



EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

20 May 2014
EMA/CHMP/SAWP/298348/2014
Procedure No.: EMEA/H/SAB/037/1/Q/2013/SME
Product Development Scientific Support Department

Qualification Opinion List of Issues

Total Kidney Volume (TKV) as a prognostic biomarker for use in clinical trials evaluating patients with Autosomal Dominant Polycystic Kidney Disease (ADPKD)

Summary

The applicant has started to work on the qualification of TKV as a predictive biomarker to enrich the ADPKD population for conducting clinical trials with the FDA in 2010 and now wishes to discuss the approach with the European agency, as well. Five data-sources have been integrated in one database. The idea is to develop a Cox-regression model in this standardized database to predict outcome (30% worsening of eGFR, 57% worsening of eGFR, transition of CKD stage 1 or 2 to stage 3, ESRD, hypertension, mortality) and then to use the respective model to predict which patients to include into a trial to arrive at a reasonable event rate in a more reasonable time frame in a clinical trial. The applicant has handed in a substantial documentation to support the qualification.

In the end, the applicant proposes to use a model including age, Baseline eGFR and baseline TKV (measured by Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scan, or ultrasound (US) imaging) to predict outcome. A number of potential trial-situations (proof of concept, dose-finding, confirmatory trial) are mentioned to illustrate the anticipated uses of the model and these may help to better understand the competence of the database that has been used to develop the model.

Questions of the applicant search concordance that (i) the context of use of the biomarker for enrichment of clinical trial populations is clearly described, (ii) the endpoints to be modeled are clinically relevant to describe progression, and (iii) the overall package is sufficient to allow qualification of the biomarker.

A formal Letter of Intent was submitted to the EMA on April 11th, 2013, followed by submission of the initial EMA Briefing Package on April 30, 2013. In response to a List of Issues provided by the EMA on the Briefing Book, a face-to-face meeting was held in London on July 9, 2013. Following questions and responses that were addressed via email during the next several months, the EMA indicated that all remaining questions could be addressed in the submission of an updated final Briefing Package. The updated package was submitted on 21/03/2014. When assessing the submission, it was felt that another set of issues has to be addressed by the applicant, before a qualification opinion can be issued. This document contains a very brief assessment of the applicant's responses up to now and asks for additional information.



Data sources

Analyses are based on observational data from five data sources (University of Colorado – Denver, Mayo Clinic, Emory University, Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease 1 (CRISP1), Consortium for Radiologic Imaging Studies of Polycystic Kidney Disease 2 (CRISP2)). In total the database contains data from 2610 patients, 2355 of whom have at least one measurement of TKV. Of these, a priority dataset has been identified consisting of 1182 patients who have at least two images for the measurement of TKV taken at least 6 months apart. This data has been described in a uniform way using CDISC standards.

Scientific Discussion

First List of Issues

Discussion of the database

1. Five data sources have been combined into one standardized database for further analysis. This process has been carefully conducted and described. A statement is missing, in how far the databases are representative for the overall PKD-population, and the coverage of the patient population in the hospitals they are associated to.

Assessment: The applicant have expressed their strong believe that the data from the respective sources are representative of the general AKPKD-population. A formal statement that centers and registries included consecutively all patients that arrived at the respective institutions cannot be made available.

2. As outlined in the description of the baseline variables in the different data-sources, differences between the patient populations (p. 77 ff., 86ff.) exist. Further arguments should be provided, why later analyses do not adjust for study / data source and in how far this may limit the generalizability of findings from the overall database.

Assessment: Differences between the different patient populations are seen as enrichment, but if all data-sources are representative for the ADPKD-population, such differences are difficult to explain. In summary the non-testable assumption has to be made that the data sources integrated for this project jointly cover the population of interest. Please comment.

3. The selection of patients from the studies for inclusion into the PKD-dataset is not transparent: why have only subgroups of the original data sources been included into the aggregate dataset? In how far does the selection of patients influence the characteristics of the patients. Criteria as proposed in STROBE should be provided for making transparent the process of integration of the datasets.

Assessment: Combination is now better understood. Still it remains puzzling that from initially more than 2000 patients only around 1100 can be used for the univariate modelling exercises. Given that only a few and quite standard explanatory variables have been included into the models, data quality seems to be an issue. It is still difficult to understand that for the analysis of 57%WoeGFR only 1140 out of 2355 patients can be included (p. 130). Please explain further.

4. Subjects included into the priority group are substantially different from those not included. How can re-assurance be provided, that this is without impact on generalizability of findings? The use of the model for predicting the event rate in a future trial obviously cannot be made dependent on the availability of a scan, because in future trials for all patients TKV would be determined.

Assessment: The applicant has elaborated on this and has renamed the populations. Issue clarified.

5. The applicant explains that a priority set has been built from those patients, where at least two images with minimum 6 months apart have been available in the aforementioned registry and that methodological investigations demonstrated that change in TKV is a less good predictor of outcome than simply using TKV at baseline. It is not fully plausible that the further modeling is restricted to

the priority set. Please compare model parameters provided to outcome from a re-analysis on the full population of patients with at least one TKV-measurement.

Assessment: A clear statement has been provided that all 2300 patients have been included in the modelling approach, however, the large number of missing values (see Q2 in the third list) requires further explanation. eGFR is an important parameter and therefore it is not totally clear why there is such high unavailability of outcome. Please comment.

Discussion on the modelling approach

6. Gender has been significant or borderline significant in some of the models predicting outcome based on TKV. It is unclear why gender has not been generally included into the prediction model (e.g. for the prediction of 30%WoEGFR gender has been significant in the univariate model but has been omitted from the multivariate model without presentation of a hazard-ratio, which still may justify adjustment). This is of particular importance because it is stated elsewhere that cyst development is age and gender specific from biology. Of note, the presentation of results with "coefficients" and hazard ratios is not easy to follow.

Assessment: Resolved.

7. A parsimonious model relating outcome to age and baseline TKV has been proposed, so that patients could be included into studies based on the predicted probability of outcome in a certain study duration. It is difficult to assess the predictive quality of the model. Please contrast the outcome against modeling including age alone, and, gender, in addition (e.g. by recalculating Table 3) and by the consistent provision of hazard ratios in univariate and multivariate models (with confidence intervals).

Assessment: eGFR has been a significant predictor of outcome, as well. There is some doubt about your modelling approach: did you add further variables only after TKV (or a transformation) has been already part of the model (explanation of residual variance)? What would be the outcome, if TKV, age and eGFR would be modelled jointly with a backwards selection algorithm to arrive at a parsimonious model? Please comment and also consider the following question about logistic regression modelling.

8. What is the conclusion about the impact of genetic mutations on outcome? There is a large number of missing values in some of the studies. Can it be assumed that missing is non-differential? Is the distribution of subtype PKD1 homogeneous in the contributing studies?

Assessment: Resolved.

9. The method for the measurement of TKV could be used as a parameter in the model to support that actually the type of measurement is irrelevant for prediction. Albeit in principle plausible, could you please add some evidence here?

Assessment: The applicant reassures that the method is of no relevance.

10. Are there other registries in other countries that could / should be used, in addition?

Assessment: Resolved.

11. The importance of % change in eGFR to be predicted in the model is not independent from the baseline eGFR. How is this captured in the modeling process?

Assessment: Baseline eGFR has been predictive. This is not implausible, because it may well be that TKV may add diagnostic certainty in early phases, whereas eGFR is a good predictor in later stages of disease. Please comment and investigate your data.

Discussion on the interpretation / qualification / usage

12. In order to better depict the later use of the model it would be useful to see sample calculations relating an age range chosen for inclusion in combination with a range of TKV at entry and then display the event rate predicted and observed for such a trial as indicated in Figure 48. Is it possible to provide some sample clinical trials situations for the candidate endpoints? Please provide also the figures from the database that are standing behind the recommended criteria for inclusion and exclusions and the respective event-rates there.

Assessment: Simulations have been provided, where however, a uniform distribution has been used to generate patient data with risk-factors. The plausibility of this is difficult to understand, as in most instances co-variables are not uniformly distributed. This leads to an important point: while Cox-Regression is perfectly situated to develop a prediction model for a certain patient. On the other hand, TKV and others are supposed to be used to enrich a patient population. Wouldn't logistic regression (to predict e.g. an outcome event within 5 or 7 years after randomization) a more plausible model? In addition ROC analyses could be used to identify optimal cut-points for influential variables to discern between high and low risk? Also model fit could be better demonstrated.

13. Question 2 asks for an estimation of the relevance of endpoints (eGFR deterioration, etc.) chosen in the prediction model. This question is not easy to be answered as still the knowledge about "correlation" between a worsening in eGFR and the classical composite endpoint of ESRD, transplantation, and death is limited. Is it possible to derive such information from your database?

Assessment: We didn't find the relevant information.

14. Please comment on the lacking ability to reproduce the impact on other disease outcomes (e.g. Hypertension, CKD transition, ESRD, Mortality) in your analysis. Is it possible that this is a consequence of the selection process? Likewise, genotype is not predictive of eGFR-worsening. Is this merely a question of power?

Assessment: Resolved.

15. Please comment on the utility of studies in biomarker-positive patients for later clinical use in all patients with ADPKD? Selection of patients with high kidney volume will increase the events in clinical trials, but will not necessarily select a population likely to respond to treatment.

Assessment: Resolved.

Second List of issues

16. Please provide for all proposed endpoints (30% worsening of eGFR, 57% worsening of eGFR, ...) the number of events observed in each of the contributing registries. Please elaborate on the question, why gender can be omitted from the modelling even though it may be plausible that gender will not be used as a criterion for inclusion / exclusion. Which other recognized potentially prognostic variables are available in your database that could be prognostic for the development of disease and should be considered in your model? Please provide some considerations why change from baseline is not predictive, whereas baseline TKV is prognostic?

Assessment: Resolved.

17. Please provide more details for the Cox-models underlying some of your argumentation. Do parameter estimates change in between full models and, e.g., after backwards selection? Please provide (at least in an appendix) parameter estimates from univariate models, full models, and models after selection of variables.

Assessment: Partially resolved, please see list below.

18. Please base your argumentation about the prognostic value of TKV primarily on established clinical endpoints in renal disease and composite endpoints formed by these? Will other baseline co-variates be of relevance, as well?

Assessment: Baseline eGFR is of relevance according to the new modeling exercise to predict certain size of eGFR decrease in a certain time-frame. The applicant did not evaluate prediction of composite endpoints.

19. Please disentangle the role of eGFR (being stable for many years) from the role of TKV. Please exclude different baseline eGFR levels (primarily, as per CKD staging) and explain the role of TKV. Please discuss in how far worsening of eGFR is predictive of later stage clinical outcome and clinically relevant hard endpoints from the longitudinal data available in your database.

Assessment: Exercise pending see list below.

20. Please discuss the impact of different methods to measure/estimate GFR and to measure TKV on conclusions about predicting changes in GFR from TKV and age.

Assessment: Resolved.

Third list of issues to be addressed in writing by June 28 to focus the discussion during the discussion meeting

1. Please substantiate the external validity of the population included in this exercise.
2. Please justify why some of the analyses have been conducted in subgroups of the total dataset. Also comment on the large amount of missing information in the registries, especially the unavailability of eGFR is a surprise.
3. There is some doubt about your modelling approach: Did you add further variables only after TKV (or a transformation) has been already part of the model (explanation of residual variance)? What would be the outcome, if TKV, age and eGFR were modelled jointly with a backwards selection algorithm to arrive at a parsimonious model?
4. It may well be that TKV may add diagnostic certainty in early phases, whereas eGFR is a good predictor in later stages of disease. Please comment and investigate your data.
5. Please consider repeating the analysis with a logistic regression model. In addition ROC-analyses could be used to identify optimal cut-points for influential variables to discern between high and low risk.
6. Please discuss the clinical relevance of 30% worsening of eGFR (or of 57% worsening of eGFR). Is it possible to assess whether this is predictive of clinical outcomes (ESRD, transplantation, death, and composite endpoints) by analyzing your datasets.
7. Please discuss thoroughly the comprehensiveness in assessing all relevant confounding factors for disease progression that are not included into the model, such as use of ACEI, ARB, hypertension control, cyst suppuration and its control.
8. Please discuss the feasibility of learning – confirming paradigm for the TKV qualification. Do you foresee confirming/updating your model using external datasets (e.g. European Registries)?