EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

04 September 2023
Case No.: EMA/SA/00000104642
Committee for Medicinal Products for Human Use (CHMP)

DRAFT Qualification opinion for GFR slope as a Surrogate Endpoint in RCT for CKD

Draft agreed by Scientific Advice Working Party (SAWP)	11 May 2023
Adopted by CHMP for release for consultation	25 May 2023
Start of public consultation	06 September 2023
End of consultation (deadline for comments)	23 October 2023

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

Keywords	Qualification of Novel Methodology, glomerular filtration rate (GFR) slope, surrogate endpoint, efficacy endpoint, Chronic Kidney Disease (CKD) clinical trials

Official address Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands

Address for visits and deliveries Refer to www.ema.europa.eu/how-to-find-us
Send us a question Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000

An agency of the European Union

Qualification Opinion as agreed by CHMP

Based on the evidence presented in the qualification opinion request and in a discussion meeting, CHMP considers that GFR slope (i.e. the mean rate of change in GFR over time) can in some trial settings - if properly specified and assessed - serve as a surrogate endpoint for CKD progression in clinical trials for standard marketing authorization and indication extension approvals.

Agreed Context of Use (CoU)

The proposed novel method, GFR slope, is intended to be used as a validated surrogate endpoint for CKD progression in randomized controlled clinical trials for standard marketing authorization and indication extension approvals.

General

The classical (currently standard) hard clinical endpoints (e.g., incidence of ESRD, renal- and overall survival) along with relative reduction of GFR (most often in the range of 40 to 57%) should generally be considered as primary efficacy endpoints. Currently, the main place for a GFR slope-based endpoint in the assessment of treatments in CKD is when trial feasibility is an issue. In such cases, the classical endpoints can be accepted as secondary endpoints, with the expectation that a trend in efficacy is shown for them. In addition, it is acceptable to use GFR slope in assessing a medicinal product during the early clinical stages of development, i.e. exploratory and dose-finding studies and to support efficacy assessment in important subgroups when classical endpoints serve as primary endpoints.

Definition

GFR slope is the mean change in GFR over time, expressed as a rate of change. Treatment effects on GFR slope would be expressed as mean difference between the GFR slope in treatment and control groups. The linearity of the GFR slope from randomization to the end of study, i.e. the acute and chronic phases, should be carefully analysed in a trial. For confirmatory studies clinically relevant differences in mean GFR slope between treatment and control groups or an appropriate non-inferiority margin need to be defined in the study protocol and may vary between different populations.

Application to diverse populations with CKD

The surrogacy of GFR slope was derived from meta-analyses of observational studies and trial level analysis. A large portion of the study participants were from studies of Type 2 Diabetes Mellitus (T2DM) and cardiovascular diseases but also included CKD and glomerular diseases. The surrogacy was demonstrated across different levels of proteinuria, GFR, T2DM and non-T2DM related diseases. CHMP anticipates that surrogacy of GFR slope can be applied to other kidney diseases where clinical composite endpoints are not feasible within a reasonable timeframe (e.g., 2-3 years). Clinical progression and clinical event rates (kidney failure with or without dialysis or transplantation; death and doubling of serum creatinine) should relate reasonably well to that of the studies that demonstrated surrogacy. The clinical endpoints and relative reduction in GFR are anticipated to support efficacy with a positive trend. For rare diseases and disease entities with less weight in the trial level surrogacy analysis considerations may differ and support of efficacy may need to be provided by additional endpoints and/or demonstration of lack of a detrimental effect.

Model

When treatment effects on GFR slope are assessed, the chosen estimand and statistical model for estimation should account for, if appropriate, acute effects, heterogeneity in GFR trajectories, intercurrent events, missing data or (informative) censoring, and any other factors potentially influencing GFR trajectories reflecting current knowledge on pathophysiology. The choice of statistical model should consider the specific circumstances of the treatment and the study population being investigated. The frequency of GFR measurements should be sufficient to characterise all phases of the GFR slope from study randomization to end of treatment.

GFR slope characterization/pre-analysis

Acceptability of slope-based analysis in confirmatory studies will depend on appropriate and robust pre-analysis of the investigational medicinal product (IMP) in the proposed or similar study/target population in earlier phases of development. These include characterization of acute effects, i.e., the size, direction and approximate timing of the knot point demarking the transition from the acute to the chronic phase, as well as the direction of the GFR slope upon discontinuation of treatment. In the presence of large acute effects GFR slope might not be suitable as a surrogate endpoint. Finally, the underlying physiology explaining the acute effect should be understood (e.g., haemodynamic, anti-inflammatory and/or other changes within the kidney microenvironment, but also extra-renal factors influencing assessment of GFR, e.g., muscle mass, inflammation etc). The expected effect shape (e.g., uniform vs. proportional) over time should be characterized.

Final Analysis

An estimate of treatment effect on GFR slope should be based on a sufficiently long-term evaluation period within a trial, preferably 3 years and usually at least 2 years. The adequate follow-up duration will also depend on the underlying disease and presence of acute effect. Reassurance for long term benefit should be defined, which reflects the chosen primary analysis. When GFR slope from study randomization is chosen for the primary analysis (i.e. includes both the acute and chronic phase of the slope) supportive evidence of efficacy could be provided by a less steep slope in the chronic phase of the IMP compared to the control arm. When GFR slope over the chronic phase is chosen for the primary analysis and the acute effect in the IMP arm is negative, the trial duration should be chosen such that the crossing of the chronic GFR slope lines can be observed to allow appropriate estimation of impact of acute effects in the pivotal trials themselves. Similarly, if acute effects lead to an increase in GFR, study duration needs to be chosen such that sufficient information is available to assess that an early improvement is still not associated with long-term deterioration (as compared to placebo) in the chronic phase.

General comments on the methods and the validation approach

The Applicant proposed that GFR slope can be used as surrogate endpoint in a broad context of use. Key aspects of validation of a surrogate endpoint were adequately addressed: Biological plausibility, individual level associations and trial level analyses. Overall, the approach to validation was appropriate.

Use of GFR as marker for kidney function was investigated in many trials and biological plausibility can be regarded as given, considering physiological and a large range of pathophysiological cases. GFR is a measure of kidney function and the main marker to define kidney function. Various GFR based endpoints are already accepted endpoints in clinical trials and the present application builds on long-

term work from the CKD Epidemiology Collaboration and the National Kidney Foundation and on results of several workshops held together with regulators (2008, 2012 and 2018).

The Applicant provided a comprehensive dataset for the validation of GFR slope as surrogate endpoint with a large number of trials included. It was acknowledged that the set of studies was based on a systematic review of the available literature. It was also noted that the randomised controlled trials covered only a limited period of follow-up for clinical endpoints and the surrogacy analysis was based on data with a median follow-up time of 35 (CI: 22-52) months.

Use of individual patient data for analysis was acknowledged and availability of these data can be considered a strength of the validation approach.

Population

The study population for use of the method was broad including four different disease categories leading to chronic kidney disease (CKD). These included diabetes with or without confirmed diabetic kidney disease, glomerular diseases, cardiovascular disease (CVD) at high risk for CKD but not selected for having CKD, and hypertension. The studies included in the analysis had to indicate progression of CKD with the number of clinical kidney failure events relative to the study size. As such, GFR slope is expected to be used for studies of secondary prevention of kidney disease progression. Overall, the clinical characteristics of the study population of the selected studies was very broad, allowing analysis of surrogacy for the various subgroups of CKD (level of GFR and proteinuria, DM vs non-DM etc).

Model based analysis for GFR slope

Regarding the analyses used for individual trial data, the extensive and well described work by the Applicant was acknowledged (e.g., CKD-EPI Consortium Technical Report in Appendix C and Vonesh E et al., Stat Med 2019). The validation approach used the same unified mixed effects model-based analysis method for GFR slope for all trials, using random effects slope and intercept terms for variability in GFR between patients. A shared parameter model was used to consider informative censoring by KFRT and death if a sufficient number of events was available. This simplified model is based on a single slope starting at 3 months post randomization and allows estimating an acute effect on GFR slope prior to three months and a chronic slope after three months. It assumes that an acute effect is lasting up to 3 months but avoids making an assumption on the shape of the GFR curve for the first 3 months. With this approach the same slope model for all trials, irrespective of acute effects, was applied.

The unified mixed effects model allows estimation of 'acute slope', chronic slope and total slope over the defined periods of 2 and 3 years (and change from baseline at 2 and 3 years). Improved performance of the model might be obtained using a more tailored approach for an individual trial. However, it likely provides a conservative estimate of trial surrogacy performance and (limited) sensitivity analysis supports this notion. The rationale for using this model for all trials was noted. However, for application in future trials the analysis model and analysis of acute effects should be tailored to the population and intervention.

Regarding the trial level surrogacy analysis, the Bayesian meta-regression was considered an appropriate method. The results were presented for "total" and "chronic" GFR slopes for 2- and 3-year periods. The sensitivity analysis and analysis for outliers was considered adequate. Factors influencing the predictive accuracy may suggest that GFR slope could not be appropriate in some trial settings. Important influencing factors are:

(1) the nature and magnitude of acute effects of the intervention

(2) rate of progression

(3) level of baseline GFR

(4) trial duration and GFR assessment schedule.

Application to future trials relies on generalisability. This was addressed by the Applicant with simulation studies in a range of scenarios for identified parameters that have an impact on the operating characteristics (see below). Regulatory acceptability of a specific slope parameter and analysis will depend on the data generated before a confirmatory trial is initiated and GFR trajectories observed in the trial. Final recommendation for analysis models in future trials could not be made at this stage, as regulatory experience with 2-slope models was missing and a simpler or otherwise optimised analysis model (e.g., to reflect physiological knowledge) may be preferable. Sponsors should use the estimand framework, justify the selected analysis model and consider how the model-based analysis in a future trial will be impacted by intercurrent events such as treatment discontinuations and missing data due to study drop-outs. Specifically, approaches to handle intercurrent events and missing data due to study drop-out should consider acute effects and their direction. Using a treatment policy strategy may not be appropriate with acute effects of an intervention. Pre-specification of supplemental estimand and analysis to address sensitivity to analysis model selection will likely be needed.

Epidemiologic cohort analysis

A meta-analysis of individual participant data from 14 cohorts and over 3 million subjects showed that a steeper eGFR decline was associated with higher risk of subsequent kidney failure with replacement therapy (KFRT), using either a mixed effects model or linear regression model to estimate slope (Grams et al, JASN 2019). This association was statistically significant in the meta-analysis over the 1-, 2-, and 3-year observation periods.

The magnitude of the relationship was assessed across patient subgroups, including those with baseline eGFR < or ≥60 mL/min per 1.73 m^2. A similar association between eGFR decline and risk of KFRT was observed within each eGFR cohort across strata of baseline age (<65/≥65 years), sex (male/female), presence of diabetes, hypertension, or history of CVD, or when adjusted for baseline use of ACEi/ARB.

The association was strongest when the GFR slope was based on 3-year observation period where the HR for KFRT associated with a 0.75 ml/min per 1.73 m2 per year change using the mixed model was 0.63 (95% CI 0.60, 0.67) in the <60 ml/min per 1.73 m2 cohort and 0.71 (95% CI 0.68, 0.73) in the ≥60 ml/min per 1.73 m2 cohort.

It can be agreed that the longitudinal cohort analyses results by Grams (Grams et al 2019) provide epidemiologic evidence to supports use of GFR slope as a surrogate endpoint for kidney failure requiring replacement therapy in clinical trials.

Trial level surrogacy analysis

The trial level analysis comprised a dataset that included data used in a previous publication by Inker and co-authors (Inker LA et al., J Am Soc Nephrol 2019), and set of new studies. The pooled databased included 66 randomised comparisons from 17 interventions, with over 187,000 participants in 4 disease categories (CKD, diabetes/diabetic kidney disease, glomerular disease, CVD). 11,558 participants reached the composite clinical endpoint of treated kidney failure (KFRT); untreated kidney

failure (eGFR < 15 ml/min/1.73m^2) or sustained doubling of serum creatinine over a median follow up of 35 (22, 52) months. This can be considered a rich dataset covering several disease areas and the heterogeneity of data can be considered an advantage for qualification purposes.

For ***total GFR slope over 3 years***, a unified analysis method was applied to data over 3 years. The observed posterior median correlation was R^2=0.98 with Bayesian credible intervals (BCI) from 0.85 to 1.00. The slope of the meta regression was -0.35 (BCI -0.42 to -0.28) ml/min/1.73m^2/year. For example, this indicates that when the treatment effect of total slope at 3 years of 0.75 ml/min/1.73m^2/year was associated with 23% lower hazard for the clinical endpoint (95% CI 19% to 27%). The intercept was close to 0 (-0.04 (95% CI-0.09 to 0.01) which indicates that when the treatment had no effect on the total slope computed at 3 years, there was a low probability of having a meaningful treatment effect on the clinical endpoint. These results support the use of total GFR slope over 3 years as a surrogate endpoint.

***Total slope over 2* years** showed lower correlation, with R^2 of 0.89 (BCI 0.68 to 0.98); slope of -0.27 (-0.33 to -0.21) ml/min/1.73m^2/year) and intercept of -0.11 (-0.16; -0,06). This indicates that the total GFR slope over 3 years is favoured over the 2-year slope.

For chronic slope, the posterior median R^2 only showed moderate association with clinical endpoints, with R^2=0.56 (CI: 0.25 to 0.78) and is lower than previously reported in the meta-analysis mentioned above (Inker LA et al., J Am Soc Nephrol 2019). The slope of the meta-regression is different from 0 (-0.32 [-0.45, -0.20]) and the intercept of the meta-regression line is close to 0. Therefore, there is only moderate agreement with clinical endpoints, but low risk of false negative or false positive conclusion on efficacy.

Predictive performance is relevant and a minimum GFR threshold to infer benefit on a clinical endpoint was derived. For example, a treatment effect of 0.75 ml/min/1.73m^2 per year predicts a median HR of 0.74. Predictive values were slightly lower for the chronic slope and had a wider CI.

Overall results for the trial level surrogacy showed that the total slope is more robust than the chronic slope for the overall population. Further, for total slope at 3 years the association between treatment effects and the clinical endpoint was well comparable across subgroups by baseline GFR, causal disease, rate of progression on control, or baseline proteinuria. For chronic slope the association between chronic slope and clinical endpoints was best for glomerular disease (R^2 0.99) and weaker for diabetes (R^2 0.78), other CKD (R^2 0.83) and CVD (R^2 0.69). The association was lowest for the subgroup with baseline GFR of < 60 ml/min/1.73m^2 (R^2 0.54). It is noted that the updated analysis for chronic slope showed weaker R^2 compared to a previous analysis (Inker 2019). Of note, the RMSE (root mean square error) was higher for the small number of studies in the CVD subgroup, indicating less precision. Overall, the total slope at 3 years outperforms the chronic slope in relation to agreement with clinical endpoints and consistency in estimation of clinical efficacy across subgroups. Impact of disease severity of a potential target population on treatment effects on GFR slope is less well understood for chronic slope and further work to understand this is desirable (Collier W et al., Clin J Am Soc Nephrol 2023).

Generalisability and application of GFR slope in future clinical trials

Analyses from longitudinal cohorts and trial level surrogacy are important validation steps. For future application of GFR slope parameters in clinical trials, the properties of GFR slopes in clinical trial settings are of importance. To explore this, the Applicant performed simulations to assess operating characteristics. As clinical endpoint in the simulations, an event was set as a 57% GFR decline, which is roughly doubling of serum creatinine, or kidney failure. Parameters varied in the simulations were:

i. acute effect (mean, attenuation, and variability),

ii. long-term treatment effect,

iii. death and renal failure event definitions, and

iv. parameters of the clinical trial setting (accrual and follow-up, measurement frequency, baseline GFR, loss to follow-up and intermittent missing data rate).

It must be noted that the simulations can only cover a limited range of the varied parameters and that there are important assumptions made for acute effects (occurrence in first weeks and resolution until study end) and the modelling of the treatment effect on the chronic slope. Still, these analyses are considered helpful for sponsors to decide on the appropriateness of using GFR slope as a primary endpoint in a clinical trial to support marketing authorisation application.

The simulations inform on the efficiency of GFR slope-based endpoints compared to time-to event endpoints (30 to 57% reduction in GFR and KFRT) as determined by the acute effect; rate of GFR progression, mean baseline GFR and GFR slope variability and the impact of length of follow up time on the required sample size to obtain 90% power. The simulations also address the impact of acute effect and rate of GFR progression on the risk of bias and type 1 error. From a regulatory perspective the analysis of Type 1 error and risk of bias are of paramount importance. Results show e.g., a dependence of false positive rate of inference of benefit and harm of acute effects, going into opposite directions for total slope and chronic slope. Therefore, the acceptability of slope-based analysis would depend on appropriate and robust characterisation of parameters important when addressing the power of future phase 3 study and provisional considerations on false positive conclusions based on available data from Phase 1 and 2. An important question is if the risk of false conclusions is in an acceptable range, and how the results for GFR slope compare to an analysis of time to GFR decline endpoints in terms of risk level and robustness of results. The planning should include the implications of intercurrent events like treatment withdrawal and study drop-outs on the risks of false conclusions. This could be specifically relevant in settings with negative acute effects that could lead to early differential withdrawal from treatment.

Conclusions

CHMP qualified GFR slope as a validated surrogate endpoint for CKD progression in clinical trials for marketing authorization and extension of indication.

The Applicant presented a comprehensive and complete validation approach for GFR slope as surrogate endpoint based on a population in four relatively common disease categories at risk of progression of kidney disease. It was acknowledged that the Applicant provided relevant discussion on the minimal clinically relevant GFR threshold, and the impact of acute effects and other parameters on endpoint efficacy and risk of type 1 error and bias.

The Applicant's proposed context of use was broad with regard to e.g., trial designs, disease settings and target populations. The appropriateness of using GFR slope as a primary endpoint for a phase 3 trial required assessment of parameters which influence the efficiency of GFR slope relative to time-to-event endpoints. These include the presence, degree and direction of acute effect, rate of progression of kidney disease of the proposed study/target population, baseline GFR, length of follow-up, as well as the risk of type-1 error and bias based on data from phase 1 or phase 2 studies and/or reference to other studies of compounds with same mechanism of action (MoA). Importantly, the study design should take into consideration the frequency of GFR assessment for reliable assessment of the linearity of the slope. Finally, in the case of acute effects, a biological rationale for the effect should be addressed based on data.

Qualification of GFR slope (total or chronic) as a validated surrogate endpoint for CKD progression is for population level analysis. Individual predictions of kidney function are not included in the Context of Use. Subgroup analysis for baseline severity is recommended based on GFR, UACR and pre-baseline GFR progression, if applicable. Secondary endpoints should be supportive. For rare diseases and disease entities with less weight in the trial level surrogacy analysis, support of efficacy may need to be provided by additional endpoints, allowing understanding of the patients' condition, and/or for demonstration of lack of a detrimental effect.