17 November 2011
EMA/879297/2011
Committee for Medicinal Products for Human Use (CHMP)

Overview of comments on 'Qualification opinion of low hippocampal volume (atrophy) by MRI for use in clinical trials for regulatory purpose – in pre-dementia stage of Alzheimer's disease' (EMA/CHMP/SAWP/763373/2011)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	Eli Lilly and Company
2	PPSB
3	Janssen Pharmaceutical Companies of Johnson & Johnson
4	Elan Pharma International Ltd

7 Westferry Circus ● Canary Wharf ● London E14 4HB ● United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu
An agency of the European Union

1. General comments – overview

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
1	The HC is the best established brain structure in terms of number of studies demonstrating a relationship between its volume and AD disease stage. Abnormally low hippocampal volume is well-established as reflecting medial temporal lobe (MTL) atrophy in the MCI/AD disease process, can be measured using a range of quantification methods and hence makes sense as the focus of this qualification application. Both the wealth of published studies reviewed by the applicants as well as the de novo analysis support the proposed context of use. Moreover, an important finding emerging both from the literature review and the de novo analysis is that the utility of hippocampal volume measures for this purpose is robust, demonstrating comparable results despite differences in methodologies and analyses (with the proviso that these are tightly controlled within each study). Loss of hippocampal volume measured by MRI is a nonspecific but well-recognized gross pathological finding consistent with neurodegeneration in AD. Because vMRI measures are not specifically limited to a single AD pathophysiological mechanism (e.g., abnormal amyloid Abeta trafficking, tau aggregation), this approach is applicable to enriching populations for trials of candidate treatments for both amyloid and non-amyloid targets.	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	The de novo analyses of ADNI data based upon multiple methods for estimating hippocampal volume are remarkably consistent. While a simple, universal regression model and associated decision mechanism would be ideal, the applicants pragmatically advise that this is not yet feasible. However, workable solutions on a trial-by-trial basis are practicable. Numerous cited examples provide support for this approach and the intended context of use. Qualification will spur further development and standardization of vMRI for this purpose. In particular, it would be useful to conduct additional de novo analyses using data from large clinical studies other than ADNI. Although many of the studies cited in the qualification document use conversion to AD dementia as an endpoint, a more important benefit of selecting patients with low hippocampal volume is identification of a cohort with a more homogeneous disease stage and hence clinical trajectory. This should result in more efficient treatment trials with disease progression (measured using continuous scales) as the outcome.	
2	1. "The proposed utility of the baseline MRI measurement of HC atrophy is in subject selection for enrolment in an AD clinical trial." This addresses one problem endemic to older clinical trials, a lower than expected rate of conversion to dementia. Subjects with greater degrees of HC atrophy will have a higher rate of conversion. The chief utility, however, should be to enrol the most informative subject. For example, one whose prior probability of a specific clinical	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	outcome—be it progression of cognitive deficit or stability is not beyond the reach of any therapeutic intervention. Therefore, the choice of a HC criterion is likely to be affected by a number of other factors, including trial duration, inclusion/exclusion criteria, primary outcome, and demographic characteristics of the subject population. The use of HC atrophy simply to boost the number of subjects expected to convert may not result in a more efficient study if the above considerations are neglected. 2. One of the primary things that the authors do not address is the effect of different post-processing pipelines upon the discriminative abilities of HCV to distinguish between people who will measurably decline over the course of a clinical trial. This may be of significant concern given Figure A2, in which the Sensitivity vs. Specificity of the cited work is much greater than was replicated in the ADNI dataset. Another reason that may lead to this discrepancy is the differing entry criteria for these patients in the trials. The proposal does not properly account for the effect that different entry criteria will have upon the discriminative ability of the entry criteria. While in general, it appears that HCV works well in determining whether a subject will exhibit progressive memory decline, the proposal does not fully address the marginal value of adding HCV on top of cognitive entry	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	criteria (i.e. what is the increase in the specificity/sensitivity when HCV is included). 3. ".. The biomarker application is therefore not dependent on the mechanism of the investigational compound." The value of HC atrophy may not be equivalent for compounds with different mechanisms of action. For instance, HC atrophy in the absence of amyloid burden (as reflected in CSF or PET) may select subjects who will be non-informative in clinical trials of anti-amyloid therapies because they lack the presumed "target pathology." The value of HC atrophy should be examined not only relative to other biomarkers, but conditional on the results of those biomarkers. For example, what is the value of HC atrophy in predicting conversion to dementia in subjects with and without a biomarker of amyloid burden? And what is the value in subjects with and without the apolipoprotein E4 allele? 4. The most detailed analysis of the operating characteristics of HC volume in prediction of conversion to dementia was performed by the sponsor using one dataset, ADNI. This analysis demonstrates comparability of results from several methods of HC volume measurement applied to this one dataset, but would application of one method to several other datasets yield similar results? The fact that sensitivity and specificity from 5 studies in the literature all fell to the	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	left of the ROC curve (Fig A1, page 29/59) raises the possibility that other datasets might differ substantially from ADNI with respect to the predictive value of HC atrophy. 5. Does the approach, validated by the sub-study analysis in the ADNI study, hold true for a more general population (memory clinic-type population) where there may be more variability observed in the baseline of HC volume? The de novo analyses of ADNI data based upon multiple methods for estimating HC volume are remarkably consistent. While a simple, universal regression model and associated decision mechanism would be ideal, the applicants pragmatically advise that this is not yet feasible. However, workable solutions on a trial-by-trial basis are practicable. Numerous cited examples provide support for this approach and the intended context of use. Qualification will spur further development and standardization of vMRI for this purpose. In particular, it would be useful to conduct additional de novo analyses using data from large clinical studies other than ADNI. 6. Although many of the studies cited in the qualification document use conversion to AD dementia as an endpoint, a more important benefit of selecting patients with low HC volume is identification of a cohort with a more homogeneous disease stage and hence clinical trajectory. This should result in more efficient treatment trials with disease progression (measured using continuous scales) as	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	the outcome. The studies in the literature, as well as ADNI, are observational. The operating characteristics of HC atrophy in predicting conversion should be evaluated in the context of its intended use (e.g., clinical trials). If these data are available from placebo arms in completed clinical trials, they should be analyzed. In addition, an important finding emerging both from the literature review and the de novo analysis is that the utility of HC volume measures for this purpose is robust, demonstrating comparable results despite differences in methodologies and analyses (with the proviso that these are tightly controlled within each study). 7. "The CHMP has given a previous positive opinion in the predementia stage of Alzheimer's disease: cerebro - spinal fluid related biomarkers for drugs affecting amyloid burden. ... Although not required from a regulatory perspective, the concomitant assessment of the two biomarkers in predementia stage of AD would be of great value." Clarification should be provided on whether it's proposed that both biomarkers should be used as selection criteria within the same trial or if both should be measured at baseline, so that their concordance and their respective operating characteristics can be examined? Would this be limited to evaluation of anti-amyloid therapies? 8. It is acknowledged in the CHMP Qualification Opinion that standardization of the measurement of low HC volume is required of all steps and that international guidelines must	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	be enforced. ("The process of measurement of low HC volume is also complex. .... International guidelines have been produced. These guidelines must be enforced.") Standardization of data acquisition, post-processing and analysis methodology is important to ensure consistent use of measurements across different clinical studies. For example the statement that the use of a central core lab will ensure consistent and reliable results means in practice that each lab would have their own definition of the hippocampus and own methodology and would not be standardized with others. More clarity is needed to determine if the CHMP is referring to a specific set of international guidelines for which a consensus exists, or to the ongoing effort to develop these guidelines as mentioned in Cliff Jack's article(e.g., Jack CF et al, Alzheimer's and Dementia 2011; 7:474-485)? 9. This qualification opinion is an important development in enabling use of baseline MRI measurement of HC atrophy to enrich subject selection in MCI trials of neuroprotection for AD. The HC is the best established brain structure in terms of number of studies demonstrating a relationship between its volume and AD disease stage. Abnormally low HC volume is well-established as reflecting medial temporal lobe (MTL) atrophy in the MCI/AD disease process, can be measured using a range of quantification methods and hence makes sense as the focus of this qualification application. 10. It is important that the opinion is based on the broader	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	scientific fact of decreased HC volume as measured by MRI, rather than specifying any particular proprietary equipment or algorithm which might be artificially restrictive in the application of the biomarker opinion in clinical trials. In this sense it is enabling rather than restrictive and is clearly pre-competitive. 11. Loss of HC volume measured by MRI is a nonspecific but well-recognized gross pathological finding consistent with neurodegeneration in AD. Because vMRI measures are not specifically limited to a single AD pathophysiological mechanism (e.g., abnormal amyloid Abeta trafficking, tau aggregation), this approach is applicable to enriching populations for trials of candidate treatments for both amyloid and non-amyloid targets.	
3	12. "The proposed utility of the baseline MRI measurement of HC atrophy is in subject selection for enrolment in an AD clinical trial." This addresses one problem endemic to older clinical trials, a lower than expected rate of conversion to dementia. Subjects with greater degrees of HC atrophy will have a higher rate of conversion. The chief utility, however, should be to enrol the most informative subject, i.e. one whose a priori probability of a specific clinical outcome (be it progression of cognitive deficit or stability) is not so high as to be beyond the reach of any therapeutic intervention. Therefore, the choice of a HC criterion is likely to be affected	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	by a number of other factors, including trial duration, inclusion/exclusion criteria, primary outcome, and demographic characteristics of the subject population. The use of HC atrophy simply to boost the number of subjects expected to convert may not result in a more efficient study if the above considerations are neglected. 13. ".. The biomarker application is therefore not dependent on the mechanism of the investigational compound." The value of HC atrophy may not be equivalent for compounds with different mechanisms of action, i.e., HC atrophy in the absence of amyloid burden (as reflected in CSF or PET) may select subjects who will be non- informative in clinical trials of anti-amyloid therapies, because they lack the presumed "target pathology". 14. The value of HC atrophy should be examined not only relative to other biomarkers, but conditional on the results of those biomarkers. For example, what is the value of HC atrophy in predicting conversion to dementia in subjects with and without a biomarker of amyloid burden? And what is the value in subjects with and without the apolipoprotein E4 allele? 15. The most detailed analysis of the operating characteristics of HC volume in prediction of conversion to dementia was performed by the sponsor using one dataset, ADNI. This analysis demonstrates comparability of results from several methods of HC volume measurement applied to this one dataset, but would application of one method to several other datasets yield similar results? The fact that sensitivity	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	and specificity from 5 studies in the literature all fell to the left of the ROC curve (Fig A1, page 29/59) raises the possibility that other datasets might differ substantially from ADNI with respect to the predictive value of HC atrophy. 16. The effect of different post-processing pipelines upon the discriminative abilities of HCV to distinguish between people who will measurably decline over the course of a clinical trial was not addressed. This may be of significant concern given Figure A2, in which the Sensitivity vs. Specificity of the cited work is much greater than was replicated in the ADNI dataset. Other reason that may lead to this discrepancy is the different entry criteria for these patients in the trials. While in general it appears that HCV works well in determining whether a subject will exhibit progressive memory decline, the proposal does not fully address the marginal value of adding HCV on top of cognitive entry criteria i.e. what is the increase in the specificity/sensitivity when HCV is included. 17. The studies in the literature, as well as ADNI, are non-interventional. The operating characteristics of HC atrophy in predicting conversion should be evaluated in the context of its intended use, i.e., clinical trials. If these data are available from placebo arms in completed clinical trials, they should be analyzed. 18. "The CHMP has given a previous positive opinion in the predementia stage of Alzheimer's disease: cerebro - spinal fluid related biomarkers for drugs affecting amyloid burden. ... Although not required form a regulatory	

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	perspective, the concomitant assessment of the two biomarkers in predementia stage of AD would be of great value." Please clarify: Is it proposed that both biomarkers should be used as selection criteria within the same trial? Or should both simply be measured at baseline, so that their concordance and their respective operating characteristics can be examined? Would this be limited to evaluation of anti-amyloid therapies? "The process of measurement of low HC volume is also complex. .... International guidelines have been produced. These guidelines must be enforced. "Please clarify: is the CHMP referring to a specific set of guidelines for which a consensus exists, or to the ongoing effort to develop these guidelines (e.g. Jack CF et al, Alzheimer's and Dementia 2011; 7:474-485)?	
4	Elan welcomes the advances by the agency in the development of tools to facilitate patient selection for early or prodromal stages of Alzheimer's disease. The company believes that with the qualification of the CSF biomarkers earlier this year, the current opinion on hippocampal atrophy and potentially other biomarkers in the future, it is important to retain flexibility in what biomarker is used. The agency refers to the relative value of the various biomarkers being worthy of future analyses. While such analyses are anticipated to be of scientific interest, the company believes it important that multiple measures do not become a requirement in clinical studies, as to do so would significantly add to the complexity of patient enrolment.	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Stakeholder no. *(See cover page)*	General comment (if any)	Outcome (if applicable)
	The company is encouraged by the comments related to the need for standardisation of methodology in relation to this and other biomarkers. The data considered in the current opinion have been generated by a variety of methodologies. The establishment and adoption of a standardised hippocampal measurement technique and the development of the Reference Standards as discussed in the draft opinion will further add to the relevance of biomarkers as patient selection methods. Finally the company queries how these biomarkers will influence final product labelling. Will the use of a biomarker to select a patient population at baseline be implemented on the final product label, as a requirement for access to the treatment?	

2. Specific comments on text

Line no.	Stakeholder no. *(To be completed by the Agency)*	Comment and rationale; proposed changes	Outcome *(To be completed by the Agency)*
35	1	Comment: The wording “searched required” does not make sense. Proposed change (if any): Reword for clarity.	Not longer applicable, new version has been updated.
920-923	1	Comment: We agree that it would be of great value to further explore the relationship between CSF Abeta and vMRI biomarkers for enrichment. However, this qualification should not be contingent upon delivery of paired data on both biomarkers. Specifically, (1) while some study participants agree to undergo both procedures, this can place an unanticipated burden on the patients and their caregivers, and increases the risk that patients will discontinue participation in the clinical trial; (2) in large, geographically-distributed Phase III trials, CSF is likely to be available at only a minority of sites, whereas MRI is a more widely available technology; (3) vMRI changes appear later in the disease stage (being most dynamic in the late-MCI thru AD stages) than CSF Abeta changes (which are most dynamic earlier in the disease process). MRI based biomarkers are thus well-suited for the task at hand, namely	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Line no.	Stakeholder no. *(To be completed by the Agency)*	Comment and rationale; proposed changes	Outcome *(To be completed by the Agency)*
		identifying patients on the cusp of dementia; (4) finally, patient selection based on CSF measures may provide a patient cohort better suited to a specific drug mechanism (i.e., A beta trafficking; tauopathy), whereas vMRI is more "neutral" in this respect. Proposed change (if any):	
924-927	1	Comment: It is not clear to what the phrase "International guidelines have been produced" refers. To date, ADNI has provided de facto standards for MRI data acquisition/post-processing and most clinical trials now adhere to these as closely as practicable. Imaging core labs supporting clinical trials have implemented analysis software internally to a standard compatible with regulatory submissions. However, further standardisation efforts are ongoing and no universally-applicable regression model and cut-off value can be proposed at this stage. The applicants emphasized the above considerations (line 523ff) and the importance of enforcing standardised processes within each specific trial using consistent, centrally-managed QA/QC, acquisition and analysis methods as best practices for the prospective application of hippocampal volume measurements to clinical trials.	Not longer applicable, new version has been updated.

Line no.	Stakeholder no. *(To be completed by the Agency)*	Comment and rationale; proposed changes	Outcome *(To be completed by the Agency)*
		Proposed change (if any): The last two sentences could Not longer applicable, new version has been updated.be reworded to refer to adherence to the "best practices" outlined by the applicants.	
35	2	Comment: The wording "searched required" does not make sense. Proposed change (if any): Reword for clarity	Not longer applicable, new version has been updated
920-923	2	We agree that it would be of great value to further explore the relationship between CSF Abeta and vMRI biomarkers for enrichment. However, this qualification should not be contingent upon delivery of paired data on both biomarkers. Specifically, 1) while some study participants agree to undergo both procedures, this can place an unanticipated burden on the patients and their caregivers, and increases the risk that patients will discontinue participation in the clinical trial; 2) in large, geographically-distributed Phase III trials, CSF is likely to be available at only a minority of sites, whereas MRI is a more widely available technology; 3) vMRI changes appear later in the disease stage (being most dynamic in the late-MCI thru AD stages) than CSF Abeta changes (which are most dynamic earlier in the disease process). MRI based biomarkers are thus	The applicant is advised to submit a new qualification advice to SAWP and further discuss the questions mention here.

Line no.	Stakeholder no. *(To be completed by the Agency)*	Comment and rationale; proposed changes	Outcome *(To be completed by the Agency)*
		well-suited for the task at hand, namely identifying patients on the cusp of dementia; 4) finally, patient selection based on CSF measures may provide a patient cohort better suited to a specific drug mechanism (i.e., A beta trafficking; tauopathy), whereas vMRI is more "neutral" in this respect. Proposed change (if any):	
924-927	2	It is not clear to what the phrase "International guidelines have been produced" refers. To date, ADNI has provided de facto standards for MRI data acquisition/post-processing and most clinical trials now adhere to these as closely as practicable. Imaging core labs supporting clinical trials have implemented analysis software internally to a standard compatible with regulatory submissions. However, further standardisation efforts are ongoing and no universally-applicable regression model and cut-off value can be proposed at this stage. The applicants emphasised the above considerations (line 523ff) and the importance of enforcing standardised processes within each specific trial using consistent, centrally-managed QA/QC, acquisition and analysis methods as best practices for the prospective application of HC volume measurements to clinical trials.	Not longer applicable, new version has been updated

Line no.	Stakeholder no. *(To be completed by the Agency)*	Comment and rationale; proposed changes	Outcome *(To be completed by the Agency)*
		Proposed change (if any): The last two sentences could be reworded to refer to adherence to the "best practices" outlined by the applicants.	