Food and Drug Administration
Center for Drug Evaluation and Research
Division of Cardiovascular and Renal Products

Date: June 3, 2015
From: Aliza Thompson, Medical Officer
Division of Cardiovascular and Renal Products, CDER

Aliza M. Thompson -S
Digitally signed by Aliza M. Thompson -S
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300423038, cn=Aliza M. Thompson -S
Date: 2015.06.03 16:04:41 -04'00'

Through: Norman Stockbridge, Director
Division of Cardiovascular and Renal Products, CDER

Norman L. Stockbridge -A
Digitally signed by Norman L. Stockbridge -A
DN: c=US, o=U.S. Government, ou=HHS, ou=FDA, ou=People, 0.9.2342.19200300.100.1.1=1300068764, cn=Norman L. Stockbridge -A
Date: 2015.06.03 17:04:53 -04'00'

To: Marianne Noone
Regulatory Project Manager
FDA/CDER Biomarker Qualification Program

Subject: Clinical Review of PKD Outcomes Consortium biomarker qualification submission

Materials Reviewed:

- Final briefing book titled "Qualification of Total Kidney Volume as a Prognostic Biomarker for use in Clinical Trials Evaluating Patients with Autosomal Dominant Polycystic Kidney Disease (ADPKD)", dated March 20, 2014
- PKDOC response, dated July 2, 2014, to FDA's June 24, 2014 information request
- PKDOC response, dated September 19, 2014, to FDA's September 12, 2014 information request
- FDA Statistical Review (Dr. John Lawrence), dated February 18, 2015
- FDA Secondary Statistical Review (Dr. Sue-Jane Wang), dated May 14, 2015
- Emails from Dr. John Lawrence, dated May 21 and 29, 2015

Executive Summary:

On March 20, 2014, the Polycystic Kidney Disease Outcomes Consortium submitted their final briefing book to support the qualification of total kidney volume as a prognostic biomarker for use in clinical trials evaluating patients with autosomal polycystic kidney disease (ADPKD). The submission contained analyses intended to support the use of baseline TKV, in combination with patient age and baseline eGFR, as a prognostic biomarker to enrich the clinical trial population in trials at all stages of ADPKD drug development, including proof of concept, dose-ranging and confirmatory trials.

In support of the proposed context of use, the consortium aggregated data from three patient registries and two longitudinal cohort studies of the natural history of ADPKD. These data were used to develop a predictive model linking baseline TKV (in combination with age and baseline eGFR) to the following outcome measures: a confirmed 30% reduction in eGFR, a confirmed 57% reduction in eGFR, and end stage renal disease (ESRD).

According to FDA's statistical reviews, including TKV in addition to patient age and eGFR in the best fit model provided a modest improvement over the best fit model using age and eGFR in predicting the risk of a confirmed 30% decline in eGFR. This finding was confirmed using cross-validation and in a separate dataset (external validation), thus supporting the use of the biomarker for prognostic enrichment. Relative to the model without TKV, the model with TKV gave a slightly higher average predicted probability for patients with events and a slightly lower average predicted probability for patients without events at year 3, indicating an improvement in discrimination over 3 years. As noted in Dr. Lawrence's review, there were too few ESRD and 57% decline in eGFR events over the time frame of a feasible clinical trial to perform meaningful analyses.

Analyses were also performed to assess the impact of using the model with TKV, as compared to the model without, on the number needed to enroll and number needed to screen to get one event. In an analysis that specified separate entry criteria for each parameter of interest (age, eGFR and TKV), using the best fit model with TKV, as compared to the best fit model without, modestly reduced the number needed to enroll to get one event but substantially increased the number needed to screen. The value of the best fit model with TKV was more apparent when a risk score was used to define entry criteria. In Dr. Lawrence's analysis, both the number needed to enroll and the number needed to screen to get one event were smaller when the best fit model with TKV, as compared to the best model without, was used to enroll patients with the top 50% of risk scores.

From a clinical perspective, we believe these data support the qualification of TKV as a prognostic biomarker for clinical trial enrichment in patients with ADPKD. Specifically, these analyses indicate that using TKV in addition to a patient's age and baseline GFR provides incremental improvement in the ability to predict a 30% decline in eGFR. However, further work is needed to understand the likely impact on trial efficiency of using TKV for prognostic enrichment or whether it varies across the stages of disease. The analyses indicate that enrolling patients with larger kidneys will result in a gain in event rates in trials of ADPKD, but this gain must be weighed against the trade-off of excluding patients who would otherwise be eligible for enrollment. As noted in the statistical review, missing data, the potential for selection bias, as well as other differences in how covariates and endpoint events were defined in this study vs. a typical clinical trial setting, may also affect the ability to generalize event rate predictions from this study population to future clinical trials.

Moving forward, we hope that further work will be done to refine and improve the use of TKV as a prognostic biomarker. We also hope that more work will be done to characterize the trade-offs of using TKV for prognostic enrichment, and, specifically, the impact of a given cut point for TKV or calculated risk score above some level on the number needed to enroll versus the number needed to screen. Analyses of screening data from past clinical trials could provide important insight into the effect of a chosen cut-point or risk score on the number needed to screen.

Disease Background and Proposed Context of Use

ADPKD is reported to be the most common hereditary kidney disease and is thought to have a "phenotypic prevalence" of 1:400 to 1:1000 based on autopsy data. The disease is characterized by

progressive enlargement of the kidneys due to cyst growth and formation, and, in some patients, the progressive loss of renal function resulting in end-stage kidney disease.

ADPKD has a variable course. It is reported that in up to half of those who are diagnosed, progressive kidney dysfunction develops. In those who progress to end stage renal disease, the average age at transplant or initiation of dialysis is ~55-56 years.

At present there are no approved therapies to treat ADPKD. Given the extended time frame for disease progression, as well as the perception among experts in the field that therapies targeting cyst growth and formation may be more effective if administered early in the course of therapy, there has been considerable interest in the use of TKV as a surrogate endpoint to establish the effectiveness of therapies for ADPKD. Although the consortium's ultimate goal is to qualify TKV as a surrogate endpoint, the current submission focuses on the qualification of TKV as a prognostic biomarker for patient selection in clinical trials.

Proposed Context of Use Statement

- **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). This biomarker 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 (for ultrasound), or by quantitative stereology or boundary tracing (for CT/MRI).

Data submitted to support qualification

Sources of data:

In support of the proposed context of use, the consortium submitted analyses using aggregated data from three patient registries and two longitudinal cohort studies of the natural history of ADPKD.

- The three patient registries/databases were from the University of Colorado-Denver, the Mayo Clinic and Emory. Data from patients with a diagnosis of ADPKD and at least one available kidney volume measurement were included in the consortium's analysis dataset.
- Data were also obtained from the Consortium for Radiologic Imaging Studies of Polycystic Kidney (CRISP) I and CRISP II, prospective, longitudinal cohort studies of the natural history of APDKD. CRISP I followed 241 patients annually for a period of three years; CRISP II extended the follow up of patients enrolled in CRISP I.
- Some patients who contributed data to the patient registries were also enrolled in the CRISP studies.

In aggregate, these registries and studies provided data on a total of 2355 patients with at least one TKV measurement. However, as might be expected, the variables of interest were not systematically captured in the registries, leading to missing data.

The imaging modalities that were used varied across the different sources of data, as shown below. For additional information on the Mayo, Emory and Colorado registries and CRISP studies, see the appendix.

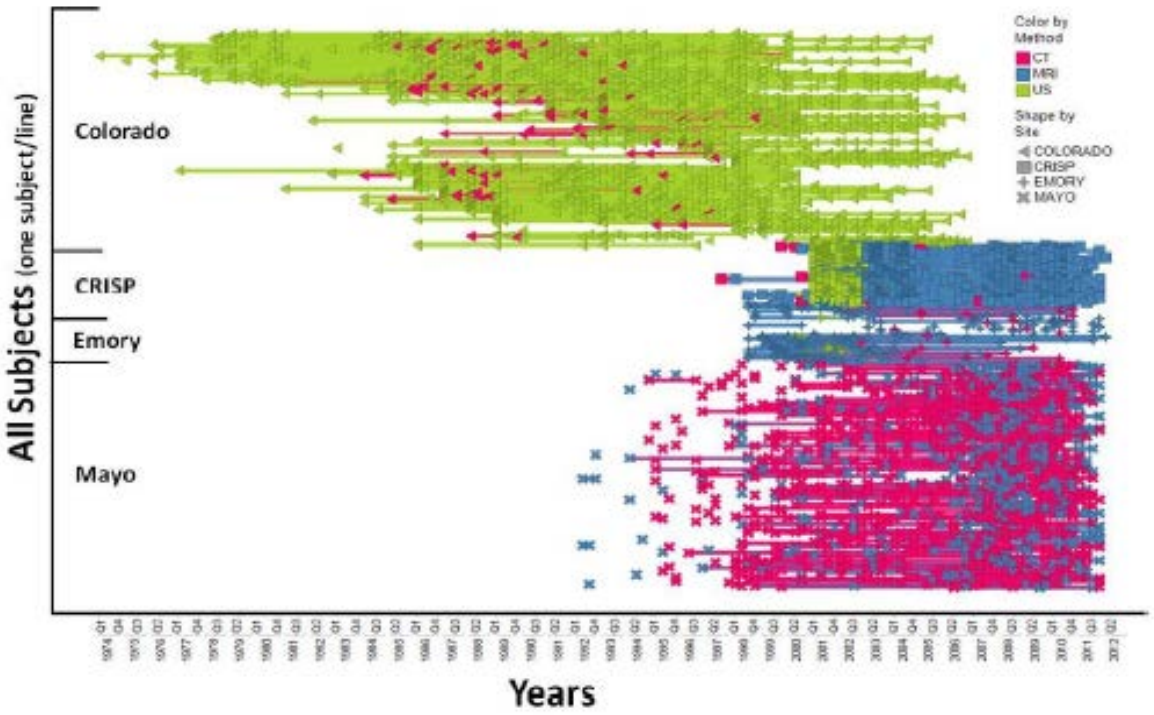

Figure 1: Time span and image modality by site/source of data

Source: Figure 23, Submitter's Briefing Book; includes all subjects with 1 or more images (n=2355)

Study population:
As previously noted, the submitter's dataset contained data on 2355 patients with ADPKD and at least one TKV measurement. The discussion that follows focuses on the baseline characteristics of the patients in this dataset (n=2355); because of missing covariate data, subsets of this population were used in particular analyses.

Age at study entry ranged from 0-84 years, with a mean and median age of 36.1 and 37.7 years, respectively. Approximately 59% of patients were female, and 82% were Caucasian. The median eGFR was 67 mL/min/1.73m^2 using the CKD-EPI equation; the mean (SD) was 70 mL/min/1.73m^2 (38).

Baseline patient characteristics, including age and renal function, varied across the sites/ sources of data. As shown in the figure below, the dataset included patients with very low eGFRs at baseline, including patients with stage 5 CKD. The submitter's dataset also included very young children (see figure in appendix).

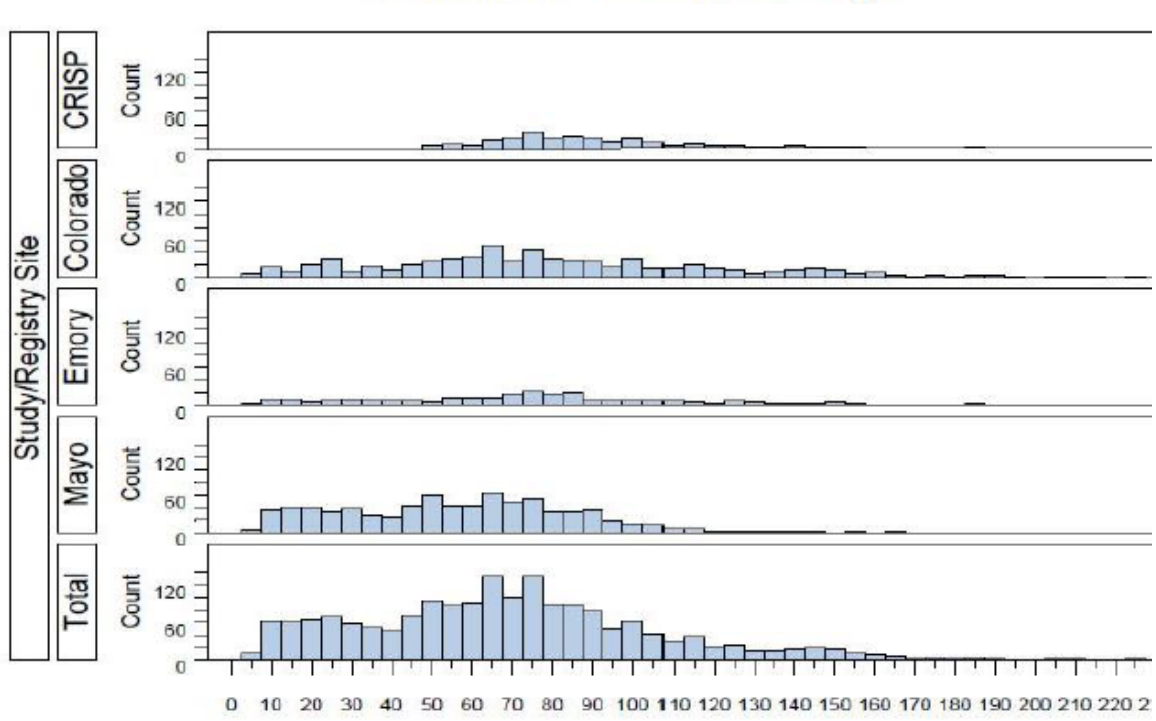

Figure 2: Distribution of eGFR at first image by site/source of data
Source: Figure 18, Submitter's Briefing Book; All eGFR values at First Image (n=1792)

The relative frequency of the endpoint events of interest (a 30% decline in eGFR, a 57% decline in eGFR and ESRD) varied across the sites/sources of data. The overall shape of the distribution of events in the CRISP cohort is what one might expect in a population with relatively preserved renal function (i.e., the number of 30% decline events is greater than the number of 57% decline events, with few ESRD events). A similar distribution of events is not seen in the registry data (i.e., Mayo, Emory, Colorado). This may be because some of the registry patients had late stage disease at the time of enrollment and/or because

follow-up serum creatinine measurements were not systematically captured at registry sites, resulting in incomplete ascertainment of eGFR decline endpoint events.

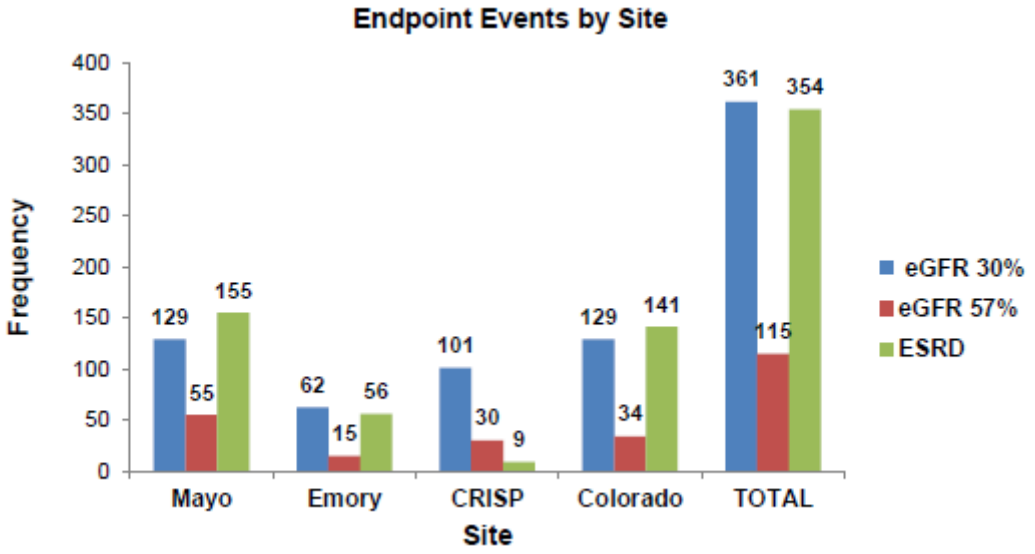

Figure 3: Number of endpoint events by site/data source
Source: Figure 22, Submitter's Briefing Book

Analytic Approach:
TKV values measured by MRI, CT, or US modalities were used in the analyses. Analyses were performed for three different outcome measures: a 30% reduction in eGFR, 57% reduction in eGFR, and ESRD (start of dialysis or kidney transplant). For the 30% and 57% reduction in eGFR, a subsequent measurement within any timeframe was required to confirm that the original decline was not transient. eGFR was estimated using the CKD-EPI equation.

As noted in Dr. Lawrence's review, FDA's approach to analyzing the data differed somewhat from the approach taken by the submitter. Herein, we highlight two of these differences.

- As discussed in the prior section, the submitter's dataset included very young patients and also patients with end-stage disease (i.e., a GFR < 15) at baseline/time of entry into the dataset. The review team felt that the analyses would be more applicable to the trial setting if the criteria for including patients in the analysis reflected, to the extent possible, the design of clinical trials. The review team raised this issue with the consortium and, based on the consortium's recommendations, the analyses were limited to patients with an eGFR ≥25 and age > 12 years.
- Some subjects had imaging performed with more than one modality.[1] Given data indicating that US measurements are less accurate and precise than MRI measurements, Dr. Lawrence used an MRI measurement if available. If no MRI measurement was available for a subject, then a CT

[1] For example, in the overall analysis dataset for the 30% decline in eGFR endpoint (n=1140 patients), 200 had baseline TKV measurements with two different modalities.

measurement was used; if no MRI or CT measurement was available, then a US measurement was used. In contrast, the submitter used the average of the measurements obtained using different modalities if more than one was available.

Findings:

30% decline in eGFR endpoint

The discussion that follows focuses on FDA's analyses, since the criteria for including patients in the analysis dataset may better reflect the study population likely to be enrolled in clinical trials.

- *Study population.* Of the 2355 patients in the database with at least one TKV measurement, 1215 patients had missing covariates, 664 of which had a missing baseline eGFR. The table below shows the baseline eGFR and age of the population used in the submitter's analysis dataset (n=1140) and FDA's (n=925) to develop a model for the 30% decline in eGFR endpoint. As might be expected, because FDA limited its analysis to patients with an eGFR $\geq$ 25 and age > 12 years, mean age and eGFR were slightly higher and TKV was slightly lower in FDA's analysis dataset.[2] There were 361 patients with a 30% decline in eGFR in the submitter's dataset and 300 in the FDA's. The proportion of patients who developed a 30% decline in eGFR was similar in the two datasets (~32%).

Table 1: Summary statistics for populations used for model development for the 30% decline in eGFR endpoint

Characteristic	Dataset used for model development	
	Submitter (n=1140)	FDA (n=925)
Age, mean (SD)	38.8 (15.8)	39.4 (13.4)
TKV, mean (SD)	1.5 (1.4)	1.3 (1.1)
eGFR, mean (SD)	70.1 (37.8)	73.4 (30.2)
eGFR <60, n (%)	453 (39.7%)	312 (33.7%)
eGFR <50, n (%)	336 (29.5%)	195 (21.1%)

Sources: Figure 25, Submitter's Briefing Book; Table 3, FDA Primary Statistical Review; email correspondence from Dr. Lawrence dated March 2, 2015

- *Selection of best fit models.* As discussed in the statistical review, Dr. Lawrence searched for the best fit model with and without TKV using the covariates age, baseline eGFR (mL/min per 1.73 m^2 or the log transformation) and the two-way interactions. Based on the Akaike information criterion (AIC), a measure of the relative quality of a statistical model for a given set of data, the best fit model without TKV included terms for age, log(eGFR) and the two-way interaction; the best fit model with TKV included log(TKV), age, eGFR and all two-way interactions. The table below shows the coefficients of the best fit model with TKV.

[2] A total of 141 subjects were excluded from FDA's analysis dataset because of an eGFR < 25 mL/min/1.73m^2.

Table 2: Coefficients for best fit model including covariates log(TKV), age, and eGFR

Term	Coefficient
log(TKV)	2.8323
Age	0.23802
eGFR	0.11326
log(TKV):Age	-0.027187
Age:eGFR	-6.0368×10^{-4}
log(TKV):eGFR	-0.012429

Source: Table 8, FDA Primary Statistical Review

- *Improvement in model fit.* Dr. Lawrence used the AIC to compare the best fit models with and without TKV. The improvement in AIC for the model with log (TKV) over the model without TKV was 86; according to Dr. Lawrence, this indicates substantial improvement in model fit.

- *Improvement in discrimination.* The C-statistic for a time-to-event outcome measures the probability of concordance between a patient's risk score and the time to the event. At 1, 2, 3 and 5 years and over the maximum period of follow-up, the c statistic was greater for the best fit model with TKV than for the best fit model without, indicating that the model with TKV was better able to discriminate patients who were likely to have a 30% decline in renal function at these time points from those who were not.

Table 3: C Statistic at different time points for 30% decline in eGFR using a model with and without TKV

Time (years)	C no TKV	C with log(TKV)	p-value
1	0.424	0.617	0.002
2	0.493	0.660	0.0002
3	0.599	0.698	0.0008
5	0.617	0.703	0.0001
Maximum	0.598	0.686	0.002

Source: Table 9, FDA Primary Statitical Review

- *Cross-validation*. Dr. Lawrence performed cross-validation using the submitter's dataset to assess the predictive ability of the modeling process. In brief, the submitter's dataset was partitioned into five mutually exclusive subsets. Four of these subsets were combined and used to determine the best fit model with and without log(TKV); the models were then tested in the remaining subset (the validation set). This process was performed five times to explore all possible combinations of the five subsets. As shown below, for nearly all iterations, the *c* statistic was higher at 2, 3 and 5 years using the models with TKV compared to the models without, indicating that the models with TKV were better able to discriminate subjects likely to develop a 30% decline at these time points.

Table 4: Cross-validation C-statistics at different time points for 30% decline

CV Iteration	Time (Years)	C no TKV	C with log(TKV)	p-value
1	2	0.331	0.529	<0.001
1	3	0.481	0.603	0.003
1	5	0.464	0.579	<0.001
2	2	0.453	0.719	<0.001
2	3	0.596	0.765	<0.001
2	5	0.628	0.757	<0.001
3	2	0.607	0.688	<0.001
3	3	0.582	0.703	0.001
3	5	0.617	0.705	<0.001
4	2	0.371	0.498	<0.001
4	3	0.640	0.635	0.720
4	5	0.673	0.689	0.583
5	2	0.491	0.683	<0.001
5	3	0.577	0.668	<0.001
5	5	0.585	0.662	0.001

Source: Table 11, FDA Primary Statistical Review

- *External validation using a separate dataset.* A separate dataset was used to evaluate whether the biomarker's predictive performance generalized to populations other than the one in which it was developed. As noted in Dr. Lawrence's review, FDA's independent validation using a separate dataset that was available internally supported the predictive performance of TKV and hence TKV's qualification as a prognostic biomarker (data not shown in FDA's statistical review).

- *Comparison with Kaplan-Meier estimated event rates.* Dr. Lawrence also compared cross-validation estimated event rates and Kaplan-Meier (K-M) estimated event rates at different time points. As shown in the figure below, the models with and without TKV performed similarly in predicting the K-M estimate of events rates.

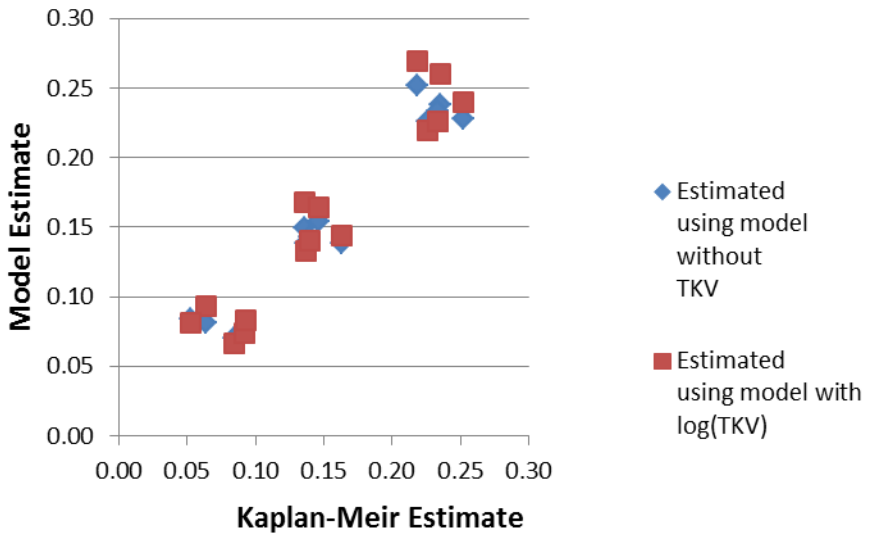

Figure 4: Estimated event rates for a 30% decline: Comparison of cross-validation Kaplan-Meier estimated event rates (x-axis) and estimated event rates using models with and without TKV (y-axis)
Source: Reviewer's plot, based on Table 10, FDA Primary Statistical Review

Reviewer's comment: Given the aforementioned analyses, we would have expected that the model with TKV would have performed better than the model without TKV in predicting the event rate; however, this was not the case.

57% decline in eGFR and ESRD events

As shown in Figure 3, there were 354 ESRD events and 115 57% eGFR decline events in the submitter's dataset. In the restricted analysis set (i.e., patients with an eGFR ≥ 25 and age >12), there were 182 ESRD events and 99 57% eGFR decline events. Of the ESRD events, 47 occurred in the first five years and 12 events occurred in the first 3 years. Of the 57% decline in eGFR events, 34 occurred with the first five years. There were too few events over the timeframe of a feasible clinical trial to perform meaningful analyses.

Reviewer's comment: It is not surprising that there were few ESRD and 57% eGFR decline events in the analysis dataset given the mean baseline eGFR of the study population and rate of disease progression.

Height-adjusted TKV

Recent studies in the published literature have focused on height-adjusted TKV as prognostic biomarker of renal disease progression in ADPKD.[3,4] According to the submitter, log-transformed baseline height-

[3] Irazabal MV, Ragnel LJ, Berstralh EJ et al: Imaging Classification of Autosomal Dominant Polycystic Kidney Disease: A Simple Model for Selecting Patients for Clinical Trials. *J Am Soc Nephrol* 26:160-172, 2015
[4] Chapman AB, Bost JE, Torres VE, et al: Kidney volume and functional outcomes in autosomal dominant polycystic kidney disease. *Clin J Am Soc Nephrol* 7: 479–486, 2012

adjusted TKV and log-transformed baseline TKV resulted in similar predictive power (ROC values at one and five years) in univariate Cox analyses for ESRD, a 30% decline in eGFR, and a 57% decline in eGFR. Log-transformed baseline TKV was used in the submitter's multivariate Cox models since it was deemed more convenient to use in a clinical setting.

Illustration of the potential utility of using TKV for trial enrichment

To illustrate the potential value of using TKV for trial enrichment, the submitter generated a table showing the predicted probabilities of a 30% worsening of eGFR by baseline age (< or ≥ 40 years), baseline TKV (< or ≥ 1 L) and baseline eGFR (< or ≥ 50 mL/min/1.73m^2). Given questions about the underlying assumptions and approach used in the submitter's analysis, a similar table was constructed relying on FDA's model. Overall, predicted event rates using the FDA model are higher with a TKV≥ 1 L, relative to TKVs < 1 L in these age and GFR subsets.

Table 5: Predicted probabilities of not having a confirmed 30% decline in eGFR in subgroups defined by baseline age, TKV and eGFR using FDA model

Followup Time (Years)	Predicted probability of not having a confirmed 30% decline in eGFR							
	TKV < 1 L				TKV ≥ 1 L			
	Age < 40 years		Age ≥ 40 years		Age < 40 years		Age ≥ 40 years	
	eGFR ≥ 50 mL/min per 1.73 m^2	eGFR < 50 mL/min per 1.73 m^2	eGFR ≥ 50 mL/min per 1.73 m^2	eGFR < 50 mL/min per 1.73 m^2	eGFR ≥ 50 mL/min per 1.73 m^2	eGFR < 50 mL/min per 1.73 m^2	eGFR ≥ 50 mL/min per 1.73 m^2	eGFR < 50 mL/min per 1.73 m^2
1	0.971 (0.005)	0.982 (0.004)	0.977 (0.005)	0.954 (0.015)	0.960 (0.008)	0.916 (0.020)	0.961 (0.007)	0.914 (0.014)
2	0.944 (0.008)	0.965 (0.007)	0.958 (0.008)	0.919 (0.024)	0.922 (0.012)	0.851 (0.026)	0.928 (0.010)	0.851 (0.021)
3	0.895 (0.010)	0.938 (0.011)	0.916 (0.013)	0.853 (0.038)	0.846 (0.016)	0.735 (0.034)	0.858 (0.016)	0.730 (0.035)
4	0.860 (0.012)	0.920 (0.015)	0.879 (0.016)	0.803 (0.045)	0.787 (0.020)	0.654 (0.039)	0.800 (0.020)	0.640 (0.043)
5	0.828 (0.013)	0.897 (0.017)	0.854 (0.018)	0.765 (0.052)	0.738 (0.022)	0.594 (0.043)	0.758 (0.023)	0.581 (0.048)

Source: Table 13, FDA Primary Statistical Review

Dr. Lawrence also constructed a similar table showing the Kaplan-Meir estimates of a 30% worsening of eGFR by baseline age (< or ≥ 40 years), baseline TKV (< or ≥ 1 L) and baseline eGFR (< or ≥ mL/min/1.73m^2). In patients with an eGFR ≥ 50, the predicted probability of having a confirmed 30% decline in eGFR appeared to be greater in the subset with a TKV ≥ 1 L relative to the subset with a TKV < 1 L at 4 and 5 years of follow-up but not during the first two to three years.

Table 6: Kaplan-Meier estimates of probabilities of not having a confirmed 30% decline in eGFR in subgroups defined by baseline age, TKV and eGFR

Followup Time (Years)	Predicted probability of not having a confirmed 30% decline in eGFR							
	TKV < 1 L				TKV ≥ 1 L			
	Age < 40 years		Age ≥ 40 years		Age < 40 years		Age ≥ 40 years	
	eGFR ≥ 50 mL/min per 1.73 m^2 (n=275)	eGFR < 50 mL/min per 1.73 m^2 (n=4)	eGFR ≥ 50 mL/min per 1.73 m^2 (n=159)	eGFR < 50 mL/min per 1.73 m^2 (n=28)	eGFR ≥ 50 mL/min per 1.73 m^2 (n=168)	eGFR < 50 mL/min per 1.73 m^2 (n=28)	eGFR ≥ 50 mL/min per 1.73 m^2 (n=128)	eGFR < 50 mL/min per 1.73 m^2 (n=135)
1	0.967	1.000	0.941	0.927	0.969	1.000	0.942	0.951
2	0.944	1.000	0.919	0.927	0.931	0.918	0.925	0.859
3	0.902	0.779	0.904	0.832	0.905	0.696	0.874	0.634
4	0.880	0.779	0.895	0.770	0.853	0.651	0.785	0.508
5	0.846	0.558	0.884	0.624	0.786	0.575	0.785	0.455

Source: Table 14, FDA Primary Statistical Review

Following the completion of Dr. Lawrence's review, further work was done to characterize the benefit of using TKV as a prognostic biomarker for trial enrichment. Some of these analyses are described below.

Average predicted probability in patients with and without events. Dr. Wang calculated the average predicted probability of having a confirmed 30% decline in eGFR in subjects who did and did not experience an event by year 3 using the models with and without TKV. The average predicted probability by event status and model is shown below. Relative to the model without TKV, the model with TKV gave a slightly higher average predicted probability for patients with events and a slightly lower average predicted probability for patients without events by year 3, indicating an improvement in discrimination over 3 years.

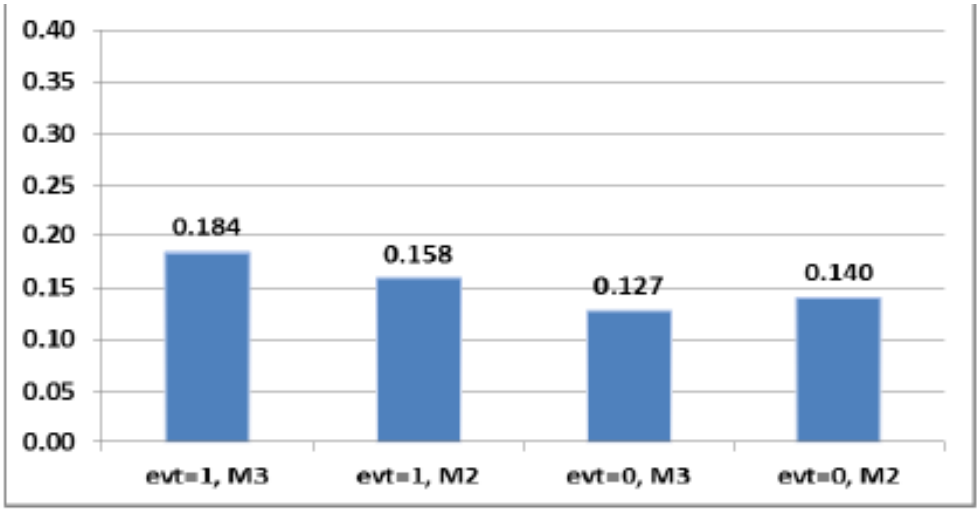

Figure 5. Average predicted probability of event status for model with TKV (M3) and without TKV (M2) at year 3; "evt=1" indicates subjects with an event (30 % decline in eGFR) and "evt=0" indicates subjects without an event.

Source: Figure 5, FDA Secondary Statistical Review

Number needed to screen and number needed to enroll. Dr. Lawrence compared the number of patients needed to produce one event and the number that would need to be screened using the best fit models with and without TKV.

In his first analysis, Dr. Lawrence assumed 3 years of follow-up and that entry criteria required an eGFR greater than 50 mL/min per 1.73 m^2 and age between 20 and 50 years. As in the analyses described above, Dr. Lawrence used a TKV threshold of > 1 L for the best fit model with TKV.

According to Dr. Lawrence, there were 676 patients in the dataset who met the aforementioned age criterion, 574 of whom also met the eGFR criterion. Of these 574 patients, 251 had a baseline TKV>1 L.

The table below shows the predicted event rate in the placebo group over 3 years, the number of subjects needed to produce one event and the number of patients that would need to be screened to enroll these subjects using the models with and without TKV. Based on the entry criteria described above, 13 patients

would need to be screened in the submitter's dataset to enroll 11 patients to get one event using the model without TKV. Using the model with TKV, approximately 25 patients would need to be screened in the dataset to enroll 9 patients to get one event.

Table 7. Predicted event rate in placebo arm over 3 years, number needed to enroll and number needed to treat to get one event using the best fit models with and without TKV

	Model without TKV‡	Model with TKV‡, using added criterion of TKV > 1 L
Predicted event rate in placebo arm over 3 years	0.0905	0.110
Number needed to enroll†	11.05	9.09
Number needed to screen	13.01 [λ]	24.5§

‡Assumes entry criteria of eGFR > 50 mL/min per 1.73 m^2 and age between 20 and 50 years; †Calculated as 1/the predicted event rate over 3 years; [λ] Calculated as 11.05 *676/574; §Calculated as 9.09 *676/251

Source: Email from Dr. John Lawrence, dated May 21, 2015

In the second analysis, Dr. Lawrence restricted the population to patients who were between 20 and 50 years of age and used the risk scores from the two models to define the entry criteria. The figure below shows the number needed to enroll for one event vs the number needed to screen as a function of the risk score. According to Dr. Lawrence's analysis, if patients with the top 50% of risk scores are selected using the best fit model without TKV, then the number needed to enroll to get one event would be 8.6 and the number needed to screen would be 17.2. If patients with the top 50% of risk scores are selected using the best fit model with TKV, then the number needed to enroll decreases to 7.8 and the number needed to screen decreases to 15.6. These findings support the clinical utility of using the model with TKV to enrich the trial population. As also noted by Dr. Lawrence, the findings indicate that using a multivariate risk score to enrich the trial population is more efficient than specifying independent entry criteria for the parameters of interest.

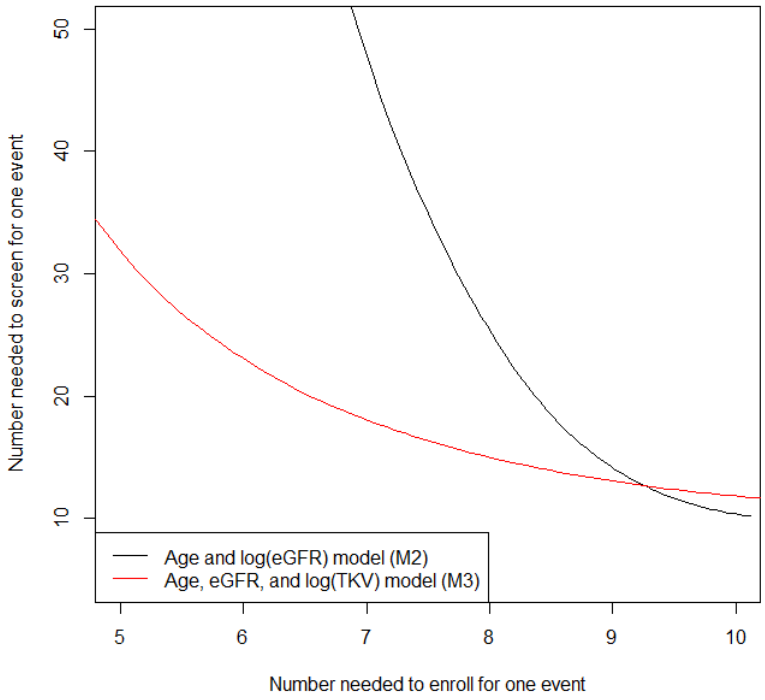

Figure 6: Number needed to enroll for one event vs. number needed to screen using the risk scores from the two models to select patients

Source: Email from Dr. John Lawrence, dated May 29, 2015

Appendix

Registry data

- The database at the University of Colorado contains data on 5,684 individuals from 1228 families with ADPKD who were recruited between 1985 and 2004 from ongoing natural history and genetic studies of ADPKD at the University of Colorado. According to the submission, one or more affected members of each family were examined during a two-day evaluation, which included measurement of serum creatinine values and also US imaging to determine kidney and liver volume and cyst counts.
- The database at Emory University contains information on over 700 phenotyped individuals from approximately 400 families. These patients were recruited for ongoing natural history and genetic studies of ADPKD at Emory University between 1998 and the present. According to the submission, patients are examined during a two-day evaluation period, which includes assessment of serum creatinine. Patients are then "tracked annually in standardized fashion with clinical information including … serum creatinine." Available renal imaging data (MRI, US and CT) are included in the database. Per the submission "Extensive renal imaging data, including MR images, US images, and CT images have been obtained."
- The database at the Mayo Clinic includes information on approximately 2,871 patients with ADPKD seen at the Mayo Clinic since 1984. Patients are evaluated during a Nephrology/General Medicine consultation at Mayo Clinic Rochester in Minnesota. Blood samples are collected for serum creatinine measurement. "Imaging studies for diagnosis of the disease, evaluation of disease severity/progression, or diagnosis/assessment of renal and extrarenal complications are acquired following each patient appropriate recommendation." The submission also indicates that "Longitudinal data concerning date of onset of ESRD or patient's death is collected and registered in the patient's medical record" and that "The Mayo PKD database is populated electronically with Mayo central clinical laboratory data, or abstracted when necessary."

Cohort studies

The two cohort studies, CRISP I and CRISP II, are prospective, longitudinal cohort studies of the natural history of APDKD. CRISP I followed 241 patients annually for a period of three years between January 2001 and August 2005. In CRISP I, serum creatinine levels were assessed at baseline and annually for three years. MRI images of the kidney were obtained at these same time points. In addition, a 24-hour urine sample was collected at baseline to determine the urine albumin to creatinine ratio.

CRISP II extending the follow-up of patients enrolled in CRISP I. The objective of CRISP I was to determine if MRI could detect small changes in renal structural involvement over a short period of time in ADPKD. The goal of CRISP II study is to continue to observe individuals from CRISP I in a prospective fashion in order to evaluate the power and reliability of MRI to predict disease progression in ADPKD including a change in both TKV and kidney function over time. In CRISP II, serum creatinine is assessed annually, whereas imaging is performed at years 1 and 3.

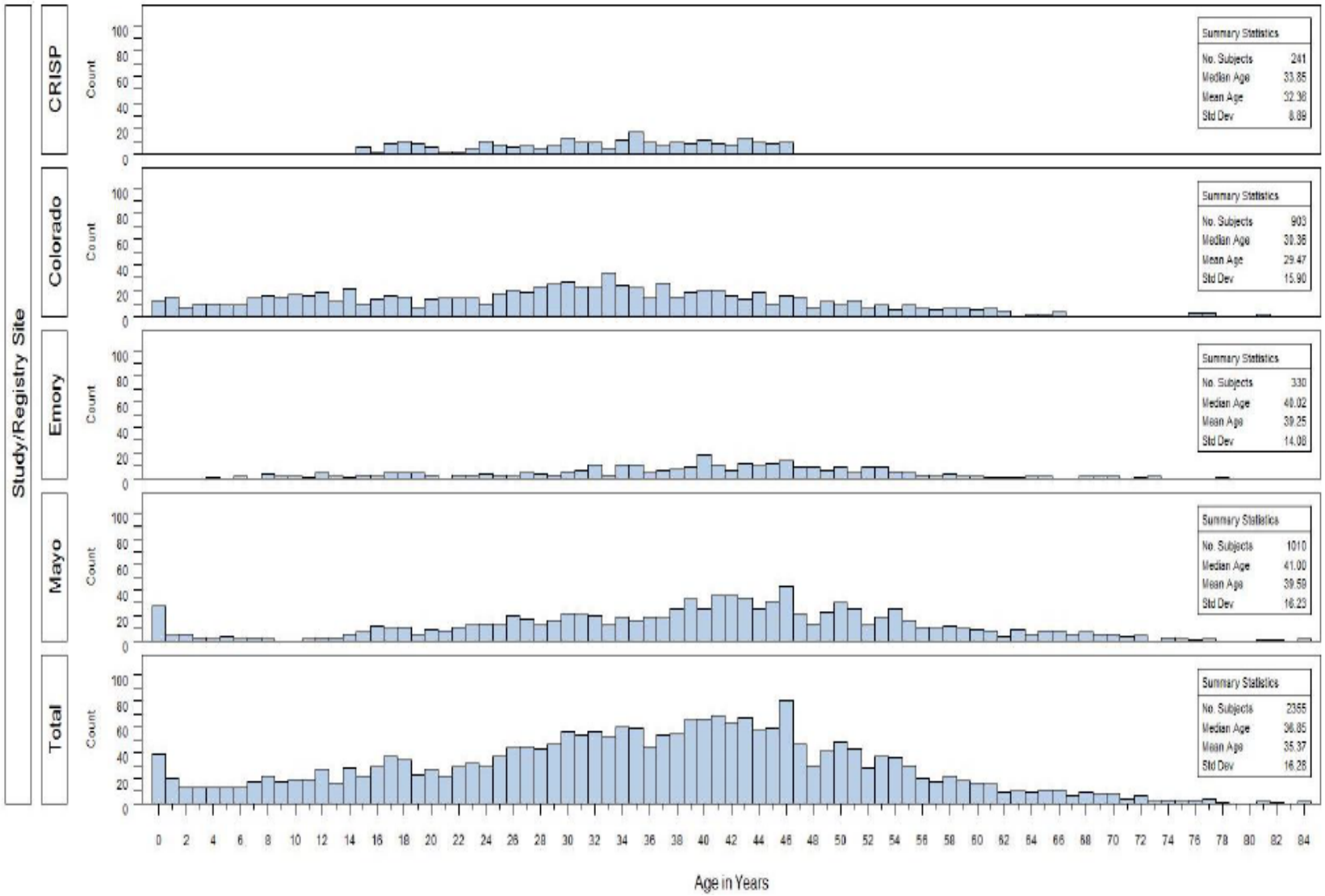

Figure 7: Distribution of Age at Study Entry by Site/Source of Data

Source: Figure 14, Submitter's Briefing Book; based on Total Population (n=2355)