Journal of Clinical Epidemiology 67 (2014) 734−744

Journal of Clinical Epidemiology

REVIEW ARTICLE

After adjusting for bias in meta-analysis seasonal influenza vaccine remains effective in community-dwelling elderly☆

Maryam Darvishian[a,b], Giedre Gefenaite[a,b], Rebecca M. Turner[c], Petros Pechlivanoglou[a,d], Wim Van der Hoek[e], Edwin R. Van den Heuvel[b], Eelko Hak[a,b,*]

[a]*Unit of PharmacoEpidemiology & PharmacoEconomics (PE2), Department of Pharmacy, University of Groningen, A. Deusinglaan 1, 9713 AV, Groningen, The Netherlands*

[b]*Department of Epidemiology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, Postbus 30 001, 9700 RB Groningen, the Netherlands*

[c]*Medical Research Council Biostatistics Unit, Institute of Public Health, University Forvie Site, Robinson Way, Cambridge. UK. CB2 0SR*

[d]*Toronto Health Economics and Technology Assessment collaborative, University of Toronto, 144 College st. Rm:685, Toronto ON M5S3M2, Canada*

[e]*Department of Respiratory Infections of the Centre for Infectious Disease Control, National Institute for Public Health and the Environment, PO Box 1, 3720 BA Bilthoven, The Netherlands*

Accepted 17 February 2014; Published online 24 April 2014

Abstract

Objective: To compare the performance of the bias-adjusted meta-analysis to the conventional meta-analysis assessing seasonal influenza vaccine effectiveness among community-dwelling elderly aged 60 years and older.

Study Design and Setting: Systematic literature search revealed 14 cohort studies that met inclusion and exclusion criteria. Laboratory-confirmed influenza, influenza-like illness, hospitalization from influenza and/or pneumonia, and all-cause mortality were study outcomes. Potential biases were identified using bias checklists. The magnitude and uncertainty of biases were assessed by expert opinion. Pooled odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated using random effects model.

Results: After incorporating biases, overall effect estimates regressed slightly toward no effect, with the largest relative difference between conventional and bias-adjusted ORs for laboratory-confirmed influenza (OR, 0.18; 95% CI: 0.01, 3.00 vs. OR, 0.23; 95% CI: 0.03, 2.04). In most of the studies, CIs widened reflecting uncertainties about the biases. The between-study heterogeneity reduced considerably with the largest reduction for all-cause mortality ($I^2 = 4\%$, $P = 0.39$ vs. $I^2 = 91\%$, $P < 0.01$).

Conclusion: This case study showed that after addressing potential biases influenza vaccine was still estimated effective in preventing hospitalization from influenza and/or pneumonia and all-cause mortality. Increasing the number of assessors and incorporating empirical evidence might improve the new bias-adjustment method.

Keywords: Meta-analysis; Bias adjustment; Observational studies; Seasonal influenza; Vaccination; Community-dwelling elderly

1. Introduction

As seasonal influenza vaccination is standard care for older adults in most of the developed countries, conducting a randomized controlled trial (RCT) to estimate its effectiveness would be considered unethical. Therefore, apart from the limited number of older RCTs [1−3], the main evidence about influenza vaccine effectiveness comes from observational studies. Such studies are prone to bias because of lack of concealed randomization and different baseline characteristics between the vaccinated and the unvaccinated groups [4,5]. It has been shown that confounding by indication (also known as selection bias or healthy user effect), if not properly adjusted for in observational studies, could lead to an invalid estimate of vaccine effectiveness [6]. Moreover, some studies gave evidence for the presence of selection bias in most of the cohort studies assessing seasonal influenza vaccine effectiveness in the elderly population [7,8]. Combining evidence from observational studies by using standard methods of meta-analysis will compound this issue [9]. For instance, the most recently conducted meta-analysis assessing influenza vaccine effectiveness in elderly population [10] found a high level of heterogeneity between studies, which could be partly explained by unadjusted

Funding: The current work is funded by the University of Groningen.
Conflict of interest: All authors declare to have no conflict of interest with regard to the work.

* Corresponding author. Tel.: (0)50-363-75-76 (direct); fax: (0)50-363-27-72.
E-mail address: e.hak@rug.nl (E. Hak).

0895-4356
http://dx.doi.org/10.1016/j.jclinepi.2014.02.009

What is new?

- After incorporating the effect of internal and external biases, between-study heterogeneity reduced considerably.
- In this case study, bias-adjustment method enables us to identify the potential biases and to arrive at more appropriate estimates, possibly at the cost of less precision.
- Standard methods of meta-analysis do not take into account the effects of biases in observational studies. Bias-adjustment methods can be used to quantify the effects of such biases in future meta-analytic case studies.

sources of biases. It has been suggested that meta-analyses of observational studies are prone to bias because they pool the results from studies of differing quality (internal bias) and relevance (external bias) [11].

Although biases could partly be addressed by using quality scores through sensitivity analysis, it has been shown that weighing the analysis by quality scores is inadequate [12,13], and sensitivity analysis is not applicable when the number of included studies is low. Furthermore, it might be possible to use meta-regression techniques to investigate possible explanations of heterogeneity. However, this is only a good strategy when a relatively large number of studies are included in the meta-analysis [4].

To resolve these limitations, a novel bias-adjustment meta-analysis method has been proposed recently by Turner et al. [14]. This method provides a technique to adjust for internal and external biases through a process of eliciting and incorporating expert opinion with the results of the included studies in the meta-analysis. To estimate seasonal influenza vaccine effectiveness in the community-dwelling elderly against influenza and influenza-related outcomes, we first conducted a conventional meta-analysis of cohort studies (which are considered high in the hierarchy of observational studies). Secondly, we applied the bias-adjustment method to quantify the potential biases in the conventional meta-analysis. Finally, we compared the performance of the 2 methodological approaches and discussed their advantages and disadvantages.

2. Methods

2.1. Conventional meta-analysis

2.1.1. Search strategy

We searched MEDLINE, EMBASE, and the Cochrane library before September 2011 to identify cohort studies assessing influenza vaccine effectiveness. The search strategy consisted the following search terms: (“Influenza Vaccines”[Mesh] OR “Influenza, Human/epidemiology”[Mesh] OR “Influenza Human/immunology”[Mesh] OR “Influenza, Human/mortality”[Mesh] OR “Influenza, Human/prevention and control”[Mesh] OR “Influenza, Human/transmission”[Mesh] OR Influenza vaccine*[tiab] OR (Influenza OR flu [tiab])) AND (Vaccine* OR immuni* OR inocul* OR efficacy OR effectiveness [tiab]) AND (old* OR age*OR elderly [tiab] OR older persons [tiab] OR senior* [tiab]) AND (Clinical Trial [Mesh] OR “Case-Control Studies”[Mesh] OR “Cohort Studies” [Mesh] OR observational studies [tiab]). Only cohort studies assessing seasonal inactivated influenza vaccine effectiveness among community-dwelling elderly on laboratory-confirmed influenza, influenza-like illness (ILI), hospitalizations from influenza and/or pneumonia, and all-cause mortality were included. In our study, laboratory-confirmed influenza was defined as influenza confirmed by viral isolation, or virus nucleic acid detected in a clinical specimen, or when influenza-specific antibody response was measured. ILI was defined as a sudden onset of high fever, cough (usually dry), headache, muscle and joint pain, severe malaise (feeling unwell), sore throat, and runny nose or a code R80 according to the International Classification of Primary Care. Hospitalization from influenza and/or pneumonia was considered as an outcome when it was coded according to the International Classification of Diseases (ICD) version-10 as J12-18, J69.0, A48.1, J10.0, J10.1, J10.8, J11.0, J11.8, according to ICD version-9 (ICD-9-CM) as 480–487 or when hospitalization because of pneumonia was reported by the patient. All-cause death was recorded when it was reported as such in the reviewed studies.

2.1.2. Data extraction

Two reviewers (MD and GG) independently extracted data on the study population, characteristics of the participants, sample size, length of follow-up, inclusion and exclusion criteria for vaccinated and unvaccinated individuals, content and antigenic match of the administered vaccines, description of viral circulation, epidemic condition, and outcomes. If information regarding the vaccine strains and epidemic condition was not available in the studies, we extracted this information from the World Health Organization (WHO) Web site [15].

2.1.3. Statistical analyses

The extracted raw data on vaccination status and outcomes from the cohort studies were entered into the Cochrane RevMan Software (version 5.2) [16]. Where applicable, the adjusted odds ratios (ORs) were used to back calculate the adjusted number of events by using the formula $r_{1adj} \approx (OR_{adj}) \times (r_2/n_2) \times n_1$, where r_{1adj} is the adjusted number of events in the intervention group, OR_{adj} is the adjusted effect size given in the original study, r_2 is the number of events in the control group, n_2 is the total number of participants in the control group, and n_1 is the

total number of participants in the intervention group. Back-calculation assumes that the OR is equal to the risk ratio, which is approximately true when outcome events are rare. After combining the raw data, the pooled ORs and their 95% confidence intervals (95% CIs) were calculated using the DerSimonian and Laird random-effects model [17].

We performed a subgroup analysis based on the epidemic condition and vaccine matching. Vaccine matching was defined as antigenic similarity between the vaccine strains and the circulating viruses. If studies reported raw data on the vaccination status and the outcomes for more than 1 influenza season, we considered each influenza season as a separate study, which had a separate dataset. To quantify heterogeneity, the measure of inconsistency I^2 (%) was calculated as proposed by Higgins et al. [18], in which the quantity I^2 of 0% indicates no observed heterogeneity, and larger values indicate increasing heterogeneity [18].

2.2. Application of bias-adjustment method

The bias-adjustment method developed by Turner et al. [14] was applied to the same cohort studies included in the conventional meta-analysis. The process of bias-adjustment included several steps: (1) describing the target setting of the meta-analysis, (2) defining an idealized version of each individual study included in the meta-analysis, (3) quantifying the effect of internal and external biases in all the included cohort studies, (4) eliciting expert opinion using bias elicitation scale and bias-checklist, and (5) conducting a bias-adjusted meta-analysis.

2.2.1. Target setting

The target setting of our study was defined with respect to: (1) the study population, which was community-dwelling elderly aged 60 years and older; (2) the intervention of the study, which was the seasonal influenza vaccine recommended by WHO for a particular influenza season when the study was conducted; (3) the control, which was no administration of the seasonal influenza vaccine during the period when the study was conducted; and (4) the following outcomes: laboratory-confirmed influenza, ILI, hospitalization from influenza and/or pneumonia, and all-cause mortality during influenza season.

2.2.2. Idealized studies and internal and external biases

In the second step, we defined an idealized version of each included study. The idealized study is an imagined repeat of each study by means of a perfect design that could eliminate all sources of internal biases [14]. We identify the internal biases affecting each study by comparing the study carried out against its idealized version. External biases can be identified by comparing the idealized version of each study with the target setting of the meta-analysis. For example, if a study investigates the effect of vaccine in a population aged 50 years and older, an external bias arises because the target population of the idealized version of that study is older, that is 60 years and older. In a mini-protocol available in the bias checklist (available on the journal's Web site), the idealized study should be defined with respect to the population, which the researchers planned to study (eg, community-dwelling elderly), their planned comparison (eg, seasonal influenza vaccination vs. no vaccination), and the outcome, which they planned to measure. The idealized study with respect to different outcomes was defined as follows: for the laboratory-confirmed influenza outcome, the idealized study was restricted to the patients who, according to European Center for Disease Prevention and Control influenza case definition [19], had at least one of the following clinical forms: ILI defined as sudden onset of symptoms, and at least one of the 4 systematic symptoms: fever, malaise, headache, or myalgia, and at least one of the 3 respiratory symptoms: cough, sore throat, or shortness of breath, and at least one of the laboratory criteria: isolation of the influenza virus from a clinical specimen, or detection of the influenza virus nucleic acid in a clinical specimen, or identification of the influenza virus antigen by direct fluorescent antibody test in a clinical specimen, or influenza-specific antibody response. For the ILI outcome, the idealized study was restricted to patients who had sudden onset symptoms, at least one of the 4 previously mentioned systematic symptoms, at least one of the previously mentioned respiratory symptoms, and a clinician judgment by a general practitioner or a nurse that the illness was due to a respiratory infection. For the hospitalization from influenza and/or pneumonia outcome, the idealized study was restricted to patients who had the diagnostic measurements of their disease such as laboratory tests and/or X-ray. For the all-cause mortality outcome, no extra specification was required for the outcome in the idealized study.

In this meta-analysis, internal biases were categorized into 5 groups: (1) selection bias, which was defined as baseline differences between the intervention and control groups; (2) performance bias, which was suspected if subjects did not receive their assigned intervention; (3) attrition bias, which resulted from loss to follow-up and missing data; (4) detection bias, which resulted from inaccurate measurement of the outcomes; and (5) suspicion of other biases defined as biases related to any other flaws [14]. External biases were categorized into population and outcome biases. Population bias resulted from the difference between age and health status of participants in each study and the target population. Outcome bias referred to the differences in the definition and timing of the idealized study outcomes as compared with the target setting outcomes [14]. Bias checklists were completed by 2 assessors (MD and GG) based on the extracted information on the presence of internal and external biases in each included study (a completed bias checklist is available at www.jclinepi.com). Before the bias elicitation meetings, all assessors read the bias checklists and the original articles.

2.2.3. Bias elicitation process

Five quantitatively trained assessors (EH, WH, GG, MD trained in epidemiology, and RT trained in medical statistics) were involved in the internal bias elicitation meeting and 4 subject-matter assessors (EH, WH, GG, and MD) were involved in the external bias elicitation meeting. At the meetings, assessors discussed each individual bias with respect to whether the magnitude of the bias was independent of the magnitude of the intervention effect (additive bias) or proportional to the magnitude of the intervention effect (proportional bias). Assessors agreed that the effects of the internal biases were independent of the magnitude of the intervention effect and that the effects of the external biases could change the magnitude of the intervention effect. Therefore, all internal biases in this meta-analysis were considered to be additive, and all external biases were considered to be proportional. For example, the internal bias that is caused by inadequate adjustment for the potential confounders is independent of the magnitude of the vaccination effect (additive bias), whereas the external bias that is caused by a study conducted during non-epidemic influenza season is proportional to the intervention effect because it underestimates the vaccination effect. After the group discussion, each assessor independently provided their opinion on the impact and uncertainty of each of the biases on the bias elicitation scales. Biases were then assessed on a relative risk scale using a 67% CI, such that the assessor assumed that the bias was twice as likely to lie inside rather than outside this range (Appendix Fig. 1 at www.jclinepi.com) [14]. Assessors judged the severity of the biases on bias elicitation scales as no bias (1), low (0.9−1), medium (0.7−0.9), or high (less than 0.7) both, in favor of the intervention and control [14].

2.2.4. Bias-adjusted meta-analysis

Using the bias elicitation scale, the internal biases and external biases from each assessor were elicited and used to calculate the means and variances of the total additive and total proportional bias for each study. The individual additive biases have been summed to find total additive bias, and the individual proportional biases have been multiplied to find total proportional bias. These data were then used to adjust the study effect estimates and standard errors. The results were pooled across assessors using the median estimate and median standard deviation [20]. Finally, fully bias-adjusted results were combined across studies using random effects meta-analysis. Data were analyzed with R statistical package [21] using the code adopted from Turner et al. [14].

2.2.5. Agreement among assessors

To assess the agreement between the assessors, we first calculated the means and variances of the internal and external biases for each assessor separately. Then, we conducted a mixed-effects model on the calculated mean values, treating the assessors as fixed effect and studies as random effects, and weighed the residuals with the reciprocal of the measured variances of the biases. Using the variance components of the studies and residuals, the intraclass correlation coefficient (ICC) [22,23] was calculated with approximate 95% CIs using the F-distribution [24] based on the Satterthwaite approach [25]. The ICC expressed as percentage can be interpreted as a measure of agreement [22,23].

3. Results

3.1. Systematic literature search

The systematic literature search resulted in 2,256 potentially relevant articles. After removing the duplicates, 1,785 titles and abstracts were screened by 2 independent reviewers (MD and GG). EH was consulted in case of disagreement. Only 14 studies met the inclusion criteria (Appendix Table 1 at www.jclinepi.com) to be included in the meta-analysis (Fig. 1). These 14 studies resulted in 19 datasets because some studies reported on multiple influenza seasons (Appendix Table 2 at www.jclinepi.com). Ten datasets measured vaccine effectiveness during epidemic years. In 15 datasets, vaccine strains had a good match with the circulating viruses. Two studies reported vaccine effectiveness against laboratory-confirmed influenza [26,27], 4 studies against ILI [27−30], 8 studies against hospitalization from influenza and/or pneumonia [28,31−37], and 5 studies against all-cause mortality [28,30,37−39].

3.2. Conventional meta-analysis

Pooling the data from the 2 studies that reported effectiveness against laboratory-confirmed influenza showed that seasonal influenza vaccine was not statistically significantly effective (OR, 0.18; 95% CI: 0.01, 3.00). The study conducted during an epidemic season [26] showed that a well-matched influenza vaccine prevented laboratory-confirmed influenza (OR, 0.05; 95% CI: 0.01, 0.35), although this was not the case during a non-epidemic season (OR, 0.64; 95% CI: 0.12, 3.31) [27] (Appendix Fig. 2 at www.jclinepi.com).

Influenza vaccination was associated with a reduction of ILI (OR, 0.55; 95% CI: 0.39, 0.78). The study conducted during an epidemic season showed that influenza vaccine was effective against ILI even when the vaccine strains did not match the circulating viruses (OR, 0.39; 95% CI: 0.21, 0.74) [28]. The studies conducted during non-epidemic seasons with a good vaccine match also demonstrated a preventive effect (OR, 0.55; 95% CI: 0.33, 0.92) [27,30]. The study conducted during a non-epidemic season and with poor vaccine match did not show a statistically significant reduction on the outcome (OR,

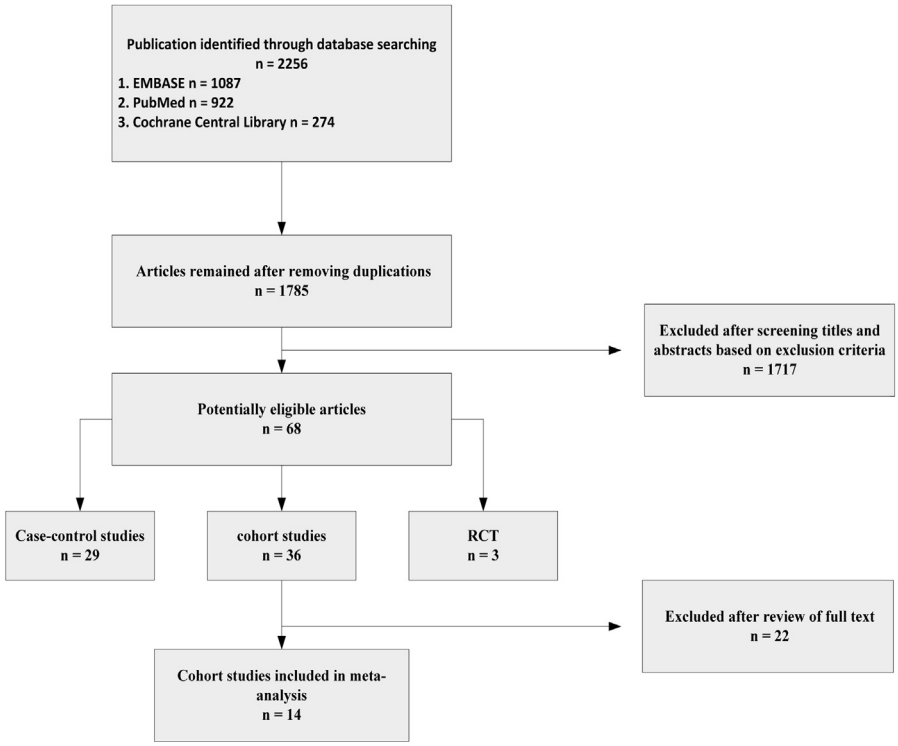

Fig. 1. Flowchart of the eligible cohort studies for inclusion in the meta-analysis.

0.78; 95% CI: 0.41, 1.47) [29] (Appendix Fig. 3 at www.jclinepi.com).

Influenza vaccination was associated with a reduction in hospitalization from influenza and/or pneumonia (OR, 0.67; 95% CI: 0.57, 0.77) (Table 3). The analysis of 6 datasets conducted in epidemic years showed a significant reduction in hospitalization from influenza and/or pneumonia in years with good vaccine match (OR, 0.70; 95% CI: 0.58, 0.84), and in years with poor vaccine match (OR, 0.58; 95% CI: 0.42, 0.79). Four datasets from non-epidemic years with good vaccine match demonstrated a reduction in the outcome as well (OR, 0.51; 95% CI: 0.39, 0.67), whereas no reduction was observed when there was a poor vaccine match (OR, 1.11; 95% CI: 0.76, 1.60) [33] (Appendix Fig. 4 at www.jclinepi.com).

Influenza vaccination was significantly effective in reducing all-cause mortality (OR, 0.51; 95% CI: 0.40, 0.65). The analysis conducted during epidemic seasons when vaccine strains were well matched with circulating viruses [37,38] showed a preventive effect (OR, 0.51; 95% CI: 0.47, 0.54), whereas no such effect was observed when the vaccine strains and circulating viruses did not match (OR, 3.41; 95% CI: 0.77, 15.03) [28]. Vaccine did not show statistically significant effectiveness against all-cause mortality during a non-epidemic season when there was a good match between the vaccine strains and the circulating viruses (OR, 0.50; 95% CI: 0.22, 1.13) [30,39] (Appendix Fig. 5 at www.jclinepi.com). For all the outcomes except ILI, the level of between-study heterogeneity was significantly high (Table 3).

3.3. Bias-adjusted meta-analysis

3.3.1. Biases identified

Appendix Table 3 (at www.jclinepi.com) summarizes the suspected internal biases affecting the 14 included studies in the meta-analysis. Selection bias was judged to affect all the included studies. Different baseline characteristics in the intervention and control group that is confounding by indication, unclear inclusion and exclusion criteria, and inadequate adjustment for potential confounders were the most common reasons for the presence of selection bias. Although the study effect estimates in 7 studies [28,30,33,34,36–38] were partially adjusted for some confounders, in most of the studies the relevant confounders such as health status and vaccine history have not been addressed. Detection bias was suspected in the studies that did not use appropriate laboratory tests to confirm influenza cases and in the studies that used self-reports to obtain information regarding ILI [26–30]. In these studies, symptomatic patients not consulting the clinics were probably not further tested for laboratory-confirmed influenza diagnosis. Furthermore, detection bias was suspected in the studies that did not use objective measures such as laboratory test and/or X-ray to confirm the hospitalization from influenza and/or pneumonia [28,31–37]. Other bias was suspected in the study where some patients had a history

Table 1. Potential external biases identified in the studies

Study	Population bias[a] (patients age)	Outcome bias
Christenson et al. [31]	65 years and older	
Christenson et al. [32]	65 years and older	Study conducted during non-epidemic year
Fleming et al. [38]	55 years and older	
Hara et al. [28]	65–79 years old	Vaccine did not match circulating viruses
Kawai et al. [27]	65–104 years old	Study conducted during non-epidemic year
Manzoli et al. [33]	65 years and older	Study conducted during non-epidemic year, vaccine did not match circulating viruses
Nichol et al. [34]	65 years and older	Study conducted during non-epidemic year (Nichol 1994 (a)), vaccine did not match circulating viruses (Nichol 1994 (c))
Nichol et al. [35]	65 years and older	
Nichol et al. [37]	65 years and older	
Nicholson et al. [26]	60–90 years old	
Ozasa et al. [29]	65 years and older	Study conducted during non-epidemic year
Shapiro et al. [39]	65 years and older	Study conducted during non-epidemic year
Voordouw et al. [30]	65 years and older	Study conducted during non-epidemic year
Wang et al. [36]	65 years and older	Study conducted in midseason period

[a] The target population of the meta-analysis was community-dwelling elderly aged 60 years and older.

of vaccination with pneumococcal vaccine [32]. Assessors agreed that this bias could overestimate the effect of influenza vaccination. Moreover, other bias was suspected in the study where there was an overlap between the inoculation period (eg, October 1, 2003 to December 31, 2003) and survey period (eg, December 1, 2003 to March 31, 2004) [28]. Assessors agreed that these biases could underestimate the effect of the vaccination.

Population bias was suspected to affect all the included studies because of the differences between the age of the study population in the included studies and the population in the defined target setting of the meta-analysis (Table 1). For example, in a study conducted by Fleming et al. [38], the influenza vaccine effectiveness was assessed among patients aged 55 years and older rather than the target age range of 60 years and older. All assessors agreed that this bias could make the intervention appear more effective because the immune response to the vaccine reduces with age and therefore, the vaccine would appear more effective in younger age than in older age. Outcome bias was expected in the studies conducted during the non-epidemic influenza seasons [32,27,33,34a,29,39,30] or when the vaccines did not match the circulating viruses [28,33,34c]. Furthermore, outcome bias was suspected in a study in which the patient recruitment (January through June 2001) started 1 month after the peak of influenza season

Table 2. Geometric averages of the internal and external biases and relative standard deviations for each outcome in the included studies in the meta-analysis

	Outcome							
	Influenza		ILI		Hospitalization		All-cause mortality	
Study	Internal bias	External bias	Internal bias	External bias	Internal bias	External bias	Internal bias	External bias
Christenson et al. [31]					0.81 (40.6)	0.96 (6.0)		
Christenson et al.(a) [32]					0.83 (49.2)	0.96 (6.0)		
Christenson et al.(b) [32]					0.83 (49.2)	0.96 (6.0)		
Christenson et al.(c) [32]					0.83 (49.2)	0.96 (6.0)		
Fleming et al. [38]							0.95 (28.4)	1.05 (6.4)
Hara et al. [28]			0.84 (50.1)	0.81 (23.8)	0.84 (50.1)	0.81 (23.8)	0.89 (43.1)	0.81 (23.4)
Kawai et al. [27]	0.98 (69.6)	0.86 (17.2)	1.05 (41.4)	0.86 (17.2)				
Manzoli et al. [33]					0.93 (50.5)	0.78 (24.1)		
Nichol et al.(a) [34]					0.92 (32.2)	1.16 (21.9)		
Nichol et al.(b) [34]					0.92 (32.2)	1.21 (19.5)		
Nichol et al.(c) [34]					0.92 (32.2)	1.01 (26.8)		
Nichol et al. [35]					0.92 (32.2)	1.21 (19.5)		
Nichol et al.(a) [37]					0.92 (32.2)	1.21 (19.5)	0.92 (31.2)	1.21 (19.5)
Nichol et al.(b) [37]					0.92 (32.2)	1.21 (19.5)	0.92 (31.2)	1.21 (19.5)
Nicholson et al. [26]	0.70 (51.8)	0.98 (23.2)						
Ozasa et al. [29]			0.99 (85.3)	1.15 (34.6)				
Shapiro et al. [39]							0.82 (58.6)	0.75 (25.2)
Voordouw et al. [30]			0.78 (68.2)	0.92 (8.1)			0.92 (48.3)	0.92 (8.5)
Wang et al. [36]					0.78 (72.7)	0.96 (56.0)		

Abbreviation: ILI, influenza-like illness.

Table 3. Pooled odds ratios from conventional meta-analysis, meta-analyses adjusted for internal biases and meta-analyses adjusted for internal and external biases

Outcome	Number of studies per outcome	Pooled OR (CI) from conventional meta-analyses	Test for heterogeneity *P*-value	Pooled OR (CI) from meta-analyses adjusted for internal biases	Test for heterogeneity *P*-value	Pooled OR (CI) from meta-analyses adjusted for internal and external biases	Test for heterogeneity *P*-value
Influenza	2	0.18 (0.01, 3.00)	$I^2 = 78\%$ $P = 0.03$	0.22 (0.02, 2.13)	$I^2 = 53\%$ $P = 0.14$	0.23 (0.03, 2.04)	$I^2 = 35\%$ $P = 0.21$
ILI	4	0.55 (0.39, 0.78)	$I^2 = 6\%$ $P = 0.36$	0.63 (0.35, 1.15)	$I^2 = 0\%$ $P = 0.87$	0.64 (0.32, 1.29)	$I^2 = 0\%$ $P = 0.87$
Hospitalization from influenza and/or pneumonia	8	0.67 (0.57, 0.77)	$I^2 = 82\%$ $P < 0.00001$	0.71 (0.57, 0.87)	$I^2 = 0\%$ $P = 0.96$	0.75 (0.60, 0.94)	$I^2 = 0\%$ $P = 0.99$
All-cause mortality	5	0.51(0.40, 0.65)	$I^2 = 91\%$ $P < 0.0001$	0.61 (0.42, 0.90)	$I^2 = 20\%$ $P = 0.29$	0.64 (0.44, 0.92)	$I^2 = 4\%$ $P = 0.39$

Abbreviations: CI, confidence interval; OR, odds ratio; ILI, influenza-like illness.

(December to March 2001) [36]. Assessors expected that these biases would reduce the effect of the intervention.

Because in the majority of the included studies, one dose of the trivalent seasonal influenza vaccine recommended by WHO for each influenza season was administered, the possible variation among studies in the dose and delivery of the influenza vaccine was regarded as diversity that would arise in practice and was not treated as bias.

3.3.2. Bias assessment

To illustrate the magnitude and direction of the biases, Table 2 represents the averages and the relative standard deviations of the total internal and total external biases (over assessors) for all outcomes and all studies. Most of the bias averages were smaller than one, meaning that assessors believed that the study effect sizes were overestimated because of likely presence of biases. On the other hand, bias averages larger than one indicate that the assessors believed that the study effect sizes were underestimated because of the biases. For instance, in the conventional meta-analysis, the measured OR for the study by Christenson et al. [31] for the outcome hospitalization from influenza and/or pneumonia was 0.89 (95% CI: 0.80, 0.99). After adjusting for the average of the total internal and external bias, the OR shifted toward no effect OR of 1.11 (95% CI: 0.48, 2.58).

3.3.3. An example of the bias-adjustment process

To explain the bias-adjustment process, we will illustrate it for the study by Wang et al. [36] (the completed bias checklists and bias elicitation form is available on the journal's Web site). The same results for the other studies can be provided to readers on request. According to the article, the influenza vaccination coverage for high-risk and low-risk elderly was 58% and 29%, respectively, and study subjects were relatively younger in the intervention group and they had more underlying diseases compared with the control group. Moreover, although the study effect estimate was partly adjusted for some confounders (age, sex, risk status, and vaccination status), other important confounding factors (eg, previous vaccination history and health status) were not addressed. In the bias-elicitation meeting assessors agreed that differences between the baseline characteristics of the intervention and control group and inadequate adjustment for confounders could lead to selection bias. Moreover, performance bias was suspected by all assessors because it was not clear how the vaccination status was ascertained. Because the article reported the attrition of 883 vaccinated elderly because of death or moving out of the county, attrition bias due to loss to follow up was also suspected.

Population bias was identified because of the differences between the target age of the participants in the study ($\geq$65 years) and the target age of the meta-analysis ($\geq$60 years). Outcome bias was suspected because the patient recruitment (January through June 2001) started 1 month after the peak of the influenza season (December to March 2001). Fig. 2A and B present the 67% ranges elicited from the assessors A–E and A–D for internal and external biases respectively. In total, there was a general degree of consistency across the assessors on the internal and external biases.

3.3.4. Agreement between the assessors

The agreement among the assessors for the outcome hospitalization from influenza and/or pneumonia for internal and external biases was 61% (95% CI: 18, 82) for the selection bias, 54% (95% CI: 25, 73) for the performance bias, 37% (95% CI: 11, 61) for the detection bias, 78% (95% CI: 50, 90) for the attrition bias, 7% (95% CI: 0.01, 30) for the population bias, and 18% (95% CI: 0.02, 57) for the outcome bias. Fig. 3 presents an example of the 67% range for selection bias elicited by one assessor in the study by Wang et al. [36].

3.3.5. Bias-adjusted meta-analysis: comparison of the results

Pooled ORs from conventional meta-analysis, meta-analyses adjusted for internal biases and meta-analyses adjusted for internal and external bias are presented in

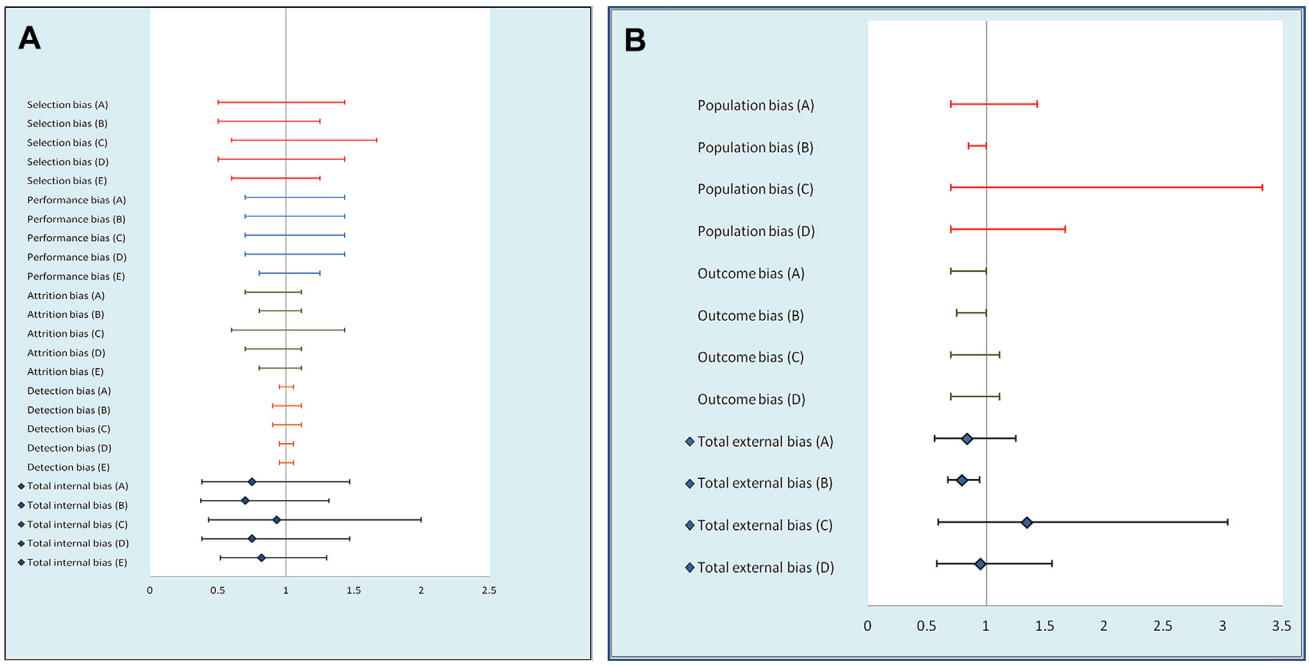

Fig. 2. (A) Internal biases in the study by Wang et al. [36]; 67% ranges elicited from assessors A–E, with means and 67% ranges for the total internal bias. (B) External biases in the study by Wang et al. [36]; 67% ranges elicited from assessors A–D, with means and 67% ranges for the total external bias.

Table 3. After incorporating the effect of internal and external biases, the pooled OR for the outcome laboratory-confirmed influenza increased to 0.23 (95% CI: 0.03, 2.04) and the level of heterogeneity reduced (I^2 = 35%, P = 0.21 vs. I^2 = 78%, P = 0.03 and τ^2 = 0.87 vs. τ^2 = 3.21) (Appendix Fig. 6 at www.jclinepi.com).

In the bias-adjusted meta-analysis of the 4 studies that reported effectiveness as reduction of ILI, influenza vaccine did not show statistically significant effectiveness against ILI (OR, 0.64; 95% CI: 0.32, 1.29). After bias adjustment, all studies point estimates shifted toward no effect, and their CIs widened. Moreover, the relative weight given to each study in the meta-analysis changed. For example, the study by Kawai et al. [27] received 14% of the weight in the conventional meta-analysis, although it received 29% of the weight in the bias-adjusted meta-analysis because the biases were judged to be smaller in this study as compared with the other studies (Appendix Fig. 7 at www.jclinepi.com).

Pooling bias-adjusted point estimates from 13 datasets showed that influenza vaccine was significantly associated with a reduction in hospitalization from influenza and/or pneumonia (OR, 0.75; 95% CI: 0.60, 0.94) with no evidence of between-study heterogeneity (I^2 = 0%, P = 0.99). Furthermore, the CI widened reflecting the uncertainty because of the biases. After incorporating the effect of potential biases, the pooled estimate of the subgroup analysis of studies conducted during the non-epidemic season when the vaccine matched the circulating virus was no longer statistically significant (OR, 0.65; 95% CI: 0.39, 1.09) (Appendix Fig. 8 at www.jclinepi.com).

The bias-adjusted meta-analysis of 6 datasets from 5 studies showed that influenza vaccine was significantly effective in preventing all-cause mortality in community-dwelling elderly. After adjusting for the internal and external biases, the pooled OR increased to 0.64 (95% CI: 0.44, 0.92), the CI widened, and the between-study heterogeneity reduced (I^2 = 4%, P = 0.39 vs. I^2 = 91%, P = 0.01 and τ^2 = 0.01 vs. τ^2 = 0.06). Moreover, in the bias-adjusted meta-analysis datasets judged to be affected by smaller biases with less uncertainty in their magnitude [37,38] tended to have more weight in the meta-analysis although datasets judged to be affected by greater biases or more uncertain biases tended to have less weight [28,30,39]. The subgroup analysis during the epidemic years when vaccine matched the circulating virus showed

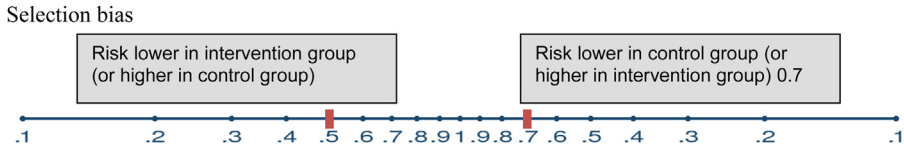

Fig. 3. Elicited 67% range for the selection bias by one assessor in the study by Wang et al. [36].

significant preventive effect for influenza vaccine (OR, 0.60; 95% CI: 0.41, 0.89) (Appendix Fig. 9 at www.jclinepi.com).

4. Discussion

In this article, we assessed the seasonal influenza vaccine effectiveness among community-dwelling elderly by using 2 different meta-analyses techniques. The results of the conventional meta-analysis showed that seasonal influenza vaccine was effective in preventing ILI, hospitalization from influenza and/or pneumonia, and all-cause mortality among individuals aged 60 years and older living in the community. Although the estimates pointed toward effectiveness, the reduction in laboratory-confirmed influenza was not statistically significant. We further found a high level of heterogeneity between the studies, which partly could be explained by failure of the conventional methods to address the potential biases (eg, inadequate adjustment for confounders) in the included studies [40].

In the next step, we therefore applied a recently developed bias-adjustment approach to the conventional meta-analysis [14]. In the bias-adjusted meta-analysis, the potential biases in the included studies were addressed through several steps: defining a target question, describing an idealized version of each study, completing the bias-checklist to identify potential biases, eliciting the bias ranges, and performing a bias-adjusted meta-analysis. After incorporating large amounts of uncertainty due to biases, influenza vaccine showed statistically significantly effectiveness in preventing hospitalization from influenza and/or pneumonia and all-cause mortality. However, the vaccine did not show statistically significant effectiveness against ILI, which was previously observed in the conventional meta-analysis. Compared with the conventional meta-analysis, the pooled effect estimates for all the outcomes moved slightly toward no effect with the largest relative difference between unadjusted and adjusted ORs for laboratory-confirmed influenza outcome (OR 0.18 vs. 0.23). Furthermore, the incorporated uncertainty because of the size of the biases increased the within study error which consequently widened the CIs and reduced the high level of between-study heterogeneity [14]. Bias adjustment also increased the relative weight of the studies with higher quality as compared with the studies with lower quality. In contrast to the conventional meta-analysis, the bias-adjustment method enabled us to interpret the results when also considering potential biases in the included studies.

There are several limitations for the bias-adjusted approach we used. Firstly, despite all the efforts to address the internal and external biases and to reduce the level of heterogeneity, the vaccine effectiveness for a non-specific outcome (all-cause mortality) was still unrealistically high [7]. Secondly, although the assessors were chosen based on their expertise, it is possible that other experts would generate different results. In fact the degree of agreement between the assessors for some biases, that is population bias and outcome bias, was low, and we believe that increasing the number of assessors could improve the elicited bias estimates and the generalizability of the results. To further improve the method of bias-adjustment, it would be desirable to incorporate empirical evidence on biases from meta-epidemiological studies [41] in addition to expert opinion [14]; however, empirical evidence is very unlikely to become available for external biases, which are specific to the target setting. Moreover, because the true magnitude of the biases remains unknown, the level of uncertainty given by assessors for some biases (eg, selection bias) is very high, which could lead to very wide CIs. However, in our opinion, compared with the conventional meta-analysis, the effect sizes are probably more appropriate after taking into account the effect of biases. Finally the bias-adjustment technique is a time-consuming process [14]. Completing bias-adjustment checklists by one assessor takes at least 2 hours. Based on our experience, the bias elicitation process for each study took at least 30 minutes. Depending on the number of included studies in the meta-analysis and outcomes of interest, this process might take even longer.

4.1. Suggestions for future research

In this particular case study, bias-adjustment method could reduce heterogeneity between the studies of seasonal influenza vaccine effectiveness. This evidence may suggest a considerable effect of biases in these non-randomized studies. Selection bias, that is confounding by indication and unclear inclusion and exclusion criteria, was the main source of bias in the studies assessing influenza vaccine effectiveness in the community-dwelling elderly. Adequate adjustment for health (co-morbidity) status, history of influenza or ILI, and history of previous influenza vaccination could reduce the selection bias [42]. To avoid detection bias, more information regarding the outcome measures (eg, types of diagnostic tests to confirm the outcome) should be provided. In the absence of randomized trials, further similar case studies using bias-adjustment techniques or similar methods need to be used to provide more consistent estimates of the influenza vaccine effectiveness [43].

5. Conclusion

In this case study, the bias-adjustment method allowed us to identify the potential internal and external biases in the included studies. When the heterogeneity was reduced, influenza vaccination of individuals aged 60 years and older living in the community, showed statistically significant effectiveness against hospitalization from influenza and/or pneumonia and all-cause mortality, but not against influenza and ILI. Under the assumption that the potential biases on average were identified by assessors, we judge

bias-adjustment method provides more appropriate effect sizes for hospitalization from influenza and/or pneumonia and all-cause mortality. Although the method provides a tool to adjust for potential biases, it could be further improved by increasing the number of assessors and incorporating empirical evidence on biases in the future studies, if this becomes available. Moreover, more case studies are needed to further evaluate this new method.

Acknowledgments

Study concept and design was by Maryam Darvishian, Giedre Gefenaite, and Eelko Hak. Analysis and interpretation of data was by Maryam Darvishian, Giedre Gefenaite, Rebecca Turner, Petros Pechlivanoglou, Edwin van den Heuvel, and Eelko Hak. Drafting of the manuscript was done by Maryam Darvishian. Critical revision of the manuscript for important intellectual content was by Giedre Gefenaite, Rebecca Turner, Petros Pechlivanoglou, Wim Van der Hoek, Edwin van den Heuvel, and Eelko Hak. All authors approved the final version of the manuscript. RT was supported by Medical Research Council grant U105260558.

Appendix

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.jclinepi.2014.02.009.

References

[1] Govaert TME, Thijs C, Masurel N, Sprenger M, Dinant G, Knottnerus J. The efficacy of influenza vaccination in elderly individuals. JAMA 1994;272:1661−5.
[2] Allsup S, Haycox A, Regan M, Gosney M. Is influenza vaccination cost effective for healthy people between ages 65 and 74 years? A randomised controlled trial. Vaccine 2004;23:639−45.
[3] Praditsuwan R, Assantachai P, Wasi C, Puthavatana P, Kositanont U. The efficacy and effectiveness of influenza vaccination among Thai elderly persons living in the community. J Med Assoc Thai 2005; 88:256−64.
[4] Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions 4.2.6 [updated September 2006]. In: The Cochrane library, Issue 4. Chichester, UK: John Wiley & Sons, Ltd; 2006.
[5] Knottnerus JA. Influenza vaccination in the elderly: current evidence and uncertainties. J Clin Epidemiol 2009;62:675−6.
[6] Hak E, Verheij TJ, Grobbee DE, Nichol KL, Hoes AW. Confounding by indication in non-experimental evaluation of vaccine effectiveness: the example of prevention of influenza complications. J Epidemiol Community Health 2002;56:951−5.
[7] Simonsen L, Viboud C, Taylor RJ, Miller MA, Jackson L. Influenza vaccination and mortality benefits: new insights, new opportunities. Vaccine 2009;27(45):6300−4.
[8] Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. Int J Epidemiol 2006;35:337−44.
[9] Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. BMJ 1998;316:140−4.
[10] Jefferson T, Di Pietrantonj C, Al-Ansary LA, Ferroni E, Thorning S, Thomas RE. Vaccines for preventing influenza in the elderly. Cochrane Database Syst Rev 2010;17:CD004876.
[11] Thompson S, Ekelund U, Jebb S, Lindroos AK, Mander A, Sharp S, et al. A proposed method of bias adjustment for meta-analyses of published observational studies. Int J Epidemiol 2011;40:765−77.
[12] Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. Biostatistics 2001;2(4):463−71.
[13] Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of meta-analyses on the basis of quality scores should be abandoned. J Clin Epidemiol 2006;59:1249−56.
[14] Turner RM, Spiegelhalter DJ, Smith GC, Thompson SG. Bias modeling in evidence synthesis. J R Stat Soc 2009;172(1):21−47.
[15] World Health Organization (WHO). WHO recommendations on the composition of influenza virus vaccines. Available at http://www.who.int/influenza/vaccines/virus/recommendations/en/index.html. Accessed May 1, 2011.
[16] Review Manager (RevMan) [computer program] version 5.2. Copenhagen, Denmark: The Nordic Cochrane Centre, The Cochrane Collaboration; 2012.
[17] DerSimonian R, Laird N. Meta-analysis in clinical trials. Controlled clin trials 1986;7:177−88.
[18] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557−60.
[19] European Center for Disease Prevention and Control. Influenza case definition. Available at http://ecdc.europa.eu/en/activities/surveillance/eisn/surveillance/pages/influenza_case_definitions.aspx. Accessed May 1, 2011.
[20] Clemen R, Winkler R. Combining probability distributions from experts in risk analysis. Risk Anal 1999;19:187−203.
[21] R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2011:ISBN 3-900051-07-0, Available at http://www.R-project.org/. Accessed May 1, 2011.
[22] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86(2):420−8.
[23] Chen CC, Barnhart HX. Assessing agreement with repeated measures for random observers. Stat Med 2011;30:3546−59.
[24] Donner A, Wells G. A comparison of confidence interval methods for the intraclass correlation coefficient. Biometrics 1986;42:401−12.
[25] Van den Heuvel ER. A comparison of estimation methods on the coverage probability of Satterthwaite confidence intervals for assay precision with unbalanced data. Commun Statistics—Simulation Comput 2010;39:777−94.
[26] Nicholson KG, Kent J, Hammersley V. Influenza A among community-dwelling elderly persons in Leicestershire during winter 1993-4; cigarette smoking as a risk factor and the efficacy of influenza vaccination. Epidemiol Infect 1999;123(1):103−8.
[27] Kawai N, Ikematsu H, Iwaki N, Satoh I, Kawashima T, Tsuchimoto T, et al. A prospective, Internet-based study of the effectiveness and safety of influenza vaccination in the 2001-2002 influenza season. Vaccine 2003;21(31):4507−13.
[28] 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.
[29] Ozasa K, Kawahito Y, Doi T, Watanabe Y, Washio M, Mori M, et al. Retrospective assessment of influenza vaccine effectiveness among the non-institutionalized elderly population in Japan. Vaccine 2006; 24(14):2537−43.
[30] Voordouw BC, van der Linden PD, Simonian S, van der Lei J, Sturkenboom MC, Stricker BH. Influenza vaccination in community-dwelling elderly: impact on mortality and influenza associated morbidity. Arch Intern Med 2003;163:1089−94.

[31] Christenson B, Lundbergh P, Hedlund J, Ortqvist A. Effects of a large-scale intervention with influenza and 23-valent pneumococcal vaccines in adults aged 65 years or older: a prospective study. Lancet 2001;357:1008−11.

[32] Christenson B, Pauksen K, Sylvan SP. Effect of influenza and pneumococcal vaccines in elderly persons in years of low influenza activity. Virol J 2008;28:52.

[33] Manzoli L, Villari P, Granchelli C, Savino A, Carunchio C, Alessandrini M, et al. Influenza vaccine effectiveness for the elderly: a cohort study involving general practitioners from Abruzzo, Italy. J Prev Med Hyg 2009;50(2):109−12.

[34] Nichol KL, Margolis KL, Wuorenma J, Von Sternberg T. The efficacy and cost effectiveness of vaccination against influenza among elderly persons living in the community. New Engl J Med 1994;331(12):778−84.

[35] Nichol KL, Wuorenma J, von Sternberg T. Benefits of influenza vaccination for low-, intermediate-, and high-risk senior citizens. Arch Intern Med 1998;158(16):1769−76.

[36] Wang CS, Wang ST, Lai CT, Lin LJ, Lee CT, Chou P. Reducing major cause-specific hospitalization rates and shortening hospital stays after influenza vaccination. Clin Infect Dis 2004;39:1604−10.

[37] Nichol KL, Nordin J, Mullooly J, Lask R, Fillbrandt K, Iwane M. Influenza vaccination and reduction in hospitalizations for cardiac disease and stroke among the elderly. New Engl J Med 2003; 348(14):1322−32.

[38] Fleming DM, Watson JM, Nicholas S, Smith GE, Swan AV. Study of the effectiveness of influenza vaccination in the elderly in the epidemic of 1989-90 using a general practice database. Epidemiol Infect 1995;115(3):581−9.

[39] Shapiro Y, Shemer J, Heymann A, Shalev V, Maharshak N, Chodik G, et al. Influenza vaccination: reduction in hospitalizations and death rates among members of ‘‘Maccabi Healthcare Services’’ during the 2000-2001 influenza season. Isr Med Assoc J 2003;5(10):706−8.

[40] Tricco AC, Tetzlaff J, Sampson M, Fergusson D, Cogo E, Horsley T, et al. Few systematic reviews exist documenting the extent of bias: a systematic review. J Clin Epidemiol 2008;61(5):422−34.

[41] Welton NJ, Ades AE, Carlin JB, Altman DG, Sterne JAC. Models for potentially biased evidence in meta-analysis using empirically based priors. J Roy Statist Soc A 2009;172:119−36.

[42] Groenwold RH, Hak E, Hoes AW. Quantitative assessment of unobserved confounding is mandatory in nonrandomized intervention studies. J Clin Epidemiol 2009;62:22−8.

[43] Nelson JC, Jackson ML, Weiss NS, Jackson LA. New strategies are needed to improve the accuracy of influenza vaccine effectiveness estimates among seniors. J Clin Epidemiol 2009;62:687−94.