order to identify which factors, beyond the simple rational ones, may explain past and future preventive behaviours. Logistic regression was applied to explain determinants of past vaccination choices whereas future behaviours were assessed considering a combination of pre-existing theoretical frameworks. In fact, future intentions about prevention were measured by a set of possible interrelated latent (unobserved) variables assessed by multiple manifest (observed) variables or indicators. To evaluate the significance of interrelationships between psychological latent constructs (i.e. Theory of Planned Behaviour, Health Beliefs Model and the Social Influence Theory) partial least square structural equation models were employed. As explained more in depth in appendix 3., this statistical approach enables researchers to answer in depth a set of linked research questions in a single, systematic and comprehensive analysis by simultaneously modelling the relationships among independent and dependent dimensions (Gefen, Straub, & Boudreau, 2000).

Before the SEM approach, exploratory factor analysis was considered to analyse the reliability of the new psychometric scales, i.e. to identify how observed measures (indicators) explained the correspondent latent dimensions.

The analysis was conducted considering two subgroups of respondents "*Vaccinated and Unvaccinated respondents that know pap-test*" and "*Unvaccinated respondents*". The first group includes respondents that declared to know the main cervical cancer screening methods.

In fact, filter questions hided the pap-test section to those respondents that were unaware of it. The second group, instead, is composed by those students that were not vaccinated. Attitudes, subject norms and self-efficacy related to future vaccination intentions were investigated for this subgroup. The group sample size was respectively of 351 and 82 respondents.

5.4.1. Psychological scales evaluation: Exploratory factor analysis

In the first analysis (Vaccinated and Unvaccinated respondents that know pap-test) a total of 17 items were included, whereas for the second subgroups ((unvaccinated respondents) the number of items considered was of 13.

The TPB and HBM questionnaire data were subject to psychometric analysis after being inspected for factorability, or suitability for factor analysis, based on Bartlett's test of sphericity ($p < .05$; Bartlett, 1954), and the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO> 0.50; Kaiser, 1970, 1974). The KMO overall measure of sampling adequacy was respectively of 0.75 and 0.8 and the Bartlett's test of sphericity was significant for both subgroups. These results suggested that the data can be appropriately used for factor analysis.
Six factors were extracted for the first subgroups and five for the second (Table…) considering a principal axis factor analysis with varimax rotation as latent constructs were not related. The number of factors extracted was established comparing the results of parallel analysis (Ledesma e Valero-Mora, 2007), the scree plot (Cattell, 1966) and the Very Simple Structure Criterion (VSS-test) (Ledesma e Valero-Mora, 2007).
The six factors for the "Vaccinated and non-vaccinated" subset explain approximately the 45% of the total variance whereas the four factors exacted in the "Unvaccinated" subgroup account for the 66% of the total variability. Referring to table 14 (Vaccinated and Unvaccinated subgroup), the first factor contains five items measuring vaccine efficacy, vaccine safety, fear of side effects, general vaccine scepticism (negative loading) and government and pharmaceutical distrust (negative loading). As consequences, the factor was named "Safety and benefits of HPV vaccination" and explains the 13% of the total variance. Meanwhile, the second factor explains the 11% of the total variance and contains three items measuring fear of pap-test and pap-test discomfort, and one measuring pap-test unnecessary after vaccination. The factor was named "Fear and unnecessary of Pap-test". The third factor contains two items measuring the perceived severity of cervical cancer (HBM constructs). As consequence, the factor was named "Perceived disease severity" and explains the 8% of the total variance. The fourth factor contains two items measuring future intentions about pap-test (Future Intentions), the fifth the two social norms items (Social Norms) and the sixth the two perceived susceptibility indicators (Perceived Susceptibility towards cervical cancer). They explain respectively, the 7% e the 3% of the total variance. However, three items measuring respectively, Perceived susceptibility, Pap-test utility and Social Norms were excluded from the analysis because their factor loadings were lower than 0.3. In addition, the perceived behavioural control dimension was undetected, i.e. the number of item considered as indicators was insufficient.

Table 14 Explorative factor analysis "Vaccinated and Unvaccinated that know pap-test"

Items (*)	Safety and benefits of HPV vaccination	Fear of Pap-test	Perceived disease severity	Future intentions	Social norms	Perceived Susceptibility of HPV infection
PercSev1 (N.32a)	0.18	-0.22	**0.81**	-0.02	0.09	-0.23
PercSev2 (N.32b)	0.02	0.04	**0.72**	0.13	0.05	0.15
AttVacc1 (N.37)	**0.53**	-0.04	0.27	0.04	0	-0.06
AttVacc2(N.37)	**-0.37**	0.06	0.11	-0.01	-0.05	0
AttVacc3(N.37)	**-0.70**	0.14	0.05	-0.01	-0.15	0.11
AttVacc5(N.37)	**0.73**	-0.07	0.18	0.05	-0.06	-0.03
AttVacc6(N.37)	**0.85**	0.01	0.09	0.06	-0.03	0.05
AttPap4 (N.37)	0.01	**0.40**	-0.16	-0.14	-0.16	-0.09
AttPap1 (N.57)	-0.1	**0.77**	0.04	-0.09	0.07	0.1
AttPap2 (N.57)	-0.04	**0.57**	-0.1	-0.18	0.08	-0.04
AttPap3 (N.57)	-0.14	**0.68**	0.06	-0.07	-0.09	-0.02
FutureInt1(N.58)	0.18	-0.47	0.17	**0.69**	0.12	0.01
FutureInt2 (N.59)	0.03	-0.26	0.04	**0.79**	0.02	-0.07
PercSusc1 (N.51)	0.05	-0.08	0.01	-0.04	-0.12	**0.53**
PercSusc2 (N.52)	-0.1	0.03	-0.02	0.01	0.15	**0.34**
Norm1 (N.62)	0.1	-0.06	0.02	-0.01	**0.37**	-0.13
Norm2 (N.63)	0.02	-0.01	0.15	0.09	**0.52**	0

(*) The number between brackets is the question number in the appendix questionnaire.

Item explanation can be found in table 16.

Table 15 summarized the factor loadings for the second respondent subgroups. Results shows that the first factor contains six items measuring fear of vaccine side effects, vaccine efficacy/safety and general vaccine scepticism. In particular, items measuring vaccine efficacy and safety loaded negatively with the latent dimension. Therefore, the factor was named "HPV vaccination scepticism". It explains the 27% of the total variance. The second factor explains the 14% of the variance and contained three items measuring the perceived susceptibility towards cervical cancer (HBM constructs). As consequence, the factor was named "Perceived Susceptibility". Finally, the third and fourth factors contain respectively, the two social norms items and the two intentions indicators. They were named Social Norms and Intentions and accounted for the 13%, 11% of the total variance. However, five items were excluded due to a low factor loading (<0.3) and the Perceived Behavioural Control dimension remained undetected.

Table 15 Exploratory factor analysis "Unvaccinated"

Item (*)	Vaccine Scepticism	Perceived Susceptibility	Social Norms	Intentions
Norms1(N.62)	-0.07	0.09	**0.81**	-0.09
Norms2(N.63)	-0.16	-0.1	**0.89**	0.07
PercSusc1(N.51)	-0.11	**0.79**	-0.2	-0.06
PercSusc2(N.53)	0.11	**0.76**	0.25	-0.23
PercSusc3(N.54)	0.26	**0.71**	0.13	0.15
Int2(N.58)	0.31	0.14	0.27	**0.57**
Int1(N.59)	0.38	0	0.24	**0.64**
AttVacc2 (N.37)	**0.67**	-0.07	0.2	-0.05
AttVacc3(N.37)	**0.80**	0.11	-0.23	-0.12
AttVacc5(N.37)	**-0.74**	-0.12	0.32	0.14
AttVacc6(N.37)	**-0.64**	0.04	0.34	0.02
AttVaccNV3(N.46)	**0.63**	0.07	0.2	-0.18
AttVaccNV4(N.46)	**0.79**	0.21	-0.04	0.03

(*) The number between brackets is the question number in the appendix questionnaire

Item explanation can be found in table 23

To evaluate model fit RMSEA and Lewis Index of factoring reliability were considered. Model fit is appropriate when the value of RMSEA drops below .05 and the Lewis Index approach 1, larger value indicating better reliability (Hu & Bentler, 1999). The RMSA and Lewis Index values were, respectively, of 0.06 and 0.9 for the first group and of 0.08 and 1 for the unvaccinated subset. Both values indicated an acceptable model fit and reliability.

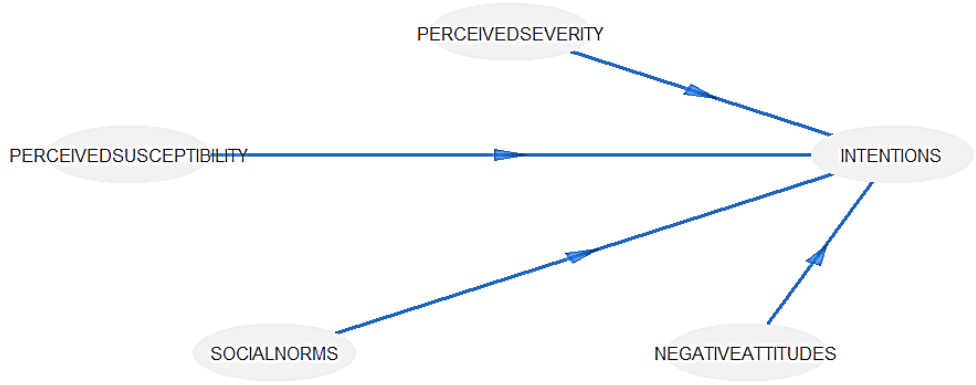
5.4.2. PLS structural equation models

As explained in appendix 3, the PLS-Path Modelling (PLS-PM) algorithm (Wold, 1975; Tenehaus et al., 2005) represents a valid alternative to covariance based structural equation models (CB-SEM), especially suited for situations when data is not normally distributed (Monecke & Leisch, 2012). PLS approach is a soft-modelling-technique with minim demand regarding measurement scales and residuals distribution (Monecke & Leisch, 2012). PLS path models consists of two main components, the measurement model and the structural model. The first also called outer model related observed variables (or manifest variables-MV) to their latent variables (LV). The structural model or inner model, instead, specifies the relationship between latent variables (LV). LVs are generally divided in two main classes, exogenous and exogenous Monecke & Leisch, 2012). Exogenous LVs do not have any predecessor in the structural models, i.e. appear only as independents, whereas endogenous are caused by one or more variables in the model and may also cause other endogenous variables, i.e. appear as dependents or dependents/independents. Based on SEM literature, LV can be modelled using either formative or reflective indicators. In the first case, manifest variables are considered as being caused by the corresponding latent variables (the arrow direction points from LVs to reflective indicators). In the second case, the latent construct is supposed to be formed by its indicators (Sanchez, 2013). In formative construct, the arrow direction points manifest variables to LVs.
The PLS algorithm estimates, through a system of independent equation based on multiple regressions, the network of relations among manifest variables and their corresponding latent variables, and among the latent variables in the model (Henseler et al., 2009; Ingrassia & Trinchera, 2008).

5.4.2.1. Future pap-test intentions PLS PM models

The dataset is composed by 351 respondents that affirmed to be knew pap-test. The theoretical framework to measure vaccination intentions with its key elements and cause-effect relationships is displayed in figure 29. Only items (manifest variables) with loaded greater than 0.3 in the exploratory factor analysis were included in the final model. One of the five theoretical constructs (perceived behavioural control) characterizing the original model was removed. As already showed in the factor analysis section, the perceived behavioural control dimension was undetected.

Figure 29 Future pap-test intentions PLS-SEM model



The model had four exogenous latent variables (negative attitudes, social norms, perceived susceptibility and perceived severity) and one dependent construct (intentions), which is also called endogenous (Hair, Hult, Ringle, & Sarstedt, 2014). The analysis was performed in R considering the PLSPM (Sanchez et al., 2017) package.

The different latent dimensions considered in the SEM model, presented in blocks of indicators with their corresponding questions are summarized in table 16. Attitudes/beliefs, intentions and perceived disease severity were measured by a 5-point Likert scale, social norms were assessed by multiple choice questions and perceived susceptibility by a probability thermometer scale (Section G Appendix 5).

Table 16 Pap-test PLS SEM latent dimensions and manifest variables

Latent dimension	***Indicators***	***Questions***	***Median***
Negative Attitudes	AttPap1(N.57)	I am scared to have a pap-test because I might have an abnormal result. (Ho paura di effettuare il pap-test perché mi spaventerebbe un eventuale risultato negativo)	2
	AttPap2(N.57)	I am too young to undergo a pap-test. (Credo di essere troppo giovane per dover effettuare l'esame del pap-test)	1
	AttPap3(N.57)	I am embarrassed when I have a pap test. (L'esame del pap-test mi mette a disagio)	2
	AttPap4(N.37)	I believe that a girl vaccinated against HPV does not need a pap-test. (Io credo che chi si vaccina contro l'HPV non abbia più bisogno di sottoporsi all'esame del pap-test in futuro)	2
Subjective Norms	Norm1(N.62)	Which person's opinion may influence your future choice about pap-test? (Se domani ti venisse riproposta la vaccinazione, il parere di quali persone potrebbe influenzare la tua scelta futura?)	4

	Norm2(N.63)	Which person would support your decision to undergo a pap-test? (Quali di queste persone sarebbe a favore di una rua eventuale vaccinazione?	4
Intentions	FutureInt1 (N.58)	How likely do you undergo a pap-test in future? (Quanto è probabile che ti sottoponga all'esame del pap-test in futuro?)	3
	FutureInt2 (N.59)	How many times do you undergo a pap-test in future? (Con quale frequenza ritieni di dover effettuare il pap-test?)	1
Perceived susceptibility to cervical cancer	PercSusc1 (N.51)	Chance of a girl to get cervical cancer in the future without vaccination and pap-test. (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che non si vaccinate e non si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	2
	PercSusc2 (N.52)	Chance of a girl to get cervical cancer in the future if vaccinated but not undergoing a pap-test. (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che sai vaccinata ma non si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	2
Perceived disease (cervical cancer) severity	PercSev1 (N.32a)	I believe HPV can be extremely harmful. (Credo che il tumore al collo dell'utero sia una malattia molto grave)	5
	PercSev2 (N.32b)	The thought of get cervical cancer in future scares me. (Mi spaventa pensare che potrei ammalarmi di cancro al collo dell'utero in futuro)	5

5.4.2.2. PLS-SEM model assessment

Similar to covariance structure analysis, applying the PLS algorithm requires an extensive model evaluation. Unlike with CB-SEM, a single goodness-of-fit criterion is not available in PLS-SEM (Hair et al., 2013). CB-SEM fit statistics are based on how closely the model-implied covariance matrix approximates the observed covariance matrix. PLS-SEM model evaluation instead, focuses on the discrepancy between the observed (for MVs) or approximated (for LVs) values of the dependent variables and the values predicted by the model in question. As consequence, the evaluation of PLS-SEM results builds on a set of non-parametric criteria and uses procedures, such as bootstrap, assessing the model's predictive capabilities (Garson, 2016; Hair et al., 2013; Chin, 2010). In PLS-SEM fit evaluation involves different types of measures according to the part of the model that should be assessed. In fact, there are three types of fit with researchers may be concerned, measurement model with reflective indicators, measurement model with formative indicators and structural model.

The extent literature lack of clear guidelines for the definition of TPB, HBM and SIT latent dimensions as reflective or formative. Among the available studies Nchise (2012) modelled all the TPB LVs as reflective whereas Hennessy et al. (2012) differentiated by direct and indirect TPB measures, modelling the firsts as reflective and the seconds as formative. Freeze & Raschke (2008) and Bolleng and Ting (2000), Coltman et al. (2007) and Chin (1998) provided some useful advices to help researchers in the definition of formative or reflective LVs. Bolleng and Ting exhorted researchers to imagine "*… a shift in the latent variable and then judges whether a simultaneous shift in all the observed variables is likely. If so, then this is consistent with an effect indictor specification. Alternatively, if the researcher imagines a shift in an observed variable as leading to a shift in the latent variable even if there is no change in the other indicators, then this is consistent with a causal indicator model* (Bollen and Ting 2000, p. 4). Reflective items are viewed as affected by the same underlying concept (i.e. the latent construct). A change in the latent dimension will affect the measurement items. An example of reflective measures might be the blood alcohol level, the driving ability and the MRI brain scan ad indicators of mental inebriation. If the amount of blood alcohol increases, i.e. there is a change in the manifest variables, it implies a change in other measures (driving ability, MRI brain scan, etc.) since they belong to the same latent concept or phenomenon (Chin, 1998). For this reason, indicators for a reflective construct should be consistent internally because all of the measures are assumed to be equally valid indicators of the underlying LV (Petter, Straub, & Rai, 2007). Thus, the uses of internal reliability measures are required to ensure the measures are reliable. In addition to that, a reflective construct should be uni-dimensional and if any measures are removed, it would not affect the content validity (Petter et al., 2007). Vice versa, formative measure would be the amount of beer, wine and hard liquor consumed as indicators of mental inebriation. If and individual increases the amount of beer consumed it affects the latent dimension but no other indicators, i.e. wine and hard liquor. Thus, while it may occur, formative indicators need not be correlated nor have high internal consistency such as Cronbach's alpha (Bollen 1984; Bollen and Lennox 1991).

For this study, all the latent dimensions were considered as reflective indicators. In fact, the "Attitudes" construct was characterized by negative and positive beliefs about the HPV vaccination (table 16.). As consequences, a change in certain belief, for example the vaccine safety or the fear of side effects, will affect all the other indicators. A similar line of reasoning was adopted for the remaining LVs, social norms, perceived susceptibility and severity. Then, a measurement model with reflective indicators and a structural model will be evaluated.

5.4.2.3. Pap-test intention PLS-SEM: Measurement model with reflective indicators assessment (Outer model)

The diagnosis of a PLS path model begins with the assessing of the quality of the measurement model. As explained in appendix, the evaluation of reflective indicators should involve three main aspects, unidimensionality (internal consistency) of indicators, the correlation between the latent variable and its indicators and correlations with other latent dimensions (Sanchez, 2013). Since PLS-PM does not rely on any distributional assumptions, resampling procedures are used to obtain information about the variability of the parameters estimates (Efron and Tibshirani 1993; Sanchez, 2013). Inner and outer model results are validated by bootstrap procedures. For this study, 5,000 bootstrap samples of 351 observations were considered (Hair et al., 2013).

Indices to evaluate unidimensionality are presented in table 17. Since reflective indicators will reflect the latent variable that they are associated with, a change in the latent construct (increase or decrease) will translate in a change of all its indicators. Unidimensionality implies that variables that are supposed to be measuring the same latent variable must be in a geometrical space of one dimension. The R package PLSPM provides three main indices to check unidimensionality, the Crobach's alpha, the Dillon-Goldestein rho and the first eigenvalue of the indicators (Appendix 3). The first is an average inter-variable correlation and measures how well a block of indicators measure their corresponding latent constructs. As a rule of thumb, a Crobach's alpha greater than 0.7 are considerable acceptable (Sanchez, 2013; Chin, 2010), but values of 0.60 to 0. 70 are acceptable in exploratory research (Hair et al., 2013; Nunally & Bernstein, 1994). The second index focuses on the variance of the sum variables in the block of interest. According to Chin (1998), Dillon-Goldstein's rho is considered to be a better indicator than Cronbach's alpha (Sanchez, 2013; Vinzi et al., 2010). In fact, each MV is assumed to be equally important in defining the latent variable (tau equivalence) (Vinzi et al., 2010). Block is considered homogenous if the DG-rho is larger than 0.7 (Sanchez, 2013; Vinzi et al., 2010; Wertz et al. 1974). Finally, a block may be considered unidimensional if the first eigenvalue of its correlation matrix is higher than 1, while the others are smaller (Kaiser's rule) (Vinzi et al., 2010). Inspecting table 17 can be noticed that the indices values are all acceptable, except for the Social Norms Cronbach's alpha value. Since the DG-rho and the first eigenvalue showed block unidimensionality, the latent construct was retained in the model as reflective.

Table 17-Measurement model indices

Latent variable	*Mode*	*MVs*	*C.alpha*	*DG.rho*	*eig.1st*	*eig.2nd*
PERCEIVED SEVERITY	Reflective	2	0.6469	0.8500	1.4781	0.5219
PERCEIVED SUSCEPTIBILITY	Reflective	2	0.7289	0.8806	1.5735	0.4265
SOCIAL NORMS	Reflective	2	0.2287	0.7217	1.1291	0.8709
NEGATIVE ATTITUDES	Reflective	4	0.6447	0.7898	1.9524	0.9022
INTENTIONS	Reflective	2	0.7706	0.8971	1.6268	0.3731

To examine the correlation between a latent variable and its indicators loadings and communalities were considered (table 18). Loadings represent the amount of correlation between a latent variable and its manifest items whereas communalities are squared correlation, i.e. squared loadings. As rule of thumb, loadings greater than 0.7 or communalities larger than 0.5 are acceptable (Sanchez, 2013; Hair et al. 2013). Nevertheless, assessment shows the range of 0.5-0.6 is acceptable for construct development as well as for the scaling construct (Chin, 1998b; Götz et al., 2010; Hair et al., 2011; Hulland, 1999; Latan & Ghozali, 2012a). In addition, the redundancy index measures the portion of variability of the manifest variables connected to the endogenous latent variable (intentions) explained by the latent variables directly connected to the block.

Table 18- Loadings, communalities and redundancy (Bootstrap estimates)

MV's	LVs	Loadings	perc.025	perc.975	Comm.	Red.
PercSev1	PERCEIVEDSEVERITY	0.6270	-0.9152	0.9829	0.3931	0
PercSev2	PERCEIVEDSEVERITY	0.5514	-0.9401	0.9716	0.3040	0
PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.6116	-0.9103	0.9424	0.3741	0
PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.4536	-0.8661	0.914	0.2058	0
Norm1	SOCIALNORMS	0.7814	0.4639	0.9245	0.6106	0
Norm2	SOCIALNORMS	0.7796	0.6113	0.9223	0.6078	0
AttPap1	NEGATIVEATTITUDES	0.8058	0.7008	0.8879	0.6493	0
AttPap3	NEGATIVEATTITUDES	0.6642	0.5229	0.7815	0.4412	0
AttPap2	NEGATIVEATTITUDES	0.7226	0.5462	0.8505	0.5222	0
AttPap4	NEGATIVEATTITUDES	0.5733	0.29	0.8591	0.3287	0
FutInt1	INTENTIONS	0.9314	0.8479	0.9812	0.8675	0.4363
FutInt2	INTENTIONS	0.8691	0.7383	0.9831	0.7553	0.3503

Considering that the psychological scales tested are new, the loadings and communality values are all acceptable except for *PercSusc2* (Table 13). Hair et al. (2013) recommended that researchers should carefully examine the effect of item removal on the composite reliability, as well as on the construct's content validity. If the elimination of weaker indicators increases

Crobach'alpha, DG-rho and first eigenvalue above of the suggested threshold values, they should be eliminated from the model. Unidimensionality indices for Perceived Susceptibility were already up to the threshold values and then weaker outer loadings were retained in the model. As indicated in the table, bootstrap confidence intervals are acceptable except for the Perceived Severity and Perceived Susceptibility indicators, for which are very wide. Sample heterogeneity, that will be analysed after model assessment, may be a possible cause of these results. In addition, data sparseness could be the second reason of high sample variability. Lastly, all construct items were newly developed and for some of them psychometric quality should be improved.

Besides checking the loadings of the indicators with their own latent variables, cross-loading should also be inspected. They represent the loads of a certain indicator with the rest of latent dimensions. In fact, the loadings of the manifest variables to their latent construct should be higher than any other loadings. Cross-loading (table 35 appendix 3) showed that assumptions are met.

5.4.2.4. Pap-test intention PLS-SEM: Structural model assessment (Inner model)

Once we have confirmed that the construct measures are reliable and valid, the next step addresses the assessment of the structural model results. This involves examining the model's predictive capabilities and the relationships between the constructs

The quality of the structural model is evaluated by examining four indices of quality metrics, the R^2 determination coefficient, the redundancy index, the goodness of fit (GOF) and the AVE (Vinzi et al., 2013; Sanchez, 2013) presented in table 19. In addition, correlation between latent constructs should be inspected.

The coefficient of determination represents the exogenous latent variables' combined effects on the endogenous latent variable (Hair et al., 2013; Vinzi et al., 2013). According to Hair et al. (2013) it is difficult to provide rules of thumb for acceptable R^2 values as this depends on the model complexity and the research discipline. For example, in consumer behaviour studies values of 0.20 are considered high. According to Sanchez (2013) and Vinzi et al. (2013) a R^2 value of greater than 0.60 (or sometimes 0.50) are considered substantial, values included between 0.30 and 0.60 are average and values smaller than 0.30 are considered weak. Similar recommendations are provided by Chin (1998b). With a value of 0.36 and a bootstrap value of 0.42 (CI: 0.2372; 0.5847) the model R^2 fall in the medium category.

The *average redundancy* instead, reflects the ability of the independent latent variables to explain the average variation of the indicators in the dependent latent variables. In this case

the average redundancy for intention implies that Perceived Severity, Perceived Susceptibility, Social Norms and Negative Attitudes predict the 27% of the variability in Intention indicators. The last column is the *Average Variance Extracted* which measures the amount of variance that a latent variable capture from its indicators in relation to the amount of variance due to measurement error. As rule of thumb AVE should be greater than 0.50, which means that 50% or more of the indicator's variance is accounted for (Sanchez, 2013; Hair et al. 2013). As indicated in table 19, Negative Attitudes presented a communality lower than 50%. The lowest value is probably caused by the presence of "weaker" manifest variables in the latent dimensions (Table 18).

A global criterion of goodness of fit has been proposed by Tenenhaus et al. (2004): the GoF (Appendix 3) index (Tenenhaus et al., 2004; Tenenhaus, Esposito Vinzi, Chatelin, & Lauro, 2005, Hair et al., 2013). Such an index has been developed in order to take into account the model performance in both the measurement and the structural model and thus provide a single measure for the overall prediction performance of the model. The model GoF is 0.48. Tenenhaus et al. (2004) and Vinzi et al. (2013) recommended a GoF value greater than 0.90, but lower values are more common in empirical studies (Ingrassia & Trinchera, 2008). In addition, the presence of non-linear relation as well as unobserved heterogeneity (paragraph 9.4.2.5.) among respondents, may cause small GoF and R^2 values.

Table 19- Structural model indices

Latent dimensions	Type	R2	Mean_Redundancy	AVE
PERCEIVED SEVERITY	Exogenous	0	0.0000	0.5978
PERCEIVED SUSCEPTIBILITY	Exogenous	0	0.0000	0.8765
SOCIAL NORMS	Exogenous	0	0.0000	0.6242
NEGATIVE ATTITUDES	Exogenous	0	0.0000	0.4646
INTENTIONS	Endogenous	0.3620	0.2722	0.7928

In addition, this study also assesses the mediation relationships that are being proposed in the research model (figure 30). The path coefficients are essentially standardized regression coefficients and they should be interpreted as the change in the dependent variable if the independent variable is increased by one and all other covariates remain constant (Henseler et al., 2015). Based on the literature, the path coefficient value needs to be at least 0.1 to account for a certain impact within the model (Hair et al., 2011; Henseler et al., 2015).

Based on the results showed in table 20 and figure 26 it is possible to see that among the four diver constructs only *Negative Attitudes* (towards Pap-test) showed a negative significative

total effect (-0.4433) on *Pap-test Intentions*. Other latent dimensions seemed not influencing the respondents' future intentions towards pap-test screening.

Figure 30- PLS-SEM global model for pap-test intentions

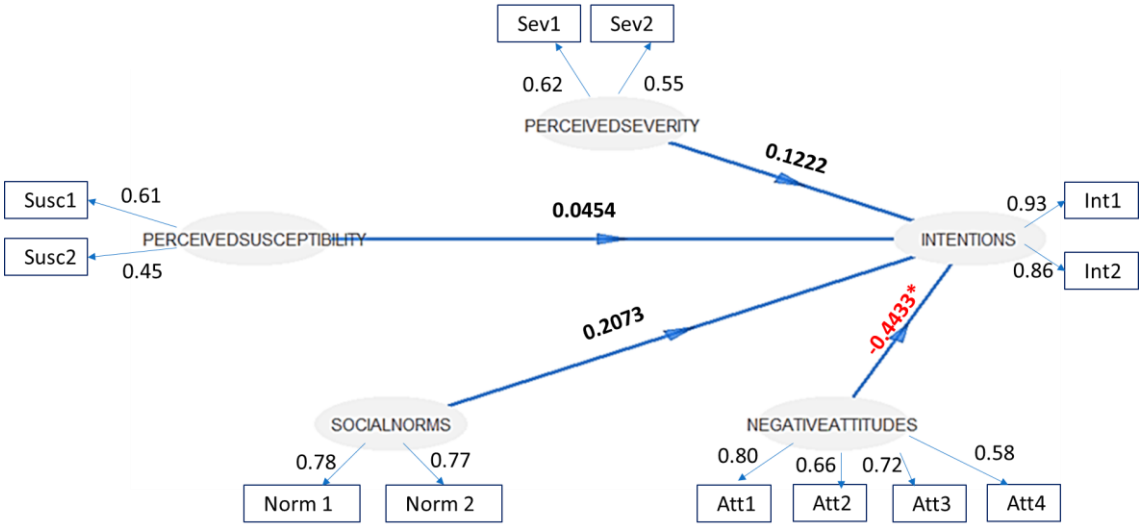


Table 20- Structural model results for the global model. Bootstrap path coefficients, SE, CI and p-values.

Exogenous LV	Path Coeff.	SE	CI:25%	CI:75%
Intercept	-	-	-	-
PERCEIVED SEVERITY	0.1222	0.0796	-0.1742	0.2881
PERCEIVED SUSCEPTIBILITY	0.0454	0.1452	-0.2532	0.2398
SOCIAL NORMS	0.2073	0.1582	-0.1788	0.4707
NEGATIVE ATTITUDES	-0.4433	0.1034***	-0.6014	-0.2821

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

5.4.2.5. Pap-test intention PLS-SEM: Unobserved heterogeneity

In this section, it is presented how to improve the prediction performance and the interpretability of the model by allowing for unobserved heterogeneity.

As indicated by Hair et al. (2013), heterogeneity of observations is another important aspect of structural model evaluation. The presence of heterogeneity can distort the results affecting the validity of the whole model (e.g., Rigdon, Ringle, Sarstedt, & Gudergan, 2011; Ringle, Sarstedt, & Mooi, 2010; Sarstedt, Schwaiger, & Ringle, 2009). In fact, in applied research is common to encounter a situation in which different parameters occur for different subpopulations, such as segments of consumers, countries, respondents. If heterogeneity is presented in the sample, two or more subgroups exhibit different underlying relationships with the latent constructs

(Hair et al., 2013). With regard to the thesis, for example, some respondents' pap-test future intentions may be influenced more by Perceived Susceptibility than by Negative Attitudes or Social Norms. There exist two types of heterogeneity, the observed and the unobserved heterogeneity (Tenenhaus et al. 2010; Hensler and Fassott 2010; Chin and Dibbern 2007). In the first case the composition of classes is known a priori, while in the second case information on the number of groups or on their composition is not available (Vinzi et al., 2013). There are not previous information about possible classes for the research topic of this thesis, in fact the extent literature lacks applications in similar fields of study especially for the Italian country. For this reason, unobserved heterogeneity was investigated.

Several methods have been proposed to deal with unobserved heterogeneity in PLS-PM framework (Hahn et al. 2002; Ringle et al. 2005; Squillacciotti 2005; Trinchera and Esposito Vinzi 2006; Trinchera et al. 2006; Sanchez and Aluja 2006, 2007; Esposito Vinzi et al. 2008; Trinchera 2007). In this thesis, the REBUS (Response Based Unit Segmentation) algorithm of Trinchera (2007) and Esposito Vinzi et al. (2008) was considered. A brief description of the algorithm is provided in A3.1.4. in appendix 3.

The REBUS approach allows to estimate at the same time the unit membership to latent classes and the class specific parameter of the local model. In fact, if a latent class exist, unit belonging to the same latent class will have similar performances as regard to the global model, i.e. similar association between latent variables.

The global model results, showed in paragraphs 9.4.2.3 and 9.4.2.4, underlined that only one significant relation was detected in the structural model, according to bootstrap CI and p-values, and factor loadings for Perceived Susceptibility were not sufficiently high for some of its indicators. In addition, GOF and R^2 values were quite modest. The application of REBUS-PLS has led to identification of two latent classes. The first latent class is composed by 99 units, i.e. 28% of the sample, and the second by 252 (72% of the sample). The local models are summarized in table 21. The table shows the differences in the structural model results between the two groups. For group 1 the only exogenous variables with a significant impact on Intentions are Negative Attitudes, with a path coefficient of -0.44, and Perceived Susceptibility, with a path coefficient of -0.21. On the other hand, in group 2 none of the exogenous variables are significantly related with the endogenous Intentions. A little improvement of structural model indices was observed for the first group, showing a R2 of 0.5585 (0.4041;0.7089) and a GoF of 0.50. In addition, the exogenous latent variable accounted for the 38% of the intentions variability. For the second group instead, structural model performances were worsted than for the global one.

Table 21 Structural models results for local models of Group 1 and Group 2. Bootstrap Path Coeff., CI and P-value.

	Endogenous LV	Exogenous LV	Path Coeff.	perc.025	perc.975
Group 1					
	INTENTIONS	PERCEIVED SEVERITY	0.2892	-0.1428	0.4892
		PERCEIVEDSUSCEPTIBILITY	-0.2138*	-0.3612	-0.0052
		SOCIALNORMS	0.1492	-0.1168	0.3916
		NEGATIVEATTITUDES	-0.4444***	-0.6486	-0.2343
		R^2	0.5585	0.4041	0.7089
Group 2					
	INTENTIONS	PERCEIVED SEVERITY	0.1259	-0.2076	0.3570
		PERCEIVEDSUSCEPTIBILITY	-0.0355	-0.2619	0.3109
		SOCIALNORMS	0.0829	-0.3543	0.4181
		NEGATIVEATTITUDES	-0.3191	-0.5604	0.5023
		R^2	0.4147	0.2889	0.5533

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

As regard to the measurement model results, showed in table 22, the two classes differ in the loadings of the manifest variables for all the latent blocks. Especially Perceived Susceptibility, Social Norms and Intentions loading, and communalities are much lower in group 2 than in the first one. In particular, the manifest variable *Int2* (frequency of future pap-test intentions) has no impact in building the latent variable Intentions in the second group. Similar results are found for *norm1*(Influential persons).

Table 22 Measurement model results for the local model computed for group 1 and group 2.

	MVs	LVs	Loading	perc.025	perc.975	Comm.
Group 1						
		DG-Rho	0.7677			
	PercSev1	PERCEIVEDSEVERITY	0.6041	-0.5920	0.9635	0.3649
	PercSev2	PERCEIVEDSEVERITY	0.6284	-0.6892	0.9816	0.3949
		DG-Rho	0.8826			
	PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.7560	-0.6999	0.9420	0.5715
	PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.8082	0.4477	0.9405	0.6532
		DG-Rho	0.7494			
	Norm1	SOCIALNORMS	0.8457	0.6426	0.9451	0.7152
	Norm2	SOCIALNORMS	0.7850	0.4978	0.9245	0.6162
		DG-Rho	0.7732			
	AttPap1	NEGATIVEATTITUDES	0.8680	0.7473	0.9303	0.7534
	AttPap3	NEGATIVEATTITUDES	0.6838	0.4829	0.8290	0.4676
	AttPap2	NEGATIVEATTITUDES	0.5534	-0.2448	0.8684	0.3063
	AttPap4	NEGATIVEATTITUDES	0.5353	-0.3509	0.9067	0.2865

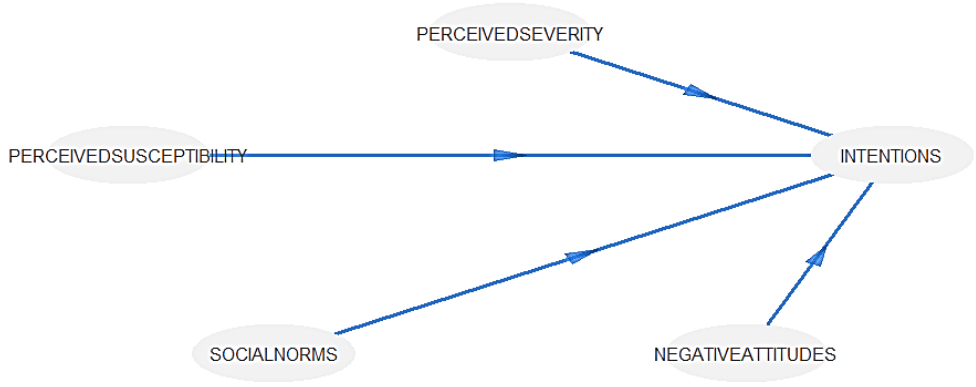
		DG-Rho	0.8273			
	FutInt1	INTENTIONS	0.9582	0.9084	0.9915	0.9181
	FutInt2	INTENTIONS	0.9379	0.7804	0.9917	0.8797
Group 2						
		DG-Rho	0.8709			
	PercSev1	PERCEIVEDSEVERITY	0.6994	-0.9657	0.9965	0.4892
	PercSev2	PERCEIVEDSEVERITY	0.6044	-0.8431	0.9969	0.3653
		DG-Rho	0.8790			
	PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.5288	-0.9064	0.9621	0.2796
	PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.5388	-0.9226	0.9639	0.2903
		DG-Rho	0.7107			
	Norm1	SOCIALNORMS	0.3115	0.0018	0.6523	0.0970
	Norm2	SOCIALNORMS	0.9268	0.7423	0.9931	0.8590
		DG-Rho	0.7788			
	AttPap1	NEGATIVEATTITUDES	0.6506	-0.8325	0.8921	0.4233
	AttPap3	NEGATIVEATTITUDES	0.5300	-0.6143	0.8018	0.2809
	AttPap2	NEGATIVEATTITUDES	0.6177	-0.7123	0.8765	0.3816
	AttPap4	NEGATIVEATTITUDES	0.5995	-0.7934	0.8899	0.3594
		DG-Rho	0.2100			
	FutInt1	INTENTIONS	0.7165	-0.8966	0.9921	0.5134
	FutInt2	INTENTIONS	0.4187	0.0840	0.9501	0.1753

5.4.3. Future vaccination intentions PLS-SEM model

The dataset considered for the analysis is composed by 82 respondents that affirmed to be unvaccinated. The theoretical framework to measure vaccination intentions with its key elements and cause-effect relationships is displayed in figure 31. Only items (manifest variables) with loaded greater than 0.3 in the exploratory factor analysis were included in the final model. One of the five theoretical constructs (perceived behavioural control) characterizing the original model was removed. The latent dimension was undetected in the explorative factor analysis because the number of items were insufficient.

The model had four exogenous latent variables (attitudes, social norms, perceived susceptibility and perceived severity) and one dependent construct (intentions), which is also called endogenous (Hair, Hult, Ringle, & Sarstedt, 2014).

Figure 31- Future vaccination intentions PLS-SEM model



The different latent dimensions considered in the SEM model, presented in blocks of indicators with their corresponding questions, are summarized in Table 23. Attitudes/beliefs, intentions and perceived disease severity were measured by a 5-point Likert scale, social norms were assessed by multiple choice questions and perceived susceptibility by a probability thermometer scale (Section G-Appendix 5).

Table 23 Future vaccination intentions model latent and manifest variables

LV	MV	Questions	Median
Attitudes/Beliefs towards vaccination	Att1	I believe that the vaccine distribution is strongly influenced by pharmaceutical companies. (Credo che la distribuzione del vaccino da parte dello stato sia influenzata dalle case farmaceutiche)	4
	Att2	I am very worried about the long-term side effects of the HPV vaccination. (Ho paura che il vaccino possa causare problemi di salute anche molto tempo dopo la vaccinazione)	3
	Att3	I believe that the protection induced by the vaccine is more important than its side effects (I credo che la protezione offerta dal vaccino sia più importante dei suoi eventuali effetti collaterali)	4
	Att4	I believe that the vaccine is safe. (Io credo che il vaccino sia sicuro)	4
	Att4	I am against vaccines. (Sono contraria a tutte le vaccinazioni)	2
	Att6	I am worried about post-vaccination side effects. (Mi spaventano i possibili effetti collaterali post-vaccinazione)	3
Perceived disease (cervical cancer) severity	PercSev1	I believe HPV can be extremely harmful. (Credo che il tumore al collo dell'utero sia una malattia molto grave)	5
	PercSev2	The thought of get cervical cancer in future scares me. (Mi spaventa pensare che potrei ammalarmi di cancro al collo dell'utero in futuro)	5
Subjective Norms	Norm1(N.44)	Which person's opinion may influence your future choice about HPV vaccination? (Se domani ti venisse riproposta la vaccinazione, il parere di quali persone potrebbe influenzare la tua scelta futura?)	2
	Norm2(N.45)	Which person would support your vaccination decision? (Quali di queste persone sarebbe a favore di una rua eventuale vaccinazione?	2
Intentions	FutureInt1	What would you do if the HPV vaccination is proposed again to you? (Cosa faresti se domani ti venisse riproposta la vaccinazione anti-HPV?)	4
	FutureInt2	Would you change your decision if you have to pay the vaccine? (Se il vaccino fosse a pagamento, cambieresti la tua decisione di vaccinarti?)	2
Perceived susceptibility to cervical cancer	PercSusc1	Chance of a girl to get cervical cancer in the future without vaccination and pap-test. (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che non si vaccinate e non si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	
	PercSusc3	Chance of a girl to get cervical cancer in the future if vaccinated but not undergoing a pap-test. (Quanto ritieni probabile che una ragazza con il tuo	4

	PercSusc4	stile di vita sessuale che non si vaccinata ma si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?) Chance of a girl to get cervical cancer in the future if vaccinated and undergoing a pap-test. (Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che sia vaccinata e si sottoponga a pap-test possa ammalarsi di tumore al collo dell'utero?)	4

5.4.3.1. Future vaccination intentions PLS-PM: Measurement model with reflective indicators (outer model) assessment

According to Chin (2000) the sample size of 82 respondents were sufficient. PLs is an iterative process that performs a series of simple OLS regressions. Thus, according to the OLS regression rule 20 cases per dependent variable are enough (Chin, 2000).

As explained above, the diagnosis of a PLS path model begins with the assessing of the quality of the measurement model. Unidimensionality of reflective indicators, loadings and cross-loadings are presented. Since PLS-PM does not rely on any distributional assumptions, resampling procedures are used to obtain information about the variability of the parameters estimates ((Efron and Tibshirani 1993; Sanchez, 2013). Inner and outer model results are validated by bootstrap procedures. For this study, 1000 bootstrap samples of 85 observations were considered (Hair et al., 2013). Inspecting table 24 can be noticed that the indices values are all acceptable, except for the Social Norms Cronbach's alpha value. Since the DG-rho and the first eigenvalue showed block unidimensionality, the latent construct was retained in the model as reflective.

Table 24 Measurement model unidimensionality indices

Latent variable	Mode	MVs	C.alpha	DG.rho	eig.1st	eig.2nd
PERCEIVED SEVERITY	Reflective	2	0.7118	0.8740	1.5525	0.4474
PERCEIVED SUSCEPTIBILITY	Reflective	4	0.7623	0.8492	2.3434	0.6626
SUBJECTIVE NORMS	Reflective	2	0.1879	0.7112	1.1037	0.8963
NEGATIVE ATTITUDES	Reflective	6	0.8903	0.9169	3.8943	0.6867
INTENTIONS	Reflective	2	0.9000	0.9522	1.8176	0.1823

Loadings, communalities and redundancy are summarized in table 25. Considering that new psychological scales and items were employed, the loadings and communality values are satisfactory. Confidence intervals size is reasonable for all manifest variables and redundancy index indicated that exogenous variables explain respectively, the 62% and the 64% of intentions indicators variance.

Table 25 Loadings, communalities and redundancy (bootstrap estimates)

MV	LV	Loadings	perc.025	perc.975	Comm	Red
PercSev1	PERCEIVED SEVERITY	0.8137	0.2465	0.9850	0.6622	0
PercSev2	PERCEIVED SEVERITY	0.8948	0.7010	0.9904	0.8007	0
PercSusc1	PERCEIVED SUSCEPTIBILITY	0.5126	0.0544	0.8290	0.2628	0
PercSusc2	PERCEIVED SUSCEPTIBILITY	0.6209	0.2584	0.8914	0.3855	0
PercSusc3	PERCEIVED SUSCEPTIBILITY	0.6623	0.3185	0.9024	0.4387	0
Norm1	SUBJECTIVE NORMS	0.7161	0.2053	0.9980	0.5129	0
Norm2	SUBJECTIVE NORMS	0.6764	0.0195	0.9841	0.4575	0
Att1	NEGATIVE ATTITUDES	0.7502	0.5812	0.8531	0.5628	0
Att2	NEGATIVE ATTITUDES.	0.8665	0.7868	0.9278	0.7509	0
Att3	NEGATIVE ATTITUDES.	0.8038	0.6734	0.9013	0.6461	0
Att4	NEGATIVE ATTITUDES	0.8118	0.6895	0.9009	0.6591	0
Att4	NEGATIVE ATTITUDES.	0.7797	0.6246	0.8920	0.6080	0
Att6	NEGATIVE ATTITUDES	0.6959	0.5052	0.5018	0.4842	0
Int1	INTENTIONS-	0.9431	0.8821	0.8923	0.8895	0.6165
Int2	INTENTIONS	0.9561	0.9328	0.9298	0.9141	0.6445

Cross-loadings of the indicators are summarized in table 36 in Appendix 3. Inspecting the results itis possible to see that loadings of manifest variables to their latent constructs are higher than any other loadings. Cross-loading assumptions are met.

5.4.3.2. Future vaccination intentions PLS-SEM: Structural model assessment (inner model)

Once we have confirmed that the construct measures are reliable and valid, the next step addresses the assessment of the structural model results. This involves examining the model's predictive capabilities and the relationships between the constructs 2013) presented in table 26. In addition, correlation between latent constructs should be inspected. With a value of 0.71 and a bootstrap value of 0.76 (CI: 0.62;0.84) the model R^2 fall in the high category (Chin, 1998b). R^2 value indicated that exogenous latent variables explain the 76% of the endogenous variance. The *average redundancy* instead, reflects the ability of the independent latent variables to explain the average variation of the indicators in the dependent latent variables. In this case the average redundancy for intention implies that Perceived Severity, Perceived Susceptibility, Subjective Norms and Negative Attitudes predict the 63% of the variability in the Intention indicators. The last column is the Average Variance Extracted which measures the amount of variance that a latent variable capture from its indicators in relation to the amount of variance due to measurement error. As rule of thumb AVE should be greater than

0.50, which means that 50% or more of the indicator's variance is accounted for (Sanchez, 2013; Hair et al. 2013). As indicated in table 24, Perceived Susceptibility presented a communality lower than 50%. The lowest value is probably caused by weaker indicators, as explained in the previous paragraph. The model GoF is 0.64. Tenenhaus et al. (2004) and Vinzi et al. (2013) recommended a GoF value greater than 0.90, but lower values are more common in empirical studies (Ingrassia & Trinchera, 2008). In addition, the presence of unobserved heterogeneity (paragraph 9.4.3.3.) among respondents, may influence GoF and R^2 values. Values of correlation among latent variables were lower than 0.20 and considered acceptable (Hair et al., 2013).

Table 26 Structural model indices

Latent dimensions	Type	R2	Mean_Redundancy	AVE
PERCEIVED SEVERITY	Exogenous	0	0.0000	0.7766
PERCEIVED SUSCEPTIBILITY	Exogenous	0	0.0000	0.2317
SUBJECTIVE NORMS	Exogenous	0	0.0000	0.5517
NEGATIVE ATTITUDES	Exogenous	0	0.0000	0.6178
INTENTIONS	Endogenous	0.7159	0.6305	0.8807

One of the study object is to assesses the mediation relationships that are being proposed in the research model (Figure 26), i.e. how psychological latent variable are interrelated. The path coefficients are essentially standardized regression coefficients and they should be interpreted as the change in the dependent variable if the independent variable is increased by one and all other covariates remain constant (Henseler et al., 2015). Based on the literature, the path coefficient value needs to be at least 0.1 to account for a certain impact within the model (Hair et al., 2011; Henseler et al., 2015).

Based on the results showed in table 27 it is possible to see that among the four diver constructs *Negative Attitudes* (towards Pap-test) showed the strongest (negative) significative total effect (-0.5904) on *Intentions towards HPV vaccination*, followed *Perceived disease severity* (0.2097) and *Subjective Norms* (0.1884). *Perceived susceptibility* seemed not influencing the respondents' future intentions towards HPV vaccination.

Figure 32- Global model for vaccination intentions

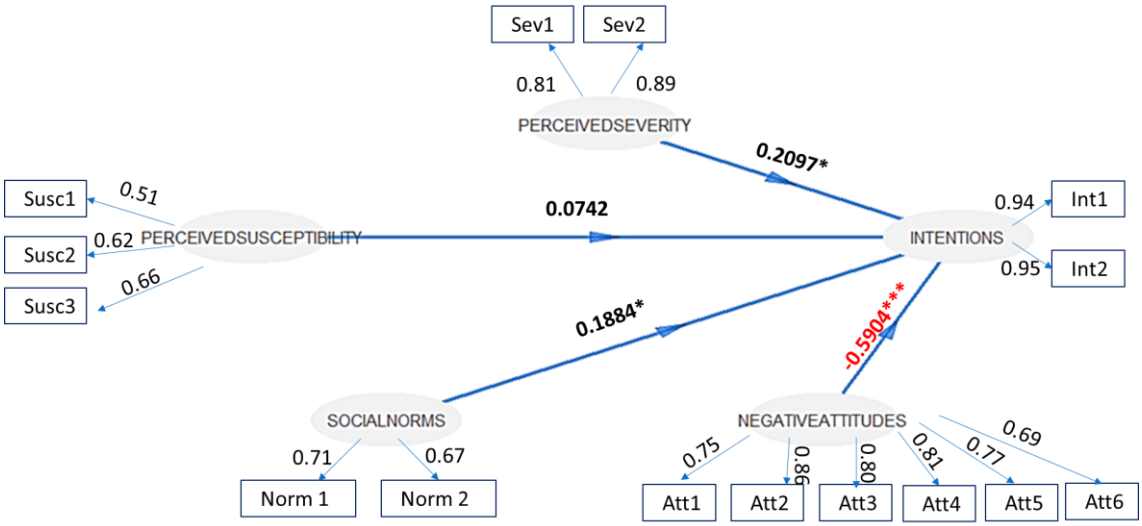


Table 27- Structural model results for the global model. Bootstrap path coefficients, SE, CI and p-values.

Exogenous LV	Path Coeff.	SE	CI:25%	CI:75%
Intercept	0.0000	-	-	-
PERCEIVED SEVERITY	0.2097	0.0911*	0.0230	0.3954
PERCEIVED SUSCEPTIBILITY	0.0742	0.1525	-0.2138	0.2864
SUBJECTIVE NORMS	0.1884	0.0725*	0.0476	0.3148
NEGATIVE ATTITUDES	-0.5904	0.0826***	-0.7638	-0.4309

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

5.4.3.3. Future vaccination intentions PLS-SEM: Unobserved heterogeneity

The unvaccinated dataset was characterized by a limited number of observations (82). For this reason, unobserved heterogeneity analysis was considered only for explorative purposes. The global model results, paragraph 9.4.3.1, underlined that only three exogenous variables were significant related with Intentions. In addition, some manifest variables showed lower loadings on the manifest model assessment. The application of REBUS-PLS has led to identification of three latent classes. The first latent class is composed by 32 units, i.e. 39% of the sample, the second by 17 (21% of the sample) and the third by 33 units (40% of the sample). The local models are summarized in table 28. Table 28 shows the differences in the structural model results between the three identified groups. An improvement of R^2 and GoF for all the three groups was observed. For group 1 a R^2 of 0.93 and a GoF of 0.81 were obtained, for group 2 of 0.85 and 0.70, and for group 3 of 0.94 and 0.76, which are greater than global model indices.

Table 28- Structural models results for local models of Group 1 and Group 2. Bootstrap Path Coeff., CI and P-value.

	Endogenous LV	Exogenous LV	Path Coeff.	perc.025	perc.975
Group1	INTENTIONS	PERCEIVED SEVERITY	0.1772	-0.0104	0.3222
		PERCEIVED SUSCEPTIBILITY	0.4343***	0.2757	0.5980
		SUBJECTIVE NORMS	0.2553**	0.0923	0.4223
		NEGATIVE ATTITUDES	-0.3147**	-0.5469	-0.1589
		R^2	0.9285	0.8736	0.9700
Group 2	INTENTIONS	PERCEIVED SEVERITY	0.0636	-0.1502	0.2125
		PERCEIVED SUSCEPTIBILITY	-0.2963*	-0.5115	-0.0984
		SUBJECTIVE NORMS	0.0689	-0.1591	0.2735
		NEGATIVE ATTITUDES	-0.6545***	-0.8102	-0.4888
		R^2	0.8471	0.7194	0.9353
Group 3	INTENTIONS	PERCEIVED SEVERITY	0.3393**	0.0905	0.5209
		PERCEIVED SUSCEPTIBILITY	0.0477	-0.2083	0.2201
		SUBJECTIVE NORMS	0.1618*	0.0212	0.3001
		NEGATIVE ATTITUDES	-0.6125***	-0.8412	-0.4113
		R^2	0.9439	0.9032	0.9748

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*'*

Inspecting the path coefficients in table 28 of the respondents in the first class, it is possible to see that three exogenous variables have a significant impact on future vaccination Intentions,

Perceived Susceptibility towards the HPV infection (0.4343:), Subjective Norms (0.2553) and Negative Attitudes (-0.3147). However, compared to the path coefficients of the other classes, respondents in class 1 perceive more the risk of infection and its consequences (cervical cancer) and the influences of others in taking future vaccination decisions. Intentions of respondents belonging to the second group are significantly and negatively influenced by two latent constructs, Perceived Susceptibility (-0.2963) and Negative Attitudes (-0.6545) toward vaccination. This implies that girls in class 2 do not perceive the risk of get cervical cancer and are mainly driven by negative attitudes. On the other hand, the Intentions of class 3 students are driven by Negative Attitudes (-0.6125), Perceived Severity of cervical cancer (0.3393) and Subjective Norms (0.1618). As regard the impact of the exogenous latent variable Perceived Severity, it is possible to notice that it has significant effect only on third group vaccination intentions.

As regard measurement model results, showed in table 29, the three classes differ in the loadings of the manifest variables for all the latent blocks. Class 1 showed the highest loadings for all the blocks of latent variables. Perceived Susceptibility and Subjective Norms and loading and communalities are lower in group 2 and 3 than in the first one. In particular, the manifest variable *Norm1* (Influential persons) has little impact in building the latent variable Subjective Norms in both, second and third groups. Similar results are found for *PercSev2* (Perceived susceptibility towards cervical cancer) for the second group of respondents. On the contrary, DG-Rho values are optimal for all the latent blocks, indicating high level of homogeneity within blocks of items (Chin, 1998).

Table 29 Measurement model results for the local model computed for group 1, group 2 and group 3

	MVs	LVs	Loading	perc.025	perc.975	Comm.
Group 1		DG-Rho	0.9076			
	PercSev1	PERCEIVED SEVERITY	0.8530	0.4834	0.9971	0.7276
	PercSev2	PERCEIVED SEVERITY	0.9237	0.8455	0.9972	0.8531
		DG-Rho	0.8806			
	PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.8769	0.7971	0.9346	0.7689
	PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.7161	0.4057	0.8884	0.5128
	PercSusc3	PERCEIVEDSUSCEPTIBILITY	0.7508	0.5388	0.8853	0.5637
		DG-Rho	0.8175			
	Norm1	SUBJECTIVENORMS	0.8572	0.6385	0.9727	0.7348
	Norm2	SUBJECTIVENORMS	0.7718	0.3741	0.9534	0.5956
		DG-Rho	0.9371			
	Att1	NEGATIVE ATTITUDES	0.7811	0.5438	0.9219	0.6100
	Att2	NEGATIVE ATTITUDES	0.8719	0.7504	0.9548	0.7602
	Att3	NEGATIVE ATTITUDES	0.8983	0.7888	0.9682	0.8069

	Att4	NEGATIVE ATTITUDES	0.8452	0.6981	0.9472	0.7144
	Att4	NEGATIVE ATTITUDES	0.8069	0.6116	0.9242	0.6510
	Att6	NEGATIVE ATTITUDES	0.8213	0.6237	0.9346	0.6746
		DG-Rho	0.9651			
	Int1	INTENTIONS	0.9688	0.9356	0.9940	0.9387
	Int2	INTENTIONS	0.9706	0.9396	0.9943	0.9420
Group 2		DG-Rho	0.8000			
	PercSev1	PERCEIVED SEVERITY	0.6180	-0.7629	0.9904	0.3819
	PercSev2	PERCEIVED SEVERITY	0.4576	-0.8282	0.9909	0.2094
		DG-Rho	0.8025			
	PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.7131	0.4870	0.8964	0.5085
	PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.8629	0.6798	0.9349	0.7446
	PercSusc3	PERCEIVEDSUSCEPTIBILITY	0.7457	0.4273	0.9177	0.5561
		DG-Rho	0.7181			
	Norm1	SUBJECTIVENORMS	0.2949	-0.7636	0.9946	0.0870
	Norm2	SUBJECTIVENORMS	0.6171	-0.9718	0.9999	0.3808
		DG-Rho	0.9027			
	Att1	NEGATIVE ATTITUDES	0.7924	0.5836	0.9099	0.6279
	Att2	NEGATIVE ATTITUDES	0.8757	0.7617	0.9529	0.7668
	Att3	NEGATIVE ATTITUDES	0.7087	0.4462	0.9165	0.5023
	Att4	NEGATIVE ATTITUDES	0.8426	0.6857	0.9565	0.7099
	Att4	NEGATIVE ATTITUDES	0.8078	0.5971	0.9276	0.6525
	Att6	NEGATIVE ATTITUDES	0.6568	0.4330	0.8666	0.4313
		DG-Rho	0.9302			
	Int1	INTENTIONS	0.9511	0.8717	0.9946	0.9046
	Int2	INTENTIONS	0.9559	0.8974	0.9948	0.9138
Group 3		DG-Rho				
	PercSev1	PERCEIVED SEVERITY	0.8870	0.5192	0.9897	0.7868
	PercSev2	PERCEIVED SEVERITY	0.7033	-0.2781	0.9805	0.4946
		DG-Rho				
	PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.6837	0.4018	0.8849	0.4674
	PercSusc2	PERCEIVEDSUSCEPTIBILITY	0.7343	0.2403	0.9313	0.5392
	PercSusc3	PERCEIVEDSUSCEPTIBILITY	0.5072	0.0829	0.8403	0.2572
		DG-Rho				
	Norm1	SUBJECTIVENORMS	0.4462	-0.2655	0.8214	0.1991
	Norm2	SUBJECTIVENORMS	0.9462	0.8944	0.9997	0.8953
		DG-Rho				
	Att1	NEGATIVE ATTITUDES	0.8474	0.6060	0.9699	0.7182
	Att2	NEGATIVE ATTITUDES	0.8692	0.6892	0.9685	0.7555
	Att3	NEGATIVE ATTITUDES	0.7051	0.4208	0.9057	0.4972
	Att4	NEGATIVE ATTITUDES	0.8097	0.5607	0.9432	0.6556
	Att4	NEGATIVE ATTITUDES	0.8279	0.4162	0.9773	0.6854

Att6	NEGATIVE ATTITUDES	0.5891	0.2048	0.8654	0.3471
	DG-Rho				
Int1	INTENTIONS	0.9734	0.9392	0.9966	0.9475
Int2	INTENTIONS	0.9761	0.9480	0.9966	0.9528

6 Discussion and conclusions

Vaccines against HPV 16 and 18 are available since 2008. Most European countries offer the vaccine to girls aged between 9 to 17 years (ECDC, 2012; ECCa, 2009; Hofman et al., 2014). However, uptake rates vary considerably between countries (range 17%-84%) (ECDC, 2012; Hofman et al., 2014). Since vaccine coverage against HPV is so low in many countries including Italy, this thesis aims to assess which factors are at work to cause the basic rational model of immunization, i.e. costs vs benefits, to fail so markedly in explaining vaccination choices. A newly questionnaire structured according to the Theory of Planned Behaviour and the Health Belief Model was developed to determine which further factors, beyond the rational ones, might be important for decision makers. The instrument was administered among young adult women aged 19-22 attending their bachelor at the universities of Pisa and Florence with the purpose of measuring the determinants of their past HPV vaccination choices as well as their attitudes, beliefs and knowledge towards HPV, HPV vaccination and cervical cancer screening practices (Pap-test).

The exploratory data analysis results showed that non-vaccinated respondents seemed more autonomy in their immunization choice. However, among them, the delay strategy was one of the main reason of being non-vaccinated. Considering that HPV is one of the most common sexually transmitted infection, the delay strategy should be considered carefully. Exploratory analysis underlined as well as lack of HPV and Pap-test knowledge among respondents. Knowledge questions highlight a slightly better performance for unvaccinated girls while the information level considerably changes across faculties. Responses quality varies according to the question level, the more detailed is the information requested, the lower is the level of accuracy.

One of the main objective of this research was to investigate factors related with realized immunization choices. The PMLE model underlined the existence of several determinants associated with declining the HPV vaccination. These included citizenship, father's age, mother's education, HPV knowledge score (number of correct answers), sources of HPV information, duration of HPV vaccine protection and attitudes towards HPV vaccination (safety and side effects). The reference category for the dependent variable (vaccination status) were the vaccinated respondents. For the analysis, the subset (N=342) of vaccinated and unvaccinated respondents that knew HPV and its associated vaccine was considered.

Being an Italian citizen was positively related with the probability of vaccination. 16(5%) respondents were foreigners and 8 of them were not vaccinated. Among the latest, 5 resided abroad at age of 16, 2 in Prato and 1 in Rome. Similar results were found by Slåttelid Schreiber

et al. (2015) for the Danish population. The non-vaccination status among foreigners may be explained in part by the timing of HPV program introduction (Gilbert et al., 2016). In fact, those that declared to reside abroad or outside Tuscany at the age of 16, had less chance to be introduced in the vaccination regional programme. Héquet e Rouzier (2017) observed that vaccination rates were inversely correlated with the prevalence of immigrants and foreigners in 8 health department in France. However, the heterogeneity of immigrant and foreign populations limited their understanding on the determinants of vaccination. In addition, some studies have shown that race/ethnicity is associated with differential levels and types of immunization concerns (Gowa et al., 2014; Kennedy et al., 2009; Committee ISR, 2001; Prislin et al., 1998).

Father's age was another significant driver of HPV vaccination. Respondents with oldest fathers (age greater than 61 years old) were less likely to be vaccinated against HPV. However, little has been reported in literature on the father's role on HPV vaccination choices. Hansen et al. (2016) found that the likelihood of adolescent HPV vaccination was lower when mothers involved their husbands/partners in the decision. Perkins et al. 2010 and Hanley et al., 2014 identified that lack of fathers' knowledge about HPV and its related consequences, was the major determinant of vaccination refusal.

In addition, the extent literature about HPV vaccination and vaccinations in general, underlined two main reasons of possible fathers' opposition. First, fathers and parents in general, may not understand why vaccination at recommended ages is important. Many studies have suggested that parents may be less willing to vaccinate younger adolescents (Hansen et al., 2016; Holman et al., 2014; Kessel et al., 2012) or that parents may refuse or delay vaccination if the daughter is not perceived at risk of HPV related diseases (Darden et al., 2008). In addition, parents' anxiety concerning vaccine safety, beliefs that vaccination encourage sexual activity, moral issues about sexuality, lack of disease specific knowledge, risk of unknown harmful side effects, and low concern for child's HPV acquisition are the most important drivers of HPV vaccination refusal (Brabin et al., 2006; Brewer & Fazekas, 2007; Constantine & Jerman, 2007; Mays, Sturm, & Zimet, 2004; Slomovitz et al., 2006). Increases in age may emphasize these negative beliefs. Another possible explanation is represented by the late fertility. Increasing in parents' age is often associated with high risk perception of negative outcomes related to children health. Several studies attested that advancing paternal and maternal age have both been associated with high risk of children's diseases, such as autisms, down syndrome, neurocognition deficiencies, rare cancer and offspring longevity and fertility (Sandin et al., 2015; Bray e Gunnell, 2006; Zhu et al., 2008; Tearne, 2015; Heidinger et al., 2016). These aspects are especially relevant for infant vaccinations, for which side effects risks are more real and tangible, but appear unfounded for

a vaccine that, according to the existent studies, should be free of severe contraindications. However, Perkins et al. (2014) found that safety concerns and uncovered desire to avoid broaching the subject of their daughter sexuality are the most commonly causes of vaccination delay. These factors are intertwined and more difficult for provides to addressed. Similarly, the researches highlighted that despite the best intentions of parents and providers, delays often result in permanent failure to initiate vaccination, due to lack or returning for well-child visits, failure to revisit the issue, or continued reluctance to deal with teen sexuality Perkins et al., 2014).

The second significant parents' variable associated with vaccination refusal was the mothers' education level. Daughters with highest (university degree or more) and lowest (None-primary-secondary) educated mothers were less likely to be vaccinated against HPV than girls with high school graduated mothers. Slåttelid Schreiber et al. (2015) found out the girls of mothers with basic education had decreased initiation for HPV vaccination compare with girls of mother with higher education. However, the implication of parental education in vaccine hesitancy is well known in literature. With respect to the lowest education level, parents with less formal education have greater distrust in the medical community, express more concerns about vaccine safety and do not believe in the necessity and efficacy of vaccines (Gowda et al., 2013; Opel et al., 2011; Shui et al., 2006; Gust et al., 2005; Prislin et al., 1998). Parents that had less than 12 years of education were more likely to report not having enough vaccination information, compared to those graduated. In addition, common distrust in medical community may induce individuals to seek out different sources of information such as media, family member or other parents (Gowda et al., 2013; Bardenheier et al., 2004; Benin et al., 2005).

As regard the highest-level class of education, it was found significant in the interaction between the Disagree class of attitudes concerning vaccine safety. These results are in line with a Canadian and Norwegian studies reporting decreased likelihood of HPV vaccination with increasing in parental education (Gilbert et al., 2016; Feiring et al., 2015; Ogilvie et al., 2010). The opposite was, however, found in a Swedish study of on-demand HPV vaccination (Leval et al., 2013). However, like income, there is a conflicting influence of education on vaccination attitudes, especially when the safety of vaccine has a determinant role in the final choice. Opel et al. (1998) found that parents with higher levels of education were nearly four times as likely to be concerned about the safety of vaccine than those from lower education levels. Our findings highlight a similar result. In fact, respondents with high educated parents that do not trust on vaccination safety are less likely to be vaccinated. In general, more educated mothers are more informed and exposed to risks-benefits information (Grignolo, 2016; Ogilvie et al., 2010; Brown et al., 2012). However, recent studies have pointed out that in

contexts with different risk information individual decision-making processes are characterized by irrational asymmetry, i.e. subjects tend to overweight the value of the high-risk judgment (Viscousi, 1997). In this perspective, an overestimation of vaccines risks may justify the low levels of vaccination among girls of more educated mothers. In addition, highly educated parents may feel more confident in their ability to interpret complex scientific and clinical health information, allowing them to ignore the advice of practitioners if contradictions exist (Patel & Berenson, 2015).

The level of HPV knowledge is negatively related with the probability of being vaccinated. i.e. the lower the level of a respondent HPV knowledge the higher is the probability of vaccination. On the other hand, the HPV source information class positively related with the probability of being vaccinated was the "Parents" one. When parents are the main source of information about HPV the daughter have high probability to receive the HPV vaccination. According to the explorative results, their influence on the immunization decision seems to be interrelated. In fact, among vaccinated, 48% got the vaccine because their parents decided it and the 49% took the decision in accordance with them. A lack on autonomy in vaccination decisions, justifiable for 11-14 girls, may translate in future absence of proper knowledge about the specific health problem. Although after the age of 18 girls should become more responsible in managing their health status, the rate at which daughter's capacity for medical decisions evolves depends on a variety of factors, including the complexity of medical conditions, the family and the beliefs system (Weithorn and Campbell,1982). Although, non-vaccinated seemed more independently in their vaccination decision, in fact 30% (24) has taken the decision about the vaccination autonomously and 42%(33) with their parents, the 54%(44) of them disagreed with the sentence "I think I can deal alone with (take a decision without the help of my parents) my health-related choices". Among vaccinated, the 57% (150) declared to disagree with it. These results seem to underline a group of students unable to regulate their own health behaviours even if most of them declared to undertake the gynaecologist visits on their own. However, little has been reported in literature on girls' role on HPV vaccination choices and its correlation with future knowledge about the virus. A Spanish study on adolescent underlined that the way adolescents look at HPV is unusual and, although they might have heard about the virus, there is little concern about it (Navarro-Illana et al., 2013).

A similar trend was observed for Pap-test knowledge. This result is confirmed by several studies which indicates that youngest women have insufficient understanding of Pap testing (Head et al., 2009) and a lack of awareness regarding cervical cancer prevention in general (Lee et al., 2001; Kim & GY, 2013; Kim, 2009).

The latest significant determinants of HPV vaccination are related to the duration of HPV vaccine protection and the attitudes items towards HPV vaccination (safety, efficacy and side effects). The attitudes items are expression of the TPB dimension Attitudes. Respondents that disagrees with the side-effects worries item (TPB attitude: I am very worried about the long-term side effects of the HPV vaccination) and thought that the vaccine protects for the whole life were in favour of vaccination. On the other hands, vaccine safety (TPB attitude: I believe that the vaccine is safe) and efficacy (TPB attitude: I believe that the vaccine is effective in preventing HPV infection) concerns were associated with refusal of the HPV vaccination. These findings are confirmed by other studies on HPV vaccination (Gefenite et al., 2012; van Keulen et al., 2013; Lenselink et al., 2007). A recent study of Larson and colleagues (2016) on general vaccinations distrust showed that European and Wester Pacific regions are characterized by more negative sentiment towards vaccinations and these are often driven by fear of vaccines side-effects and doubts about their safety.

To evaluate the significance of interrelationships between psychological latent constructs (i.e. Theory of Planned Behaviour, Health Beliefs Model dimensions) partial least square structural equation models were employed. The structural equation results showed that only *Negative Attitudes* (towards Pap-test) has significative total effect (β = -0.44; $p < 0.05$) on *Pap-test Intentions*. The negatives attitudes towards pap-test screening that mainly contribute to the psychological dimension are those related with fear of negative results, embarrassing and young-age. Other latent dimensions seem not influencing the respondents' future intentions towards pap-test screening. A similar study, employing the TPB model to study 18 to 26-year Singaporean women intentions towards HPV vaccination and pap-test, showed different results (Chirayil et al., 2014). Chirayil and colleagues found out that subjective norms were a strongest predictor of respondents' intentions toward pap-test and HPV vaccination, while the attitudes dimension was not significantly related with the endogenous variable. Their results are consistent with past research showing that subjective norms are the strongest predictor of prevention intentions for both, pap-test screening (Duffett-Leger et al., 2008) and HPV vaccination (Teitelman et al., 2011). However, Jennings-Dozier (1999) and Linton (2009) found that the strongest predictor of pap-test uptake were attitudes towards the behaviour, which accounted respectively for the 58% and 83% of the total intentions variance. A second study, on Korean unmarried university students who had never undergone a Pap test, underlined the subjective norms as the most important predictors of woman's intention to undergo a pap-test (Kim, 2014).

HBM constructs instead, were considered by Ma et al. (2013) that found perceived susceptibility, but not perceived severity, as significant predictor of cervical cancer screening.

The Tanner-Smith and Brown review, (2010) of studies predicting Pap-test behaviours using HBM, underlined that among the six studies that included perceived severity of cervical cancer only two suggested high levels of perceived severity of cervical cancer. Results inconsistency may be due to different reasons. In respect to HBM dimensions, university women may have different perception of the risk of contract cervical cancer in future and of the severity of it. For example, the 19% (65) of the respondents were not sexually active and among those who had sexual intercourses, the 61% (215) had 1 or 2 fixed partners in the whole life. Having a fixed relation or a few numbers of partners may reduce the perception of the risk of infection. For not sexually active girls, instead, the risk of infection approximates zero. In addition, young girls may feel themselves not a risk to develop cervical cancer in the future. Besides, subjective influence and HBM/TPB dimensions are often strictly connected with the culture in which they originate and operate (Cialdini & Trost, 1998) and this can explain studies differences. Insignificant interrelationship between latent variables may depend as well as on the sample heterogeneity. There are not existing studies that explored the presence of unobserved heterogeneity in Pap-test intentions among university women. The application of REBUS-PLS for unobserved heterogeneity detection, has led to identification of two latent classes. For group 1 Negative attitudes ($\beta = -0.44$; $p < 0.05$) and Perceived Susceptibility ($\beta = -0.21$; $p < 0.05$) has been identified to have negative influence on members' intentions. They accounted for the 37% of intentions variability. The 73%(72) of respondents belonging to the first group are sexually active and 68% (42) of them had one or two fixed partners in the whole life. The 21% (15) of sexually active girls had undertook a pap-test in the past. The main reasons of examination were prevention and health issues. 50% of the respondents had and high level of knowledge about pap-test, while 50% ignore completely the basics. On the other hand, in group 2 none of the exogenous variables are significantly related with intentions.
Girls younger than 25 years are not included in the national cervical cancer screening and with high probability did not undergo a pap-test. Their inadequate knowledge and lack of awareness about consequences of HPV infection may lead to fear and embarrassment feelings towards cervical screening. In addition, poor knowledge about HPV risks may increase their ignorance about risky behaviours and decrease their levels of susceptibility. These reasons might explain the negative influences of perceived susceptibility and attitudes on future behaviour intentions.

PLS-SEM for vaccination intentions showed that, among the four latent constructs, Negative Attitudes (towards Pap-test) showed the strongest (negative) significative total effect ($\beta=-0.5904$; $p<0.05$) on Intentions towards HPV vaccination, followed Perceived disease severity

(β=0.2097; p<0.05) and Subjective Norms (β=0.1884; p<0.05). Perceived susceptibility seemed not influencing the respondents' future intentions towards HPV vaccination. Our TPB results are consistent with previous studies (Keulen et al., 2013; Askelson et al., 2010; Fahy et al., 2009; Gerend et al., 2009). These studies found the attitudes and subjective norms appeared strongly related to HPV vaccination intention. The most-dominant attitudes/beliefs were those related with safety and effectiveness of HPV vaccine (Keulen et al., 2013). The thesis results are consistent with Keulen and colleagues. In fact, Negative Attitudes towards HPV vaccination are mostly driven by items concerning safety and efficacy. Perkins and colleagues (2012) showed that safety concerns and low perceived severity of HPV consequences played an important role among factors that discourage vaccination. There are not existing studies that explored the presence of unobserved heterogeneity in vaccination intentions among university women. The application of REBUS-PLS for unobserved heterogeneity detection, has led to identification of thee latent classes. Given the small group size the results have only explorative purposes. For group 1 three exogenous variables have a significant impact on future vaccination Intentions, Perceived Susceptibility towards the HPV infection (0.4343:), Subjective Norms (0.2553) and Negative Attitudes (-0.3147). However, compared to the path coefficients of the other classes, respondents in class 1 perceive more the risk of infection and its consequences and the influences of others in taking future vaccination decisions. Intentions of respondents belonging to the second group are significantly and negatively influenced by two latent constructs, Perceived Susceptibility (-0.2963) and Negative Attitudes (-0.6545) toward vaccination. This implies that girls in class 2 do not perceive the risk of cervical cancer and are mainly driven by negative attitudes. On the other hand, the Intentions of class 3 students are driven by Negative Attitudes (-0.6125), Perceived Severity of cervical cancer (0.3393) and Subjective Norms (0.1618). Perceived Severity has significant effect only on third group's intentions.

This thesis represents one of the first Italian studies to examine determinants of past HPV vaccination choices and possible barriers to future prevention intentions. The study revealed that different factors had influenced the past respondents' HPV vaccination choices. These include socio-demographic variables, such as fathers' age and mothers' education, as well as psychological dimensions mainly connected with fear of HPV vaccine side-effects and safety. In addition, the research underlines that different psychological constructs may influence respondents' future intentions towards HPV vaccination and cervical cancer screening. The study was one of the first to examine unobserved heterogeneity in the sample showing that different groups of young women are influenced by distinct psychological dimensions. These results need to be confirmed by future studies that will involve at least a probabilistic

sampling design. However, it is important that future communication will address social psychological variables that appear to be related to HPV vaccination (i.e. attitude, beliefs, subjective norms, habit and perceived susceptibility). For example, future communication could address important beliefs about the safety of the vaccine and could emphasize the importance of receiving the HPV vaccination before daughters become sexually active. Furthermore, future communication strategies about the HPV vaccination should be targeted at different groups of recipients. Individuals heterogeneity should be taken into account by policy makers in the definition of public information campaigns.

Study limitation

The present study had some limitations that are worth mentioning. First, it was based on a non-probabilistic sample, which makes it impossible to draw definite conclusions about causality of the associations found. Second, even if some baseline characteristics were comparable to the general Italian population the sample seemed to mirror the source population of the Tuscany. Third, in order to fully test the importance of intentions, data on future screening behaviours should be collected. Another limitation includes the lack of standardization for the psychological dimension involved. These measures are not standardized since the scale items depend on the behaviour and the population of interest. In addition, study items were newly developed and for some of them validity and reliability conditions are not met.

.

Appendix 1: Statistical models

The theory in this chapter can be found in Agresti (2002), Hosmer and Lemeshow (2004), Hastie et al. (2008), Moons et al. (2004), unless otherwise indicated.

A1.1. Logistic regression

Logistic regression is widely used to model the outcomes of a categorical dependent variable. Let X be an $n\,x\,(p\,+\,1)$ matrix of covariates, with rows $x_i = (x_{i0};\, x_{i1};\, \ldots\ldots\,;\, x_{ip})$, where x_{i0}=1 is the intercept term, and let $Y = (Y_1;;\ldots\ldots;\, Y_n)$ be a random vector of binary outcomes, i.e. $Y \in (0,1)$. In our case, X is the covariate matrix with all the independent variables and Y is the response vector where 1 indicates a vaccinated respondent and 0 an unvaccinated.

Given that Yi is a Bernoulli distributed random variable with success probability $\pi(x_i)$, the logistic regression model is given as,

$$\pi(x_i) = P(Y_i = 1|x_i) = \frac{exp(x_i\beta)}{1 + exp(x_i\beta)} \tag{1}$$

where $\beta = (0; 1; \ldots\ldots; p)$ is the vectors of coefficients. Equivalently, the log odds, called the logit, has the linear relationship

$$logit[\pi(x_i)] = log\frac{\pi(x_i)}{1 - \pi(x_i)} = x_i\beta = \sum_{j=0}^{p}\beta_j x_{ij} \tag{2}$$

The parameter β_j in equation (2) represents the multiplicative effect on the odds for the vaccinated obtained from a 1-unit increase in the jth covariate when keeping the levels of all other covariates fixed. When $\beta_j > 0$, $\pi(x_i)$ increases as x_i increases. When $\beta_j < 0$, $\pi(x_i)$ decreases as x_i increases.

A1.2. Maximum likelihood estimation of the logistic regression model

The method of maximum likelihood (ML) yields value for the unknown parameters which maximize the probability of obtaining the observed set of data. To apply the methods, the likelihood function expressing the probability of the observed data as a function of the unknown parameters is needed. The ML estimators of these parameters are the value that maximize the specified likelihood function. Thus, the resulting estimators are those which agree most closely with the observed data.

Since the Y_i's are conditionally independent and $Y_i \sim Bernoulli\left(\pi(x_i)\right)$, $i = 1,..n$ the likelihood function of the above logistic model is obtained as follow,

$$l(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i}\left(1 - \pi(x_i)\right)^{1-y_i} \tag{3}$$

The loglikelihood is thus,

$$L(\beta) = ln[L(\beta)] = \sum_{i=1}^{n} \{y_i ln[\pi(x_i)] + (1 - y_i) ln[1 - \pi(x_i)]\} \tag{4}$$

ML are use to obtain β_j values, i.e. those that maximize the $L(\beta)$. Differentiating $L(\beta)$ with respect to β_j and setting the resulting expressions to zero, yields the likelihood equation

$$\sum_{i=1}^{n} x_i' [y_i + \hat{\pi}_i] = \mathbf{X}'[\mathbf{y} - \hat{\boldsymbol{\pi}}] = 0 \tag{5}$$

Since these $p + 1$ equations are non-linear in β they require an iterative solution, i.e. the Fisher scoring or the Newton-Raphson algorithm, which are asymptotically equivalent for the logistic regression model. The algorithm starts at an initial guess for the parameter values that maximize the likelihood function. Successive approximations produced by the algorithms tend to fall closer to the ML estimates. The Fisher scoring algorithm for doing this was first proposed by R. A. Fisher for ML fitting of probit models. For binomial logistic regression and Poisson loglinear models, Fisher scoring simplifies to a general-purpose method called the Newton–Raphson algorithm.

A1.3. Sequential variable selection strategy: backward selection

Many model selection procedures exist, no one of which is always best. In this thesis, backward selection procedure was considered. Backward elimination begins with a complex model and sequentially remove terms. At each stage, it selects the term for which its removal has the least damaging effect on the model (e.g. the largest P-value) The process stops when any further deletion leads to significantly poorer fit (Agresti, 2002).
The final model was selected according to the Akaike Information Criteria (AIC), which balances the goodness of fit of the model, as measured by the loglikelihood, with its complexity. The model with the lower criterion value is preferable. Thus, the optimal model is the one that tends to have fit closet to reality. Give a sample, AIC selects the model that minimized,

$$AIC = -2(\text{maximized log likelihood- number of parameters in the model}) \quad (6)$$

A1.4. Penalized methods

The sensitivity of the results was explored evaluating different methodological approaches. Traditional stepwise selection methods, such as forward and backward, suffer from high variability and low prediction accuracy, especially when the number of possible covariates is large and/or the level of multicollinearity is notable (Hastie, Tibshirani, and Friedman 2001). In such situation, penalized regression methods, such as LASSO, Elastic Net and penalized maximum likelihood estimation (PMLE), combine the advantages of selection procedure preserving the model prediction accuracy. In this paragraph, a brief description of the penalized methods considered is presented.

A1.4.1. Penalized maximum likelihood

Penalized MLE (PMLE)51~53 is a general technique for shrinking (stabilizing) regression fits. Instead of maximizing the log-likelihood, PMLE maximizes a penalized log-likelihood which is the sum of the ordinal model log-likelihood and a penalty, resulting in

$$logL - \frac{1}{2}\lambda \sum_{i=1}^{n} (s_i \beta_i)^2 \quad (7)$$

where L is the maximum likelihood of the fitted model, λ a so-called penalty factor, β_i the estimated regression coefficient for each predictor i in the model, and $s_1, s_2, \ldots\ldots\ldots, s_p$ are scale factors chosen to make $s_i\beta_i$ unitless (Harrell, 2001; Houwelingen,2001). Accordingly, the estimated regression coefficients and predictive accuracy measures are directly (i.e., during the model fit) adjusted for overoptimism. This direct adjustment is the major advantage of PMLE. PMLE is developed for logistic regression models and is a generalization of the ridge regression method used to obtain more stable parameters for linear regression models (Draper & Smith, 1998). The usual methods (for example, Newton-Raphson) are used to maximise equation (7).

The main problem in using PMLE is how to choose the optimal penalty factor λ. Many investigators use cross-validation methods to estimate the penalty factor. However, maximizing the modified Akaike's information criterion (AIC) method is a more efficient method (Harrell, 2001; Houwelingen,2001). Finding the optimum penalty factor for penalized maximum likelihood estimation requires use of a modified AIC, defined as

$$\chi_{LR}{}^2 - 2p \times df_{eff} \tag{8}$$

where $\chi_{LR}{}^2$ is the likelihood ratio χ^2 (i.e., the difference in log likelihood between the model with and without predictors or covariates) of penalized model and df_{eff} is the degrees of freedom after penalizing the fitted predictors. In ordinary logistic regression, the degree of freedom is equal to the number of predictors in the model; The higher the number of predictors, the higher the degree of freedom and more likely the model is overfitted. Due to the penalization, the deegree of freedom effectively used in PMLE is lower than the actual number of predictors, descreasing the potential for overfitting. More technically, df_{eff} is deriveed from an equation that describes the reduction in variance od the penalized parameter estimates in comparison to the variance of ordinary paramenter estimates (Harrell, 2001; Houwelingen,2001). To find the optimal λ, a trial and error process os assplied in qhich the corrisponded modified AIC is estimated for a variety of $\lambda - values$. The λ that maximizes the modified AIC is the optimal penalty factor.

A1.4.2. Predictive accuracy indices

Internal validation was considered to assess what Harrell and colleagues call overfitting. This occurs when a model has too many parameters and smaller sample sizes, which can cause it to

overestimate the data (Harrell, 2001). Model discrimination was assessed considering four indices, that are described below.

ROC Curves and Harrell's C index. A classification table cross-classified the binary response with a prediction whether $y = 0$ or 1. The prediction is $\hat{y} = 1$ when $\hat{\pi}_i > \pi_0$ and $\hat{y} = 0$ $\hat{\pi}_i \leq \pi_0$, for some cut-off π_0. Most classification tables use $\pi_0 = 0.5$ and summarize predictive power by

$$Sensitivity = P(\hat{y} = 1|y = 1) \; and \; specificity = P((\hat{y} = 0|y = 0)$$

In statistics, a receiver operating characteristic curve, i.e. ROC curve, is a graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied. ROC is a plot of sensitivity as a function of sensitivity of $(1 - specificity)$for the possible cutoff π_0. This curve usually has a concave shape connecting the points $(0,0)$ and $(1,1)$. The higher the area under the curve, the better the prediction.
The area under the ROC curve is identical to the value of another measure of predictive power, the concordance index. Consider all pairs of observation (i, j) such that $y_i = 1$ and $y_j = 0$. The concordance index estimates the probability that the predictions and the outcomes are concordant, the observation with the larger y also having the larger $\hat{\pi}$ (Harrell et al., 1984). ROC and C values ranges between 0.5 and 1: a value of 0.5 indicates that the model has no ability to discriminate between low and high-risk subjects, whereas a value of 1 indicates that the model can perfectly discriminate between these two groups.

Somer's Delta. Harrell's *rms* package reports these using Somers' D statistic, which is defined as the difference between the two conditional probabilities of concordance and discordance. Harrell defines Somers' D = 2C −1, where C is the above-mentioned Harrell's C (Harrell, 2001).

For model calibration, Brier's score and calibration plots were considered.

Brier's score. The Brier score was proposed by Glenn W. Brier in 1950 and is a proper score function that measures the accuracy of probabilistic predictions (Brier, 1950). Also known as Quadratic Loss (Hvattum & Arntzen, 2010), the Brier Score measures the average squared deviation between predicted probabilities for a set of their events and their outcomes. For an individual match i the score is (Brier, 1950)

$$BS = \sum_{i=1}^{N} \left(p_{ij} - e_{ij}\right)^2 \tag{9}$$

and the cumulative score,

$$BS = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{e} \left(p_{ij} - e_{ij}\right)^2 \tag{10}$$

where p_{ij} are the predicted probabilities and e_{ij} the actual outcome (0 if happen and 1 if not happen). The Brier score ranges between 0 and 1 and the lower the score, i.e. the difference between predicted probabilities and actual outcome, the better the predictions are calibrated.

Calibration plot. A calibration plot displays the relationship between the predicted and the true probabilities, and a departure from the diagonal line indicates poor calibration.

A1.4.3. Penalized regression

Traditional stepwise selection methods, such as forward and backward, suffer from high variability and low prediction accuracy, especially when the number of possible covariates is large and/or the level of multicollinearity is notable (Hastie, Tibshirani, and Friedman 2001). In such situation, penalized regression methods, such as LASSO, Elastic Net combine the advantages of selection procedure preserving the model prediction accuracy. The lasso (Tibshirani, 1996) is a popular method for regression that uses an L1 penalty to achieve a sparse solution. The LASSO approach used a shrinkage method based on L1 penalty that produce sparse solutions (Hastie et al. 2003). Another penalization term L2 is used in ridge regression and it provides coefficients for every variable (Hastie et al. 2003). $\lambda > 0$ is used as a tuning parameter,

$$L1: \lambda \sum_{j=1}^{p} |\omega[j]| = \lambda \|\omega\|_1 \qquad L2: \lambda \sum_{j=1}^{p} (\omega[j])^2 = \lambda \|\omega\|_2^2 \tag{11}$$

the λ parameter control the amount of shrinkage and higher values for λ sets some coefficients exactly to zero in case of L1. Elastic Net were introduced by Zou and Hastie (2005) as a linear combination of L1 and L2,

$$\lambda\|\omega\|_1 + \lambda\|\omega\|_2^2 \tag{12}$$

LASSO and Elastic Net perform variable selection and coefficient estimation simultaneously, given the research purposes, they were preferred over ridge. Penalized regression methods introduce bias in coefficient estimation by continuously shrinking the regression coefficients. However, this shrinkage provides a decrease in variance. This is called the *bias-variance trade-off.*

In Equation 4 the log-likelihood is maximized to distinguish optimal solution for logistic regression. After introducing two regularizations terms we have the following maximization,

$$F_1(\lambda) = -min_{(w\in R^p)} \sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1-y_i) \ln[1-\pi(x_i)]\} + \lambda \sum_{j=1}^{p} |\omega[j]| \tag{12}$$

$$F_2(\lambda_1\lambda_2) = -\min_{(w\in R^p)} \sum_{i=1}^{n} \{y_i \ln[\pi(x_i)] + (1-y_i) \ln[1-\pi(x_i)]\} + \lambda\|\omega\|_1 + \lambda\|\omega\|_2^2 \tag{13}$$

For each value of λ can outer loop is created which computes the quadratic approximation about the current parameters. Then the coordinate descendent are used to solve the penalized weighted least-square problem (Friedman et al.,2008),

Selection of the tuning parameter is very important as it can have a big influence on the performance of the estimator. Cross-validation is considered the simplest and most widely used method for minimisation the prediction error (Hastie et al., 2001; Chand, 2012). K-fold cross validation was considered for the estimation.

Appendix 2: Determinants of vaccination supplementary results

Table 30 Odds ratio CI complete and reduced dataset

Variable	Odds Ratio Complete dataset		Odds Ratio Reduced dataset [-2;2]	
	2.5	97.5	2.5	97.5
Intercetta	0.346	1.4432	0.0080	0.9383
Father's age (Ref:"51-60") **				
Father's age: 41-50	0.3757	1.6620	0.2203	1.5497
Father's age: 61+	0.0625	0.6657	0.0074	0.1919
Father's age Don't live with father	0.3397	2.5390	0.0980	1.2343
Citizenship (Ref: Abroad) *				
Citizenship: Italy	1.6192	27.3377	3.6544	178.6058
Mother's education (Ref:" High school") .				
Mother's education: None, primary, secondary	0.1913	1.1483	0.0266	0.3404
Mother's education: Graduate or professional degree	0.2615	2.4865	0.0611	1.3444
Attitude 1 (Ref: "Agree") *** [1]				
Att1: Disagree	0.1054	0.9122	0.0272	0.4413
TPB Attitude 2 (Ref: "Agree") **[2]				
Att2: Disagree	0.1960	1.3949	0.0426	0.6680
TPB Attitude 3 (Ref: "Agree") **[3]				
Att3: Disagree	1.0447	5.7075	1.7023	14.7339
Score HPV knowledge	0.6234	0.8218	0.4669	0.7047
Sources of HPV information (Ref:" School/University")**				
Source: Others (Health workers, internet, radio and tv, books, etc.)	0.2486	3.5290	0.3916	17.6084
Source: Parents	1.4657	8.8530	3.5392	39.5208
Source: Medical doctor/Gynaecologist	0.6543	5.1136	0.7320	8.9490
Vaccine duration of protection (Ref:" I don't know") *				
Protection: 10 years	0.7183	3.0677	0.7820	4.8153
Protection: Lifetime	1.4559	8.3724	3.2138	41.1965
Interactions: Mother's education: Attitude 2 (Ref =" High school:Agree")*				
MotherEd*Att2: None_primary_secondary: Disagree	0.2486	4.8111	0.5330	23.0292
MotherED*Att2: Graduate or professional degree: Disagree	0.0185	0.7013	0.0060	0.6420

Table 31 Complete dataset alpha values, cross-validation error, Brier's score and ROC Area

Alpha(Lambda.1se)	CVE	Brier'score	Area ROC
0.1	0.7504	0.2103	**0.7339**
0.2	0.7408	0.2108	0.7221
0.3	0.7279	0.2103	0.7196
0.4	0.7345	**0.2099**	0.7116
0.5	**0.7233**	0.2100	0.7124
0.6	0.7262	0.2105	0.7066
0.7	0.7309	0.2100	0.7067
0.8	0.7245	0.2102	0.7105
0.9	0.7326	0.2101	0.6998
1.0	0.7283	0.2100	0.70493

Table 32 Reduced dataset alpha values, cross-validation error, Brier's score and ROC Area

Alpha(Lambda.1se)	CVE	Brier'score	Area ROC
0.1	0.7304	0.1803	**0.7339**
0.2	0.7208	0.1808	0.7221
0.3	0.7279	0.1803	0.7196
0.4	0.7345	**0.1750**	0.7116
0.5	**0.7033**	0.1800	0.7124
0.6	0.7050	0.1805	0.7066
0.7	0.7107	0.1800	0.7067
0.8	0.7200	0.1802	0.7105
0.9	0.7230	0.1801	0.6998
1.0	0.7250	0.1800	0.70493

Table 33 Complete dataset: variables selected and estimates by different statistical models

	Elastic Net	Penalized ML	Logistica
(Intercept)	1.5621	0.5747	0.2308
Father's age (Ref:"51-60") **			
Father's age: 41-50	1.0000	0.7974	0.7867
Father's age: 61+	0.6840	0.3901	0.2067
Father's age Don't live with father			
Citizenship (Ref: Abroad) *	1.3794	0.8676	0.8908
Citizenship: Italy			
Mother's education (Ref:" High school") .	1	1.1417	0.8822
Mother's education: None, primary, secondary	1.5222	2.9858	0.92741
Mother's education: Graduate or professional degree	1	1.6270	3.5934

Attitude 1 (Ref: “Agree”) *** [1]			
Att1: Disagree	1	1.3271	1.4713
TPB Attitude 2 (Ref: “Agree”) **[2]	1.1317	2.4210	3.3830
Att2: Disagree			
TPB Attitude 3 (Ref: “Agree”) **[3]	0.6044	0.4198	0.3124
Att3: Disagree			
Score HPV knowledge	0.8753	0.6028	0.5208
Sources of HPV information (Ref:” School/University”)**			
Source: Others (Health workers, internet, radio and tv, books, etc.)	1.4868	2.0324	2.4484
Source: Parents	0.9088	0.7593	0.7200
Source: Medical doctor/Gynaecologist			
Vaccine duration of protection (Ref:” I don’t know”) *	1.0000	0.8267	1.0883
Protection: 10 years	0.4097	0.1508	0.12043

Table 34 Reduced dataset: variables selected and estimates by different statistical models

	Elastic Net	Penalized ML	Logistica
(Intercept)	1.0547	0.9579	0.0939
Father’s age (Ref:”51-60”) **			
Father’s age: 41-50	1.0000	0.7788	0.5863
Father’s age: 61+	0.4581	0.3902	0.0404
Father’s age Don’t live with father	1.0000	0.6930	0.3415
Citizenship (Ref: Abroad) *	2.082	2.4274	23.1767
Citizenship: Italy	0.9407	0.5059	0.1006
Mother’s education (Ref:” High school”) .	1.0000	0.6479	0.2815
Mother’s education: None, primary, secondary			
Mother’s education: Graduate or professional degree	1.0000	1.5165	2.4702
Attitude 1 (Ref: “Agree”) *** [1]	1.8107	3.3390	11.3537
Att1: Disagree	1.0000	1.5171	2.5183
TPB Attitude 2 (Ref: “Agree”) **[2]			
Att2: Disagree	1.0000	1.2472	1.9059
TPB Attitude 3 (Ref: “Agree”) **[3]	1.5113	2.5574	10.5330
Att3: Disagree			
Score HPV knowledge	0.4755	0.4088	0.1137
Sources of HPV information (Ref:” School/University”)**			
Source: Others (Health workers, internet, radio and tv, books, etc.)	0.7526	0.5212	0.1760
Source: Parents			

Source: Medical doctor/Gynaecologist	1.8722	2.2359	4.8910
Vaccine duration of protection (Ref:" I don't know") *	0.8754	0.7456	0.5823
Protection: 10 years			
Protection: Lifetime	1.0000	0.8331	3.3688
Interactions: Mother's education: Attitude 2 (Ref =" High school:Agree")*	0.3052	0.1253	0.0664

Appendix 3: PLS-SEM theory and tables

The theory in this chapter can be found in Vinzi et al. (2010), Hair et al. (2014), and Sanchez (2013) unless otherwise indicated.

A3.1. Partial least square path model

Structural Equation Models (SEM) represent a statistical methodology aimed to estimate a network of casual relationships, based on a theoretical framework in which two or more latent constructs are liked to each other and each is measured through a number of observable indicators, i.e. items questionnaire (Bollen 1989; Kaplan 2000). The basic idea is that a complex system of relationship can be analysed taking into account the casual network among latent components, called Latent Variables (LV), each measured by several observed items often indicated as Measurement Variables (MV). In this sense SEM can be viewed as a joint-point between Path Analysis and (Tukey 1964; Alwin and Hauser 1975) and Confirmatory Factor Analysis (CFA) (Thurstone 1931).

PLS path models consists of three main components, the structural model, the measurement models and the weighting scheme. Whereas structural and measurement parts are presented in any kind of SEMs with latent constructs, the weighting scheme is a typical PLS component.

The structural model, also called inner model, relates the LVs with each other according to underlying theory. LVs can be exogenous or endogenous. Exogenous (independent latent variables) LVs do not have any predecessors, i.e. their values are wholly casually independent from other variables in the system. All others are endogenous (dependent latent variables). The equation describing the relation among LVs in the in the structural model is the following,

$$\xi_j = \beta_{0j} + \sum_{q=1}^{Q} \beta_{qj}\, \xi_q + \zeta_j \qquad (1p)$$

In the specific case, ξ_j represents intentions towards HPV vaccination and pap-test (endogenous variable) whereas ξ_r (exogenous variables) are respectively, Subjective Norms, Negative Attitudes, Perceived Susceptibility and Perceived Disease Severity.

The measurement model or outer model instead, relates observed variables (MVs) to their latent variables (LVs). Often observed variables are indicated as manifest variables, indicators or items and latent variables as factors. In the PLS framework one manifest variable can only

be related to one LV. There are two ways in which a block of MVs can be related to a LV, reflective and formative. In the reflective way (mode A in the plspm package) each bloc of MVs reflects its LV and can be written as the multivariate regression:

$$x_{pq} = \lambda_{p0} + \lambda_{pq}\xi_q + \varepsilon_{pq} \qquad (2p)$$

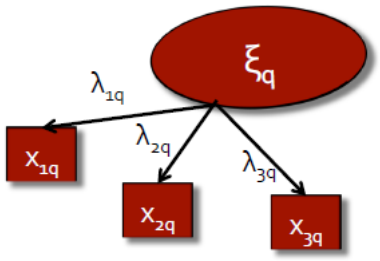


Figure 33 Reflective measurement model

where the weights λ_{pq} represent the OLS simple regression coefficients of each block MV with the corresponding LV ξ_q.

For the formative way (Mode B in the plspm package) the LV is considered to be formed by it MVs following a multiple regression model

$$\xi_q = \sum_{p=1}^{p_q} w_{pq}x_{pq} + \delta_q \qquad (3p)$$

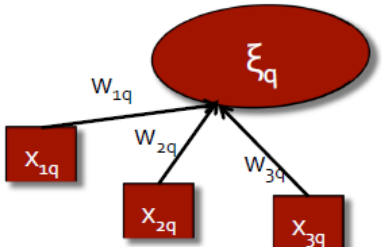


Figure 34 Formative measurement model

where the weights w_{pq} are the OLS regression coefficients of each LV on the corresponding block of MVs.

For this study, all the latent dimensions were considered as reflective indicators. In the specific case, the measurement model will be characterized by a different number of equations according to the number of MVs that characterized the specific LV. Therefore, in the

vaccination PLS model we will have respectively, two OLS regression equations for the Perceived Severity latent dimension, three for the Perceived Susceptibility, two for the Subjective Norms, six for Negative Attitudes and two for Intentions toward HPV vaccination.
In the pap-test PLS model we will have respectively, two OLS regression equations for the Perceived Severity latent dimension, two for the Perceived Susceptibility, two for the Subjective Norms, four for Negative Attitudes and two for Intentions toward HPV vaccination.
All relations between latent variables (LVs) can be drawn as a *Path diagram* in which ellipses or circles represent the latent variables and rectangles or squares refer to the manifest variables. Arrows show causations among the variables (either latent or manifest), and the direction of the array defines the direction of the relation, i.e. variables receiving the array are to be considered as endogenous variables in the specific relationship (Trinchera & Russolillo, 2010).

The weight scheme instead, relates the virtual LVs and the material MVs. In PLS-PM, latent variable are estimated as linear combination of their manifest variables. Moreover, an estimated LV_i is called a score and is denoted as

$$LV_j = Y_j = \sum_{q=1}^{Q} \omega_{jq}\, x_{jq} \qquad (4p)$$

In fact, this the very reason why PLS-PM is called component-based approach because latent variables are calculated as a weighted sum of their indicators.
The Partial Least Square (PLS) approach to SEM offers an alternative to covariance-based estimation procedures which is especially suited for situations when data is not normally distributed. PLS path modelling is referred to a *soft-modelling-technique* with minimum demand regarding measurement scales, sample sized and residual distributions (Monecke & Leish,2012). This interesting characteristic is optimal in those application fields where such assumptions are not tenable, at least in full. On the other hand, confidence intervals and hypothesis testing procedures based on resampling methods, such as jackknife and bootstrap, are needed to obviate the lack of a classic parametric inferential framework (Chin 1998; Tenenhaus et al. 2005). It also leads to less ambitious statistical properties for the estimates,
e.g. coefficients are known to be biased but consistent at large (Cassel et al. 1999, 2000). Finally, PLS-PM is more oriented to optimizing predictions (explained variances) than statistical accuracy of the estimates. PLS is an iterative algorithm that estimates partial model relationships separately solving out the blocks of measurement model and then, in a second

step, estimates the path coefficients of the structural model. Therefore, PLS-PM is claimed to explain at best the residual variance of the latent variables and, potentially, also of the manifest variables in any regression run in the model (Fornell and Bookstein 1982). For this reason, PLS is considered more as an exploratory approach than a confirmatory one.

A3.1.1. The PLS algorithm

The PLS algorithm aims at estimating the values for LVs by an iterative procedure. The basic idea is to construct a specific LV by the sum of its MVs. PLS-PM is made of a system of interdependent equations based on simple and multiple regressions. Such system estimates the network of relations among the latent variables as well as the links between the manifest variables and their own latent variables. In PLS-PM an iterative procedure permits to estimate the outer model's weights and the latent variables score. The first stage of the algorithm involves the computation of the weights to get the score of the latent variables. In fact, the iterative procedure of the PLS algorithm proceeds as follows:

- Start: the iterative process starts by assigning "arbitrary" values to outer weights.

- Step 1: Once the arbitrary weights are computed latent variables are computed by external approximation, i.e. as a weighted sum of their indicators or MVs.

- Step 2: Once the initial scores of the LVs are computed, inner weights (structural model weights) can be obtained. LVs as computed. The goal of this step is to recalculate scores but instead of getting the score of a LV as a linear combination of its indicators (MVs), score are computed as a linear combination of its associated variables. In other words, the connection among constructs in the inner model are taken into account in order to obtain a proxy of each latent variable calculated this time as a weighted aggregate of its adjacent LVs. There are three options to calculate inner weights, *centroid scheme*, *factor scheme* and *path scheme*. The path scheme was chosen for the analysis. However, according to the authors (Sanchez, 2013; Vinzi et al.,2010) the choice of the weighting scheme has little relevance on the estimation process and on the results significantly.

- Step 3: Once the inner weights have been estimated the internal approximation of latent variables can be computed.

- Step 4: Once the inside approximation is done the internal estimates of LVs must then be considered with regard their indicators. This is done by updating the outer weights.

The algorithm steps are repeated until convergence of outer weights is reached.

A3.1.2. Unidimensionality

The diagnosis of a PLS path model begins with the assessing of the quality of the measurement model. The evaluation of reflective indicators should involve three main aspects, unidimensionality (internal consistency) of indicators, the correlation between the latent variable and its indicators and correlations with other latent dimensions (Sanchez, 2013). In fact, as the reflective block reflects the (unique) latent construct, it should be homogeneous and unidimensional. The MVs belonging to a certain block are assumed to measure the same unique underlying concept. The main index for checking the block homogenenity are summarized as follows:

(a) *Cronbach's alpha*: this is a classical index of reliability and it traditionally used by the SEM community as a measure of internal consistency. As a rule of thumb, a block is considered homogeneous is the index value is greater than 0.7 for confirmatory studies. The index is expressed as:

$$\alpha = \frac{\sum_{p \neq p} cor(x_{pq}.x_{p'q})}{P_q + \sum_{p \neq p} cor(x_{pq}.x_{p'q})} \times \frac{P_q}{P_q - 1} \qquad (5p)$$

where P_q is the number of manifest variables in the q-th block.

(b)*Dillon-Goldstein's rho*: better known as composite reliability. As rule of thumb a block is considered homogeneous if the index value is larger than 0.7

$$\rho = \frac{\left(\sum_{p=1}^{P_q} \lambda_{pq}\right)^2}{\left(\sum_{p=1}^{P_q} \lambda_{pq}\right)^2 + \sum_{p=1}^{P_q}(1 - \lambda_{pq}^2)} \qquad (6p)$$

According to Chin (1998) Dillon-Goldstein rho is considered a better indicator than Cronbach's alpha. In fact, the latter give the same importance to each manifest variable (tau equivalence), while the second is based on the results from the models (loadings) rather than on the correlation observed between MVs. The Cronbach's index actually provides a lower bound estimate of reliability.

(c) *Principal component analysis of a block*: According to the Kaiser's rule, a block can be considered unidimensional if the first eigenvalue of its correlation matrix is higher than 1, while others are lower.

A3.1.3. The quality indexes (structural and measurement model assessment)

PLS Path Modeling lacks a well identified global optimization criterion so that there is no global fitting function to assess the goodness of the model. Furthermore, it is a variance-based model strongly oriented to prediction. Thus, model validation mainly focuses on the model predictive capability. According to PLS-PM structure, each part of the model needs to be validated: the measurement model, the structural model and the overall model. That is why, PLS Path Modeling provides three different fit indexes: the *communality index*, the *redundancy index* and the *Goodness of Fit (GoF)* index. A brief description of the index is provided as follows.

(a) *Communality index*: For each q-th block in the model with more than one MV the quality of measurement model is assessed by the *communality* index:

$$Com_q = \frac{1}{P_q}\sum_{P=1}^{P_q} corr^2\left(x_{pq,\xi_q}\right) \;\forall q: P_q > 1 \tag{7p}$$

The index measures how much of the manifest variables variability in the q-th block is explained by their own latent variable scores ξ_q. Moreover, the communality index for the q-th block is nothing but the average of the squared correlations between each manifest variable in the q-th block and the corresponding latent variable scores.

(b) *Average communality*. The quality of the whole measurement model can be assessed by means of the *average communality* index:

$$\overline{Com} = \frac{1}{\sum_{q:P_q>1} P_q} \sum_{:P_q>1} P_q Com_q \tag{8p}$$

This is a weighted average of all the Q block-specific communality indexes with weights equal to the number of manifest variables in each block.

(c) A common measure to establish convergent validity on the construct level is the average variance extracted (AVE). This criterion is defined as the grand mean value of the squared loadings of the indicators associated with the construct (i.e., the sum of the squared loadings divided by the number of indicators). Therefore, the AVE is equivalent to the communality of a construct.

(d) The most commonly used measure to evaluate the structural model is the coefficient of determination (R2 value). This coefficient is a measure of the model's predictive accuracy and is calculated as the squared correlation between a specific endogenous construct's actual and predicted values. The coefficient represents the exogenous latent variables' combined effects on the endogenous latent variable.

(f) In view of linking the prediction performance of the measurement model to the structural one, the redundancy index computed for the j-th endogenous block, measures the portion of variability of the manifest variables connected to the j-th endogenous latent variable explained by the latent variables directly connected to the block. The index is:

$$Red_j = Com_j \times R^2\left(\hat{\xi}_j, \hat{\xi}_{q:\xi_q \to \xi_j}\right) \tag{9p}$$

(g) A global quality measure of the structural model is also provided by the *average redundancy* index, computed as

$$\overline{Red} = \frac{1}{j}\sum_{j=1}^{J} Red_j \tag{10p}$$

(h) The global criterion of goodness of proposed by Tenenhaus et al. (2004), the GoF index. Such an index has been developed in order to take into account the model performance in both the measurement and the structural model and thus provide a single measure for the

overall prediction performance of the model. For this reason, the GoF index is obtained as the geometric mean of the average communality index and the average R^2 value:

$$GoF = \sqrt{\overline{Com} \times \overline{R^2}} \tag{11p}$$

where the average R^2 value is obtained as:

$$\overline{R^2} = \frac{1}{J} R^2(\hat{\xi}_j, \hat{\xi}_{q:\xi_q \to \xi_j}) \tag{12p}$$

A3.1.4. Rebus algorithm

A new method for unobserved heterogeneity detection in PLS-PM framework was recently presented by Trinchera (2007) and Esposito Vinzi et al. (2008). REBUSPLS is an iterative algorithm that permits to estimate at the same time both the unit membership to latent classes and the class specific parameters of the local models. The core of the algorithm is a so-called *closeness measure* (CM) between units and models based on residuals. The idea behind the definition of this new measure is that if latent classes exist, units belonging to the same latent class will have similar local models. Moreover, if a unit is assigned to the correct latent class, its performance in the local model computed for that specific class will be better than the performance of the same unit considered as supplementary in the other local models.
The CM used in the REBUS-PLS algorithm represents an extension of the distance used in PLS-TPM by Trinchera et al. (2006), aiming at taking into account both the measurement and the structural models in the clustering procedure. In order to obtain local models that fit better than the global model, the chosen *closeness measure* is defined according to the structure of the Goodness of Fit (GoF) index, the only available measure of global fit for a PLS Path Model.

A3.1.5. Cross-loadings tables

Table 35 Cross-loadings Pap-test intentions PLS-PM

MV	LV	PERSEV	PERCSUSC	SUBNORMS	NEGATT	INT
PercSev1	PERCEIVEDSEVERITY	0.8292	0.0265	0.0978	-0.1093	0.1494
PercSev2	PERCEIVEDSEVERITY	0.7460	0.0397	0.0322	0.0200	0.1254
PercSusc1	PERCEIVEDSUSCEPTIBILITY	0.0655	0.9144	0.0851	-0.0135	0.1431
PercSusc2	PERCEIVEDSUSCEPTIBILITY	-0.0070	0.8117	-0.0141	-0.0207	0.0992
Norm1	SUBJECTIVE NORMS	0.0420	0.0801	0.7927	-0.1675	0.2770
Norm2	SUBJECTIVE NORMS	0.1430	-0.0102	0.7205	-0.2187	0.2522
AttPap1	NEGATIVEATTITUDES	-0.0561	-0.0067	-0.2387	0.7966	-0.4645
AttPap3	NEGATIVEATTITUDES	-0.0261	0.0230	-0.1990	0.6962	-0.3553
AttPap2	NEGATIVEATTITUDES	-0.0228	-0.0186	-0.1737	0.7297	-0.3836
AttPap4	NEGATIVEATTITUDES	-0.0902	-0.0767	-0.1068	0.4948	-0.2205
FutInt1	INTENTIONS	0.1588	0.1334	0.3930	-0.5372	0.9202
FutInt2	INTENTIONS	0.1455	0.1145	0.1942	-0.3663	0.8123

Table 36 Cross-loadings vaccination intentions PLS-PM

MV	LV	SEVERITY	SUSCEPTIBILITY	SOCIALNORMS	SEVATTITUDES	INTENTIONS
PercSev1	SEVERITY	0.8559	0.3271	0.1240	-0.4308	0.4637
PercSev2	SEVERITY	0.9059	0.2289	0.2177	-0.4618	0.5664
PercSusc1	SUSCEPTIBILITY	0.1975	0.0966	0.0294	-0.1056	0.1298
PercSusc2	SUSCEPTIBILITY	-0.1026	0.2331	-0.0860	-0.1099	0.0814
PercSusc3	SUSCEPTIBILITY	0.0466	0.5834	0.0257	-0.1740	0.2428
Norm1	SUBJNORMS	0.1165	-0.1099	0.7710	-0.1432	0.2818
Norm2	SUBJNORMS	0.1821	0.0116	0.7133	-0.1702	0.2561
Att1	SEVATTITUDES	-0.2215	-0.1017	-0.1616	0.7478	-0.4752
Att2	SEVATTITUDES	-0.3698	-0.3186	-0.2144	0.8604	-0.7031
Att3	SEVATTITUDES	-0.5139	-0.4099	-0.0710	0.8001	-0.6387
Att4	SEVATTITUDES	-0.6041	-0.4661	-0.1973	0.8172	-0.7644
Att4	SEVATTITUDES	-0.2476	-0.1384	-0.1816	0.7777	-0.5214
Att6	SEVATTITUDES	-0.3254	-0.2494	-0.1601	0.7034	-0.5207
Int1	INTENTIONS	0.4243	0.3059	0.4036	-0.7158	0.9280
Int2	INTENTIONS	0.6621	0.4611	0.2871	-0.7581	0.9488

Appendix 4: Psychological theory supplementary materials

A4.1. Conformity

The purpose of this section is to review the theories and illustrative research that have developed since that time regarding conformity and independence.
The initial interest in conformity behaviour was rooted in the earliest social psychological research on suggestibility (Asch, 1948; Moscovici, 1985). Charcot's hypnosis work with hysterical patients in the late 1800s had been transported into the realm of social behaviour, and evidence accrued showing that people held positions that were not based on sufficient information, acted in ways that were contrary to the obvious facts of reality, and followed sometimes outlandish group behaviour, even when it contradicted the individual's personal position (Asch, 1948). Definitions of conformity have been varied and contested (see Allen, 1965; Levine & Russo, 1987; Nail, 1986; Willis, 1963, 1965). Most of the literature on conformity distinguishes it from normative behaviour by applying the criterion of movement from one's own position to a contradictory position; that is, the individual's personal position is contrary to that expressed by a comparison other or group (Asch, 1956). Thus, we conform to others when perceived or real pressure from them causes us to act differently from how we would act if alone (Kiesler & Kiesler, 1969; Myers, 1996). Given our focus on goal-directed behaviour, it is important to discuss why we conform. As noted earlier in the section "Social Norms," Deutsch and Gerard (1955) distinguished between two different motivations for conforming: informational influence represents conformity to others' positions when the concern is to make accurate and valid judgments; normative influence, on the other hand, represents conformity when the concern is to seek social approval from others or social harmony with others. In addition, Deutsch and Gerard note that conforming to our own self-expectations can enhance feelings of self-esteem or self-approval, while nonconformity can lead to feelings of anxiety or guilt. In the sections that follow, we will consider the evidence concerning when people conform and how conformity can facilitate achievement of the goals of behaving effectively, building and maintaining social relationships, and managing self-concept (Cialdini & Trost, 1998: 162).

A4.1.1. The goal of effective action

Although conformity typically carries a negative connotation in Western culture, it can be an effective and time-saving strategy. Group consensus is typically the most direct route to goal

attainment (Festinger, 1950). Moreover, consensus provides an easy heuristic about how to act, because consensus implies correctness (Cialdini, 1993). Research following Sherifs (1936) autokinetic studies had been aimed primarily at demonstrating "slavish submission to group pressure" (Asch, 1956:2). The autokinetic effect is a perceptual illusion created by shining a small point of light on a wall in a completely darkened room. In the absence of an external frame of reference, the light appears to move erratically, even though it is held perfectly still. This novel, ambiguous stimulus provided "objectively unstable situations that would permit themselves to be structured in several ways, depending upon the character of the subjectively established reference points" (Sherif, 1936, p. 91). The author compared the behaviour of group members when they were alone and when they were tested in group. Three experimental conditions were created, an individual session (subjective norm) an individual session followed by the group session (forming the group norm) and group session followed by an individual session (conservation group norm). On the basis of the experimental results, Sherif (1936) concluded that unstable situations evoke confusion and uncertainty. Under such circumstances, people assume that "the group must be right" (p. 111) and look to the group to establish a common norm. Sherif (1936) concluded that in the absence of an objective rule for behaviour, people were most likely to behave according to the group consensus. Sherif's research showed how contact with others influences our immediate perceptions of reality. This influence can also be internalized; Rohrer ct al. (1954) found that people who were retested individually as much as a year after the norm induction still reported the group norm. On the other hand, Ash's experiments on autokinetic stimulus showed that participants gone along with the majority because they felt the majority had to be right and they had to be wrong. In other words, those who conformed did so to meet both social and accuracy goals. Asch (1956) himself interpreted the results as evidence of conformity due to an informational conflict between one's own senses and the perceptions of others. Research into the nature of conformity arose out of Asch's (1956) remarkable finding that people with normal vision would ignore their own eyes to agree publicly with an obviously inaccurate group judgment. Subsequent researchers have uncovered other factors that influence the level of conformity exhibited in similar studies, unanimity of the majority, group cohesion, private responses, prior commitment, individual differences and the size of the group.

Unanimity of the majority is a critical component; having just one other participant give an accurate alternative some or all of the time significantly decreases errors made by the targets (Asch, 1955; Morris & Miller, 1975). At the same time, manipulating the perception of *group cohesion* affects errors; for example, participants try to support the group response when a reward is offered for the best group performance (Deutsch & Gerard, 1955). On the other

hand, a person who is not accepted by the group exhibits less conformity when a reward is offered for group performance.

Private responses also elicit less conformity than public, face-to-face, or simulated group responses (Argyle,1957; Asch, 1956; Campbell & Fairey, 19S9; Deutsch & Gerard, 1955; Levy, 1960; Mouton, Blake, & Olmstead,1956). Deutsch and Gerard (1955) had some of their participants write down their own perceptions before being exposed to the inaccurate group response. *Prior commitment*, especially when the participants expected the experimenter to see the initial judgment, dramatically reduced the number of errors. Even writing the individual response on an anonymous piece of paper or an erasable "magic pad" decreased the error rate over a no-commitment condition (Cialdini & Trost, 1998). Similarly, *task difficulty* can affect reliance on others' input for making an accurate judgment: greater difficulty elicits more conformity to the group position (Coleman, Blake, & Mouton, 1958). Also, a target who has been made to feel incompetent in his or her judgments prior to group exposure will rely more on the group's judgment than will a competent target (Hochbaum, 1954).

The *size of the group* opposing the individual's judgment would also be expected to affect conformity if others were providing informational influence: several independent sources giving the same response should be more informative than one. Accordingly, Asch (1956) found that the amount of conformity increased dramatically as the number of opposing voices grew from one to three, although the influence of additional group members (more than three, up to fifteen), was minimal (a leveling effect replicated by Rosenberg, 1961). Gerard, Wilhelmy, and Conolley (1968) highlighted the independence of the group members' responses, and reported a linear increase in the level of conformity as the group size increased.

A4.1.2. Majority and minority social influence.

To counter the "conformity bias" that had characterized social influence research well into the 1960s, Moscovici and his colleagues began a program of research to demonstrate that social influence is a reciprocal process in which both the minority target and the majority source are agents and receivers of influence. From this perspective, a minority is not simply a passive target accepting the influence of the more numerous majority, it also has the ability to challenge the status quo, creating conflict and the possibility of innovation as opposed to stagnation." According to Moscovici (1976), a dissenting minority breaks the consensus of the majority, challenging the validity of the majority position and creating a dilemma between

risking deviance by accepting the minority's informational advantage (conversion) or maintaining the social support provided by complying with the majority position (compliance). Arguing that minorities and majorities elicit different outcomes, Moscovici (1980, 1985) predicted that they would also trigger distinct social influence processes. Moscovici's (1980, 1985) dual-process model proposes that a contrary majority elicits a comparison process for the target of influence, who focuses on the discrepancy between his or her position and that advocated by the more numerous majority. This focus outward on the majority message triggers the need for consensus, producing compliance with (but not internalization of) the majority position. On the other hand, a contrary minority elicits a validation process for the target, who is compelled to examine the accuracy of the minority position critically if the minority presents a realistic, consistent alternative viewpoint. Since the target is scrutinizing a reasonable message, this examination should elicit internalized change. Resolving the conflict between minority and majority influence should take the path of least resistance (Moscovici, 1980). That is, when confronted with a disagreeable majority, it is easier to maintain one's public image by publicly complying with the majority but not changing one's internal altitude. Publicly agreeing with a deviant minority, however, can damage one's reputation (Mugny, 1982); therefore, it may be easier to change one's internal opinion without voicing it in public. An extensive body of literature has developed in the past twenty-five years supporting Moscovici's contention that numerically inferior minorities can exert influence (Levine & Moreland, 1998, in this Handbook; Levine & Russo, 1987; Maass & Clark, 1984; Maass, West, & Cialdini, 1987; Moscovici, 1980, 1985; Moscovici, Mucchi-Faina, & Maass, 1994; Nemeth, 1986; Wood et al., 1994). Wood et al.'s (1994) meta-analysis indicates that majorities tend to have more influence on public measures and private direct measures, while minorities have more influence on private, indirect measures (although minority influence has also been found on direct measures; sec Maass & Clark, 1983; Trost, Maass, & Kenrick, 1992). Subsequent research has shown that consistency is an essential criterion for eliciting minority influence (Nemeth, Swcdlund, & Kanki, 1974): a consistent minority of two is more influential than either a consistent individual or an inconsistent minority (Moscovici & Lage, 1976), and a minority that shifts to a less extreme position has little influence (Kiesle & Pallak, 1975; Paicheler, 1976,1977). Although the minority may instigate an informational conflict, when people are instructed to be accurate they tend to rely on the consensus heuristic and the majority position. Moscovici and Lage (1978) found that as the accuracy goal became more salient, the majority's influence increased (and vice versa for the originality norm). They argued that the *majority elicits convergence* on the *group norm*, whereas a dissenting *minority elicits innovative responses* (Cialdini& Trost, 1998:165).

A4.1.3. Conformism and vaccination decision making process

The conformism concept can play an important role in the acceptance of a specific vaccination plan, influencing and limiting the individuals' decision-making process. Let is consider for example the most common paediatric vaccinations, so-called "mandatory". Is it more likely that parents decide to vaccinate their children after having gathered information and read up about the risks and benefits of vaccines (this process requires some efforts in terms of cost and time that someone may prefer to avoid) or they may be influenced by the masses, and in particular by those people who represent their reference point? This consideration may suggest that conformism and social norms, may represent an independent or alternative mechanism of behaviour in respect to the benefit-cost approach. By following the masses, the decision comes without any effort and it may look reassuring. This mass behaviour in the long term may reduce the general vaccination goals and upsides' awareness and lead to vaccination refusal.

A4.1.4. Minority social influence of antivaccination groups

An example of minority social influence may be represented by the increase in the number of anti-vaccination people. Paediatric anti-vaccination groups obviously represent a social minority: this is proved by the fact that, except in special situations, the European vaccination coverage is generally more than 90% for all vaccines included in the national immunisation plans; this happens even in those countries where vaccinations are not compulsory. However, in the remaining 10%, very active anti-vaccination groups exist. Some of them acts in closed communities, as they are religious minorities (for example the so-called "dutch anti-vaccination belt"), or specific philosophical or life styles' groups (for example in Trentino Alto-Adige, with a vaccination coverage of 25-30%). The other anti-vaccination groups use web and social network to broadcast their message with a higher impact. Anti-vaccination groups can be differentiated according to the reason of their position: religion, medical theories or conspiracy reason; therefore, there is not a unique approach to be followed for all of them. Approaches to understand the origin and the diffusion of anti-vaccination groups should refer to the Moscovici theory, as they are a minority in the society and they have a strong diachronic and synchronic coherence. Being and anti-vaccinator means to have a very active person, with respect to the vaccinators counterpart.

A4.2. Weakness of TPB

Measurement and definition of TRA/TPB constructs is often characterized by some level of difficulties. The absence of clear guidelines and/or the different research contexts may result in biased results or make dimensions uninterpretable. In particular, Ajzen and Fishbenin (1975, 1980, 1991) indicated that is necessary a high degree of analogy (correspondence) between measures of attitude, norm, perceived control, intention, and behaviour in terms of action, target, context and time. The "principle of correspondence" (Ajzen & Fishbein, 1977;Fishbein & Ajzen, 1975) indicates that all the model variables (attitude, norm, perceived control, intention, and behaviour) should be matched with the four elements (action, target, context, and time) in order to maximize the predictive power of the theory. A low level of correspondence between them will result in a low correlation rate between TRA/TPB constructs. Often in empirical researchers the principle of correspondence might not been met (Rutter & Quine, 2002; Glanz et al., 2008).

According to Fishbein (2000), each behaviour is substantively unique. In fact, the influence of the different model constructs on the behaviour in question depends on the population or culture under study. For a certain culture, some behaviours are influenced by attitudes, others by subjective norms, and still others by perceived behavioural control (Rutter & Quine, 2002). In fact, a weakness aspect of the theory is the lack of any sort of guideline that clearly explain how the influence of attitude, subjective norm and behavioural control may change across different subjective contexts (Glanz et al, 2008).

Another issue related with the theory is total absence of actual behaviour control measurements. More generally, if a study fails to measure the impacts of actual behaviour control, the estimation of the perceived behavioural control-behaviour casual effect may be biased. That occurs unless perceived control is considered as accurate counterpart of actual control (i.e. that perceived and actual control are perfectly correlated and this correlation arises from an direct causal effect of actual on perceived control) (Glanz et al, 2008). In addition, the absence of a theoretical hypothesis that assumes a relationship (and correlation) between actual behaviour control, behaviour and normative beliefs (that may exist empirically) it would have implications for the interpretation of regression analyses from which actual control was omitted. In particular, intention is expected to have a stronger influence on behaviour the greater the degree of actual control the person has over the action in question. Even though the TPB is a deliberative processing model, it is important to bear in mind that, in many situations the generation of beliefs, attitude and intentions is automatic and based on an "unconsciousness" mental process (Ajzen & Fishbein, 2000; Fishbein & Ajzen,

1975). In our opinion, the underlie process by which pro-vaccination parents formulate their belief, attitude and intentions about the HPV vaccination choice, is almost entirely unconsciousness. However, although it seems plausible that automatic processes control the formation and change of beliefs, attitudes and intentions, for most health-related behaviours, it seems less plausible to suggest that behaviour itself is automatically elicited.

In TPB, behaviour determinants are separate in two main categories, the proximal determinants specified by the theory and, all other causes, left unspecified and suppose to influence the behaviour only through the previous. For this reason, the theory is often considered a "sufficient" behaviour model. Thus because, in many situations, external factors, such as socio-demographic variables, may be direct cause of a certain behaviour. The model needs a theoretical support that account for possible relationship between external variable, proximal determinants and behaviour.

Appropriate statistical and psychology tools are needed to evaluate behavioural intentions. To increase statistical power and measurement reliability, questionnaires should include a variety of items for intention question related. The role of intention in leading to behaviour has been rejected by some recent studies. Although longitudinal studies show a high correlation between intentions and behaviour (Sheeran, Trafimow, Finlay, and Norman, 2002), experimental design and observational researches have indicate a less powerful relationship (Webb and Sheeran, 2006). Additionally, behaviours may be both, overstate or understate, by intentions. Finally, some studies and reviews assert that the model postulate by which attitudes can affect behaviours only through intentions is not always met. Analysis that control for past behaviours often find that intentions in no longer a significant predictor of behaviour (Hagger, Chatzisarantis, and Biddle, 2002).

A4.3. Weakness of the HBM

There are no clear guidelines on how model dimensions should be combined to influence behaviour or how different beliefs should have been connected one another. In fact, Rosestock did not specified how HBM constructs should be liked or related to predict behaviours. The absence of specific guidelines has led to variation in HBM applications and as a result, different studies have used different combinations of variables or treated/interpreted them differently during the data analysis (Rutter & Quine, 2002). For example, whereas many studies have attempted to establish each of the major dimensions as independent, others have tried multiplicative approaches (Glanz et al., 2008). In addition, some researchers proposed that the perceived susceptibility may only become relevant once perceived severity reaches a

certain threshold. (Weinstein, 1988; Lewis, 1994). Others decided to combine benefits and barrier in a single index (e.g. Giannetti et al, 1985) ignoring the theoretical differences between the two dimensions (Weinstein, 1988; Cambridge book). In an empirical point of views, in fact, the benefits dimension could be characterized by different components, i.e. the efficacy of the behaviour in achieving an outcome, the psychological benefits, the subjective approval, etc. Similarly, the barriers may include various aspects, such as physical limitations, psychological costs, etc. These differences show how a single index could be inadequate in accounting for all aspects of each dimensions (Glanz et al., 2008; Cambridge book). Finally, cues to action and health motivation have been relatively neglected in empirical tests of the HBM. Neither Janz and Becker (1984) nor Harrison et al. (1992) included these components because of the paucity of relevant studies. One reason for researchers' failure to operationalize these components may be the lack of clear construct definitions. Analytical approaches to identifying model construct relationships are needed to further the utility of the HBM in predicting behaviour (Glanz et al., 2008).

Appendix 5: The Questionnaire

Prevenzione dell'infezione da HPV e del tumore del collo dell'utero

Gentile Studentessa,

chiediamo la tua collaborazione per poter conoscere quali siano le opinioni e gli atteggiamenti delle giovani donne relativamente alla prevenzione della salute, con particolare riferimento alla vaccinazione contro il virus del papilloma umano (o virus HPV). Il tuo contributo sarà molto importante per far progredire le conoscenze sui fattori che maggiormente influenzano le decisioni di salute delle giovani donne che a partire dalla vostra età iniziano a dover scegliere autonomamente come salvaguardarla.

Il questionario tratterà vari argomenti, riguardanti la prevenzione dell'infezione da virus del papilloma umano (HPV), il vaccino ad esso associato, le percezioni di rischio e la sfera sessuale. Ti preghiamo di notare che non esistono risposte giuste o sbagliate: **siamo interessati unicamente alle tue opinioni**. Se vi fossero domande che ti mettono in imbarazzo, delle quali ci scusiamo, sentiti libera di non rispondere.

Il questionario è completamente anonimo e la compilazione richiede tra i 25 e i 35 minuti

Scopo della ricerca
Lo scopo della seguente indagine è quello di individuare quali siano state le determinanti principali che possono aver influenzato la tua decisione o quella dei tuoi genitori (o tutori legali) in merito alla vaccinazione contro il papilloma umano. Inoltre, ampliando il contesto d'indagine, siamo interessati a capire le strategie delle giovani donne in tema di prevenzione della salute.
La ricerca non intende discriminare né giudicare le scelte realizzate, ma solo interpretare obiettivamente le ragioni che le accompagnano.

Trattamento dei dati personali
Come richiesto dal Codice in materia di protezione dei dati personali (D.Lgs. 196/2003, "Testo Unico sulla Privacy"), ti informiamo che i tuoi dati personali saranno utilizzati esclusivamente per scopi di ricerca scientifica e solo da personale autorizzato. Le informazioni raccolte saranno strettamente confidenziali e saranno usate in modo aggregato ed anonimo per la preparazione di rapporti scientifici nei quali tu non sarai in alcun modo identificabile. Sarà nostra cura segnalarti le future iniziative di diffusione o presentazione dei risultati della ricerca che saranno promosse alla fine dell'indagine.

Ti ringraziamo molto per la preziosa collaborazione.

Per eventuali chiarimento potete rivolgervi al responsabile dell'indagine, il prof. Piero Manfredi, Dipartimento di Economia e Management, Università di Pisa, email: piero.manfredi@unipi.it.

Sezione 1: Informazioni socio/demografiche

1. In quale anno accademico ti sei iscritta all'università per la prima volta?

1 2015/2016
2 2014/2015
3 2013/2014
4 2012/2013
5 2011/2012
6 2010/2011
7 2009/2010
8 2008/2009
9 2007/2008
10 2006/2007
11 2005/2006
12 2004/2005
13 2003/2004
14 2002/2003
15 2001/2002

2. Da chi è composto il tuo nucleo familiare?

	Età (in anni compiuti)
☐ Io	
☐ Madre	
☐ Padre	
☐ Fratello 1	
☐ Fratello 2	
☐ Fratello 3	
☐ Sorella 1	
☐ Sorella 2	
☐ Sorella 3	
☐ Altri (specificare_____)	
☐ Altri (specificare_____)	
☐ Altri (specificare_____)	

3. Qual è la tua cittadinanza?
1 Italiana
2 Altro (specificare_____)

4. In che provincia italiana abitavi quando hai compiuto 16 anni?
1 _____
2 Risiedevo all'estero (specificare lo stato _____)

5. Qual è il titolo di studio di tua madre
1 Nessuno
2 Licenza elementare
3 Licenza media
4 Diploma/Titolo di istruzione secondaria superiore
5 Laurea
6 Master post-laurea/Dottorato di ricerca

7 Altro(specificare_____)
8 Non so/non rispondo

6. Qual è il titolo di studio di tuo padre?

1 Nessuno
2 Licenza elementare
3 Licenza media
4 Diploma/ Titolo di istruzione secondaria superiore
5 Laurea
6 Master post-laurea/Dottorato di ricerca
7 Altro(specificare_____)
8 Non so/non rispondo

7. Qual è la tua religione?

1 Cattolica
2 Ortodossa
3 Protestante
4 Musulmana
5 Buddista
6 Induista
7 Testimone di Geova
8 Agnostica
9 Atea
10 Altro(specificare_____)
11 Non so/non rispondo

Solo per chi ha indicato una risposta da 1 a 7 nella domanda n.7

8. Sei praticante?

1 Sì
2 No

9. Qual è la religione della tua famiglia?

1 Cattolica
2 Ortodossa
3 Protestante
4 Musulmana
5 Buddista
6 Induista
7 Testimone di Geova
8 Agnostica
9 Atea
10 Non praticante
11 Altro(specificare_____)
12 Non so/non rispondo

Solo per chi ha indicato una risposta da 1 a 7 nella domanda n.9

10. I tuoi genitori sono praticanti?

1 Sì, entrambi

2 Solo mio padre
3 Solo mia madre
4 Nessuno dei due

11. Sei una fumatrice?
1 Sì **(Prosegui con la domanda 12)**
2 No **(Prosegui con la domanda 13)**

12. Quante sigarette al giorno fumi?
Indicare il numero _____

13. In futuro il gruppo di ricerca svolgerà altre indagini volte a conoscere l'opinione di ragazze giovani come te. Se sei interessata a partecipare, puoi lasciarci una tua email?

Email:
_____
____

Sezione 2: Sessualità, conoscenze e prevenzione

In questa sezione ti chiediamo alcune informazioni di carattere personale che riguardano la tua sfera sessuale e le tue strategie di prevenzione.

14. Hai mai avuto rapporti sessuali completi?

1 Sì **(Prosegui con la domanda 15)**
2 No **(Prosegui con la domanda 20)**

15. A che età (in anni compiuti) hai avuto il tuo primo rapporto sessuale completo?

_____

16. Quanti partner (persone con cui hai avuto rapporti completi) hai avuto negli ultimi 12 mesi?

1 Nessuno **(Prosegui con la domanda 18)**
2 1 **(Prosegui con la domanda 18)**
3 2-3 **(Prosegui con la domanda 17)**
4 4-5 **(Prosegui con la domanda 17)**
5 >5 **(Prosegui con la domanda 17)**

17. Potresti specificare il numero esatto?

_____

18. Quanti dei partner avuti negli ultimi 12 mesi potevano essere considerati occasionali?

1 Nessuno
2 Una piccola parte Una piccola parte
3 Circa la metà Circa la metà
4 Più della metà Più della metà
5 Tutti

19. Quanti partner (persone con cui hai avuto rapporti completi) hai avuto nell'arco della tua vita?

1 Nessuno
2 1
3 2-3
4 4-5
5 6-10
6 >10

20. Quali metodi contraccettivi utilizzi/utilizzeresti tra quelli sotto elencati? (Sono possibili più risposte)

1 Coito interrotto
2 Profilattico
3 Pillola anticoncezionale
4 Pillola "del giorno dopo"
5 Altro (specificare_____)
6 Nessuno dei precedenti

21. Quali metodi contraccettivi, tra quelli sotto elencati, possono proteggere dal contagio con malattie sessualmente trasmissibili durante un rapporto sessuale? (Sono possibili più risposte)

1 Coito interrotto
2 Profilattico
3 Pillola anticoncezionale
4 Altro (specificare_____)
5 Nessuno dei precedenti

22. Con quale frequenza utilizzi il profilattico con

Partner occasionali	☐Sempre	☐Spesso	☐Ogni tanto	☐ Mai
Partner non occasionali	☐Sempre	☐ Spesso	☐Ogni tanto	☐ Mai

23. A chi ti sei affidata sino ad oggi per ottenere informazioni sulla sessualità e sui metodi contraccettivi? (Sono possibili più risposte)

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli
4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)

24. Ti è mai capitato di parlare in famiglia della prevenzione dei tumori femminili?

1 Sì
2 No
3 Non so/Non ricordo

25. Escludendo casi di malattia, con che frequenza vai dal ginecologo?

1 Più di una volta all'anno **(Prosegui con la domanda 26)**
2 Ogni anno **(Prosegui con la domanda 26)**
3 Ogni due/tre anni **(Prosegui con la domanda 26)**
4 Non ci sono mai stata **(Prosegui con la domanda 27)**

26. Facendo riferimento alla tua ultima visita dal ginecologo

1 L'ho eseguita di mia iniziativa
2 L'ho eseguita su consiglio del medico curante
3 L'ho eseguita su consiglio dei miei genitori
4 Altro (Specificare_____)

27. Hai mai effettuato il "pap-test"?

1 Sì **(Prosegui con la domanda 28)**
2 No **(Prosegui con la domanda 29)**
3 Non lo conosco **(Prosegui con la domanda 29)**

28. Facendo riferimento al tuo primo pap-test

1 L'ho eseguito di mia iniziativa
2 L'ho eseguito su consiglio del medico/ginecologo per prevenzione
3 L'ho eseguito su consiglio del medico/ginecologo per motivi di salute
4 L'ho eseguito su consiglio dei miei genitori
5 Altro (Specificare_____)

29. Quali tra le seguenti vaccinazioni ti ricordi di aver effettuato durante l'infanzia? (Sono possibili più risposte)

1 Difterite
2 Epatite virale B
3 Morbillo
4 Parotite (Orecchioni)
5 Pertosse
6 Poliomielite
7 Rosolia
8 Tetano
9 Altro(specificare_____)
10 Non so/non ricordo

Sezione 3: conoscenza del virus HPV e del tumore al collo dell'utero

30. Hai mai sentito parlare del Virus del Papilloma Umano (anche indicato con l'acronimo inglese HPV, Human Papilloma Virus)?

1 Si
2 No **(Prosegui con la sezione 8)**

31. Per ognuna delle seguenti affermazioni spunta la casella corrispondente al tuo parere personale

- Il virus HPV può causare il tumore al collo dell'utero
  ☐ Sì ☐ No ☐ Non so

- Il cancro del collo dell'utero è uno dei tumori più frequenti nelle donne
  ☐ Sì ☐ No ☐ Non so

- L'infezione da HPV si trasmette tramite rapporti sessuali
  ☐ Sì ☐ No ☐ Non so

- L'infezione da HPV può trasmettersi anche in seguito ad un rapporto sessuale non completo
  ☐ Sì ☐ No ☐ Non so

- L'infezione da HPV può essere contratta solo dalle donne
  ☐ Sì ☐ No ☐ Non so

- Sono necessari svariati anni perché l'infezione da HPV si trasformi in tumore
  ☐ Sì ☐ No ☐ Non so

- La maggior parte delle infezioni da HPV può guarire spontaneamente senza alcun trattamento medico/farmacologico
  ☐ Sì ☐ No ☐ Non so

- L'utilizzo del preservativo garantisce una protezione quasi totale dal virus HPV
  ☐ Sì ☐ No ☐ Non so

- Avere più partner sessuali aumenta il rischio di contrarre il virus HPV
  ☐ Sì ☐ No ☐ Non so

- Effettuare il Pap-test con regolarità riduce il rischio di insorgenza del tumore al collo dell'utero
  ☐ Sì ☐ No ☐ Non so

- Solo alcuni tipi di virus HPV sono in effetti ad alto rischio di generare il cancro
  ☐ Sì ☐ No ☐ Non so

- L'HPV è l'unica causa conosciuta del tumore al collo dell'utero
  ☐ Sì ☐ No ☐ Non so

- L'HPV è un'infezione sessualmente trasmissibile molto diffusa in Italia
  ☐ Sì ☐ No ☐ Non so

- L'infezione da HPV è in genere asintomatica
  ☐ Sì ☐ No ☐ Non so

32. Quanto ti trovi in accordo o in disaccordo con le affermazioni ai punti a e b? (Spunta la casella corrispondente al tuo parere personale)

a. Credo che il tumore al collo dell'utero sia una malattia molto grave

Completamente in disaccordo	Non d'accordo	Neutrale (Indifferente)	D'accordo	Completamente d'accordo
☐	☐	☐	☐	☐

b. Mi spaventa pensare che potrei ammalarmi di cancro al collo dell'utero in futuro

Completamente in disaccordo	Non d'accordo	Neutrale (Indifferente)	D'accordo	Completamente d'accordo
☐	☐	☐	☐	☐

33. Dove/da chi hai ottenuto informazioni sull'HPV, sulle sue conseguenze e sul vaccino ad esso associato? (Sono possibili più risposte)

1 Medico di famiglia/ ginecologo
2 Operatori sanitari (ASL, consultorio, etc.)
3 Genitori
4 Altri familiari (sorelle, fratelli, etc.)
5 Amici
6 Libri/articoli scientifici
7 Scuola/Università
8 Televisione/radio
9 Internet
10 Brochure/opuscoli
11 Quotidiani/riviste
12 Pubblicità da parte di aziende farmaceutiche
13 Associazioni nazionali per la lotta al cancro
14 Altro (specificare_____)
15 Non ho cercato informazioni

Sezione 4: conoscenza vaccino anti-HPV e stato vaccinale

34. Hai già sentito parlare del vaccino anti-HPV?

1 Si
2 No **(Prosegui con la sezione 8)**

35. In base alle tue conoscenze il vaccino oggi disponibile protegge:

1 Per l'intera durata della vita contro tutti i tipi di HPV in grado di generare il tumore
2 Per l'intera durata della vita contro i più comuni tipi di HPV ad alto rischio di generare il tumore
3 Per almeno dieci anni contro tutti i tipi di HPV in grado di generare il tumore
4 Per almeno dieci anni contro i più comuni tipi di HPV ad alto rischio in grado di generare il tumore
5 Altro (specificare_____)
6 Non so

36. Secondo te l'efficacia del vaccino si riduce se fatto ad una ragazza sessualmente attiva

1 Si
2 No
3 Non so

37. Quanto ti trovi in accordo/disaccordo con le seguenti affermazioni? (Spunta la casella corrispondente)

	Completamente in disaccordo	Non d'accordo	Neutrale (Indifferente)	D'accordo	Completamente d'accordo
Credo che il vaccino sia efficace nel proteggere contro l'infezione da HPV	☐	☐	☐	☐	☐
Credo che la distribuzione del vaccino da parte dello Stato sia influenzata dalle case farmaceutiche	☐	☐	☐	☐	☐
Ho paura che il vaccino possa causare problemi di salute anche molto tempo dopo la vaccinazione	☐	☐	☐	☐	☐
Credo che la vaccinazione fornisca una protezione parziale dal rischio di contrarre l'HPV in eventuali rapporti non	☐	☐	☐	☐	☐

protetti					
Credo che la protezione offerta dal vaccino anti HPV contro il tumore al collo dell'utero sia più importante dei suoi eventuali effetti collaterali	☐	☐	☐	☐	☐
Credo che il vaccino sia sicuro	☐	☐	☐	☐	☐
Credo che chi si vaccina contro l'HPV non abbia più bisogno di sottoporsi all'esame del pap-test in futuro	☐	☐	☐	☐	☐
Credo di poter gestire da sola (scegliere senza l'ausilio dei miei genitori) le scelte che riguardano la mia salute.	☐	☐	☐	☐	☐

38. Sei stata vaccinata contro il virus HPV?

1 No, non sono stata vaccinata **(Prosegui con la sezione 5)**
2 Sì, sono stata vaccinata **(Prosegui con la domanda 39)**
3 Non so/Non ricordo **(Prosegui con la sezione 7)**

39. A che età sei stata vaccinata? (Rispondi e prosegui con la sezione 6)

1 Età (in anni compiuti) _____
2 Non so/Non ricordo

Sezione 5: Decisioni e atteggiamenti

Solo per chi ha risposto "No, non sono stata vaccinata" alla domanda 41

40. Non sei stata vaccinata:

1 Perché si è preferito posticipare la vaccinazione
2 Perché si è preferito non effettuare la vaccinazione
3 Perché non ci si è mai posti il problema
4 Altro (specificare_____)

41. Chi ha preso la decisione finale di non farti vaccinare?

1 Ho deciso autonomamente
2 Ho preso la decisione insieme ai miei genitori (o tutori legali)
3 Hanno deciso i miei genitori (o tutori legali)
4 Non ricordo

42. Cosa faresti se domani ti venisse riproposta la vaccinazione anti-HPV?

1 Sicuramente mi vaccinerei **(Prosegui con domanda 43)**
2 Probabilmente mi vaccinerei **(Prosegui con domanda 43)**
3 Neutrale
4 Probabilmente non mi vaccinerei **(Prosegui con domanda 44)**
5 Sicuramente non mi vaccinerei **(Prosegui con domanda 44)**

43. Se il vaccino fosse a pagamento, cambieresti la tua decisione di vaccinarti?

1 Sì
2 No

44. Se domani ti venisse riproposta la vaccinazione, il parere di quali persone potrebbe influenzare la tua scelta?

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli
4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)

45. Quali di queste persone sarebbe a favore di una tua eventuale vaccinazione?

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli
4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)
7 Nessuno dei precedenti

46. Quanto ti trovi in accordo/disaccordo con le seguenti affermazioni? (Spunta la casella corrispondente al tuo parere personale)

	Completamente in disaccordo	Non d'accordo	Neutrale (Indifferente)	D'accordo	Completamente d'accordo
La mia religione non consente i rapporti prematrimoniali quindi non ho bisogno del vaccino	☐	☐	☐	☐	☐
Credo che dovrei informarmi meglio prima di prendere una decisione sulla vaccinazione anti HPV	☐	☐	☐	☐	☐
Sono contraria a tutte le vaccinazioni	☐	☐	☐	☐	☐
Mi spaventano i possibili effetti collaterali post-vaccinazione	☐	☐	☐	☐	☐

Sezione 6: Decisioni e atteggiamenti

Solo per chi ha risposto "Sì, sono stata vaccinata" alla domanda 41

47. Chi ha preso la decisione finale di farti fare la vaccinazione anti HPV?

1 Ho deciso autonomamente
2 Ho preso la decisione insieme ai miei genitori (o tutori legali)
3 Hanno deciso i miei genitori (o tutori legali)
4 Non so/non ricordo

48. Ti ricordi se qualcuna delle persone e/o dei gruppi sotto elencati hanno condizionato la tua decisione o quella dei tuoi genitori (o tutori legali) di vaccinarti contro l'HPV? (Sono possibili più risposte)

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli
4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)

49. Hai avuto effetti collaterali legati alla vaccinazione HPV?

1 Sì **(Prosegui con la domanda 50)**
2 No **(Prosegui con la sezione 7)**

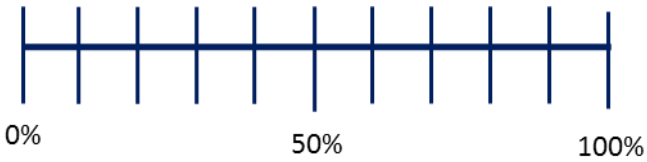
50. Quali effetti collaterali hai avuto?

Sezione 7(Per tutti): Percezioni del rischio associato al virus HPV

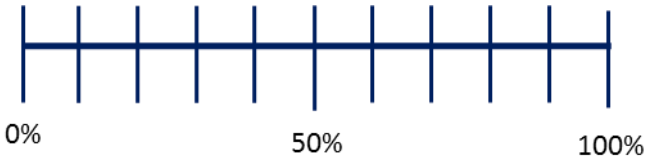
In questa sezione ti preghiamo di indicare con una crocetta, in una scala percentuale da 0% a 100%, quanto ritieni probabile che un determinato evento si verifichi (0% indica un evento ritenuto impossibile e 100% un evento che si verifica sempre).

NOTA BENE: se non conosci né la vaccinazione né il pap-test non rispondere alle domande seguenti

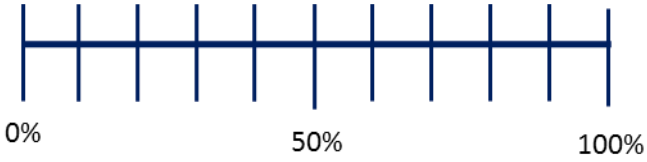
51. **Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che NON SI SIA VACCINATA e NON SI SOTTOPONGA A PAP-TEST possa ammalarsi di tumore al collo dell'utero?**



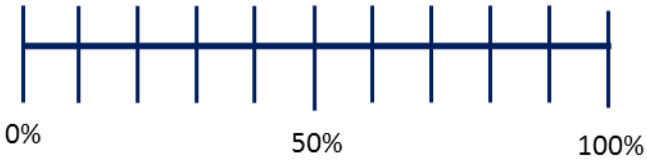
52. **Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che NON SI SIA VACCINATA ma SI SOTTOPONGA A PAP-TEST possa ammalarsi di tumore al collo dell'utero?**



53. **Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che SI SIA VACCINATA ma NON SI SOTTOPONGA A PAP-TEST possa ammalarsi di tumore al collo dell'utero?**



54. **Quanto ritieni probabile che una ragazza con il tuo stile di vita sessuale che SI SIA VACCINATA e SI SOTTOPONGA A PAP-TEST possa ammalarsi di tumore al collo dell'utero?**



Sezione 8(Per Tutti): Atteggiamenti e prevenzione del cancro al collo dell'utero

NOTA BENE: se non conosci il pap-test non rispondere alle domande seguenti

55. Quante delle tue amiche più strette hanno già effettuato il pap-test?

1 Tutte
2 Più della metà
3 Circa la metà
4 Una piccola parte
5 Nessuna
6 Non lo so

56. Per ognuna delle seguenti affermazioni spunta la casella corrispondente al tuo parere personale

- Il pap-test prevede il prelievo di cellule dalla cervice uterina
☐ Sì ☐ No ☐ Non so

Un pap-test negativo indica l'assenza dell'HPV
☐ Sì ☐ No ☐ Non so

- Il pap-test è raccomandato a partire dai 25 anni di età, ogni 3 anni fino ai 65 anni d'età
☐ Sì ☐ No ☐ Non so

- Se il pap-test è positivo è necessario sottoporsi ad ulteriori indagini diagnostiche (colposcopia, visita ginecologica, etc.)
☐ Sì ☐ No ☐ Non so

57. Quanto ti trovi in accordo/disaccordo con le seguenti affermazioni? (Spunta la casella corrispondente)

Se non conosci il vaccino non rispondere alle domande seguenti

	Completamente in disaccordo	Non d'accordo	Neutrale (Indifferente)	D'accordo	Completamente d'accordo
Effettuare il pap-test regolarmente fa diminuire la possibilità che mi ammali di cancro cervicale in futuro	☐	☐	☐	☐	☐
Ho paura di effettuare il pap-test perché mi spaventerebbe un eventuale risultato negativo	☐	☐	☐	☐	☐
Credo di essere troppo giovane per dover effettuare l'esame del pap-test	☐	☐	☐	☐	☐
Credo che grazie al	☐	☐	☐	☐	☐

pap-test eventuali forme pre-cancerogene possano essere trattate precocemente					
L'esame del pap-test mi mette a disagio	☐	☐	☐	☐	☐

58. Quanto è probabile che tu ti sottoponga all'esame del pap-test in futuro?

1 Sicuramente effettuerò il pap-test
2 Probabilmente effettuerò il pap-test
3 Neutrale
4 Probabilmente non effettuerò il pap-test
5 Sicuramente non effettuerò il pap-test

59. Con che frequenza ritieni di dover effettuare il pap-test?

1 Ogni anno
2 Ogni tre anni
3 Ogni cinque anni
4 Non so

60. In aggiunta al pap-test, che altro intendi fare in futuro per prevenire il tumore al collo dell'utero? (Puoi selezionare più risposte)

1 Visite dal ginecologo **(Prosegui con la domanda 61)**
2 Utilizzo contraccettivi in grado di prevenire l'infezione dal virus HPV **(Prosegui con la domanda 62)**
3 Altro(specificare_____) **(Prosegui con la domanda 62)**

61. Con che frequenza intendi fare le visite in futuro?

1 Ogni anno
2 Ogni tre anni
3 Con frequenza maggiore di tre anni

62. Il parere di quali persone potrebbe influenzare le tue scelte di prevenzione futura (esame pap-test) del tumore al collo dell'utero?

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli
4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)

63. Quali di queste persone sarebbero d'accordo con te se tu effettuassi il pap-test in futuro?

1 Ginecologo/medico di famiglia
2 Genitori
3 Sorelle, fratelli

4 Amiche/Amici
5 Operatori sanitari (ASL, consultorio, etc.)
6 Altro (specificare_____)

References

Abrahamse W, Steg L, Gifford R, Vlek C. Factors influencing car use for commuting and the intention to reduce it: A question of self-interest or morality? Transp Res Part F Traffic Psychol Behav. 2009;12(4):317-324. doi:10.1016/j.trf.2009.04.004.

Agresti A. An Introduction to Categorical Data Analysis. Vol 7. Second. John Wiley & Sons; 2008. doi:10.1002/pst.339.

Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process. 1991;50(2):179-211. doi:10.1016/0749-5978(91)90020-T.

Allen D, Coronado G, Williams R, et al. A Systematic Review of Measures Used in Studies of Human Papillomavirus (HPV) Vaccine Acceptability. Vaccine. 2010;28(24):4027-4037. doi:10.1016/j.vaccine.2010.03.0631.

Allen JD, Coronado GD, Williams RS, et al. A systematic review of measures used in studies of human papillomavirus (HPV) vaccine acceptability. Vaccine. 2010;28(24):4027-4037. doi:10.1016/j.vaccine.2010.03.063.

Asch SE. The doctrine of suggestion, prestige and imitation in social psychology. Psychol Rev. 1948;55(5):250-276. doi:10.1037/h0057270.

Ayers S, Baum A, McManus C, et al. Cambridge Handbook of Psychology , Health and Medicine. Second. Cambridge University Press; 2007.

Bandura A. Health Promotion by Social Cognitive Means. Heal Educ Behav. 2004;31(2):143-164. doi:10.1177/1090198104263660.

Bandura A. Social Cognitive Theory: An Agentic Perspective. Annu Rev Psychol. 2001;52(1):1-26. doi:10.1146/annurev.psych.52.1.1.

Bartholomew D., Knott M. Latent Variable Models and Factor Analysis.; 1999.

Berenson AB, Laz TH, Hirth JM, McGrath CJ, Rahman M. Effect of the decision-making process in the family on HPV vaccination rates among adolescents 9-17 years of age. Hum

Vaccines Immunother. 2014;10(7):1807-1811. doi:10.4161/hv.28779.

Berkowitz A D. Applications of social norms theory to other health and social justice issues. Soc norms approach to Prev Sch Coll age Subst Abus A Handb Educ Couns Clin. 2003;(March):259-279.

Bish A, Sutton S, Golombok S. Predicting uptake of a routine cervical smear test: A comparison of the health belief model and the theory of planned behaviour. Psychol Health. 2000;15(1):35-50. doi:10.1080/08870440008400287.

Bollen KA. Structural equation models that are nonlinear in latent variables: A least-squares estimator. Sociol Methodol. 1995:223-251. http://www.irss.unc.edu/content/pdf/Bollen (1995 Sm).pdf.

Bollen KA. Structural Equations with Latent Variables. Vol 35.; 1989. doi:10.1037/028856.

Bonanni P, Levi M, Latham NB, et al. An overview on the implementation of HPV vaccination in Europe. Hum Vaccin. 2014;7 Suppl(February):128-135. doi:10.4161/hv.7.0.14575.

Bowyer HL, Forster AS, Marlow LA V, Waller J. Predicting human papillomavirus vaccination behaviour among adolescent girls in England: results from a prospective survey. J Fam Plan Reprod Heal Care. 2014;40(1):14-22. doi:10.1136/jfprhc-2013-100583.

Breitkopf C, Pearson H. A Theory-Based Approach to Understanding Follow-up of Abnormal Pap Tests. J Health Psychol. 2009;14(3):361-371. doi:10.1177/1359105308101674.A.

Brick J, Kalton G. Handling missing data in survey research. Stat Methods Med Res. 1996;5(3):215-238. doi:10.1177/096228029600500302.

Britt RK, Hatten KN, Chappuis SO. Perceived behavioral control, intention to get vaccinated, and usage of online information about the human papillomavirus vaccine. Heal Psychol Behav Med. 2014;2(1):52-65. doi:10.1080/21642850.2013.869175.

Burd E. Human papillomavirus and cervical cancer. Clin Microbiol Rev. 2003;16(1):1-17. doi:10.1128/CMR.16.1.1.

Busse JW, Walji R, Wilson K. Parents' experiences discussing pediatric vaccination with healthcare providers: A survey of Canadian naturopathic patients. PLoS One. 2011;6(8):e22737. doi:10.1371/journal.pone.0022737.

Canada PHA of. Canadian Human PapillomavirusVaccine Research PrioritiesWorkshop.; 2006.

Cassel C, Hackl P, Westlund AH. Robustness of partial least-squares method for estimating latent variable quality structures. J Appl Stat. 1999;26(4):435-446. doi:10.1080/02664769922322.

Chatterjee A, Lahiri SN. Asymptotic properties of the residual bootstrap for lasso estimators. Proc Am Math Soc. 2010;138(12):4497-4509. doi:10.1090/S0002-9939-2010-10474-4.

Chin WW. Bootstrap cross-validation indices for PLS path model assessment. In: Handbook of Partial Least Squares. ; 2010:171-193. doi:10.1007/978-3-540-32827-8.

Chin WW. Commentary: Issues and Opinion on Structural Equation Modeling. MIS Q. 1998;22(1):1. doi:Editorial.

Chin WW. How to Write Up and Report PLS Analyses. In: Handbook of Partial Least Squares. ; 2010:655-690. doi:10.1007/978-3-540-32827-8_29.

Chin WW. The partial least squares approach to structural equation modeling. In: Modern Methods for Business Research. ; 1998:295-336. doi:10.1016/j.aap.2008.12.010.

Chirayil EI, Thompson CL, Burney S. Predicting Human Papilloma Virus Vaccination and Pap Smear Screening Intentions Among Young Singaporean Women Using the Theory of Planned Behavior. SAGE Open. 2014;4(4):1-10. doi:10.1177/2158244014554961.

Cialdini R, Trost M. Social influence: Social norms, conformity and compliance. Handb Soc Psychol Vol 2. 1998:151-192. doi:10.2307/2654253.

Cialdini RB, Goldstein NJ. Social Influence: Compliance and Conformity. Annu Rev Psychol. 2004;55(1):591-621. doi:10.1146/annurev.psych.55.090902.142015.

Cialdini RB, Kallgren CA, Reno RR. A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. Adv Exp Soc Psychol.

1991;24(C):201-234. doi:10.1016/S0065-2601(08)60330-5.

Cialdini RB, Reno RR, Kallgren CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. J Pers Soc Psychol. 1990;58(6):1015-1026. doi:10.1037/0022-3514.58.6.1015.

Coltmana T, Devinneyb TM, Midgleyc DF, Venaikd S. Formative versus reflective measurement models: Two applications of erroneous measurement. J Bus Re. 2008;61(12):1250-1262. doi:10.1016/j.jbusres.2008.01.013.

Coniglio MA, Platania M, Privitera D, Giammanco G, Pignato S. Parents' attitudes and behaviours towards recommended vaccinations in Sicily, Italy. BMC Public Health. 2011;11(1):305. doi:10.1186/1471-2458-11-305.

Conner M, Armitage CJ. Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. J Appl Soc Psychol. 1998;28(15):1429-1464. doi:10.1111/j.1559-1816.1998.tb01685.x.

Conner M. Health Behaviors. Int Encycl Soc Behav Sci. 2015;1(1996):582-587. doi:10.1016/B978-0-08-097086-8.14154-6.

Daniels M, Hogan J. Missing Data in Longitudinal Studies: Strategies for Bayesian Modeling and Sensitivity Analysis. Vol 51. Chapman & Hall/CRC; 2008. doi:10.1080/00107510903093086.

de Bekker-Grob EW, Hofman R, Donkers B, et al. Girls' preferences for HPV vaccination: A discrete choice experiment. Vaccine. 2010;28(41):6692-6697. doi:10.1016/j.vaccine.2010.08.001.

Dong Y, Peng C-YJ. Principled missing data methods for researchers. Springerplus. 2013;2(1):222. doi:10.1186/2193-1801-2-222.

Draper NR, Smith H. Applied Regression Analysis. Technometrics. 1998;47(3):706. doi:10.1198/tech.2005.s303.

Draper NR, Smith H. Applied Regression Analysis.; 1998.

Duffett-Leger L, Letourneau N, Croll J. Cervical Cancer Screening Practices among University Women. J Obs Gynecol Neonatal Nurs. 2008;37(5):572-581.

http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=ovftj&NEWS=N&AN=00001240-200809000-00008.

Enders CK. Applied Missing Data Analysis. The Guilford Press; 2010. doi:10.1017/CBO9781107415324.004.

European Centre for Disease Prevention and Control. Introduction of HPV Vaccines in European Union Countries – an Update.; 2012. doi:10.2900/60814.

Feiring B, Laake I, Molden T, et al. Do parental education and income matter? A nationwide register-based study on HPV vaccine uptake in the school-based immunisation programme in Norway. BMJ Open. 2015;5(5):e006422-e006422. doi:10.1136/bmjopen-2014-006422.

Fernandez AC, Paiva AL, Lipschitz JM, et al. Disease prevention without relapse: Processes of change for HPV vaccination. Open J Prev Med. 2013;3(3):301-309. doi:10.4236/ojpm.2013.33041.

Fernández de Casadevante V, Gil Cuesta J, Cantarero-Arévalo L. Determinants in the Uptake of the Human Papillomavirus Vaccine: A Systematic Review Based on European Studies. Front Oncol. 2015;5(June):141. doi:10.3389/fonc.2015.00141.

Ferrer HB, Trotter C, Hickman M, Audrey S. Barriers and facilitators to HPV vaccination of young women in high-income countries: a qualitative systematic review and evidence synthesis. BMC Public Health. 2014;14(1):700. doi:10.1186/1471-2458-14-700.

Fourn L, Haddad S, Fournier P, Gansey R. Determinants of parents' reticence toward vaccination in urban areas in Benin (West Africa). BMC Int Health Hum Rights. 2009;9 Suppl 1:S14. doi:10.1186/1472-698X-9-S1-S14.

Friedman J, Hastie T, Tibshirani R. Applications of the lasso and grouped lasso to the estimation of sparse graphical models. Tech Rep. 2010:1-22.

Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010;33(1). doi:10.18637/jss.v033.i01.

Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. Biostatistics. 2008;9(3):432-441. doi:10.1093/biostatistics/kxm045.

Fues H. Comparison of logistic and ridge regression in genetic association studies Matematiska institutionen. 2015.

Gamble HL, Klosky JL, Parra GR, Randolph ME. Factors influencing familial decision-making regarding human papillomavirus vaccination. J Pediatr Psychol. 2010;35(7):704-715. doi:10.1093/jpepsy/jsp108.

Gefenaite G, Smit M, Nijman HW, et al. Comparatively low attendance during Human Papillomavirus catch-up vaccination among teenage girls in the Netherlands: Insights from a behavioral survey among parents. BMC Public Health. 2012;12(1):498. doi:10.1186/1471-2458-12-498.

Geoff L, Ryan G, Hefny SA, Sutherland D, Wang G. Lecture 2 : August 30. 2012:1-8.

Gerend MA, Shepherd JE. Predicting human papillomavirus vaccine uptake in young adult women: Comparing the health belief model and theory of planned behavior. Ann Behav Med. 2012;44(2):171-180. doi:10.1007/s12160-012-9366-5.

Gerend MA, Shepherd JE. Using message framing to promote acceptance of the human papillomavirus vaccine. Heal Psychol. 2007;26(6):745-752. doi:10.1037/0278-6133.26.6.745.

Geshnizjani A, Jozkowski KN, Middlestadt SE. Factors Influencing the Intention of Getting the HPV Vaccine among College Women : An Application of the Reasoned Action Approach. 2013;11(2):1-11.

Gilbert NL, Gilmour H, Dubé È, Wilson SE, Laroche J. Estimates and determinants of HPV non-vaccination and vaccine refusal in girls 12 to 14 y of age in Canada: Results from the Childhood National Immunization Coverage Survey, 2013. Hum Vaccines Immunother. 2016;12(6):1484-1490. doi:10.1080/21645515.2016.1153207.

Giorgi Rossi P, Ricciardi A, Cohet C, et al. Epidemiology and costs of cervical cancer screening and cervical dysplasia in Italy. BMC Public Health. 2009;9:71. doi:10.1186/1471-2458-9-71.

Giraldi G, Martinoli L, de Luca d'Alessandro E. The human papillomavirus vaccination: A review of the cost-effectiveness studies. Clin Ter. 2014;165(6):e426-e432. doi:10.7417/CT.2014.1787.

Göbl CS, Bozkurt L, Tura A, Pacini G, Kautzky-Willer A, Mittlböck M. Application of penalized regression techniques in modelling insulin sensitivity by correlated metabolic parameters. PLoS One. 2015;10(11):1-19. doi:10.1371/journal.pone.0141524.

Goeman AJ, Meijer R, Chaturvedi N, Lueder M, Rcpp I, Rcpp L. Package " penalized ." 2017;1. https://cran.r-project.org/web/packages/penalized/penalized.pdf.

Gonçalves AK, Cobucci RN, Rodrigues HM, de Melo AG, Giraldo PC. Safety, tolerability and side effects of human papillomavirus vaccines: a systematic quantitative review. Brazilian J Infect Dis. 2014;18(6):651-659. doi:10.1016/j.bjid.2014.02.005.

Gottvall M. Introduction of School-Based HPV Vaccination in Sweden.; 2014.

Gowda C, Dempsey AF. The rise (and fall?) of parental vaccine hesitancy. Hum Vaccines Immunother. 2013;9(8):1755-1762. doi:10.4161/hv.25085.

Grant CC, Turner NM, York DG, Goodyear-Smith F, Petousis-Harris HA. Factors associated with immunisation coverage and timeliness in New Zealand. Br J Gen Pract. 2010;60(572):180-186. doi:10.3399/bjgp10X483535.

Gunes F. Penalized Regression Methods for Linear Models in SAS/STAT ®. :1-14.

Guvenc G, Akyuz A, A??ikel CH. Health Belief Model Scale for Cervical Cancer and Pap Smear Test: Psychometric testing. J Adv Nurs. 2011;67(2):428-437. doi:10.1111/j.1365-2648.2010.05450.x.

Haesebaert J, Lutringer-Magnin D, Kalecinski J, et al. French women's knowledge of and attitudes towards cervical cancer prevention and the acceptability of HPV vaccination among those with 14 – 18 year old daughters: a quantitative-qualitative study. BMC Public Health. 2012;12(1):1034. doi:10.1186/1471-2458-12-1034.

Hair JFJ, Hult GTM, Ringle C, Sarstedt M. A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Vol 46.; 2014. doi:10.1016/j.lrp.2013.01.002.

Han J, Kamber M. Data Mining: Concepts and Techniques. Vol 54. (Elsevier, ed.).; 2006. doi:10.1007/978-3-642-19721-5.

Hanley SJB, Yoshioka E, Ito Y, et al. An exploratory study of Japanese fathers' knowledge of and attitudes towards HPV and HPV vaccination: Does marital status matter? Asian Pacific J Cancer Prev. 2014;15(4):1837-1843. doi:10.7314/APJCP.2014.15.4.1837.

Hansen BT, Kjær SK, Arnheim-Dahlström L, et al. Human papillomavirus (HPV) vaccination and subsequent sexual behaviour: Evidence from a large survey of Nordic women. Vaccine. 2014;32(39):4945-4953. doi:10.1016/j.vaccine.2014.07.025.

Harmsen I a., Ruiter R a. C, Paulussen TGW, Mollema L, Kok G, de Melker HE. Factors That Influence Vaccination Decision-Making by Parents Who Visit an Anthroposophical Child Welfare Center: A Focus Group Study. Adv Prev Med. 2012;2012:1-7. doi:10.1155/2012/175694.

Harrell F. Package "rms" Title Regression Modeling Strategies. 2017. http://biostat.mc.vanderbilt.edu/rms.

Harrell F. Regression Modeling Strategies. Version 5.1-0. Inhttp://biostat.mc.vanderbilt.edu/rm.

Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3(2):143-152. doi:10.1002/sim.4780030207.

Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treat Rep. 1985;69(10):1071-1077. http://www.ncbi.nlm.nih.gov/pubmed/4042087.

Harrell FE. Ordinal Logistic Regression. Springer Ser Stat. 2015:311-325. doi:10.1007/978-3-319-19425-7_13.

Harrell FE. Regression Modeling Strategies. With Applications to Linear Models, Logistic Regression, and Survival Analysis.; 2001. doi:10.1007/978-1-4757-3462-1.

Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics) (9780387848570): Trevor Hastie, Robert Tibshirani, Jerome Friedman: Books. In: The Elements of Statistical Learning: Data Mining, Inference, and Prediction; 2011:501-520.

Hastie T, Tibshirani R, Wainwright M. Statistical Learning with Sparsity: The Lasso and Generalizations. Crc. 2015:362. doi:10.1201/b18401-1.

Hastie T, Tibsshirani R, Friedman J. The Elements of Statistical Learning. Springer; 2009. doi:10.1007/b94608.

Henseler J, Hubona G, Ray PA. Using PLS path modeling in new technology research: updated guidelines. Ind Manag Data Syst. 2016;116(1):2-20. doi:10.1108/IMDS-09-2015-0382.

Héquet D, Rouzier R. Determinants of geographic inequalities in HPV vaccination in the most populated region of France. PLoS One. 2017;12(3):1-12. doi:10.1371/journal.pone.0172906.

Herald zur, Hausen, Francoise B-S. The discoveries of human papilloma viruses that cause cervical cancer and of human immunodeficiency virus. Nobel Prize. 2008;(8):1-26.

Hosmer DW, Lemeshow S. Applied logistic regression. Wiley Ser Probab Stat. 2000:373. doi:10.1198/tech.2002.s650.

ICO HPV Information Centre. Human Papillomavirus and Related Diseases.; 2017. www.hpvcentre.com.

Jopson A. The Role of Norms in Mode Choice. In: 3rd ICTTP Conference. ; 2004. doi:10.1039/c3ib40166a.

Kaplan D. Structural Equation Modeling.; 2000. doi:10.1300/J073v21n04_05.

Kaplan D. Structural Models for Categorical and Continuous Latent Variables. Struct Equ Model Found Extensions. 2009:181-206. doi:10.4135/9781452226576.

Karni E. Savages' Subjective Expected Utility Model. Retrieved March. 2005:1-12. doi:10.1057/9780230226203.1474.

Kim HW. Awareness of Pap testing and factors associated with intent to undergo Pap testing by level of sexual experience in unmarried university students in Korea: results from an online survey. BMC Womens Health. 2014;14(1):100. doi:10.1186/1472-6874-14-100.

Kim JJ, Andres-Beck B, Goldie SJ. The value of including boys in an HPV vaccination programme: a cost-effectiveness analysis in a low-resource setting. Br J Cancer. 2007;97(9):1322-1328. doi:10.1038/sj.bjc.6604023.

Kornfeld J, Byrne MM, Vanderpool R, Shin S, Kobetz E. HPV knowledge and vaccine acceptability among hispanic fathers. J Prim Prev. 2013;34(1-2):59-69. doi:10.1007/s10935-013-0297-0.

Kumar VM, Whynes DK. Explaining variation in the uptake of HPV vaccination in England. BMC Public Health. 2011;11(1):172. doi:10.1186/1471-2458-11-172.

KWAN T. Human Papillomavirus Testing in Cervical Cancer Screening : Potential Harms and Implications for Intervention Tracy Tak Ching KWAN. 2011.

Latan H, Ramli NA. The Results of Partial Least Squares-Structural Equation Modelling Analyses (PLS-SEM). SSRN Electron J. 2013;(January). doi:10.2139/ssrn.2364191.

Lavanchy G, Strehler M, Llanos Roman MN, et al. Handbook of Partial Least Squares. Vol 70.; 2010. doi:10.1007/978-3-540-32827-8.

Lechuga J, Swain GR, Weinhardt LS. The cross-cultural variation of predictors of human papillomavirus vaccination intentions. J Womens Health (Larchmt). 2011;20(2):225-230. doi:10.1089/jwh.2010.1993.

Little R, Rubin M. Statistical Analysis with Missing Data. John Wiley & Sons; 1987.
Lockhart R, Taylor J, Tibshirani RJ, Tibshirani R. A significance test for the lasso. Ann Stat. 2014;42(2):413-468. doi:10.1214/13-AOS1175.

Lowry PB, Gaskin J. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. IEEE Trans Prof Commun. 2014;57(2):123-146. doi:10.1109/TPC.2014.2312452.

Luca DL, Lubet AS, Onarheim KH, et al. Benefits and Costs of the Health Targets for the Post-2015 Development Agenda. Vol 19 January.; 2015.

Ma GX, Gao W, Fang CY, et al. Health beliefs associated with cervical cancer screening among Vietnamese Americans. J Womens Health (Larchmt). 2013;22(3):276-288. doi:10.1089/jwh.2012.3587.

Mackie G, Moneti F, Shakya H, Denny E. What Are Social Norms? How Are They Measured?; 2015.

Manhart LE, Burgess-Hull AJ, Fleming CB, Bailey JA, Haggerty KP, Catalano RF. HPV vaccination among a community sample of young adult women. Vaccine. 2011;29(32):5238-5244. doi:10.1016/j.vaccine.2011.05.024.

McElroy A. BioculturalModelsinStudiesofHumanHealth.pdf. Med Anthropol Q. 1990;4(3):243-265.

McRee A-L, Brewer NT, Reiter PL, Gottlieb SL, Smith JS. The Carolina HPV Immunization Attitudes and Beliefs Scale (CHIAS): Scale Development and Associations With Intentions to Vaccinate. Sex Transm Dis. 2009;37(4):1. doi:10.1097/OLQ.0b013e3181c37e15.

Mencarini L, Vignoli D, Gottard A. Fertility intentions and outcomes. Implementing the Theory of Planned Behavior with graphical models. Adv Life Course Res. 2015;23:14-28. doi:10.1016/j.alcr.2014.12.004.

Molenberghs G, Verbeke G. Models for Discrete Longitudinal Data. Vol 20.; 2005. doi:10.1007/0-387-28980-1.

Monecke, A. and Leisch F. semPLS : Structural Equation Modeling Using Partial Least Squares. J Stat Softw. 2012;48(3):1-32. doi:10.1108/EBR-10-2013-0128.

Moons KGM, Donders ART, Steyerberg EW, Harrell FE. Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: A clinical example. J Clin Epidemiol. 2004;57(12):1262-1270. doi:10.1016/j.jclinepi.2004.01.020.

Morris J, Marzano M, Dandy N, O'Brien L. Theories and models of behaviour and behaviour change. For Sustain Behav Behav Chang Theor. 2012:1-27.

MOSCOVICI S, Moscovici S. Experiment and Experience: An Intermediate Step from Sherif to

Asch. J Theory Soc Behav. 1991;21(3):253-268. doi:10.1111/j.1468-5914.1991.tb00197.
Muller TR. Understanding Young Women' s Cervical Cancer Screening Attendance. 2016;(November).

Munro S, Lewin S, Swart T, Volmink J. A review of health behaviour theories: how useful are these for developing interventions to promote long-term medication adherence for TB and HIV/AIDS? BMC Public Health. 2007;7(1):104. doi:10.1186/1471-2458-7-104.

Navarro-Illana P, Diez-Domingo J, Navarro-Illana E, Tuells J, Alemán S, Puig-Barberá J. "Knowledge and attitudes of Spanish adolescent girls towards human papillomavirus infection: where to intervene to improve vaccination coverage." BMC Public Health. 2014;14(1):490. doi:10.1186/1471-2458-14-490.

Nchise AC. An Empirical Analysis of the Theory of Planned Behavior: A Review of Its Application on E-democracy Adoption Using the Partial Least Squares Algorithm. JeDEM - eJournal eDemocracy Open Gov. 2012;4(2):171-182. http://www.jedem.org/article/view/129.

Noar SM, Zimmerman RS. Health Behavior Theory and cumulative knowledge regarding health behaviors: Are we moving in the right direction? Health Educ Res. 2005;20(3):275-290. doi:10.1093/her/cyg113.

O'Connor AM, Tugwell P, Wells GA, et al. A decision aid for women considering
Ogilvie G, Anderson M, Marra F, et al. A population-based evaluation of a publicly funded, school-based HPV vaccine program in British Columbia, Canada: Parental factors associated with HPV vaccine receipt. PLoS Med. 2010;7(5). doi:10.1371/journal.pmed.1000270.

Ogilvie GS, Remple VP, Marra F, et al. Parental intention to have daughters receive the human papillomavirus vaccine. Cmaj. 2007;177(12):1506-1512. doi:10.1503/cmaj.071022.

Opel DJ, Taylor JA, Mangione-Smith R, et al. Validity and reliability of a survey to identify vaccine-hesitant parents. Vaccine. 2011;29(38):6598-6605. doi:10.1016/j.vaccine.2011.06.115.

Ostini R, Nering ML. Polytomous IRT models (# 144, Remo Ostini and Michael L. Nering) Jorge Tendeiro 16 April 2014. 2014;(April).

Panatto D, Amicizia D, Tanzi E, et al. Prevalence of human papillomavirus in young Italian

women with normal cytology: how should we adapt the national vaccination policy? BMC Infect Dis. 2013;13:575. doi:10.1186/1471-2334-13-575.

Patel PR, Berenson AB. Sources of HPV vaccine hesitancy in parents. Hum Vaccines Immunother. 2013;9(12):2649-2653. doi:10.4161/hv.26224.

Perez S, Tatar O, Shapiro GK, et al. Psychosocial determinants of parental human papillomavirus (HPV) vaccine decision-making for sons: Methodological challenges and initial results of a pan-Canadian longitudinal study. BMC Public Health. 2016;16(1):1223. doi:10.1186/s12889-016-3828-9.

Perez S. Parents' human papillomavirus vaccine decision-making: Theory, measurement and models. 2016;(December). http://digitool.library.mcgill.ca/webclient/DeliveryManager?pid=146789&custom_att_2=direct.

Perkins RB, Clark JA, Apte G, et al. Missed Opportunities for HPV Vaccination in Adolescent Girls: A Qualitative Study. Pediatrics. 2014;134(3):e666-e674. doi:10.1542/peds.2014-0442.

Perkins RB, Mehta PK, Langrish SM. Fathers' intentions to accept human papillomavirus vaccination for sons and daughters: Exploratory findings from rural Honduras. Int J Public Health. 2012;57(1):143-148. doi:10.1007/s00038-011-0271-7.

Pokropek A. Missing by Design: Planned Missing-Data Designs in Social Science. Ask Res Methods. 2011;20(January 2011):81-105.

Prochaska JO, Velicer WF, Rossi JS, et al. Stages of change and decisional balance for 12 problem behaviors. Heal Psychol. 1994;13(1):39-46. doi:10.1037/0278-6133.13.1.39.

Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Heal Promot. 1997;12(1):38-48. doi:10.4278/0890-1171-12.1.38.

Ratanasiripong NT, Chai KT. A Concept Analysis of Attitude toward Getting Vaccinated against Human Papillomavirus. Nurs Res Pract. 2013;2013. doi:10.1155/2013/373805.

Reiter PL, Brewer NT, Gottlieb SL, McRee AL, Smith JS. Parents' health beliefs and HPV vaccination of their adolescent daughters. Soc Sci Med. 2009;69(3):475-480.

doi:10.1016/j.socscimed.2009.05.024.

Reno RR, Cialdini RB, Kallgren CA. The transsituational influence of social norms. J Pers Soc Psychol. 1993;64(1):104-112. doi:10.1037//0022-3514.64.1.104.

Reyna VF, Farley F. Risk and rationality in adolescent decision making: Implications for theory, practice, and public policy. Psychol Sci Public Interes Suppl. 2006;7(1):1-44. doi:10.1111/j.1529-1006.2006.00026.

Roncancio AM, Ward KK, Fernandez ME. Understanding Cervical Cancer Screening Intentions Among Latinas Using An Expanded Theory of Planned Behavior Model. Behav Med. 2013;39(3):66-72. doi:10.1080/08964289.2013.799452.

Rosenstock IM, Strecher VJ, Becker MH. Social Learning Theory and the Health Belief Model. Health Educ Q. 1988;15(2):175-183. doi:10.1177/109019818801500203.

Russolillo G. Partial Least Squares Methods. 2009. doi:doi:10.1201/9781584887706.ch1.

Sampson R, Wong L, MacVicar R. Parental reasons for non-uptake of influenza vaccination in young at-risk groups: A qualitative study. Br J Gen Pract. 2011;61(588):e386-e391. doi:10.3399/bjgp11X583155.

Sanchez G. PLS Path Modeling with R.; 2013. doi:citeulike-article-id:13341888.
Sartori S. Penalized Regression : bootstrap confidence intervals and variable selection for high dimensional data sets . 2010.

Saulle R, Miccoli S, Unim B, et al. Validation of a questionnaire for young women to assess knowledge, attitudes and behaviors towards cervical screening and vaccination against HPV in Italy. Epidemiol Biostat Public Heal. 2014;11(2):1-12. doi:10.2427/8913.

Saunders KA. Factors that influence HPV vaccinations in young women residing in Canada. ProQuest Diss Theses. 2013. http://search.proquest.com/docview/1294305230?accountid=12339%5Cnhttp://mcgill.on.worldcat.org/atoztitles/link?sid=ProQ:&issn=&volume=&issue=&title=Factors+that+influence+HPV+vaccinations+in+young+women+residing+in+Canada&spage=&date=2013-01-01&atitle=Fac.

Schaefer K. Beliefs and attitudes regarding HPV vaccination among college-age women: an application of the Health Belief Model. 2010. http://drum.lib.umd.edu//handle/1903/11235.

Schmitt H-J, Booy R, Aston R, et al. How to optimise the coverage rate of infant and adult immunisations in Europe. BMC Med. 2007;5(1):11. doi:10.1186/1741-7015-5-11.

Scorcia-Wilson T. The relationship between knowledge and beliefs about human papillomavirus, acceptance of the human papillomavirus vaccine, and intentions to practice safer sex behaviors among female college students. Diss Abstr Int Sect B Sci Eng. 2011;72(1-B):219.

Sherif M. The psychology of social impact. Am Psychol. 1936;36(4):343-356. doi:10.1037/0003-066X.36.4.343.

Sherif M. The Psychology of Social Norms. Vol 41.; 1936. doi:10.1111/j.1468-5914.2011.00472.x.

Smith PJ, Humiston SG, Marcuse EK, et al. Parental delay or refusal of vaccine doses, childhood vaccination coverage at 24 months of age, and the Health Belief Model. Public Health Rep. 2011;126 Suppl:135-146. doi:10.1177/00333549111260S215.

SNIPES D. Bridging the gap: Using the theory of planned behavior to predict HPV vaccination intentions in men. 2013. https://dizzyg.library.vcu.edu/handle/10156/4147.
Stafne T. Human Papilloma Virus Awareness , Knowledge and Vaccine Acceptance among Norwegian Adolescents. 2014.

Steele F, Skrondal A, Rabe-Hesketh S. Generalized Latent Variable Models: Multilevel, Longitudinal, and Structural Equation Models. Vol 17. Chapman & Hall/CRC; 2004. doi:10.1007/BF02295939.

Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models : A framework for some traditional and novel measures. Epidemiology. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2.Assessing.

Stockwell MS, Rosenthal SL, Sturm LA, Mays RM, Bair RM, Zimet GD. The effects of vaccine characteristics on adult women's attitudes about vaccination: A conjoint analysis study. Vaccine. 2011;29(27):4507-4511. doi:10.1016/j.vaccine.2011.04.031.

Sutton S. Health Behavior, Psychosocial Theories. Int Encycl Soc Behav Sci. 2015:577-581. doi:10.1016/B978-0-08-097086-8.14153-4.

Taylor D, Bury M, Campling N, et al. A Review of the use of the Health Belief Model (HBM), the Theory of Reasoned Action (TRA), the Theory of Planned Behaviour (TPB) and the Trans-Theoretical Model (TTM) to study and predict health related behaviour change. Heal (San Fr. 2006;(June):1-215.

Taylor VM, Yasui Y, Burke N, et al. Pap testing adherence among Vietnamese American women. Cancer Epidemiol Biomarkers Prev. 2004;13(4):613-619.

Tenenhaus M, Vinzi VE, Chatelin YM, Lauro C. PLS path modeling. Comput Stat Data Anal. 2005;48(1):159-205. doi:10.1016/j.csda.2004.03.005.

Tenenhaus M. PLS path modeling. Comput Stat Data Anal. 2005;48(1):159-205. doi:10.1016/j.csda.2004.03.005.

The Lancet. HPV vaccination: a decade on. Lancet. 2016;388(10043):438. doi:10.1016/S0140-6736(16)31206-5.

Thurstone LL. Influence of Motion Pictures on Children's Attitudes. J Soc Psychol. 1931;2(3):291-305. doi:10.1080/00224545.1931.9918977.

Trinchera L, Russolillo G. Role and treatment of categorical variables in PLS Path Models for Composite Indicators. Proc 6th …. 2009:1-10. http://hbanaszak.mjr.uw.edu.pl/TempTxt/TrincheraRussolillo_2006_Role and Ttreatment of Categorical Variables in PLS Path Models for Compoisite Indicators.pdf.

Vaccaro CM, Manacorda T, Coletta V. Italian women and HPV prevention . Knowledge , fears , uncertainty on Human Papillomavirus and the relative vaccination : dual research approach. 2012;9(1):55-63.

van Buuren S. Flexible Imputation of Missing Data. Chapman & Hall/CRC; 2012.

van Houwelingen JC. Shrinkage and Penalized Likelihood as Methods to Improve Predictive

Accuracy. Stat Neerl. 2001;55(1):17-34. doi:10.1111/1467-9574.00154.

van Keulen HM, Otten W, Ruiter RA, et al. Determinants of HPV vaccination intentions among Dutch girls and their mothers: a cross-sectional study. BMC Public Health. 2013;13(1):111. doi:10.1186/1471-2458-13-111.

Velan B, Boyko V, Lerner-Geva L, Ziv A, Yagar Y, Kaplan G. Individualism, acceptance and differentiation as attitude traits in the public's response to vaccination. Hum Vaccines Immunother. 2012;8(9):1272-1282. doi:10.4161/hv.21183.

Vinzi E, Trinchera L, Amato S. Chapter 2 PLS Path Modeling: From Foundations to Recent developments and Open Issues for Model Assessment and Improvement. In: Handbook of Partial Least Squares, Springer Handbooks of Computational Statistics. ; 2010:47-82.

Vinzi VE, Ringle CM, Squillacciotti S, Trinchera L. Capturing and Treating Unobserved Heterogeneity by Response Based Segmentation in PLS Path Modeling. A Comparison of Alternative Methods by Computational Experiments. SSRN Electron J. 2007;(JuLY):1-24. doi:10.2139/ssrn.2394044.

Vinzi VE, Trinchera L, Amato S. Handbook of Partial Least Squares.; 2010. doi:10.1007/978-3-540-32827-8.

Vinzi VE, Trinchera L, Squillacciotti S, Tenenhaus M. REBUS-PLS: A response-based procedure for detecting unit segments in PLS path modelling. Appl Stoch Model Bus Ind. 2008;24:439-458. doi:10.1002/asmb.

Viscusi WK. Alarmist Decisions with Divergent Risk Information. Econ J. 1997;107(445):1657-1670. doi:10.1111/1468-0297.00248.

Walker CL. Attitudes , Practices , and Beliefs About Human Papillomavirus Vaccine Among Young Adult African-American Women : Implications for Effective Implementation By Chapel Hill. 2009;(August). http://dc.lib.unc.edu/cdm/ref/collection/etd/id/2849.

Wei F, Mullooly JP, Goodman M, et al. Identification and characteristics of vaccine refusers. BMC Pediatr. 2009;9(1):18. doi:10.1186/1471-2431-9-18.

Weinstein ND. Testing four competing theories of health-protective behavior. Heal Psychol. 1993;12(4):324-333. doi:10.1037/0278-6133.12.4.324.

Weinstein ND. The precaution adoption process. Heal Psychol. 1988;7(4):355-386. doi:10.1037/0278-6133.7.4.355.

White KM, Hogg MA, Terry DJ. Improving Attitude-Behavior Correspondence Through Exposure to Normative Support From a Salient Ingroup. Basic Appl Soc Psych. 2002;24(2):91-103. doi:10.1207/S15324834BASP2402_2.

WHO. Human papillomavirus vaccines: WHO position paper, October 2014. World Heal Organ Wkly Epidemiol Rec. 2014;89(43):465-492. doi:10.1186/1750-9378-2-15.Voir.
Winters U, Roden R, Kitchener H, Stern P. Progress in the development of a cervical cancer vaccine. Ther Clin Risk Manag. 2006;2(3):259-269. doi:10.2147/tcrm.2006.2.3.259.

Wong LP, Wong YL, Low WY, Khoo EM, Shuib R. Knowledge and awareness of cervical cancer and screening among Malaysian women who have never had a Pap smear: A qualitative study. Singapore Med J. 2009;50(1):49-53. doi:10.1080/10705500802365490.

Wray-lake L, Crouter AC, Mchale SM. Developmental Patterns in Decision-Making Autonomy across Middle Childhood and Adolescence: European American Parents' Perspectives. Child Dev. 2011;81(2):636-651. doi:10.1111/j.1467-8624.2009.01420.x.Developmental.

Yuan Y. Sensitivity Analysis in Multiple Imputation for Missing Data. SAS Inst Inc. 2014:1-12.

Zadkarami M. Bootsrapping: A Nonparametric Approach to Identify the Effect of Sparsity of Data in the Binary Regression Models. J Appl Sci. 2008;8(17):2991-2997.

Zhuang QY, Wong RX, Chen WMD, Guo XX. Knowledge, attitudes and practices regarding human papillomavirus vaccination among young women attending a tertiary institution in Singapore. Singapore Med J. 2016;57(6):329-333. doi:10.11622/smedj.2016108.

Zimet GD, Weiss TW, Rosenthal SL, Good MB, Vichnin MD. Reasons for non-vaccination against HPV and future vaccination intentions among 19-26 year-old women. BMC Womens Health. 2010;10(1):27. doi:10.1186/1472-6874-10-27.

Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Ser B Stat Methodol. 2005;67(2):301-320. doi:10.1111/j.1467-9868.2005.00503.x.

zur Hausen H. Papillomaviruses in the causation of human cancers - a brief historical account. Virology. 2009;384(2):260-265. doi:10.1016/j.virol.2008.11.046.