DRIVE D7.6 Brand-specific influenza vaccine effectiveness in Europe season 2018/19

EFPIA comments, 2nd round (23/08/2019) – reply grid

No	Comments from EFPIA	ISC reply	WP7 reply
1	The rationale for stratifying all results by age and setting should be added. It had been thoroughly discussed but does not appear in the report. A few lines would bring added value.		Added to background: Age is an effect modifier in IVE studies. In addition, patients presenting to different healthcare setting (primary care vs. hospital) are expected to have different levels of disease severity. To reduce clinical heterogeneity, estimates in the 2018/19 season were stratified by age and setting wherever possible.
2	For transparency do we want to make the R program also available as the SAP and report?		Ideally yes. But we would prefer to prepare this for next year after all proper documentation and code cleaning has taken place. The way it is now, it is very difficult to follow the code unless you have the databases. We could share this year some code examples. Still, code could be made available if requested.
3	If the analyses are performed centrally based on individual level data (which was not the initial assumption), should we explain why not pooling the data into a single dataset and run the pooled analysis instead of using the meta-analysis approach? →An explanation in the stat description section why a two-stage pooling (use of individual effect estimates) was performed and not a one stage pooling (use of individual subject data)?		Individual-level data was centrally analysed also in the pilot year. In the pilot year a sensitivity analysis was done showing equivalence between one-stage pooling and two-stage pooling. The advantage of two-stage pooling is that it enables the incorporation of data from sites that are only willing to share aggregated data, as well as combining results from TND studies and cohort studies.

			Added to section 3.13.2.3: Meta-analysis is preferred over individual-level data pooling for DRIVE, as it enables the (future) incorporation of data from sites that are only willing to share aggregated data, and allows pooling of results from TND and cohort studies. Equivalence of the two approaches was demonstrated in the pilot year.
4	How do you determine that a IVE is outlying/influential? It is explained in the SAP, but a few lines should be consider for the report to avoid continuous back and forth with the SAP		This is already written in the report, see last 2 paragraphs in section 3.13.2.3 "Studentized deleted residuals r were used to identify outliers in the meta-analysis. Site-specific IVE estimates were considered outlying from meta-analysis when $\|r\| > 2.5$, where $\| r \|$ indicates the absolute value of the residual."
5	The reference to I-MOVE should be further clarified. →Either, I-Move is considered as the gold standard in Europe and this should be clarified and other ref should be added (in particular because the I-MOVE data are interim estimates).		Renamed this "another European network" as officially the I-MOVE funding has stopped. No end-of-season 2018/19 IVE estimates from European countries have been published in peer-reviewed literature. Added provisional end-of-season UK estimates from PHE report
6	On Table 3, the number of subjects in TND studies/registered based cohort is presented by case/control status together with the vaccine coverage. It would be informative to have the vaccination update among cases and controls and not overall.		Added vaccine coverage among cases and control separately.
7	Considering a 8 days period between symptom onset and swab collection increases significantly the risk of false negative outcome and bias the VE. A rationale should be added for this choice.		An 8-day period was chosen to maximize sample size. A sensitivity analysis excluding swabs taken 4 days or more after ILI/SARI onset was done (see section 4.4.2.4 and Annex 15). The results for the sensitivity analysis are in the

			same line as the ones for the main analysis. The CI of the sensitivity analysis results are wider as less ILI-cases were kept when requiring a shorter period between symptom onset and swab collection.
8	Why excluding patients who underwent under antiviral therapy instead of adjusting for it? ➔ A rationale should be added		It was the choice of the Romania NIID site to exclude patients who received antiviral therapy. This was not a general exclusion criterion/covariate.
9	From the list of confounders presented in the report, based on the variables considered, the only variable that needs to be "forced" a priori in all models in the smooth function of age. ➔ An explanation to consider all covariates a priori should be briefly discussed ➔ For next year, the other covariates to be defined based on their level of influence per strata.		The SAP is followed for the analyses. Note that last year no covariates were forced into the model and were selected by backwards model selection. Following comments on this approach this was changed to forcing all covariates into the model. Strategies for confounder-adjustment can be rediscussed at the time of SAP development.
	Additional specific comments (mostly on new content highlighted in blue)		
10	Page 6 Exec Sum - The DRIVE platform is still expanding, and not all vaccine brands used in Europe are covered during the 2018/19 season. - Add: Nor was sufficient sample size achieved to estimate brand specific VE for all brands.		Added.
	Page 8 Ten vaccines were licensed in Europe in 2018/19 and seven brands were included in the DRIVE dataset (Table 2). To clarify whether "included" means – exposures were identified, but not necessarily was VEffect calculated?		Changed 'included' to 'identified'.

	Similar for Table 2 Is this intended to reflect the brands for which exposure was identified or for which VEffect was estimated? If the latter than it is not clear why vaccines are listed for which no VE was estimated (i.e. Afluria).		
11	page 8 just above table 2: In most countries, type-specific vaccine recommendations were in place for specific risk groups. Add: and age groups (i.e. those 75 years and over for example).		Added “and age groups”.
12	Page 9 top - Strain circulation: ECDC defines dominance as follows: **Dominant virus** reports on the dominant influenza virus type and/or subtype/lineage in the MS. • *The dominant influenza virus type, subtype or lineage is reported when 10 or more influenza-positive results per week (or weeks) are available, with the type (A or B) defined as a minimum return. The threshold for dominance is set at 60% and the threshold for co-dominance is set between 40% and 60%. The report of subtypes or lineage also requires a minimum of 10 positive viruses sub-typed or ascribed to a lineage.* Suggest to align the text and take caution of the %s which are at right at the borderline of dominance.		Edited to say Spain HUVH and Spain FISABIO are also approximately equal (ca. 56% A/H1N1). Kept terminology of A/H1N1 dominance in Finland HUS (61.4%) and Italy CIRI (60.3%), nuance is given as the % are presented.
13	Page 9 *IVE estimates: Pooled TND* Three robust confounder-adjusted pooled VE estimates were obtained; other estimates were non-robust and should be interpreted with caution. Suggest to rephrase: *IVE estimates: Pooled TND*		Added “for any vaccine exposure”

	Three robust confounder-adjusted pooled VE estimates for any vaccine exposure were obtained; other estimates were non-robust and should be interpreted with caution.		
14	Page 9 Limited amount of data captured per vaccine brand, distributed over appropriate-yet multiple strata (age, setting, and type of outcome) resulted in non-robust IVE estimates with wide to very wide confidence intervals. Brand-specific IVE estimates in children aged 6m-17y were available for 4 brands (4 in primary care setting, 3 in hospital setting), 5 brands in adults 18-64y (4 in primary care setting, 4 in hospital setting), and 5 brands in elderly aged 65+y (3 in primary care setting, 5 in hospital setting). Similarly, type-specific IVE estimates were non-robust. The term "available" is misleading since these were non-robust. Suggest to replace with Calculated Suggest to rephrase: Limited amount of data captured per vaccine brand, distributed over appropriate-yet multiple strata (age, setting, and type of outcome) resulted in non-robust IVE estimates with wide to very wide confidence intervals (CI). Brand-specific IVE estimates and 95% CI in children aged 6m-17y were calculated for 4 brands (4 in primary care setting, 3 in hospital setting), 5 brands in adults 18-64y (4 in primary care setting, 4 in hospital setting), and 5 brands in elderly aged 65+y (3 in primary care setting, 5 in hospital setting). Similarly, type-specific IVE estimates were calculated but non-robust.		Added "calculated" and "were calculated".
15	Page 10 All IVE estimates obtained from the THL register-based cohort were robust. These could not be pooled with the TND results because stratification by setting (primary care vs hospital) was not available. The reason is also because it is not clear whether cohort and TND can be pooled at all from a statistical perspective.	I think EFPIA's comments refer here to the presentation by Jos Nauta at the annual forum. I did not find this presentation very coherent and would not wish to accept the view he	We disagree with this comment. TND and cohort data can be pooled. The presentation by Jos Nauta at the AF cited as reason not to pool the generalizability of the results. While the Finnish cohort includes the whole population, and the TND only a sample of the total population, we are still only capturing subjects

	Following the presentation at the annual forum on this topic, subsequent discussions took place about differences between TND and cohort and appropriateness to pool.	expressed that pooling of TND and cohort VE estimates should not be done – at least not without much more methodological discussion. It was unfortunate that the methodological meeting planned for earlier this year was cancelled, I understand due to late withdrawals by I-MOVE participants. I think it is very important that the SC pursue this objective notwithstanding the blinkered stance of some potential I-MOVE attendees as exemplified in their article in the EU IMPACT web-based publication.	that decide to go to a doctor/hospital because of influenza in both study designs. Discussion to be continued.
16	Additional points to add to the Exec summary: For most brands in each age stratum, data could only be derived from a single source.		Not added because this is not true. For any influenza: for 10 brand-specific age/setting strata there is a single source, for 13 brand-specific age/setting strata there is more than 1 source.
17	Add to limitations in exec summary and 5.5.1: confounder adjustment was fixed and did not account for specific local aspects. Confounding may have applied differently in settings where multiple vaccines were available for a given age group.		See reply to comment #24
18	Add to recommendations: further considerations for confounder adjustment and optimization.		See reply to comment #28 for text added to the main report

19	Page 44: Symptom onset time in days since start of the study was included to account for changes in the risk of infection over the season Add – and differences in strain circulation.		Added “and differences in strain circulation”
20	Page 47 For the interpretation of IVE point estimates, D4.6 “Guideline for interpretation of IVE results” was used. VE point estimates of 0-30% are interpreted as ‘low’, 31-50% as ‘moderate’, 51-75% as ‘good’ and 76-100% as ‘very good’. Change to – since only robust estimates are interpreted: For the interpretation of robust IVE point estimates, D4.6 “Guideline for interpretation of IVE results” was used. VE point estimates of 0-30% are interpreted as ‘low’, 31-50% as ‘moderate’, 51-75% as ‘good’ and 76-100% as ‘very good’.		Added “robust”
21	Page 52 Please check the links – not all appear to work. Best would be to have these recommendations in the Annex in case online information is no longer available. For example Austria: https://www.sozialministerium.at/cms/site/attachments/0/0/6/CH4062/CMS1538134077648/empfehlung_zur_jaehrlichen_influenza-impfung-version_8.2.pdf		Website updated (appears to have been moved when 2019 recommendations were posted). An appendix has been added where the webpages the links refer to can be found.
22	Table 16 to 19 “Number of subjects” – suggest to clarify that this is overall and not exposed cases – or better would be to add the exposed cases between brackets as the driver of the power. Otherwise the reader may be put on the wrong foot with the high sample sizes were exposed cases are still low.		Column title edited to “Total N subject (n vaccinated cases)” . Number of vaccinated cases added.
23	Page 67 Population characteristics for each vaccine brand are provided in ANNEX 5 Since brand specific VE is the primary objective, for any following annual report the brand descriptives should be available in the body text.		Need to balance including all relevant information vs overloading the report body. Can be discussed when the mock report is developed, what to keep in report body vs annexes.

24	5.5.1 Limitations To add: confounder adjustment was fixed and did not account for specific local aspects. Confounding may have applied differently in settings where multiple vaccines were available for a given age group.		We do not understand what is meant by specific local aspects. Do you mean confounding by indication? Otherwise see response to question 9 on confounder adjustment
25	In addition: review of the annual report should include a review of the full report by the participating study sites		To start as of next year
26	Page 113: *Study population characteristics by brand were compared for subjects 65+y included at study sites in Italy to explore confounding by indication.* Add: (data not shown)?		Added "(data not shown)"
27	Conclusions The primary objectives were not met in the 2018/19 season due to insufficient sample size per strata. Few robust IVE estimates were obtained. Ways to increase sample size should be further explored for next season. Change to: The primary objectives were not met in the 2018/19 season due to insufficient sample size per strata. Few robust IVE estimates were obtained, but outside the register based cohort study, no robust IVE was obtained at the brand level. Ways to increase sample size should be further explored for next season.		Added "and outside the register-based cohort study, no robust brand-specific IVE estimates were obtained"
28	5.8 recommendations: Add to recommendations: further considerations for confounder adjustment and optimization.		Added "including method of adjustment" to the last bullet points in the recommendations.
	<u>Typos/errors/clarifications</u>		
29	P27, it should be explicit that only one site Spain HUVH performed a matched analysis and explain why and/or why not for the others.		This was the choice of the study site, as it is less labor-intensive as they identified subjects from electronic records.
30	Table 3 – what does R-b mean (above cohort)?		Register-based, fully written

31	Table 12: more accurate would be to use ≥ not +.		Changed to "≥"
32	Table 14. Study population characteristics, hospital TND studies, 2018/19 characteristics missing: Number of.... Hospitalisations; Influenza vaccination... status		Part of the field was hidden, edited so that the full field is visible
33	Table 15, it should read person year instead of Person for the 65+		Edited
34	Please check the estimates p89 and 112 end of 5.2. discrepancy between text and tables/figures		We do not know what the discrepancy is (besides rounding).
35	p112, the sentence should read: *"For Vaxigrip Tetra, the IVE was estimated at 54% (95%CI: 43-62) in children aged 6m-6y and 30% (95%CI: 25-35) in elderly 65+y"*		Added "%" and "95%CI"

General ISC reply to EFPIA comments (2nd round):

ISC1: I have no specific comments but found myself questioning the rather stringent definition of robust - namely "Robust VE estimates were defined as VE estimates with a CI width of <40%." I know this was in the SAP but given the additional stratification by brand, age and whether primary care or hospital admitted I wondered whether the bar has been too high in relation to this project and the EMA expectations. While it is of interest to look at protection by setting and age, in relation to meeting EMA's requirements is this necessary for the primary analyses? Thus the report reads as if this year was a failure as per the sentences below from the exec summary:

"Limited amount of data captured per vaccine brand, distributed over appropriate-yet multiple strata (age, setting, and type of outcome) resulted in non-robust IVE estimates with wide to very wide confidence intervals. …Similarly, type-specific IVE estimates were non-robust….. The primary objectives were not met in the 2018/19 season due to insufficient sample size per strata, particularly at the brand level. Few robust IVE estimates were obtained".

Also I am not sure why "non-robust estimates should be interpreted with caution" if the method is obtained by the less biased TND method, the point estimate is positive and the CIs don't include zero. The robust criterion as defined for the DRIVE project is rather more stringent that applied by many others to the results of observational epi studies and also I think compared with the former immunogenicity criteria used by EMA for annual release of seasonal vaccines.

I would therefore appreciate a discussion of this concept of "robust" which I feel may be unduly stringent when applied to stratified analyses. What was its origin as it wasn't suggested by the ISC.

I thought that overall the comments of EFPIA on the report (as listed in the email from Cedric dated 02/08/2019) were reasonable though non-material as Cedric acknowledged and could be incorporated with the exception of the one on pooling (see below).

ISC2: I have no comments.

ISC3: Table 2 lists 9 vaccines, while the text mentions 10 vaccines licensed, and seven included?