EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

22 July 2015
EMA/CHMP/SAWP/473433/2015
Product Development Scientific Support Department

Draft qualification opinion

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

Draft agreed by scientific advice working party	04 June 2015
Adopted by CHMP for release for consultation	25 June 2015
Start of public consultation	22 July 2015
End of consultation (deadline for comments)	02 October 2015

Comments should be provided using this template. The completed comments form should be sent to qualification@ema.europa.eu

Keywords	polycystic kidney disease, kidney volume, patient selection, prognostic biomarker

30 Churchill Place ● Canary Wharf ● London E14 5EU ● United Kingdom
Telephone +44 (0)20 3660 6000 **Facsimile** +44 (0)20 3660 5555
Send a question via our website www.ema.europa.eu/contact

Introduction

The Critical Path Institute's Polycystic Kidney Disease Outcome Consortium (PKDOC) intends to qualify Total Kidney Volume (TKV) as a prognostic biomarker (i.e. predictive for the outcome with the current standard of treatment) to enrich the ADPKD population with the aim to conduct clinical trials more efficiently. The applicant started to work with the FDA in 2010 and initiated the discussion with the European Medicines Agency in 2013. Five observational studies including long-term outcome regarding the change in TKV over time, with or without various therapeutic interventions (such as diet, blood pressure control, cytostatics etc.) have been integrated into one database according to Clinical Data Interchange Standards Consortium (CDISC) standards.

The primary goal was the development of a Joint Model linking the trajectory of TKV, utilizing this standardized database, to predict clinical outcome variables (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 this model to predict which patients should be included into a future trial to arrive at a reasonable event rate in a more reasonable time frame in order to formally prove efficacy and positive benefit/risk of a new medical treatment. While all six endpoints were examined, there were not sufficient data or results to use CKD stage transition, hypertension, and mortality. The applicant has submitted substantial documentation to support the qualification. The development of the joint model, requiring at least two measurements of TKV per patient lead to a substantial loss in observations from the database. In addition, it turned out that baseline TKV per se is a similarly well predicting co-variate. Simple Cox-regression is thus a suitable tool to model the impact of TKV on the aforementioned endpoints.

The proposed Cox regression model includes 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 discussed to illustrate the anticipated uses of 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 of disease, 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 Agency indicated that all remaining questions could be addressed in the submission of an updated briefing package. The updated package was submitted on 20 March 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. The list of issues was sent on 20 May 2014. Response has been provided on 27 June, 2014 and a teleconference was planned on the 7th of July 2014. An additional request for data has been submitted to enable re-analyses for a better understanding of the competence of the database and the model.

TKV is a plausible predictor of clinical outcome with a relatively unspecified background of interventions and provided data allow qualifying TKV as a biomarker for enrichment of a potential trial population. However, the database as presented leaves some questions open with respect to the ability of the biomarker to efficiently enrich towards a trial population. These are discussed below. The proposed models could help studying the impact of certain criteria for inclusion or exclusion of patients on the

risk-profile in a future trial. Nevertheless, the provision of the raw data that have been used in the modelling process directly would enable trialists to select a population matching the one to be enrolled in a planned trial.

The current qualification opinion addresses these issues and provides recommendations to be eventually addressed in a forthcoming follow-up procedure or that might be directly implemented.

Context of use statement (as proposed by the applicant in the 20 March 2014 briefing book)

General area: Clinical trial enrichment in Autosomal Dominant Polycystic Kidney Disease (ADPKD)

Target population for use: Patients with ADPKD

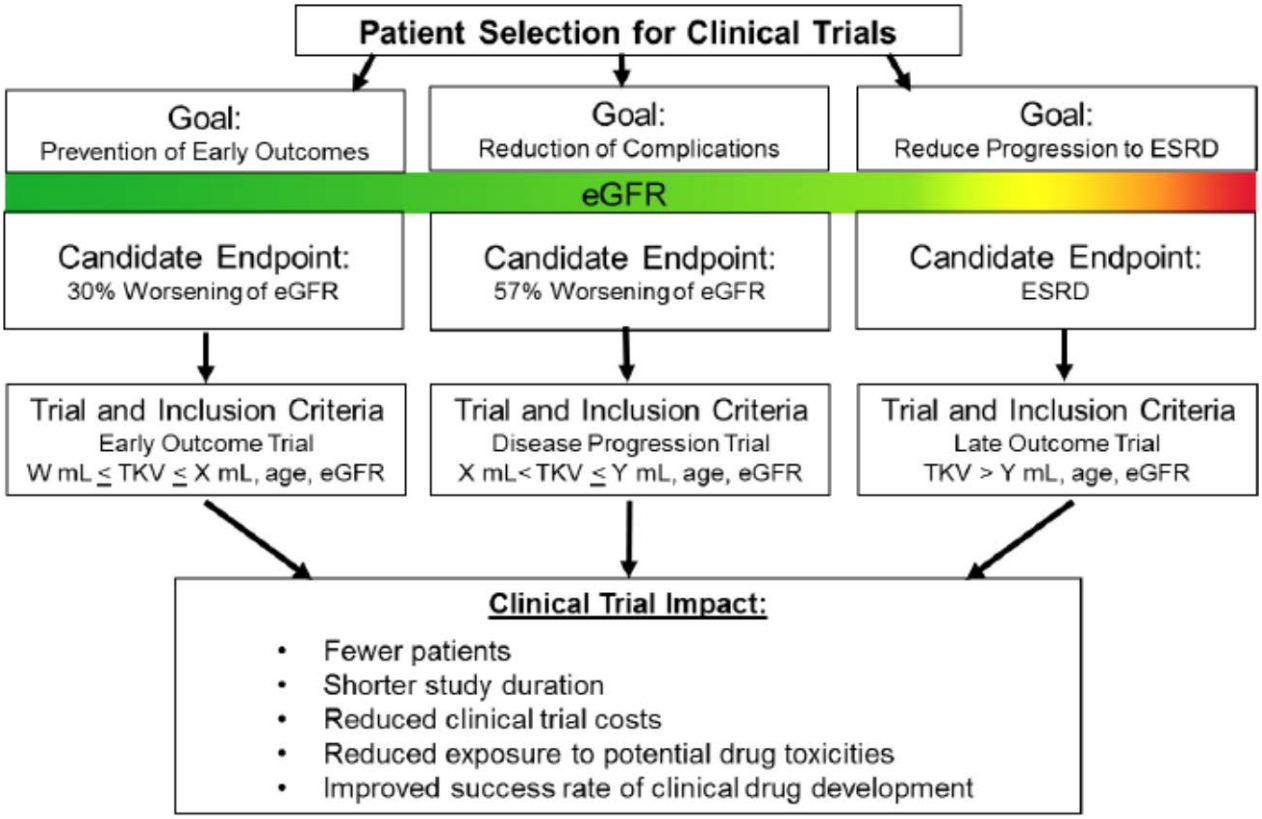
Stage of drug development for use: All clinical stages of ADPKD drug development, including proof of concept, dose-ranging, and confirmatory clinical trials.

Intended application: Baseline TKV can be applied as a prognostic biomarker that, in combination with patient age and baseline estimated Glomerular Filtration Rate (eGFR), can be used to help identify those ADPKD patients who are at the greatest risk for a substantial decline in renal function defined as (1) 30% worsening of eGFR, (2) 57% worsening of eGFR (equivalent to doubling of serum creatinine), or (3) End-Stage Renal Disease (ESRD, defined as dialysis or transplant). Baseline TKV will be used as an inclusion criterion in clinical trials to identify patients likely to show a clinically relevant decline in kidney function during the duration of the trial. Data are provided showing the calculated risk of each of these outcomes of declining renal function depending on age, total kidney volume, and baseline eGFR. Tables will be used by clinical trial researchers to determine the inclusion criteria to help select patients who are likely to reach the clinical endpoint of interest within a timeframe practical for the trial. These criteria include the optimum age, TKV, and eGFR for selecting subjects to be enrolled in the clinical trial.

TKV can be measured by Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scan, or ultrasound (US) imaging, and the volume calculated by a standard methodology, such as an ellipsoid volume equation, or by quantitative stereology or boundary tracing (for CT/MRI).

Using the same analysis and modelling approach described in section 5 of the briefing package, PKDOC also examined two other potential biomarkers, the longitudinal change in TKV and the rate of TKV growth. The longitudinal change in TKV did not improve prognostic performance beyond that provided by baseline TKV and age. Additionally, the rate of change of TKV requires longitudinal measurements making it an impractical biomarker for use as a clinical trial enrichment criterion. Therefore, these potential biomarkers were not included in this submission.

The figure below demonstrates an approach using TKV in drug development to enrich patient population, as proposed by the applicant in the 20 March 2014 briefing book.



Based on the qualification team report the CHMP gave the following answers

Question 1

Do the FDA and EMA agree that the Context of Use clearly describes how TKV will be used by applicants as a prognostic biomarker to enrich clinical trial population in clinical trials at all stages of ADPKD drug development, including proof of concept, dose-ranging, and confirmatory trials?

Applicant's position

PKDOC believes that the Context of Use as described in section 3.3 provides clinical trial researchers with a tool to select Baseline TKV, Baseline eGFR, and age cut-off values for use as inclusion criteria in clinical trials. Clinical trial researchers can use the tables supplied to understand how doing so will increase the probability of enrolling patients in the trial who are most likely to progress to a stage of renal disease that will meet the clinical endpoint of interest (see section 6 of briefing book).

CHMP answer

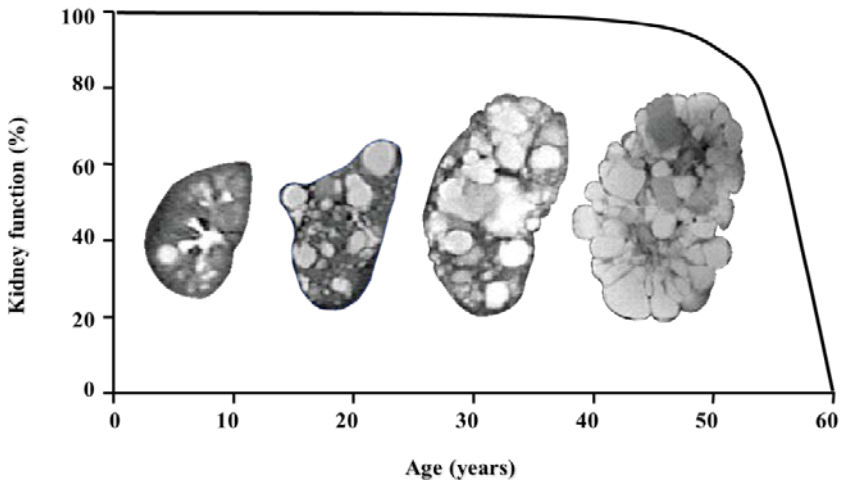
The applicant has formulated and investigated prediction models (Cox-regression models, and so-called Joint Models for time to event outcome variables (a linear mixed-effect model with a random intercept (baseline ln-transformed TKV) was used to fit ln-transformed TKV values over time)) that appropriately fit the clinical data and has derived cut-points for age and total kidney volume to predict outcome probabilities within a certain time-frame. During the discussions with regulatory agencies it has been further elaborated, that the models should include, in addition, baseline eGFR, which led to

an improvement of model fit. In consequence, cut-points in three-dimensions have to be derived to estimate the impact of patient selection on the probability of outcome in a certain time-frame.

Sample situations have been provided to illustrate the impact of these variables on the estimated event rate in a clinical trial for situations that have been envisaged for certain phases of drug development of drugs to treat ADPKD. However, whether the provided tables are really helpful to understand the structure of the forthcoming trial population is not yet sufficiently clear.

From the data provided it is reasonable to expect that baseline TKV can predict disease progression and is a biomarker valuable for risk stratification. The consortium is encouraged to further elaborate on the value of this biomarker with or without its dynamics during certain period of disease progression as the potential tool for enrichment.

The applicant stated that a key perspective that has come out of this project was the understanding that there are times (earlier in the disease continuum) when baseline TKV (in combination with age and eGFR) is clearly superior to eGFR alone in predicting future disease progression. On the other hand, there would be later phases in the disease, where eGFR (in combination with age and TKV) was equal or better than TKV. Neither was the best option during all stages of disease progression, however, in the context of enrichment then the aim should be to develop a model to predict from the initial stages of the disease which patients are at an increased risk for deterioration. The applicant has provided the following conceptual graphic, which clarifies that the clinical course of ADPKD is marked by a decades-long period of stable kidney function, as measured by eGFR, despite the relentless expansion of total kidney volume (TKV) due to growth of cysts.



As a tool in drug development and for planning clinical trials, the provision of the dataset (as presented to regulatory authorities) will provide a useful tool to study the impact of variations in the inclusion and/or exclusion criteria of patients on the expected event rate for a number of endpoints within different time-frames.

Question 2

Do the FDA and EMA agree that the following are clinically relevant endpoints of ADPKD and are adequate to track disease progression?

- **30 % Worsening of eGFR**

- **57% Worsening of eGFR (selected based on equivalence to doubling of serum creatinine)**
- **End-Stage Renal Disease (ESRD)**

Applicant's position

PKDOC believes that each is a relevant clinical endpoint in a PKD clinical trial, and that TKV can be used as an enrichment biomarker in a trial using any of these as an endpoint. See Sections 3.4.6, 5, and 6 of the briefing book.

CHMP answer

Agreement exists that a 30% worsening of eGFR, or a 57% worsening of eGFR are useful endpoints to develop a prediction model to identify patients at increased risk of worsening disease. Models for these endpoints may be developed to identify and include patients into a clinical trial that have an increased risk of developing ESRD, kidney failure or death that are established endpoints in kidney disease.

A 50% worsening of eGFR has been mentioned as an example for a change in eGFR that might serve as an endpoint in the draft guideline on the clinical investigation of medicinal products to prevent development/slow progression of chronic renal insufficiency (EMA/CHMP/355988/2014). The use of a 30% worsening of eGFR as an endpoint is still controversial, because correlation with clinical endpoints is less firmly established, and may be affected by acute drug effects on eGFR. Moreover it is not clear whether drug induced changes of this size predict ESRD or death. Its application should therefore be restricted to those situations (e.g. phase II dose-finding), where independent replication in a phase III clinical trial with more robust endpoints is foreseen and rare disease may not be the most appropriate place to provide further evidence for surrogacy.

It is noted that the applicant has not provided specific information regarding the surrogacy of the aforementioned changes in eGFR for clinical endpoints in this qualification procedure. During this biomarker validation procedure it has been requested to investigate whether the correlation between the surrogates and the clinical outcome of ESRD, and mortality can be formally demonstrated from the collected registry data. The applicant didn't provide the respective information eventually because of paucity of the data in its longitudinal aspect.

The CHMP reiterates that its position to accept sole 30% and 57% decrease in eGFR as markers for the proposed prediction models should not be seen as an acceptance of these surrogate endpoints for clinical studies, as this was not part of the current submission, and as the applicant failed to provide data linking these surrogate markers with hard clinical endpoints in the target population.

Question 3

Do the FDA and EMA agree that the totality of data accumulated and the scientific evidence generated through the execution of the PKDOC Research plan, is sufficient in supporting the qualification of Baseline TKV, in combination with age and baseline eGFR, as a prognostic biomarker in ADPKD patients?

Applicant's position

PKDOC believes that the rich source of longitudinal data from three academic registries and two observational trials provide both sufficient quantity and diversity of data to support the qualification, and that the modelling and validation approach are state-of-the-art and in agreement with what was

previously discussed as an approach to use. The results of the analysis show a strong correlation between baseline TKV and the likelihood of renal disease progressing to one of the three endpoints above, and can reliably be used as an inclusion criterion.

CHMP answer

CHMP agrees. A number of prediction models have been proposed and described in this application. These models relate TKV, age, and eGFR at baseline to the aforementioned clinical outcomes. These models suggest that TKV is predictive for an increased risk in patients and can be used to investigate the impact of certain thresholds for criteria on inclusion and exclusion on the probability of an event within a certain time-frame, and thus support concepts of enrichment.

However a number of limitations hold true on the concept of enrichment, the model building dataset and the utility of the predictive models.

Impact of enrichment

The impact of enrichment deserves further consideration in the following aspect: recruitment of an enriched population may take longer so that some (or all) of the effect of enrichment (a higher event rate in a shorter period of time to demonstrate a larger treatment effect) is compensated by a generally longer duration of recruitment for the trial. This aspect was discussed with the applicant, as well, and it is recommended that it is carefully investigated whether under the specific conditions for the disease under investigation, enrichment is of real benefit for the conduct of future clinical trials.

The applicant confirmed that enrichment may be of value even if recruitment duration would be prolonged because expensive investigations could be spared. They also confirm that each applicant in designing the trial for their particular compound will balance any trade off of enrichment versus enrolment based on their own internal decision. They explain that the availability of a qualified prognostic biomarker does not force its use on any application; however, it offers the applicant better options to aid in designing the trial that best suits their needs.

The second effect of enrichment, namely the investigation of the treatment effect in a sub-population of the patient population again comes at the price that benefit/risk of treatment in a broader population may differ from what is seen in the enriched population. As the whole point of enrichment is to facilitate demonstration of a treatment effect that can be investigated with a smaller sample-size it will often be the case that in the complement the treatment is the same, or smaller than in the enriched population. Therefore, the benefit/risk-ratio will require re-discussion. Enrichment may also lead to a population that is less or more amenable for a treatment effect and extrapolation will be particularly difficult if this cannot be excluded. It is important to note that without clinical data and particularly in a rare condition it may be difficult to find information to allow bridging from the enriched to the non-enriched population.

This consequence of enrichment needs to be discussed at the time-point of licensing.

External validity of the model building dataset

The applicant has made an enormous attempt to collect all available systematic evidence and to include this in a systematic and structured way into one database using CDISC standards. This is for rare disease a highly appreciated undertaking. If model parameters are estimated from this data-source it has to be assumed that the database in totality is representative of the ADPKD population also regarding its quantitative composition and epidemiological aspects.

The applicant is convinced that this is the case and states that all patients who entered the clinics during the timeframe of data collection and had a diagnosis of ADPKD were included in the database, unless they were already on dialysis or had a kidney transplant. The applicant states that all of the studies underwent national recruitment efforts and included all races and ethnicities. The applicant re-iterates that data from multiple, longitudinal, well-characterized observational registries maintained by PKD investigators at leading American academic medical institutions extending over seven decades were utilized. The applicant is convinced that while each individual dataset on its own may not be entirely representative due to some population differences, the characteristics of the aggregated datasets are very representative of the general ADPKD population.

From the raw data provided it is, however, not possible to understand the impact of each of the different data sets on the outcome. Nevertheless the applicant articulates its strong believe that the provided dataset is representative of the ADPKD population.

For various reasons the set of observations finally included into the modelling process is substantially restricted against the combination of the original datasets. It has to be assumed, in addition, that these restrictions do not impact on the ability to properly develop a risk model. This aspect should deserve thorough consideration particularly as it leads to non-testable assumptions about nature and reason for missing data, or the availability of additional measurements and information. This again increases the importance of an independent replication step.

The applicant has explained why some observations needed to be excluded and expressed their strong belief that these exclusions do not bias the final conclusions in how far TKV predicts outcome.

Need for external validation

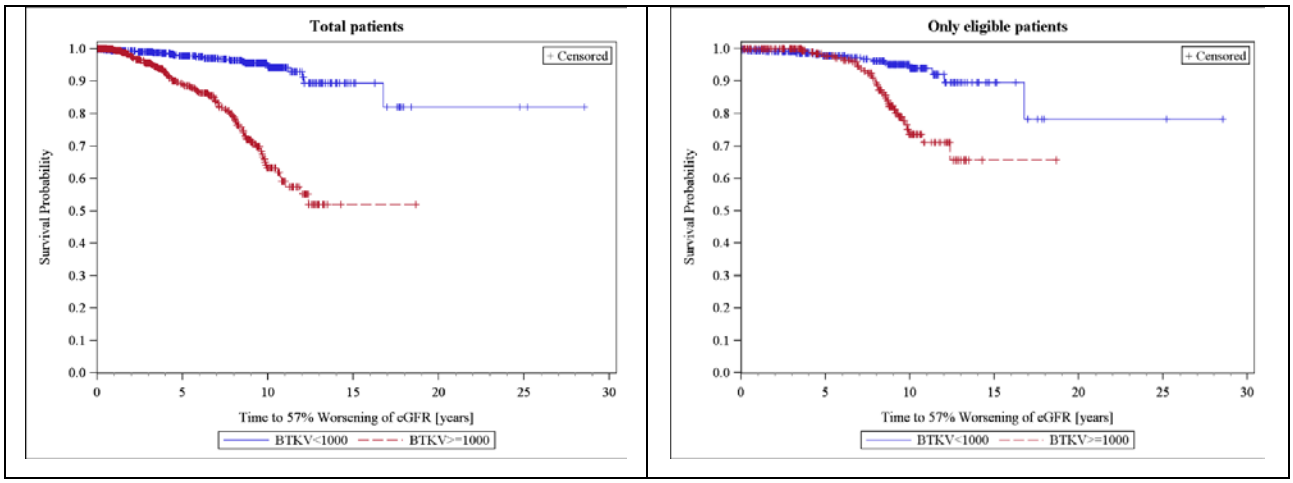
For understandable reasons all the information in the source databases has been used in the model building approach and therefore no independent replication of the modelling process is available. Cross-validation techniques have been used to ameliorate this aspect, but these cannot formally replace an independent verification step. PKDOC represents a number of world pharmaceutical companies and it is currently unclear, whether any data from randomized clinical trials are available to the consortium that could be used to test / validate the model or to demonstrate that with appropriately chosen criteria for inclusion and exclusion a population within the PKDOC dataset can be identified that is structurally similar to the trial population.

Meanwhile a randomized clinical trial (TEMPO 3/4) has been published and the respective dataset could be used to independently replicate some aspects of the PKDOC dataset under investigation. The data could be used to test / validate the model or to demonstrate that with appropriately chosen criteria for inclusion and exclusion a population within the PKDOC dataset can be identified that is structurally similar to the trial population (as has been done in some of the additional analyses provided below).

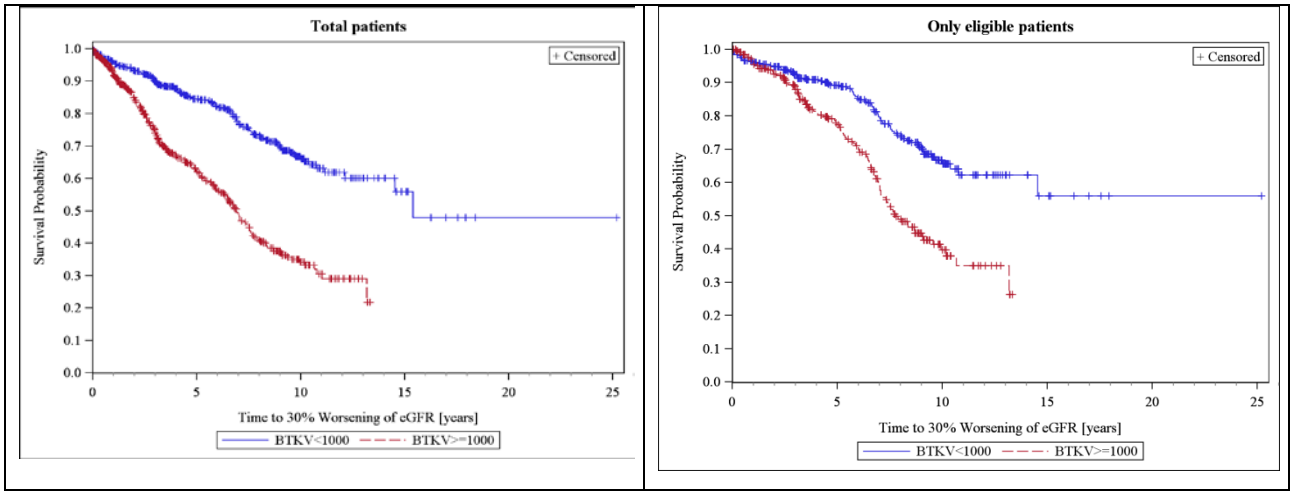
Model building dataset vs. eligible population for clinical trials

Trials in ADPKD were checked and it was found that most of them investigate patients from the age of 18 years and older and have and baseline eGFR between 50 and 200 mL/min/1.73m^2. A dataset has been provided from PKDOC for the qualification team. A re-analysis of the provided data has been conducted and predicted outcomes in a subset of the dataset matching the clinical trial eligible population.

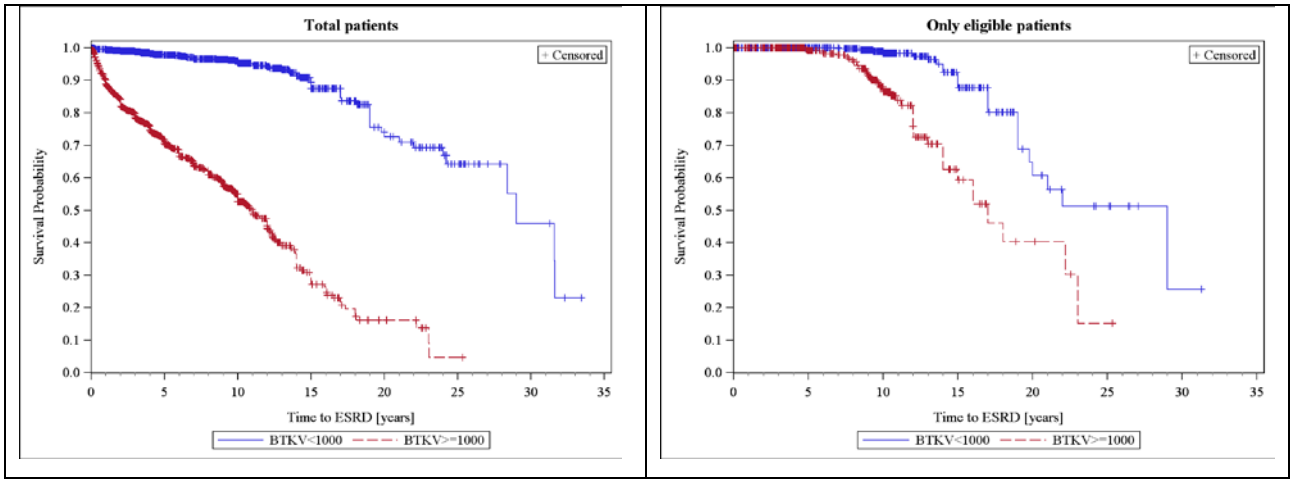
Results for a 57% worsening of eGFR (below) demonstrate that the effect of TKV in selecting a population at higher risk is grossly overestimated in the full population, but visible also in the “eligible” population.



266 Similar findings can be demonstrated for 30% worsening of eGFR, though curves separate earlier:



267 End stage renal disease curves are again more strongly overlaying in the eligible population and
268 separate lately.



269 Implications for study planning are obvious: the time until the development of traditional endpoints is
270 still substantial even if population enrichment is taken into consideration. Arguments as outlined above
271 are of importance to properly estimate the value of enrichment.

Utility of predictive models for study planning

One of the aims of this qualification is also the provision of novel tools for study planning. The direct utility of the model as presented or the derived tables are questioned: while these tables have been calculated from the developed models, they only reflect interactions between model parameters to the extent that they are captured in the model. Of note, causal (parsimonious) models may be inferior in predicting outcome as compared to models including more co-variates. Researchers utilizing these models would have to simulate them with the anticipated population fraction to properly reflect the forthcoming study population. This also requires an understanding of the distribution of the co-variates and potential interactions beyond what is available from the causal models. A more elegant approach is the provision of the (appropriately anonymized) datasets so that a direct restriction of the population will be possible for the interested researcher trying to understand the impact of changes in the criteria for inclusion and exclusion on the event rate in the control group of a future trial.

The applicant has agreed to provide a completely anonymised dataset so that interested parties can apply potential sets of criteria for inclusion and exclusion of patients to a potential trial population and then in detail investigate the structure of this patient population and the applicant confirms that this step is under consideration.

Qualification opinion

CHMP support baseline total kidney volume, in combination with patient age and eGFR as a prognostic biomarker to identify patients likely to experience a progressive decline in renal function, as characterized by a decline in eGFR or progression to end-stage renal disease.

CHMP encourage ADPKD trialists, to request access to the anonymised PKDOC dataset, in order to investigate the impact of inclusion/exclusion criteria, and more specifically baseline TKV, age and eGFR, on clinical outcomes. It is envisaged that access to PKDOC dataset will help optimizing clinical trials in terms of population to be enrolled, study duration and expected placebo effect. As discussed the impact of enrichment should be weighed against enrollment times. In addition as changes in eGFR and TKV occur at different disease stages, it is crucial to consider the mechanism of action and anticipated treatment benefit when selecting a population for clinical trials. Finally the relevance of the benefit risk demonstrated in an enriched population, to the wider PKD population, needs to be justified on a case by case basis.

Regulators encourage further data sharing activities and analyses to replicate the findings from PKDOC and are open to follow up qualification discussions.

Annexes

- PKDOC final briefing book (20 March 2014)
- 3rd list of issues (20 May 2014)
- Written responses from applicant to 3rd list of issues (27 June 2015)
- Applicant's presentation (03 July 2015)