Basic principles of test-negative design in evaluating influenza vaccine effectiveness

CrossMark

Wakaba Fukushima [a,b,*], Yoshio Hirota [c,d]

[a] Department of Public Health, Osaka City University, Graduate School of Medicine, Osaka, Japan
[b] Research Center for Infectious Disease Sciences, Osaka City University, Graduate School of Medicine, Osaka, Japan
[c] College of Healthcare Management, Miyama, Japan
[d] Clinical Epidemiology Research Center, Medical Co. LTA, Fukuoka, Japan

ARTICLE INFO

ABSTRACT

Based on the unique characteristics of influenza, the concept of "monitoring" influenza vaccine effectiveness (VE) across the seasons using the same observational study design has been developed. In recent years, there has been a growing number of influenza VE reports using the test-negative design, which can minimize both misclassification of diseases and confounding by health care-seeking behavior. Although the test-negative designs offer considerable advantages, there are some concerns that widespread use of the test-negative design without knowledge of the basic principles of epidemiology could produce invalid findings. In this article, we briefly review the basic concepts of the test-negative design with respect to classic study design such as cohort studies or case-control studies. We also mention selection bias, which may be of concern in some countries where rapid diagnostic testing is frequently used in routine clinical practices, as in Japan.

© 2017 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

It is widely accepted that the best study design for obtaining conclusive findings on prophylactic or therapeutic effects in human population is the randomized controlled trial (RCT). Such a concept can be also applied in assessing efficacy/effectiveness for almost all vaccines. With regard to the influenza vaccines, however, even a large and well-conducted RCT would simply provide a time-, place-, and subject-specific observation because: (1) epidemic strains of influenza differ by time and place; (2) the proportion of those having pre-existing antibody titers differ by time, place and age group; (3) vaccine strains differ by time (i.e., season) [1]. Together with the ethical consideration that influenza vaccination is recommended for wide-ranging high risk groups [2], the concept of "monitoring" the influenza vaccine effectiveness (VE) across the seasons using the same observational study design has been developed.

During the last decade, a test-negative design, which is a modified case-control study, has been introduced to assess VE against influenza. The design enables us to estimate VE in the early, mid, and end of the influenza season in a timely manner. Several countries including the US [3], Canada [4], Europe [5], Australia [6] and New Zealand [7] have applied the method for monitoring the annual VE. Because the test-negative design is practically easier to conduct than other study designs, a growing number of reports have been recently published. However, there are some concerns that widespread use of the test-negative design without knowledge of the basic principles of epidemiology would introduce invalid findings. In this article, we briefly review the basic concepts of the test-negative design with respect to classic study design such as cohort studies or case-control studies. We also discuss selection bias, which may be introduced when results from clinician-ordered laboratory testing is used as an outcome measure. This may be particularly of concern in some countries, including Japan, where rapid diagnostic testing for influenza is frequently used in routine clinical practice.

2. Rationale for applying the test-negative design in evaluating influenza VE

At present, the test-negative design seems to be very useful in evaluating VE against influenza. Using laboratory-confirmed influenza as an outcome measure, we can reduce disease misclassification. Furthermore, the design enable us to minimize confounding due to health care-seeking behavior. For a better understanding

of the latter advantage, the basic principles in cohort studies should be referred.

In cohort studies, both vaccinees and non-vaccinees should be followed-up with "equal intensity" to identify the occurrence of the outcome [8,9]. If influenza-like illness (ILI) is used as an outcome measure, equal intensity of follow-up would be achieved via telephone or questionnaire survey for all subjects on a weekly or monthly basis to obtain information on onset of the disease (i.e., active surveillance) [10–12]. In contrast, when using outcome of laboratory-confirmed influenza, a more strictly defined outcome, there is a concern that bias due to health care-seeking behavior becomes an issue because: (1) the outcome is usually confirmed only after the subjects visit medical institutions due to symptoms (i.e., passive surveillance); (2) vaccinees and non-vaccinees are inherently different in the likelihood of a medical visit (Fig. 1). Given these issues relating to health care-seeking behavior, the basic principle of following the vaccinees and non-vaccinees with equal intensity is difficult to satisfy when laboratory-confirmed influenza is used as an outcome measure in cohort studies. It is still possible to comply with the principle, as noted in a previous RCT among children [13]. In that study, the investigators contacted all subjects on a weekly basis to obtain the information on ILI onset, and once they confirmed that a subject had developed ILI, they attempted to collect his/her respiratory specimens within a couple of days. Obviously, such procedures require significant efforts and costs. Other exceptions may include a VE study based on antibody efficacy, in which all subjects received vaccine and medical visits for respiratory illnesses were compared between those with and without protective level of hemagglutination inhibition titer [14]. As subjects were not aware of their post-vaccination antibody level, the distortion due to health care-seeking behavior would be non-differential. Although antibody efficacy is expected to be an accurate index of VE [15], the estimates are strain-specific and interpretation of the results is sometimes complicated. Thus, it is considered a reasonable alternative for researchers to accept ILI as an outcome measure in cohort studies, which ensures achievement of equal intensity of follow-up resulting in higher feasibility and validity [10–12].

The test-negative design has a notable strength in controlling for afore-mentioned health care-seeking behavior (Fig. 2). Typically, study subjects are patients who visit medical institutions due to ILI during the influenza season. Subjects with positive test results for influenza are classified into cases, while subjects with negative results are classified as controls, and then vaccination status during the season can be compared between cases and controls. As the subjects are likely to visit a medical institution soon after ILI onset, both cases and controls are considered to be similar in their health care-seeking behavior. Therefore, the test-negative design can minimize confounding by health care-seeking behavior in evaluating influenza VE even though the outcome measure is laboratory-confirmed influenza, which is expected to resolve the dilemma in cohort studies.

Some articles have discussed the theoretical issues of the test-negative design [16–19]. VE against influenza is supposed to be the same in those who do seek care for ILI and who do not [17], although the test-negative design is limited by visitor attendance at the medical institution. An important factor relating to seeking of care may be the disease severity because disease severity is also expected to be associated with vaccination status. For example, it is possible that non-vaccinees are likely to develop severe ILI once
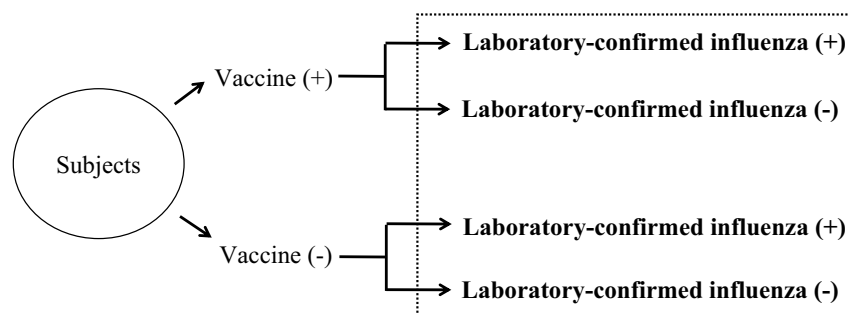
Fig. 1. Design of a cohort study to evaluate influenza vaccine effectiveness against laboratory-confirmed influenza. "Health care-seeking behavior" can introduce bias because (1) the outcome is usually confirmed only after the subjects visit medical institutions and (2) vaccinees and non-vaccinees are inherently different in the likelihood of their medical visit.
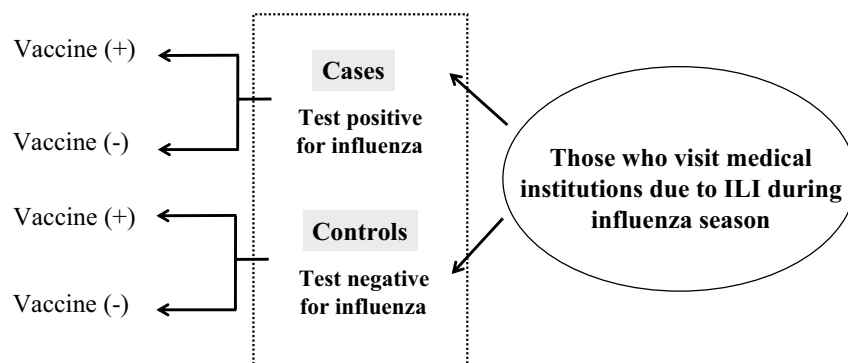
Fig. 2. A test-negative design to evaluate influenza vaccine effectiveness against laboratory-confirmed influenza. ILI denotes influenza-like illness. The test-negative design can minimize confounding by health care-seeking behavior even though the outcome measure is laboratory-confirmed influenza because "health care-seeking behavior" is likely to be similar between cases and controls.

they get infected by influenza, and those with severe ILI are likely to seek care. Thus, an appropriate adjustment for disease severity in analyses will be required to obtain a valid VE estimate [18].

3. Several principles must be satisfied in controls in the test-negative design

In the test-negative design, researchers are not aware of subjects' case/control status at recruitment, but later classify the subjects into cases or controls according to the test results. However, they should satisfy the same basic principles as for the classic case-control studies.

First, controls should be drawn from a source population, which generates the cases (i.e., study base principle). This condition may be inherently met in the test-negative design because both cases and controls are subjects who visited the same institution due to ILI.

The second principle is that both cases and controls are likely to have the same extent of experience in their exposure to influenza virus (i.e., a necessary cause in disease etiology). Recruitment of cases and controls when influenza is not circulating should be avoided, which translates to avoiding recruitment of the subjects who were not at risk of the disease in cohort studies. This is straightforward because case-control studies provide findings that mirror what could be learned from cohort studies [20].

Finally, controls should be selected independently of the exposure status. In test-negative design assessing influenza VE, the risk of non-influenza ILI that places the subjects into controls should be independent of influenza vaccination status. Controls in the test-negative design potentially consist of two types of ILI patients: negative for influenza per se but positive for other respiratory virus (other respiratory virus [ORV] positive controls), and negative for all respiratory virus tested (pan-negative controls). Recently, an argument regarding "appropriate controls" has been discussed.

The issue was pointed out for the first time in a study from Australia, in which VE against trivalent inactivated influenza vaccine (IIV3) was evaluated among young children aged ≤5 years in 2008 [21]. The study unexpectedly found that the proportion of vaccination was higher among ORV positive controls than pan-negative controls, resulting in higher VE using ORV positive controls. Nasal swabs were used as respiratory specimens, which were logistically difficult to obtain from young children. As pan-negative controls would include some false negatives for influenza, they interpreted that ORV positive controls were more appropriate in ensuring adequate sample collection.

The phenomenon of higher vaccination rates in ORV positive controls compared with pan-negative controls was further discussed. A viral interference known as "temporary non-specific immunity" has been suggested [22]. This is a biological mechanism that involves a respiratory virus infection, which induces immunity not only against the same viruses but also for other viruses over a short time. Those who receive influenza vaccine would miss two opportunities: to be infected with influenza and to acquire temporary non-specific immunity to other respiratory viruses through natural infection of influenza. In the test-negative design, such vaccinated subjects would be classified into ORV positive controls, and contribute to higher vaccination rates among all controls. Control selection irrespective of vaccine status may be violated, and VE using all controls or ORV positive controls would be greater in comparison to that using pan-negative controls.

Some test-negative studies of IIV3 using different controls showed inconsistent results. Reports from Japan [23] and Portugal [24] found considerable VE variation, whereas no difference was observed in studies in the US [25] or Australia [26]. One small RCT in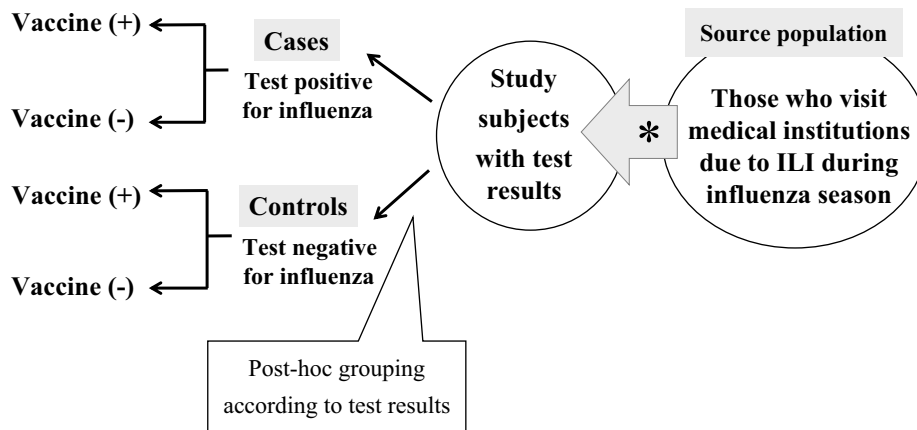 Hong Kong children reported that those who received IIV3 had an increased risk of non-influenza infections during the pre-pandemic period in 2009 [27]. On the other hand, a validation study using datasets from 4 published, double-blind RCTs found no meaningful association between live attenuated influenza vaccine and increased risk for non-influenza respiratory episodes [28]. A recent simulation study indicated that the effect of temporary non-specific immunity was significant when the attack rate of influenza was elevated to pandemic levels (>50%) but just marginal in typical influenza seasons (<20%) [29]. This simulation also suggested that combined data across the multiple influenza seasons may conceal the variation in attack rate, which may partly account for the inconsistency in the previous findings. To date, no recommendation regarding the most appropriate controls has been provided. Further discussion including what is meant by ORV positive controls or pan-negative controls is required.

4. Cautions for applying the test-negative design in routine clinical practice

Practically, test-negative designs would be easier to conduct by clinicians in comparison to classic case-control studies. Although real-time reverse transcription polymerase chain reaction (RT-RCR) or viral culture are desirable in defining case/control status in test-negative studies, results based on rapid diagnostic testing for influenza can be used as an outcome measure. In some countries where rapid diagnostic testing is widely available in routine clinical practice, such as Japan, test-negative studies can be readily employed using clinician-ordered testing results. The dataset would be huge if the information from many institutions is combined. However, such careless use of the test-negative design would result in some repercussions.

First, using rapid diagnostic testing results as an outcome measure has been demonstrated to underestimate VE due to imperfect sensitivity and specificity in comparison to RT-PCR or viral culture. A simulation study examined the extent of underestimation, in which the sensitivity and specificity of the rapid test was set at 80% and 90%, respectively. When rapid testing results were used in the test-negative design, the true VE of 90%, 70% and 50% was decreased to approximately 72.6%, 57.0% and 41.1%, respectively [16]. Another simulation also showed that when true VE was set at 70% for young children and 50% for all ages, use of a rapid test with the same sensitivity and specificity (i.e., 80% and 90%, respectively) in test-negative studies resulted in a VE estimate of 53% for children and 37% for all ages, respectively [30]. It is notable that lower specificity of the laboratory test for influenza was expected to contribute more to underestimation of VE than a lowering of sensitivity, if one value (e.g., sensitivity) was fixed at 1.0 and the other value (e.g., specificity) was changed from 0.8 to 1.0 [30]. Since specificity of rapid diagnostic tests is usually high, the influence of applying a rapid test for estimation of VE in test-negative studies might not be meaningful. However, as previously mentioned, the combination of imperfect sensitivity and specificity would greatly affect the VE even if rapid test misclassification was compensated by its high specificity. An approximate 20% reduction in effect estimates are considerable in influenza VE studies.

Second, enrolling the study subjects within a routine clinical setting can introduce selection bias. As shown in Fig. 3, the source population for the study is the patients with ILI who visit medical institutions. A certain proportion is then sampled as study subjects from the source population. The study subjects should have their test results for influenza because they have to be classified into either cases or controls thereafter. If the study subjects are limited to those who received the clinician-ordered test in a routine clinical setting, application of the test would depend on the likelihood

Vaccine (+)

Cases

Test positive for influenza

Vaccine (-)

Vaccine (+)

Controls

Test negative for influenza

Vaccine (-)

Study subjects with test results

Source population

Those who visit medical institutions due to ILI during influenza season

*

Post-hoc grouping according to test results

Fig. 3. A test-negative design to evaluate influenza vaccine effectiveness against laboratory-confirmed influenza. ILI denotes influenza-like illness. Asterisk (*) indicates a point where selection bias may occur. If the study subjects are limited to those who received the clinician-ordered test in the routine clinical setting, the application of the test would be related to (1) the likelihood of having influenza (outcome) or (2) influenza vaccination status (exposure), resulting in biased sampling (non-representativeness) of the study subjects from the source population.

of having influenza (outcome) or influenza vaccination status (exposure), resulting in biased sampling (non-representativeness) of the study subjects from the source population. For example, if clinicians order the diagnostic test for those with severe ILI and those who did not receive the vaccine, the proportion of non-vaccinees among cases is likely to increase, resulting in overestimation of VE. This translates to selection bias and it is impossible to estimate its extent or direction once such a bias is introduced. A report from the US pointed out that clinician-ordered rapid diagnostic testing could be a potential source of bias in influenza VE studies using the test-negative design [31]. The study showed that VE estimates based on rapid diagnostic testing results in the routine clinical setting were considerably underestimated and significant VE would have been missed. This study emphasized the importance of active recruitment of ILI patients according to the pre-defined standardized criteria.

With respect to possible selection bias in recruitment of the subjects, some researchers claim that during the influenza epidemic, clinicians would be too busy to develop their idea regarding application of the test. However, we cannot completely deny the possibilities that selection bias arise unconsciously. In order to avoid selection bias as far as possible, it is essential to recruit study subjects systematically from the source population according to pre-defined criteria. In effect, research on test-negative designs should be employed separately from routine clinical practice.

5. Conclusion

The methodology of VE studies is evolving. The test-negative design, a modified case-control study, has notable advantages in estimating influenza VE. Given that principles of case-control studies are more complicated than that of cohort studies or RCTs, collaboration, or consultation with epidemiologists would be useful. It should also be noted that reflecting on the basic concepts of epidemiology is always worthwhile. Accumulation of evidence from appropriately conducted test-negative designs will provide valid and universal estimates of VE against influenza.

Conflict of interest

None declared.

References

[1] Hirota Y, Kaji M. Principles and methods of influenza epidemiology: with special reference to field evaluation of vaccine efficacy. Kansenshogaku Zasshi 1994;68(11):1293–305 [in Japanese].
[2] World Health Organization. Vaccines against influenza: WHO position paper – November 2012. Weekly Epidemiol Rec 2012;87(47):461–6.
[3] Treanor JJ, Talbot HK, Ohmit SE, Coleman LA, Thompson MG, Cheng PY, et al. Effectiveness of seasonal influenza vaccines in the United States during a season with circulation of all three vaccine strains. Clin Infect Dis 2012;55 (7):951–9.
[4] Skowronski DM, Janjua NZ, De Serres G, Dickinson JA, Winter AL, Mahmud SM, et al. Interim estimates of influenza vaccine effectiveness in 2012/13 from Canada's sentinel surveillance network, January 2013. Euro Surveill 2013;18 (5).
[5] Kissling E, Valenciano M, Falcao J, Larrauri A, Widgren K, Pitigoi D, et al. "I-MOVE" towards monitoring seasonal and pandemic influenza vaccine effectiveness: lessons learnt from a pilot multi-centric case-control study in Europe, 2008–9. Euro Surveill 2009;14(44).
[6] Carville KS, Grant KA, Sullivan SG, Fielding JE, Lane CR, Franklin L, et al. Understanding influenza vaccine protection in the community: an assessment of the 2013 influenza season in Victoria Australia. Vaccine 2015;33(2):341–5.
[7] Turner N, Pierse N, Huang QS, Radke S, Bissielo A, Thompson MG, et al. Interim estimates of the effectiveness of seasonal trivalent inactivated influenza vaccine in preventing influenza hospitalisations and primary care visits in Auckland, New Zealand, in 2014. Euro Surveill 2014;19(42).
[8] Hirota Y, Fukushima W, Fujieda M, Ohfuji S, Maeda A. Essential tools for assessing influenza vaccine efficacy in improperly conducted studies: a Japanese perspective. Vaccine 2008;26(50):6455–8.
[9] Ozasa K. The effect of misclassification on evaluating the effectiveness of influenza vaccines. Vaccine 2008;26(50):6462–5.
[10] Fujieda M, Maeda A, Kondo K, Kaji M, Hirota Y. Inactivated influenza vaccine effectiveness in children under 6 years of age during the 2002–2003 season. Vaccine 2006;24(7):957–63.
[11] Hara M, Sakamoto T, Tanaka K. Effectiveness of influenza vaccination in preventing influenza-like illness among community-dwelling elderly: population-based cohort study in Japan. Vaccine 2006;24(27–28):5546–51.
[12] Ochiai H, Fujieda M, Ohfuji S, Fukushima W, Kondo K, Maeda A, et al. Inactivated influenza vaccine effectiveness against influenza-like illness among young children in Japan–with special reference to minimizing outcome misclassification. Vaccine 2009;27(50):7031–5.
[13] Belshe RB, Mendelman PM, Treanor J, King J, Gruber WC, Piedra P, et al. The efficacy of live attenuated, cold-adapted, trivalent, intranasal influenzavirus vaccine in children. N Engl J Med 1998;338(20):1405–12.
[14] Fukushima W, Ohfuji S, Deguchi M, Kawabata K, Hatayama H, Yoshida H, et al. Effectiveness of an influenza A (H1N1) 2009 monovalent vaccine among Japanese pregnant women: a prospective observational study assessing antibody efficacy. Vaccine 2012;30(52):7630–6.

[15] Hirota Y, Kaji M, Ide S, Kajiwara J, Kataoka K, Goto S, et al. Antibody efficacy as a keen index to evaluate influenza vaccine effectiveness. Vaccine 1997;15 (9):962–7.
[16] Orenstein EW, De Serres G, Haber MJ, Shay DK, Bridges CB, Gargiullo P, et al. Methodologic issues regarding the use of three observational study designs to assess influenza vaccine effectiveness. Int J Epidemiol 2007;36(3):623–31.
[17] Jackson ML, Nelson JC. The test-negative design for estimating influenza vaccine effectiveness. Vaccine 2013;31(17):2165–8.
[18] Foppa IM, Haber M, Ferdinands JM, Shay DK. The case test-negative design for studies of the effectiveness of influenza vaccine. Vaccine 2013;31(30):3104–9.
[19] Haber M, An Q, Foppa IM, Shay DK, Ferdinands JM, Orenstein WA. A probability model for evaluating the bias and precision of influenza vaccine effectiveness estimates from case-control studies. Epidemiol Infect 2015;143(7):1417–26.
[20] Rothman KJ. Epidemiology: an introduction. New York: Oxford University Press; 2002.
[21] Kelly H, Jacoby P, Dixon GA, Carcione D, Williams S, Moore HC, et al. Vaccine effectiveness against laboratory-confirmed influenza in healthy young children: a case-control study. Pediatr Infect Dis J 2011;30(2):107–11.
[22] Cowling BJ, Nishiura H. Virus interference and estimates of influenza vaccine effectiveness from test-negative studies. Epidemiology 2012;23(6):930–1.
[23] Suzuki M, Minh le N, Yoshimine H, Inoue K, Yoshida LM, Morimoto K, et al. Vaccine effectiveness against medically attended laboratory-confirmed influenza in Japan, 2011–2012 Season. PLoS One 2014;9(2):e88813.
[24] Nunes B, Machado A, Guiomar R, Pechirra P, Conde P, Cristovão P, et al. Estimates of 2012/13 influenza vaccine effectiveness using the case test-negative control design with different influenza negative control groups. Vaccine 2014;32(35):4443–9.
[25] Sundaram ME, McClure DL, VanWormer JJ, Friedrich TC, Meece JK, Belongia EA. Influenza vaccination is not associated with detection of noninfluenza respiratory viruses in seasonal studies of influenza vaccine effectiveness. Clin Infect Dis 2013;57(6):789–93.
[26] Blyth CC, Jacoby P, Effler PV, Kelly H, Smith DW, Robins C, et al. Effectiveness of trivalent flu vaccine in healthy young children. Pediatrics 2014;133(5): e1218–25.
[27] Cowling BJ, Fang VJ, Nishiura H, Chan KH, Ng S, Ip DK, et al. Increased risk of noninfluenza respiratory virus infections associated with receipt of inactivated influenza vaccine. Clin Infect Dis 2012;54(12):1778–83.
[28] De Serres G, Skowronski DM, Wu XW, Ambrose CS. The test-negative design: validity, accuracy and precision of vaccine efficacy estimates compared to the gold standard of randomised placebo-controlled clinical trials. Euro Surveill 2013;18(37).
[29] Suzuki M, Camacho A, Ariyoshi K. Potential effect of virus interference on influenza vaccine effectiveness estimates in test-negative designs. Epidemiol Infect 2014;142(12):2642–6.
[30] Jackson ML, Rothman KJ. Effects of imperfect test sensitivity and specificity on observational studies of influenza vaccine effectiveness. Vaccine 2015;33 (11):1313–6.
[31] Coleman LA, Kieke B, Irving S, Shay DK, Vandermause M, Lindstrom S, et al. Comparison of influenza vaccine effectiveness using different methods of case detection: clinician-ordered rapid antigen tests vs. active surveillance and testing with real-time reverse-transcriptase polymerase chain reaction (rRT-PCR). Vaccine 2011;29(3):387–90.