VALUE IN HEALTH 16 (2013) 604–609

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Comparative Effectiveness Research/Health Technology Assessment (HTA)

Estimating the Effectiveness of HPV Vaccination in the Open Population: A Bayesian Approach

Willem Woertman, PhD, Gert Jan van der Wilt, PhD*

Department for Health Evidence, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

ABSTRACT

Objectives: Estimation of the effectiveness of human papillomavirus (HPV) vaccination in the open population on the basis of published data from various sources. **Methods:** A Bayesian approach was used to reanalyze the data underlying a guidance by the Dutch National Health Insurance Board about the quadrivalent HPV vaccine Gardasil. Several studies document the vaccine's effectiveness in preventing cases in different subpopulations. None of these (sub)populations, however, is representative of the actual target population that the vaccination program will be applied to. We used a Bayesian approach for restructuring the data by means of reweighting the subpopulations by using HPV prevalence data, to estimate the effectiveness that can be expected in the actual target population. **Results:** The original data show an effectiveness of 44% in the entire population and an effectiveness of 98% for women who were compliant and were HPV-free at the start of the study. In the study population, the HPV prevalence was below 4%. In the relevant target population, however, the actual prevalence could be very different. In fact, some publications find an HPV prevalence of around 10%. We used Bayesian techniques to estimate the effectiveness in the actual target population. We found a mean effectiveness of 25%, and the probability that the effectiveness in the target population exceeds 50% is virtually zero. The results are very sensitive to the HPV prevalence that is used. **Conclusions:** A supplementary analysis can put together the bits and pieces of information to arrive at more relevant answers. A Bayesian approach allows for integrating all the evidence into one model in a straightforward way and results in very intuitive probability statements.

Keywords: Bayesian analysis, decision support techniques, evidence-based medicine, Gardasil, health insurance reimbursement.

Introduction

Policy decisions concerning reimbursement of drugs have crucially important implications for access to medical treatments for patients. Therefore, the available evidence on clinical effectiveness and cost-effectiveness of drugs should be carefully considered and synthesized when reimbursement decisions are made. This task is frequently hampered because the available evidence is incomplete or inconsistent. Also, the evidence may originate from multiple, heterogeneous sources, including randomized controlled trials, cohort studies, record reviews, registries, laboratory studies, and clinical and patients' experiences. It is rarely the case that there exists evidence that directly answers the questions that are most relevant for decision makers. Instead, there are usually bits and pieces of information available that answer subquestions that are relevant to the policy decision under consideration. Therefore, it often seems that there is a mismatch between the questions that policymakers are grappling with and the answers that science typically offers. Here, we argue that a supplementary analysis that combines all the relevant pieces of information (in the sense of a multiparameter evidence synthesis [1] or as in the confidence profile method [2]) might come closer to actually answering the main questions that decision makers are dealing with, and could therefore be very helpful in the policymaking process. Furthermore, we argue that Bayesian methods are well suited for such a supplementary analysis. Most of the literature on multiparameter evidence synthesis or the confidence profile method works within a Bayesian framework [1]. Moreover, it has often been suggested that a Bayesian approach to data analysis may be better suited than the standard frequentist methods for answering policy questions (e.g., [3–11]). There are two main reasons for this. First, a Bayesian approach offers a natural way for combining evidence from different sources in a systematic and transparent way, even when dealing with heterogeneous sources of evidence. In Bayesian statistics, unlike in frequentist statistics, it is very common to consider the new information that is gathered from an experiment together with the information that was available before the experiment. The Bayesian approach offers a formal

* *Address correspondence to:* Willem Woertman, Department for Health Evidence, University Medical Centre Nijmegen, P.O. Box 9101, 6500HB Nijmegen, The Netherlands.

E-mail: W.Woertman@ebh.umcn.nl.

1098-3015/$36.00 – see front matter

http://dx.doi.org/10.1016/j.jval.2013.01.001

model for combining prior information with newly available information, so that previously held judgments are updated. Second, Bayesian statistics has important conceptual advantages over frequentist statistics, making the outcomes easier to interpret and understand for relative laypersons (i.e., members of appraisal committees). For instance, the frequentist concept of the *P*-value gives an estimate of the probability of obtaining an outcome equal to or more extreme than the observed outcome, under the null hypothesis of no effect. This *P*-value, however, does not provide a direct statement about how unlikely the null hypothesis in fact is, nor how likely any alternative hypothesis. Arguably, however, this is precisely the sort of statement that the various stakeholders would like to be able to make: what is the probability that intervention *x* will produce an effect of *y* (or larger), given the observed results *z*? Bayesian analyses do produce such probabilities. Therefore, when used as a supplement to the standard frequentist results, perhaps Bayesian statistics could aid policymakers in comprehending and assessing what the data have to say about the questions that are most relevant to the problems they face.

In spite of these potential advantages, the Bayesian approach is relatively unfamiliar and relatively little used in the context of supporting policy decisions.

To put the alleged advantages of Bayesian methods for policymaking to the test, we performed a Bayesian reanalysis of an actual reimbursement advice that was drafted in 2009 by the National Health Insurance Board of The Netherlands (College voor Zorgverzekeringen [CVZ]), and compared the outcomes with the original results.

The Case of Gardasil

Gardasil is a prophylactic quadrivalent vaccine that prevents anogenital diseases associated with human papillomavirus (HPV) types 6, 11, 16, and 18. Infection with HPV is sexually transferable and can cause genital warts, intraepithelial neoplasia, and invasive cancers [12]. Of these diseases, cervical cancer is particularly important as it is the second most common cancer in women [13]. HPVs cause virtually all cervical cancers, and HPV types 16 and 18 cause approximately 70% of all HPV-related cervical cancers worldwide [13]. HPV types 6 and 11 cause most genital warts [12].

In 2007 and 2008, CVZ issued two advises to the Dutch Minister of Health about the reimbursement of Gardasil [14,15]. In 2007, CVZ recommended that Gardasil should not be reimbursed for 13- to 26-year-old women and girls [14]. CVZ acknowledged the therapeutic added value of Gardasil, but was not convinced of its cost-effectiveness. Shortly thereafter, the Dutch Health Council recommended including HPV vaccination in the national vaccination program for 12-year-old girls and that girls who were then 13 to 16 years old would also be eligible for vaccination [16]. After that, in 2008, the manufacturer of Gardasil asked for a reassessment, with the request to reimburse the vaccine for 17- and 18-year-old girls as well. For the same reasons as in 2007, CVZ advised not to reimburse Gardasil [15]. In both advices, one of CVZ's main points of critique regarding the cost-effectiveness model supplied by the manufacturer was the effectiveness of the vaccine that was used in the model. In the first advice, the cost-effectiveness model used the per-protocol susceptible effectiveness from one of the phase 3 trials. CVZ considered this assumption to be unrealistic and overly optimistic. In the second advice, in the new cost-effectiveness model supplied by the manufacturer, attempts were made to correct for existing HPV-16/18 prevalence, but CVZ maintained that the assumptions were not sufficiently supported by data.

In both advices, the evidence came from a variety of sources, and was analyzed with standard frequentist methods. The main articles that the advices refer to were a small phase 2 trial [17] and two large placebo controlled phase 3 trials (FUTURE I AND II [12,13]). The largest of these studies, the FUTURE II study, uses (the surrogate outcome measure of) HPV-16/18–related cervical intraepithelial neoplasia grade 2 or 3, adenocarcinoma in situ, or cervical cancer as the primary end point [13]. In the FUTURE II study, both the placebo arm and the vaccine arm participants received injections at day 1, month 3, and month 6 of the follow-up period. The main measure of effectiveness that was used is the proportion of events that are prevented through vaccination, given by one minus the vaccine event rate divided by the placebo event rate. The FUTURE II study contained three main analyses, corresponding to three different populations: 1) the per-protocol susceptible population of women and girls, who were uninfected with HPV-16/18 until 1 month after the third and final injection, who received all the injections at approximately the right moment, and who had no other protocol violations; 2) an unrestricted susceptible population of women who were uninfected with HPV-16/18 at the day of the first injection; and 3) an intention-to-treat population of all participating women in the study. All participants who belong to the first population also belong to the second population, and all participants who belong to the second population also belong to the third. Women were eligible to participate in these studies if they were not pregnant, if they did not report abnormal results on a Pap smear, and if they had a lifetime number of no more than four sex partners. Moreover, subjects were asked to use effective contraception during the vaccination period (day 1 through month 7).

Clearly, not all women and girls were eligible to participate in these studies. Considering the exclusion criteria, it seems likely that among the women who were excluded the HPV prevalence would be higher than among the women who were included in the study. Indeed, a higher number of lifetime sex partners is strongly associated with HPV prevalence [18]. Also, it is known that HPV vaccination has no therapeutic benefit for women who are already infected with HPV before vaccination. In contrast, for women who were uninfected at the time of vaccination, Gardasil is highly effective in preventing events with effectiveness near 100% [13]. Therefore, the HPV-16/18 prevalence before vaccination in the population that is to be vaccinated will be a major determinant of the effectiveness that will eventually be found. Because even in the intention-to-treat populations the HPV-16/18 prevalence prior to vaccination will probably be lower than the HPV-16/18 prevalence in the open population, the intention-to-treat effectiveness is also likely to be an overestimate of the effectiveness in the actual target population.

Therefore, if we could estimate the effectiveness that can be expected in the target population, based on the above-mentioned effectiveness estimate and on information about the HPV prevalence, we would be much closer to a satisfactory answer to the most relevant question: "What will be the effectiveness—and therefore cost-effectiveness—among all girls who would in fact be eligible for vaccination?"

Methods

Our supplementary analysis is mainly based on the data from the FUTURE II study [13], which provides estimates for the vaccine's effectiveness in preventing cases in three different populations. This article also reports the numbers of subjects and the number of cases (having at least one primary end point event) underlying the effectiveness estimates in each of these three populations.

We started our supplementary analysis by restructuring the three populations that the FUTURE II study considers (populations 1, 2, and 3 from the previous section) into three other, newly formed groups of participants, which we will denominate groups A, B, and C. Group A exactly equals population 1 (the per-protocol susceptible population) from the FUTURE II study. Group B consists of subjects who were included in population 2 (the unrestricted susceptible population), but who were *not* included