DRIVE
Development of Robust and Innovative Vaccine Effectiveness

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Title page | The results presented here are based on a limited number of sites using partially differing study protocols. They arise from a pilot season whose main objective was to build the DRIVE study platform for estimating brand-specific IVE in Europe. Due to the pilot nature of the study, the brands have been anonymized. The results should not be used to inform medical or regulatory decision-making. | Textual comment, proposed revised text: The results presented here are based on a limited number of sites using partially differing study protocols. They arise from a pilot season whose main objective was to build the DRIVE study platform for estimating brand-specific IVE in Europe not to obtain robust estimates of (brand-specific) vaccine effectiveness. Due to the pilot nature of the study, the brands have been anonymized. The results should not be used to inform medical or regulatory decision-making. | (1) Agree.<br>(2) Not needed.<br>(3) Agree, for this year, given limitations, results should not be used to inform regulatory or public health decision-making. | | | We don't think this is essential information for a lay summary, which we tried to keep as simple as possible. Personally (Kaat), I think a lay summary should be more simple than this one. How to write lay summaries is a topic for WP5 on communication. |
| 13 | Responsible parties - Report review (EFPIA) | The European Federation of Pharmaceutical Industries and Associations (EFPIA) partners of DRIVE (Abbott, GSK, Sanofi Pasteur & Seqirus) have all provided written comments. | All individuals are listed except EFPIA partners (company name). The review of the report does not necessarily represent the positon of the companies. We would suggest adding the name of the persons (scientists) who review the report (For Sanofi Pasteur Cédric Mahé, Laurence Torcel-Pagnon, Hélène Bricout, Clotilde El Guerche Seblain) | (2) Ok. | | | 'Governance model' doesn't seem plain English. |
| 15 | Lay Summary | Five study sites from four countries (Austria, Finland Italy, and Spain) participated in this first study. | To be into context I'd probably explain why there is only 5 sites participating for the pilot study and also explain why the sites did not used harmonise the protocol (essentially the short timeframe after project endorsement by IMI which allow limited flexibility to prepare in depth the study start). | (1) Yes, agree, useful context.<br>(2) Ok, but not sure this is needed. | | | Agree, it was unclear. We deleted the sentence as the distinction between TND and cohort is probably not understood by lay people. |
| 15 | Lay Summary | Although DRIVE is a public-private partnership which includes partners from both public sector and vaccine manufacturers, all studies were done in a separate working group consisting of organizations other than manufacturers. The results are also evaluated by an Independent Scientific Committee (ISC) which first convened in January 2018 | The statement may trigger some questions as such. I'd suggest a more neutral statement and provide some background information.<br><br>"DRIVE is a public-private partnership which includes partners from both public sector and vaccine manufacturers.<br>As per governance model, and in order to reassure public health instates partners, the study conduct was carried out in a separate working group, where only non-manufacturers were involved. The latter provided comments in written through a traceable review process, on the draft study reports.<br>The results are also evaluated by an Independent Scientific Committee (ISC) which first convened in January 2018 | (2) Ok, but not sure this is needed. | | | We agree there was a mismatch between the 'piloting objective' of the study and the objecitves as per SAP. We tried to make that distinction clear by adding the 'overarching objective' of the pilot season. "The overarching and primordial objective of this pilot study was to test the different operational aspects of the project including the IT infrastructure, the DRIVE governance for conducting IVE studies and streamlining key processes such as data collection, statistical analyses and dissemination of study results. (see page 19)" . [3]: We would stay away from presenting the results very different from what was agreed in the SAP. This would require an amendment of the SAP and it has been agreed to not to so (and rather spend our time on the preparation of next season). We agree that presenting the results across ages and settings is problematic (but done nonetheless otherwise there was not much to pilot). Your suggestions will be taken on board for next year. Thanks. |
| 15 | Lay Summary | Overall, 4999 patients participated in the studies, | Here is it unclear what you are referring too.<br>All study settings or only case control studies? For the sake of clarity, I'd explicitly distinguish and state:<br><br>1. TND<br>2. Cohort | (1) Agree, unclear.<br>(2) Agree. | | | I don't understand the comment. It is presented as a covariate for stratification. |
| 15 | Exec summary - discussion | Brand-specific IVE estimates could be obtained for some brands, | Brand-specific IVE estimates could be obtained for some brands, but not in a sufficiently robust manner to allow their interpretation and understand the factors that drive vaccine effectiveness as this was not the intend of the pilot season. | (1) Agree.<br>(2) Is logical consequence of being a pilot to build the platform. | | | This has been reformulated and better clarified in Figure 1 and Table 3. |
| 19 | Objectives | Primary and secondary objectives | It is unclear how these objectives are assessed in the analyses. There is no endpoint described in the document and it should be added to explain how each objective are expected to be evaluated. | (1) ISC already noted mismatch between formal primary and sceondray objectives in document with objective as stated in Background and lay summary.<br>(2) See previous comments: needs to be consistent throughout.<br>(3) Putting together such dissimilar results from a diversity of sites is a challenge. Thus, for this and future years, I would propose the following objectives: Overall VE, by outpatient and hospitalizations/ER visits, by age, and by brand (if possible). In regard to tables and figures, I would suggest: vaccine effectiveness (VE): (A) OVERALL: 1. Overall VE against outpatient visits, 2. Overall VE for inpatients(hospitalizations/Emergency room visits); 3 Overall for all outcomes (less useful, though) ; 4. Each of the above by age (age groups for the pilot year are fine, but for subsequent years, hopefully with more power, I would use 6 months to <9 years (because of the two dose recommendation), 9 to <18 yrs, 18 to 49, 50 to 64, and 65+; (B) VE by type/subtype: 1. A(H1N1), 2. A(H3N2), 3. B/Yamagata lineage; 4. B/Victoria lineage; 4. all A, 5. all B; then each one again stratified by age, hospitalized and outpatient. For subsequent years, if power is sufficient, add analyses by Vaccine type and Brand, including each of the subanalyses I mention above. Analyses by prior year vaccination and by time since vaccination might be nice additions. | | | Added 'lineage'. In Austria the Centre of Virology characterizes influenza viruses beyond subtype/lineage for their surveillance system. |
| 19 | Secondary objectives | To estimate seasonal overall IVE by any influenza vaccine, stratified by host-related covariates:<br>• age group (6 months – 14 years, 15 – 64 years and 65+ years)<br>• presence of at least one chronic condition (yes versus no, see also Section 9.8.2)<br>• vaccination status in previous season (yes versus no) | In the document this is unclear how the prior vaccination is assessed. It is described as a covariates and a confounding factors but here it is presented as an outcome to be studied. | (1) Agree.<br>(2) Presented as a covariate here. | | | Please refer to Figure 2. |
| 19 | 2.2 Secondary objectives | | **Clarify that the piloting is not complete and further development work remains.**<br><br>The pilot study is not complete in all aspects. The report should make clear that while the brand specific VE were obtained, many limitations still apply to getting at robust and interpretable VE estimates; further and specifically the conclusion should reinforce [or, repeat] that.<br>One major change we note has occured in the final applied SAP in that the "by brand and type" as well as the strat analysis by design were removed from all the sec objectives. As a result - assessing the feasibility of being able to perform analysis that are key to understand the driving factors of VE at the type and brand level as well as the differences by design has not be achieved. | (2) Ok. | | | When the information on the ILI symptoms was available (not available for all study sites), we checked whether the ILI case definition was indeed met. |
| 19 | 2.2 Secondary objectives | Presence of at least one chronic condition (yes versus no, see also Section 9.8.2) | Presence of chronic conditions is of major interest because these persons have an indication for influenza vaccination. This "yes versus no" is to be deleted in future reports, for two reasons: 1: the category "no chronic conditions" is of no particular interest. 2: The comparison "chronic conditions" versus "no chronic conditions" suggests that the comparison is informative on the effect of chronic conditions on the risk of influenza. It is not. For example, the two groups are very different with respect to the age distribution. A proper comparison would require a network meta-analysis approach, which is very complicated. For this reason "versus" presentations are to be avoid. | (1) Needs discussion - not sure I understand this point.<br>(2) Similar to the discussions/comments we had in Valencia, agree.<br>(3) Agree with the comment. "Yes vs no" chronic conditions is not very informative. Also, the detection of a chronic condition (particularly among the elderly) is sometimes just a reflection of health-seeking behaviors. | | | |
| 19 | 2.2 Secondary objectives | Vaccination status in previous season (yes versus no) | Vaccination in the previous season is a know confounder of IVE. It is surprising that Finland was not identified as an influential estimate. | (2) ? | | | No, this refers to missing/incomplete vaccination records. Partially vaccinated children were also excluded (see section 3.9.1) but that is a seperate matter. Clarified in the text that 2017/2018 refers to the season studies and the other years to prior seasons (these are two distinct variables in the data). |
| 21 | 3.1 Overview study site characteristics | Figure 1 | What is the difference between swab systematic and Swab all? | (1) Agree.<br>(2) Same questions, to be clarified. | | | There is no strong rationale for having the cut-off point at 14 years; a cut-off was needed to separate children from adults. |
| 22 | 3.2 Catchment population | Influenza virus positive nasopharyngeal swab samples are further analysed to identify the type, subtype and strain. | B-lineage as well?<br><br>What do you mean by strain? | (1) Agree.<br>(2) To be clarified. | | | Added that this information is available for all sites. Full details on data can be found in Appendix 9.1. |
| 23 | 3.3 Study period | For the register-based cohort study, the study period was defined as starting from the early stage of the epidemic until the end of the study (Figure 2). | The study period for the register based cohort is not clear. | (1) Don't understand - the study period for Finland it is shown in Fig 2 along with the other study sites using the TN method - clear for me.<br>(2) To be clarified. | | | Indeed, ideally it should have been done against all strains, and this will be considered for next season's analysis. |
| 25 | 3.5.2 Case definition verification | ILI case definition could be verified based on symptoms for Spain Valencia and Spain La Rioja. | What do you mean by verifying ILI case definition based on symptoms? | (2) Seems clear to me. | | | The covariates kept in the final site-specific regression models are summarized in Figure 9. For next year, we will do this differently and force some covariates in the final model (see p65; experiences re: data analysis). |
| 26 | Exclusion criteria | Note: a patient could be selected several times as long as he/she did not have a previous laboratory confirmed influenza for the current season | What does it mean pragmatically?<br>That a control can serve a control for several cases? Or something else? | (1) Agree, some further explanation here would be useful.<br>(2) Seems clear to me. | | | Agree, therefore the results on the crude VE are only presented in the Appendix. However, I still think it is informative to be able to informally investigate the impact of confounder adjustment by simply visually comparing crude and adjusted IVE |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 26 | Exclusion criteria | • Subjects with incomplete vaccination records for 2017-2018 and 2014-2015, 2015-2016 and 2016-2017 | Does it mean that the subject is excluded if he/she has partial vaccination records?<br><br>Why on one hand 2017/18 & 2014/15 and on the other hand 2015/2016 & 2016/2017? | (2) To be clarified. | | | The text has been modified to capture the aspect of sample size as well. |
| 31 | 3.8.1 Age groups | Age was categorized into the age groups 6 months to 14 years, 15-64 years | Could you please remind the rationale for those age range for kids and younger adults? | (2) Ok, but not sure this is needed.<br>(3) Age is an important effect modifier that should be included in all models, if possible. Ideal age groups for flu studies are 6 months to 8 years (because of the recommended 2 dose schedule), 9 to 18, 19 to 49, 50 to 64, 65+, although I assume this was not feasible in this first year, which is ok. Also, for future years maybe a separate category for "pregnant women" can be included (if power is sufficient). | | | Agree, these are all important topics for further discussion. We will have plenty of future opportunities to discuss these topics (EFPIA brainstorming, SAP next season). |
| 31 | 3.8 Risk groups, confounding factors and effect modifiers, other variables | The following covariates were used: age group, sex, presence of at least one chronic condition, number of hospitalizations in previous 12 months, vaccination status in previous season | Not clear why these covariates were selected, confounder, effect modifier, both? | (2) To be clarified. | | | They are available on SharePoint (at least for the ISC members). Next time, all appendices will be shared by mail as well |
| 32 | 3.8.4 Vaccination in the previous season | Influenza vaccination in the previous season was categorized as yes/no. | Which sites did have this information? | (1) Agree - summarise this in main report.<br>(2) Should be clear for all key variables.<br>(3) Agree, prior year vaccination is a covariate/confounder of interest. Nonetheless, my fear is that not all suitable sites will have this information, and the existing studies on the effect of prior vaccination are not conclusive. Thus, for future years, I would not include this (interesting) covariate as a study participation requirement, it should be interesting to have it, but it is not a must. | | | Agree, these are all important topics for further discussion. These topics will be rediscussed during the EFPIA brainstorming and the review of the SAP for next season. |
| 32 | 3.11.3 Step 2: Meta-analysis | We conducted a random effects meta-analysis.... | The weighting method should be mentioned. The analyses here use inverse-variance weighting. If this is the best weigting scheme is to be disccused for the next SAP. | (2) Ok. | | | We fully agree. However, we were not having the names of the EFPIA reviewers. We will collect that information from the EFPIA partners directly. |
| 33 | 3.10 Sample size considerations | NA | Sample size considerations are not well addressed in the report – it does not address what is needed to display a VE. All VE are computed whatever the heterogeneity and the size. | (1) Yes may require a disclaimer about this being done to test the pooling method and that results are not statistically valid due to heterogeneity.<br>(2) Not essential for this pilot. | | | The topic of minimal sample size will be re-discussed during the EFPIA brainstorming. The following disclaimer is on the front page of the report: *DISCLAIMER: The results presented here are based on a limited number of sites using partially differing study protocols. They arise from a pilot season whose main objective was to build the DRIVE study platform for estimating brand-specific IVE in Europe. Due to the pilot nature of the study, the brands have been anonymized. The results should not be used to inform medical or regulatory decision-making.* |
| 33 | 3.11.2 Step 1: Site-specific estimates - Differences in covariate adjustment (pp. 33 - 45) | NA | The logistics regression should include the same minimal core variable for all sites (e.g. age, chronic condition and nbr hospitalization) for biological plausibility | (1) Agree.<br>(2) Not essential for this pilot. | | | Agree, this will be done for next season. However, it was not done for this season as it was unclear which covariate information would be available. |
| 33 | 3.10 Sample size considerations and 5.4.1 Considerations for results interpretation (pp. 33, 44) | For details please refer to the Annex 3 of the SAP (Appendix 9.2).<br>AND<br>The small number of studies included in the meta-analysis in the pilot year limited the number of possible brand-specific estimates and stratifications (e.g. by both brand and age, or brand and setting). | **Sample size requirements and criteria for robust VE are inadequate as currently defined**<br><br>The SAP for 2017-18 refers to Annex 3 for the sample size considerations and in Annex 3 the approach to the min. detectable VE is described - however not calculated. It is not clear why these are not accounted for. The report and the response document to EFPIA comments states in certain places that analyses were not performed based because because of "too wide" confidence intervals - whereas in the report VE are presented with very wide confidence intervals of >50%. This inconsistency is not acceptable and highlights the limitations of the ambiguity around sample size requirements. | (2) Presumably CI can be wider than the ones reported. | | | Agree - it would have been better to have the information on the study sites in the forest plots analysis-specific (now it is only site-specific). It seems like a little job, but requires quite some re-programming. We will do this for next year. Thanks. |
| 34 | 3.11.5 Sensitivity analyses | No sensitivity analysis for outlying and influential estimates was conducted, since none were identified | It is surprising that Finland was not identified as an influential estimate. | (2) ? | | | I (Kaat) disagree – I would not use the I2 statistic to decide on pooling or not. A high I2 statistic is a reason for using a random effects model (and not a fixed effects model) and is not a reason for pooling or not. IVE estimates can be very heterogeneous for reasons that cannot be controlled or are difficult to control (differences in healthcare use between countries, circulating strains). Still, we would be interested in an overall estimate that reflects the heterogeneity - this could be achieved using a random effects meta-analysis. Also the other way around, a low I2 statistic is not a guarantee for pooling. Assume the IVE happens to be similar across healthcare settings for a given season (maybe unrealistic?), would you then pool? The I2 criteria is a measure that reflects heterogeneity between sites (relative to the within-site heterogeneity), and as such an informative measure, but I would not use it to decide on pooling or not (though it has been presented like this in the past). Rather, I would decide based on 'epidemiological' arguments whether to pool or not (not pooling across healthcare settings, age groups, ...), use the I2 statistic as a measure to reflect heterogeneity and always use random effects meta-analysis (even in case the I2 would be low, as in that case, the random effects meta-analysis will boil down to a fixed effect analysis anyways) |
| 34 | 3.11.5 Sensitivity analyses | | With the recent introduction of Estimands, the view of what or proper sensitivity analyses and what are not has changed. For the next SAP: what are propoer sensitivity analyses. | (2) Ok. | | | We agree that the results are 'unexpected' and that it would be nice to explore the topic of waning protection further. For now, we followed the SAP and explored methodology. |
| 35 | 3.11.6 Deviations from the SAP | The exploratory objective on the waning of the vaccine effect is performed based on VE against AH1N1 because the highest VE was observed for this strain. | This is not a scientifically relevant argument to limit the analysis to VE against H1N1 strain. | (1) Given the increase in VE with time since vaccination it might be of interest to look at the other strains in a similar way.<br>(2) To be clarified.<br>(3) The dominant strains were B/Yamagata and A/H3N2. Ideally, they should be the ones analysed for waning immunity (if power permits), but current results suggests low to very low VE, so I am not clear if this particular season is suitable for an analysis of waning immunity for the dominant strains.. Maybe this analysis could be done instead in a more suitable season (with higher VE against the main circulating strains? | | | We indeed didn't want to embark on a 'epidemiological' discussion of the results as the main objective of this pilot system was setting-up the system. Therefore, we often used as reply "the report has been substantially revised" as we acknowledge that the previous version of the report was too much focussed on the epidemiolobgy whereas the focus should have been the 'setting up the system'. Many of the comments raised or good suggestions that will be taken on board for next year. For example, next year we will not pool across health care settings and we are trying to disentangle the information on primary care and hospital cases (e.g. Austria was a mixed setting this year, will be able to provide information on primary care and hospital cases next year). Our focus is improving the data collection for next year, not re-analysing the data of the pilot season. |
| 45 | Differences in covariate adjustment | | It is unclear how the covariates in the final model were retained or dropped the stat explanation is missing. | (2) To be clarified. If I remember correctly, it is mentioned in the methods, and in the discussions, but in the results this was a bit isolated and hard to grasp immediately.<br>(3) For the final model, I would retain essential covariates regardless, including age group. | | | I agree that continous improvements will have to be made throughout the course of the DRIVE project. The 'by type and brand' analysis were always primary objectives. The 'by design' objective was never truly considered (maybe still part of the wrongly shared initial version of the SAP) as we only have one cohort study (and potentially excluding one estimates was part of the outlying/influential analysis). |
| 45 | 5.4.1 Considerations for results interpretation | Differences in covariate adjustment. | In the report the distinctions confounder versus effect modifier more or less disappeared. This is very confusing. Number of hospitalizations is a marker for which confounder? In TND studies and nested case-control studies calender time is usually a confounder. Inclusion to be discussed in next SAP. | (2) Ok. | | | The discussion on thresholds for analyzing/reporting IVE estimates is still ongoing. I (Kaat) personally disagree to pre-specify a minimal sample size or CI width. The sample size calculations clearly demonstrate that it is practically impossible indeed. The sample size depends on many unknown factors, including attack rate, brand-specific coverage and between-study heterogeneity – all factors that are impossible to control, educated guesses are still possible, but even then, within the range of plausible educated guesses, sample sizes vary widely. I also fail to understand why it is needed as confidence intervals reflect the statistical uncertainty, and the IVE estimates should always be interpreted jointly with their CIs. I agree that CIs are poorly understood by lay people, so I would not communicate widely about results with a wide CIs, but still, I feel results with wide CIs can be mentioned in a report on a pilot study testing methodology. I also agree there is some inconsistency here - although we calculated the IVE for all brands, the CIs for the brands that were not reported were either extremely wide (let's say width > 200%) or the CIs could not be estimated. For next year, as a matter of transparency, I would opt to present all brand-specific results irrespective of the width of the CI (but this need to be further discussed). |
| 46 | 5.4.2 IVE by any vaccine and by vaccine brand -> IVE by healthcare setting (pp. 46-57) | NA | Forest plots: the %vacc in child, adulty and elderly should be updated based on the analysis performed (it is always the same per site). The #LCI per analysis/line should be included to better reflect the size/precision (what is the minimal #LCI to allow a results display?) | (2) To be clarified. | | | Table 15 is purely qualitative, and therefore not included in the SAP. The brand-specific information was given in the SAP, though it was indeed not specified we would use pie charts. We will be more explicit about how we will present the brand-specific information next year. |
| 46 | 5.4.2 IVE by any vaccine and by vaccine brand -> IVE by healthcare setting (pp. 46-57) | NA | The heterogeneity I2 in forest plots is not taken into account. Results are pooled whatever the I2 result | (1) See above - needs a disclaimer.<br>(2) Not essential for this pilot. | | | Added a section 'heterogeneity across studies' to the discussion. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 46 | 5.4.2 IVE by any vaccine and by vaccine brand | | I like the presentations by age agroup. In fact, with one exception, all presentations should be by age group, because 1: age is a known effect modifier, and 2) the youngest and the oldest age groups are age groups of special interest (for example, being 60+ is in many countries an indication for influenza vaccination.) Forest plots should also be by age group, overall IVE estimates should not be given. | | | | Not satisfactory for any of us. [The ISC were asked to indicate whether they agree with EFPIA comments and in the case of disagreement provide their rationale. The review process will be reassessed for the future.] |
| 47 | 5.4.3 IVE by vaccine antigen (live attenuated, inactivated) | IVE by vaccine antigen (live attenuated, inactivated) | To be skipped in future reports. Such comparisons are not requested in the EMA Guideline. Furthermore, the comparison is not based on a statistical analysis, and thus biased. See also my comment in line 13. | | | | Agree - I-MOVE reports similar measures of heterogeneity as we do (despite years of harmonization by I-MOVE). Random effects meta-analysis (complemented with meta-regression in case of a sufficient number of studies) is a nice way of dealing with heterogeneity as well as being more selective in the studies [edited for brevity]. Many things to discuss still.<br>An introductory sentence was added to the discussion section on generic protocols mentioning the limitation. Added a section 'heterogeneity across studies' to the discussion. |
| 47 | 5.4.4 IVE by vaccine antigen (subunit, split virion) | | Same comment as previous. | | | | There was a wrong version of the SAP shared once, explaining the inconsistencies. Agree that time since vaccination/time during the season should be furhter investigated. |
| 47 | 5.4.2 - 5.4.4 | | Too many comparisons, only a few are relevant for our mission. Relevant: 5.4.2, 5.5.1, 5.5.4. | | | | Agree - but not much more we could do the first year. Propensity scores will be investigated for the next seasons. |
| 48 | 5.4.6 IVE by vaccine type (adjuvanted, non-adjuvanted) | | Same comment as previous. Furthermore, the adjuvanted vaccine is a subunit vaccine. Here the comparison is subunit + adjuvant versus subunit, split, etc. Valid comparison. | | | | We suggest to keep the disclaimer as is: "The results presented here are based on a limited number of sites using partially differing study protocols. They arise from a pilot season whose main objective was to build the DRIVE study platform for estimating brand-specific IVE in Europe. Due to the pilot nature of the study, the brands have been anonymized. The results should not be used to inform medical or regulatory decision-making." |
| 57 | 5.5.5 | IVE by healthcare setting. | Section number missing. | (2) Ok. | | | Rephrased: 'Preliminary brand-specific IVE estimates could be obtained…' |
| 57 | 5.5.5 | IVE by healthcare setting. | To be discussed for nest SAP: can results of different studies be pooled? | (2) Ok. | | | yes vs no': We will replace with 'yes or no', as we don't intend to make a comparison of both groups, only adjust for chronic conditions. The presence of chronic conditions is indeed acknowledged to be a major confounder. A working group has been established to better define chronic conditions. Also, we are investigating wheter some vaccine brands are preferentially used for some risk groups. |
| 58 | 5.6.1 Comparison of 1-stage and 2-stage pooling approaches | The results of the comparison of pooling approaches were in line with expectations based on statistical theory. | Nice and interesting exercise, but given the statistical theory needed in a report like this? | (2) ? | | | We checked the influential estimates and confirm that Finland was not identified as an outlying influential estimate. This is most likely explained by the lack of statistical power of these tests in case of a limited number of studies. We acknowledge this lack of power, but implemented these tests nonetheless as more study sites are expected to be included over the next years. |
| 61 | Table 16. Influenza vaccine effectiveness against AH1N1, crude and adjusted estimates | | The reliability of such analysis is questioned (+ See comment p35) | | | | These are the covariates identified among a longer list of covariates as being part of the minimum data requirements in the protocol. They were chosen as they are the ones commonly adjusted for in the analysis. They were treated as possible confounders (age was also treated as an effect modifier in the analysis stratified by age). Nr of hospitalizations in past 12 months is used as a proxy for severity of chronic disease. |
| 61 | 5.6.2 Time since vaccination | NA | The analysis on time since vaccination should be dropped. It does not account for flu circulation timing and the results are not biologically plausible. | (1) Don't agree - it is piloting the method. Maybe a bit more discussion about the results is warranted though.<br>(2) Not essential for this pilot. | | | We applied standard meta-analysis using random effects inverse variance weighted averages with a moment estimate of the between-study variances. Added in section 3.13.3. I understand the reason for deviating from the standard approach, though it requires a very good rationale and 'objective' alternative weighting criteria. To be discussed, though it will be challenging to deviate from the 'standard'. |
| 61 | 5.6.2 Time since vaccination | | **Time since vaccination analysis is not biologically plausible.**<br><br>The results from this analysis is not biologically plausible and is too minimally described in the SAP to assess the robustness of this approach. Also the approach to the analysis was removed at some point from the SAP, but then reappeared in the final SAP and thus could not be properly reviewed. | | | | We checked the influential estimates and confirm that Finland was not identified as an outlying influential estimate. This is most likely explained by the lack of statistical power of these tests in case of a limited number of studies. We acknowledge this lack of power, but implemented these tests nonetheless as more study sites are expected to be included over the next years. |
| 61 | 5.6.2 Time since vaccination | Time since vaccination. | Important issue but much to complex for a report like this? | (1) Depends on the audience for this report - for which I remain unclear.<br>(2) Can stay. | | | Recent introduction of estimands? Not clear what you mean - better understanding of what should be our estimands? Agree that we should re-discuss the sensitivity analysis. |
| 64 | Experiences and next steps: brand-specific information | sites with a high diversity of brands should be preferentially included whenever possible. | Having a broad rep. of brands is one aspect but the counterpart is to have sufficient subjects to run the brand specific VE estimates and this is not captured here.. | (2) Ok. | | | Agree - to be re-discussed. |
| 65 | Conclusion | | What might need to be potentially discussed is:<br><br>- Whether or not data from different healthcare setting is recommended.<br>- Any emphasis to recruit further age groups? Outputs by age are not very informative<br>- Should we consider time since vaccination or calendar time?<br>- Explain why we opted to not match the cases and control despite it would certainly help to account for part of the heterogeneity by sites (this has been raised several times, and it can be tackled separately, but it may interesting areas for discussion).<br>- Do we want also to mention the a-priori level of precision to run or not run separate analysis? Number of subjects, wild confidence interval<br><br>In the context of the analyses based on a priori criteria, it would have been useful to have the appendices attached together with the document in particular the SAP for pooled analysis and SAP site specific as well as the additional tables and figures<br><br>As well, it is important to acknowledge that during the SAP development no a priori thresholds have been defined to decide upfront whether analysis could be performed or not. | (2) Don't understand the 'we'? | | | Agree - for next year. |
| 65 | 6. Discussion | The ISC decides on integration of EFPIA comments or justifies non-integration. This year, ISC and EFPIA review occurred in parallel, however an update to the process has been proposed. | **Improved process for review and comments**<br><br>In addition to the limitations described, we note that the minority of the comments have received a response from the ISC. In addition, we should consider standards for comments. For example textual comments can be critical for interpretation but overall the language in the reports falls short on quality - yet we are requested to refrain from such comments. | (1) This is because we were instructed to only respond to EFPIA comments that we disagreed with. The role of the ISC in relation to reviewing the EFPIA comments still needs better definition as the current process is not satisfactory from either the EFPIA or ISC's perspective in my view. | | | Agree -to be re-discussed. |
| 69 | 9 Appendix list | | In the context of the analyses based on a priori criteria, it would have been useful to have the appendices attached together with the document in particular the SAP for pooled analysis and SAP site specific as well as the additional tables and figures | (1) Strongly agree!<br>(2) Ok. | | | Agree here too. |
| 30/43 | Table 9, Table 15, Figure 6 | | **Brand data presentation beyond the purpose of the pilot season and as per SAP should be removed.**<br><br>Table 15 and figure 6 are not in line with the SAP or the descriptive analysis described in section 3.1.1 of the report. The purpose of the pilot to test the framework. These tables (including table 9) and figures go beyond this purpose and beyond the SAP and should be removed. Considering the feasibilty to capture brand VE for as many brands as possible, the interest is to know accross all sites how many brands could be captured, but this purpose is already achieved in tables 13 and 14 - which are in line with the SAP.<br>In addition - as a consequence of presenting this data the de-identification of the brand is incompleted as for some vaccine types there is only one vaccine and in some countries there are limited vaccine brands available. This also applies to the presentation of VE by vaccine type. | (3) For future seasons, if power is sufficient, I would suggest identifying the brands.. | | | Agree here too. Historical note: the main reason of having a 'layered' primary objective (1.by vaccine brand, 2a. by influenza vaccine type: by vaccine antigen (live attenuated, inactivated, subunit, split virion),2b. by valency (number of vaccine virus strains) and by adjuvant (adjuvanted vs. non-adjuvanted),3 by any influenza vaccine) was the concern of not having sufficient sample size to do the brand-specific IVE estimates (as we often didn't know which brands were going to be used). However, I can also imagine that some of the objectives are also useful for other stakeholder groups (but should not be the DRIVE primary objective). We agree that for DRIVE, the brand-specific VE should be the only primary objective. The objectives will be re-discussed next year. |
| General /Discussion | General/Discussion | | **Challenges and limitations of the single stratified analysis presentation to be reflected as lesson learned**<br><br>There is considerable limitations to single stratified analysis for the interpretation which can lead to mis- or overinterpretation, especially where strata are not mutually exclusive or where stratified analysis is a must (i.e. presenting brand-specific VE independent of their age indication/or use is not appropriate - VE simply differs by age groups and different recommendations exist in these groups.).<br>In addition we experience the challenge to reflect the differences between the individual studies in the results presentation. These points deserve to be recognized in the discussion. | (1) Might be appropriate for year 2.<br>(2) Ok. | | | Agree |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NA | NA | NA | General comment: this report has been extremely improved and well presents now the first season as a pilot with lessons learnt and expectations for the next season. Nevertheless, there is still an important issue that affects the credibility of the analysis and the project added value compared to similar initiatives: minimal sample size and level of heterogeneity are not accounted for before producing estimates. This is a key component of any robust SAP as per industry standard. | (1) This comment may reflect the mismatch between the broad objectives of piloting the method and providing VE estimates since pooling data even if not statistically justified was done I assume in order to pilot the method.<br>(2) Not essential for this pilot.<br>(3) Agree. | | | Added. |
| Several locations incl. page 14, 15, 16, 62, 65 | Several sections, but specifically summary and discussion/ conclusion | | **Protocol harmonization will not be able to deal with all differences - – considerations for interpretation of VE in presence of differences and additional considerations for the approach to the analysis should be equally considered.**<br><br>The text reflects general over-expectation from the harmonization of the protocols which is planned in the next season. The next season may be better aligned accross some sites, but other sites are introducing new "heterogeneity" from a population perspective (i.e. HCW, pregnant women etc). Several differences between the sites/locations will always remain outside of our control. There should be equal attention to how we deal with these differences in the analysis - i.e. appropriate analysis by subgroups/applying sensitivity/ applying restrictions etc. Criteria for pooling data in general (if differences are so extensive and VE can be estimated at the local level - should the data be pooled if that leads to comparing apples to oranges?). | (1) Good point.<br>(2) Agree, to be clarified. | | | Agree, better to not pool across healthcare settings. For next season, we are trying to disentangle primary care cases from hospital cases for those sites who couldn't do this (mixed healthcare settings). Next year, with more study sites participating, we will be probably able to stratify all analysis by healthcare setting |
| Throughout sections | Methods and discussion | | **The report is incomplete wrt to the description of the limitations and lessons learned.**<br><br>We note that the report has been significantly improved to reflect the pilot nature of the 2017-18 season. However, we noted that there are a number of important limitations, lessons learned and challenges which have been noted in the review (and similarly highlighted at the Annual Forum presentation) but have not been noted as limitations/lessons learned in the report. These include at min.:<br>- Analysis of data in line with the recommendation and age indications of the vaccines<br>- The ability to stratified by multiple factors to allow appropriate interpretation of VE<br>- To give considerations to the match between the circulating strains against those included in the vaccine in the calculation of VE, at the individual vaccine type/brand level.<br>- Criteria for pooling data - more attention to the question of the legitimacy of pooling (or lack thereof) depending on the heterogeneity of data, but not only from a statistical perspective. Specifically appropriateness and approach to pooling hospital and GP data should be further vetted.<br>- Relatively crude approach to the adjustments for potential confounders (dichotomization of variables, use of proxies for frailty etc) and understanding the applicable confounding and the potential differences by vaccine type and brand - starting with the presentation of the descriptives by brand. No investigation of residual confounding as applied.<br><br>Many of the previous comments on these aspects have been replied with that "the report has been substantially revised" or that the "analysis followed the SAP". The alternative and more valuable approach would be to incorporate applicable caveat/limitation/learning in the updated report. | (1) Report's primary and secondary objectives are the problem here I think since this first year was a pilot of the methods.<br>(2) Can be included in discussion. | | | We agree that the results are nicely demonstrating statistical theory, and hence, not so informative for statisticians. Nonetheless, it was a nice exercise and good to have all code ready for next seasons. We think such an exercise fits well in a report on a 'pilot study' to test methods. |
| | | | **Minimal set for confounder adjustment**<br><br>The analysis should include the same minimal core variable for all sites. Also the approach to the confounder adjustment is crude: dichotomization of variables, use of unvalidated proxies for frailty etc. Descriptives should support the understanding of presence of potential confounding and the potential differences by vaccine type and brand. Investigation of residual confounding should be considered. | (1) Agree. | | | We think it fits well in a report on a pilot study to test methods. We agree that waning vaccine protection/changing VE over time is a very interesting topic and believe it is an essential part of understanding IVE and the disease burden prevented by vaccination. To discuss further. |