DRIVE D7.4 ISC comments – reply grid

Thanks again for all your comments and helping us to improve the studies. Your comments are very helpful and always very much appreciated. There are still some open questions indicated in the document (indicated in green). Feel free to share your views on this or anything else you think we should consider for future studies. Thanks again!

Anke and Kaat, on behalf of Marga (holidays) and the WP7 team

Reviewer	Number	Comment	Reply
1	1	The main objective of the 2017/18 pilot season as stated in the background to the 2017/18 SAP (item 16.2 in the Appendix list) was to test the different functional aspects of the DRIVE project. Given the limited amount of data and heterogeneity between sites in data collection methods and surveillance populations I do not feel it helpful to make detailed comments on the actual results which are broadly in line with what has already been reported by the I-MOVE project for the 17/18 season. I have therefore focussed on more generic comments.	We have changed the focus of the report to emphasize the achievements and lessons learnt from this pilot year rather than the epidemiological results. The most important changes are: • The addition of a primary objective (pilot) • Presentation of the results in the context of system testing and without interpretation • Substantial revision of the discussion to describe the experience, lessons learned and next steps • Addition of the results from the exploratory objectives
1	2	The authors are to be congratulated on a detailed and thorough report of the first season's results from the 5 sites participating in year 1 of the DRIVE project. The report demonstrates the feasibility of analysing centrally individual data extracts from sites using the test negative design (TND) to obtain a pooled estimate of effectiveness for two vaccine brands. Only one site using cohort methodology provided data so the feasibility of obtaining a pooled brand-specific estimate by this method could not be assessed.	NA
1	3	In paragraph 9.11.5 of the report it is stated that "In the 1-stage pooling approach, the individual-level data were pooled". However, this combined analysis does not seem to have been completed as yet. The descriptions of figures 2-11 all refer to	The two exploratory analyses (1 vs. 2-stage pooling; effect of time since vaccination) have now been completed and added to the report.

		"pooled adjusted IVE estimates" which I assume were obtained by the 2-stage pooling method as the captions mention meta-analyses. Until 1-stage pooling results are provided for review, the overall feasibility of this approach to generate brand specific VE estimates cannot be assessed, neither can the exploratory objective relating to comparison of results with 1-stage and 2-stage pooling be completed. The exploratory analysis relating to the effect of time since vaccination was also not included. These missing analyses are potentially important in informing the approach to be taken in year 2.	
1	4	In Annex 1 of the core protocol for case control studies (D7.1 protocol version 1.0 dated 27/4/2018) there is a detailed list specifying the minimum set of data items for the pooled analysis but this was not referred to in the SAP for the pooled analysis for the 2017/18 season. For the one stage pooling, and for assessing the individual site-specific results, it would be helpful to know how the data items were structured in each of the extracts. I understand that for this initial season the participating sites were using their existing protocols but there must have been sufficient commonality in the key fields that were extracted to allow site-specific analyses according to the site-specific SAP (item 16.3 in the Appendix list). There is a list of common variables required for the pooling given on page 12 of the SAP for the pooled analysis but it is unclear whether the data extracts from each site provided the variables as listed or whether some were derived centrally, e.g. age group from sample dates and birth dates which would be a more challenging task. The inclusion criteria specify that cases should meet the ILI/SARI case definitions which are precise and require detailed symptom information (as listed under 6.1.1 and 6.1.2 in the site-specific SAP). Was this verified and if so, how?	• A data management report for data cleaning and transformation procedures conducted centrally has been added (Appendix 9.7) • Two sites (La Rioja, Valencia) had information on ILI symptoms, allowing the case definition to be checked. This information has been added to the report (section 3.5.2). During the site visits, it will be verified whether it is feasible to obtain information on the ILI symptoms to allow checking whether the ILI/SARI case

		Did the data extracts provided by each site contain all the variables necessary to apply the exclusion criteria? Based on the exclusions in the flow diagrams for each site some exclusions seem to have been applied centrally after the data were received (e.g. swab>8 days after onset) while others were either NK or had zero exclusions, suggesting some data cleaning prior to extraction. The site-specific protocols provided as annexes to the 2017/18 pooled analysis do not provide this type of information as they just describe the existing surveillance system at each site. The SAP plan for the 17/18 pooled analysis mentions 3 sites in Spain and describes the data available from each (Table 1) but results are only provided for 2 Spanish sites. It was only noted in a footnote in the SAP for the pooled analysis that the data for the 3rd Spanish site in Canaria was not yet available. It would have been helpful if the incomplete parts of the planned analysis had been indicated at the beginning of the report.	definitions are met. Based on our first experiences with the site visits, it seems that not all sites will be able to provide that information, but many will. We will likely turn the ILI symptoms into 'strongly recommended to have' variables in the Minimal Data Requirements. • Exclusion criteria: Exactly, for some sites the exclusion criteria were already applied before receiving the data, for others not. It is unclear what would be the best approach. On the one hand, it is better to check the exclusion criteria centrally, but on the other hand, the GDPR requires that minimal data are collected. This will also be discussed during the site visits. Your views on this would be appreciated. • Indeed 3 sites from Spain were expected but one of them could not provide all the required data (apparently no data on controls) and was therefore excluded. This has now been added to section 9.1. • Noted for future reports to specify up front in the report any incomplete analysis that would be added later
1	5	I note that the number of co-variates included in the final model for each site-specific analysis varied considerably between sites with age group not apparently included as a covariate in the final models for the two Spanish sites, nor number of hospitalizations in the previous 12 months at the Italian site (see table 14, section 11.2, the latter variable was not available for Austria). However, under the paragraph on "Meaning of the study" (page 44) it is stated that "The main confounders identified as significant across all sites were the age group and number of hospitalizations in the past 12 months". Can this apparent discrepancy be clarified?	The paragraph "meaning of the study" has been deleted, this was incorrect. • Confounder adjustment: The confounder adjustment was done site-specific deleting all non-significant terms from the model starting from a full model that includes all potential confounders. This approach was driven by a concern of sample size and statistical efficiency, however it turned out complicated to explain that some of the adjusted VE estimates were adjusted for a particular confounder whereas others were not. Maybe a better approach would be to force some of the confounders in the regression model (even when not significant), such as age

			and sex. Maybe we should keep all common confounders in the model (provided they are not too many) to facilitate the epidemiological interpretation. Your views on this would be appreciated.
1	6	Given that this first season's analysis was essentially a pilot of the methods to be used in year 2 when more sites are participating and providing data according to an agreed common protocol, it would have been helpful if the report had focussed more on problems encountered and lessons learned rather than on interpreting and commenting on the actual results for which there was generally insufficient power for making comparisons between vaccines, influenza subtypes or age groups. While there is a section in the discussion on Strengths and Weaknesses this largely deals with difficulty in interpreting results due to lack of harmonization of protocols and the small number of participating sites. It is a shame that the more challenging aspects of this pilot analysis, namely achieving 1-stage pooling results, comparing these with the 2-stage pooling and deriving estimates of effectiveness since time from vaccination were not included in this report even though it has the word "final" in the title. I assume that another "final" report including these analyses will be made available shortly?	• The discussion has been rewritten to focus on the problems encountered and the lessons learned and not on interpreting and commenting the pooled IVE estimates. • The final in the report name comes from the deliverable name as stated in the Description of Action, but this report was only a first draft. The title has been changed to a more meaningful one. • Exploratory analyses have now been included in the report.
1	7	Reference is made in the 17/18 Report to the DRIVE Generic SAP: combining information on Influenza Vaccine Effectiveness across study sites (D4.4) and also to the Data management plan (D4.2). These documents may fill in some of the missing detail outlined in paragraph 4 above. While these documents may be available on the SharePoint? I have been unable to access the SharePoint for some time. It would be helpful if it could be ensured that all referenced documents were made available for the ISC to inform their review. In all there are a large number of documents that have been generated so far for each DRIVE Work Package (as detailed in the ppt shared with the ISC at the Valencia meeting)	The ISC members should have access to the DRIVE SharePoint. Apologies for this. If you continue having issues accessing the DRIVE SharePoint, let us know.

		some of which would help provide useful background for the ISC even though produced for other Work Packages. Can the ISC access these?	
2	8	-p10: why not include in methods/results the observed heterogeneity in vaccines used, case definitions, coverage, in/exclusion? Overall it is sometimes confusing to what extent this report focuses on a feasibility pilot, or IVE results? The results mention adjusted pooled estimates decreasing with age: so these were age-stratified analysis, rather than age-adjusted? Should the mismatch not be mentioned in the results already, rather than emerging suddenly in the conclusions?	• Wording has been added to the different sections in methods to highlight differences between sites. Figure 2 has been added to illustrate the most important differences between sites. • Information on vaccine match has been added to section 7.1 and is further illustrated in Figure 7. • As part of the secondary objectives, IVE was stratified by host-related covariates. For every stratified analysis, the variable stratified upon was excluded from the adjustments. For example, confounder adjusted stratified by age means adjusted for other confounders excluding age. This has been added in the methods section, 3.11.2 (p33).
2	9	-p12-14: can the descriptions of each of the sites be more harmonised/standardised? E.g. some mention vaccination coverage, others don't; some provide absolute numbers, others incidence; some % with decimal others without; etc.	The descriptions have been more standardized based on the information we had available. A lesson learned (added to the discussion) is that for the next report we should be more specific on the information we expect from the sites for these descriptions to enable better standardization.
2	10	-p14: study objectives: clarify these relate to this first pilot season only, but are to be related to the overall objectives. Also, objectives on feasibility?	We have added a primary objective on the piloting governance model/infrastructure/tools/procedures developed in the pilot year.
2	11	-p15: do the Austrian and/or Finnish data capture the origin of the data (primary vs. inpatient), to allow regrouping?	Not in the data for this season. We have added a variable to the minimum data requirements for next season to distinguish between data collected from GPs vs. hospitals. At this stage, it is unlikely that Finland will be able to disentangle GP and hospital cases. For Austria, this will be discussed during the site visit (planned for Oct 22).
2	12	-p17: table 2: consistently include size (and representativeness) of catchment population	The information is presented in Table 3 of the current report. We were unable to obtain this information from the sites. For next season, several measures will be taken to improve the site engagement and responsiveness.
2	13	-p18: case definition of Valencia should be labeled: modified EU ILI case definition. What % of cases where not sudden? Is it possible to do a sensitivity analysis, excluding them?	• The term 'modified' has been added. • We were not able to obtain a clear answer to question on the % of cases that where not sudden. This is a topic will be discussed during the site visits for the 2018/2019 season.

			• A sensitivity analysis was not planned for this and due to the pilot nature of this season no amendments to the SAP will be made.
2	14	-p18: not sure subtyping from sentinel data is a strong proxy to 'explain' IVE. Will this change (e.g. add subtyping in individual cohort members) in future years?	Wording modified. Adding subtyping for individuals in the cohort study is currently not planned. Nonetheless, we believe having information on influenza subtypes is important from a regulatory/MAHs point of view (and hence important for DRIVE). We plan to have this discussion soon.
2	15	-p19: 2nd exclusion criterium not clear (twice 'and'??)	Separated into 2 criteria for clarity.
2	16	-p21/22: was vaccine brand used associated with indication? Or random?	Section 9.7.2 has been modified to reflect specific vaccine type use for specific target populations (as recommendations were based on type and not brand), or otherwise have indicated that choice was free for the vaccine provider
2	17	-p23: pregnancy not included as chronic condition (for obvious reasons), but registered separately in Finland and Italy: was this in the analysis used to construct a new variable chronic including pregnancy, for all sites?	For some sites, the chronic conditions contained pregnancy. For others, pregnancy was separately reported. For the pilot season, we therefore combined pregnancy with other conditions to harmonize the case definition of 'chronic conditions' for the pilot year. For next seasons, pregnancy will be separately reported.
2	18	-p24: chr[onic] neurological in Austria? La Rioja has no option >5 admissions? Or no cases in that category? Is the option >1 or 1-5?	• Neurological conditions filled in for Austria • La rioja had no cases with more than 5 hospitalizations. Wording has been added to clarify that.
2	19	-p26: stratified adjusted analysis by almost all common confounders (except stratification by number of previous hospitalisation) are presented in 11.2 and 11.3, which is not clearly described or motivated. Also, not clear why only no stratification by previous hospitalisation?	Number of hospitalizations was only intended for adjustment and not for stratified analysis since it is used as a proxy for the severity of chronic conditions.
2	20	-p27: for Italy no information is given that approval was granted, nor on composition of committee, nor on duration to obtain permission?	This information has been added.
2	21	-p30: very relevant figure, but no data?? n=?	Apologies, corrected.
2	22	-p36: the declining adjusted VE stratified by age is reported several times, but no data is provided to support this is a significant trend! As the CI of the 65+ group includes the point estimate of the youngest group, this may not be likely. Yet not even a comment is made, whereas for other different estimates (e.g. [vaccine type redacted] vs. [vaccine type redacted]) there are repeated	• All text on interpretation of the results has been removed. • No trend analysis was done. Significance between estimates was also not formally tested, not planned in the SAP and not in study objectives due to the pilot nature of this study (the pilot study was deliberately kept purely descriptive).

		statements of probable not significant. Why differential? Because decline with age expected and accepted? Was significance between estimates tested? If yes, how; if no why not?	• The SAP for next year still needs to be written, and hypothesis tests/trends tests will likely be included but this discussion still need to happen.
2	23	-p40: data 11.4 to follow when? (we did receive most recent version this time I trust;))	We do understand your worry. You did receive the correct version. Results exploratory analyses were not included due to time constraints/holidays.
2	24	-p41: please provide numbers for sensitivity analysis: how many patients excluded?	Included.
2	25	-p42: the discussion is more a repetition of results than a discussion of validity and implications? Maybe this is considered premature in view of data limitations, or too sensitive within the consortium? Even then, the discussion could/should add and discuss critically outcomes, feasibility and lessons learned etc., more than just repeating/summarising results.	Discussion has been substantially revised.
2	26	-p42: the IVEs mentioned for [vaccine type redacted] vs. [vaccine type redacted] for the TND sites only, seem not to match with the data presented on p39?	Results were rounded up in the discussion to facilitate the reading.
2	27	-p42: pooled estimates lower in hospital settings: I agree, suggest this could be included as a separate stratification, esp. if the sites with mixed designs can make this distinction. There has been discussion that VE might be suboptimal, but as long as severe infections were prevented/milder, this was still important. Please discuss (also a statement is given on page 43)	• Austria and Finland could not provide the data separately for outpatient and hospitalization, which left us with only 2 sites to pool for outpatient and 1 for hospitalization. • Ideally for future analysis with an increased number of sites, analysis should always be presented separately per setting. • We have chosen to no longer discuss the results.
2	28	-p43: the differential effect of a stricter case definition (swab within 4 days) is remarkable. Please discuss if this is considered chance, real, possible explanations/implications?	The difference here is driven by the data from Valencia, where many individuals were swabbed more than 4 days from symptom onset. Subjects were swabbed in hospital as soon as possible following admission, so it may be related to delays in seeking healthcare, or subjects seen first in primary care with ambulatory treatment that did not improve or got worse.
2	29	-p43: not sure you can state all TND sites had same EU case definition	This has been corrected, the Valencia used a modified case definition.
2	30	-p43: in view of relevant comment on observed confounding by nr of hospitalisations: data are not shown, why not?	We present IVE stratified by presence of chronic conditions. Hospitalizations is related to this and used as proxy for chronic condition severity.

2	31	-p43: last paragraph strenghts and weaknesses: indeed, how much variation is real, how much is methodological bias. Please discuss how to assess and minimize.	We have chosen no longer to discuss the results but focus instead on the experiences and lessons learnt in the pilot year. It is indeed the challenge of DRIVE to reduce the between-site heterogeneity to the extent possible and trying to understand the remaining heterogeneity.
2	32	-p44: meaning study: more than half of isolates identified were mismatch: unclear if this statement refers to B/Vict only or also to A/H3N2	This paragraph has been removed.
2	33	-p44: update on sites for 2018/19? Should be known by now!	Sites for 2018/2019 have been added to the discussion.
2	34	-p44: in view of the very low VE, I'm not sure that more precise estimates of different vaccines will provide much support to PH decision makers.	All comments on the results have been removed.
2	35	-p45: IVE against A/H3 was not null or low! Rather than protect, vaccination after adjustments was a statistically significant risk!! Better to acknowledge, discuss if this could be real or biased, not to hide or ignore. Statements on higher (actually, less low) protection among children etc. need results on significance of these differences.	We have looked further into this. The negative IVE for A/H3N2 is driven by Austria and Valencia. We had to rerun the analysis since infants less than 6 months were not excluded for Austria since age was given as 0 years and we had assumed that <6m had already been take out of the dataset that was received. In the new analysis, the significance is not there anymore. We agree that it is an unexpected finding to see a negative (although no longer significant) VE against AH3N2. Similar results were seen in the UK. We are wondering whether these results might suggest 'bias by indication' in the IVE studies. Your views on this would be appreciated.
3	36	Page 2: Transparency is a priority for DRIVE. Roles and responsibilities for participants should be well clarified in the list of contributors. For example, the role of OPBG (not included in list of abbreviations), UNIFI and SURREY for the present report and studies is not inferable or indicated in chapters 2 and 3.	The contributions have been spelled out: p 11 –p13
3	37	Page 10. In the summary is stated that "This report presents the IVE estimates in the 2017/18 season as a result of the first pilot studies of the DRIVE consortium". However as stated later on, "For the DRIVE consortium, the influenza season 2017/18 was considered a pilot season. The main objective of the pilot was to	The summary has been substantially revised.

		test the different operational aspects of the project including the governance and streamlining key processes such as data collection, statistical analyses and dissemination of study results". In the present report, emphasis is given to the IVE estimates compared to the pivotal operational aspects to be strengthened in subsequent seasons.	
3	38	Page 11: in the summary conclusion the statement "Finally, IVE was higher in primary care settings (28%, 95%CI 3-47%) than in hospital settings (-6%, 95%CI -33-15)" is misleading. As far as I understand the difference was in severity of prevented cases and not in the settings themselves.	The summary has been substantially revised and no longer includes IVE estimates.
3	39	Pag. 13 Chapter 7: inhomogeneous description of background circulation of influenza. A common set of information should be provided for each study site including estimates of population-based incidence.	Addressed where possible. Please also see comment #9
3	40	Pag 15: Chapter 9 Providing an estimate of proportion of population included in each study compared to the general local population in table 9.3 would be useful to understand the actual size of the studies.	This information has been added – Table 3 current report
3	41	Pag. 22: Table 8. Please clarify what "Vaccine brands (% of total at site*)" means. Is it a vaccination coverage?	It's the proportion of all vaccinated subjects at the site that received each brand, so no, it's not vaccination coverage. Clarified in table to make it explicit and avoid confusion.
3	42	Chapter 10: Ethical approval unclear description about the approval of the current studies (e.g. Italy)	This information has been added.
3	43	Pag. 29: the higher number of test-positive cases compared to controls in some study sites is strange. I would have expected an inverse proportion. May any unstated selection of test negative cases have occurred?	All sites had more cases than controls except for Valencia. It is not unusual to get more controls than cases (has also been observed in I-MOVE studies). Valencia is the only hospital-based study. Closely following the SARI case definition might result in many patients with cardiac failure, COPD exacerbations without any signs of infections. We do not completely understand the phenomenon, and attention will need to be paid in 2018/19 to ensure adherence to the mutually agreed case definition.