The association between albuminuria and long-term renal risk
Kröpelin, Tobias Felix

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER 4

Number and frequency of albuminuria measurements in clinical trials in diabetic nephropathy

Kröpelin TF
de Zeeuw D
Andress DL
Bijlsma MJ
Persson F
Parving H-H
Lambers Heerspink HJ

Abstract

Background and objectives:
Albuminuria change is often used to assess drug efficacy in intervention trials in nephrology. The change is often calculated using a variable number of urine samples collected at baseline and end of treatment. Yet more albuminuria measurements usually occur. Because albuminuria shows a large day-to-day variability, this study assessed to what extent the average and the precision of the antialbuminuric drug effect varies with the number of urine collections at each visit and the number of follow-up visits.

Design, setting, participants and measurements:
This study used data from three randomized intervention trials (Aliskiren Combined with Losartan in Type 2 Diabetes and Nephropathy, Selective Vitamin D Receptor Activation for Albuminuria Lowering, and Residual Albuminuria Lowering with Endothelin Antagonist Atrasentan) including patients with type 2 diabetes and macroalbuminuria. Albuminuria-lowering drug effects were estimated from one, two, or three urine collections at consecutive days before each study visit and reported as albuminuria change from baseline to end of treatment or the change over time considering an average of all follow-up albuminuria measurements.

Results:
Increasing the number of urine collections for an albuminuria measurement at baseline and end of treatment or using all study visits during follow-up did not alter the average drug effect. The precision of the drug effect increased (decreased SEM) when the number of study visits and the number of urine collections per visit were increased. Using all albuminuria measurements at all study visits led to a 4- to 6-fold reduction in sample size to detect a 30% albuminuria-lowering treatment effect with 80% power compared with using baseline and end-of-treatment albuminuria measurements alone.

Conclusion:
Increasing the number of urine collections per study visit and the number of visits over time does not change the average drug effect estimate but markedly increases the precision, thereby enhancing statistical power. Thus, clinical trial designs in diabetic nephropathy using albuminuria as an end point can be significantly improved, leading to smaller sample sizes and less complex trials.
Introduction

Albuminuria is often used as clinical trial end point to establish drug efficacy in early stages of drug discovery. In these trials the change in albuminuria is often determined between two pre-defined time-points: randomization and end of treatment.1-3 However, many more urine samples for albuminuria measurement are usually collected during the trial that are not used to determine drug efficacy. The use of only baseline and end-of-treatment albuminuria values to determine drug efficacy may be problematic because albuminuria is subject to large day-to-day variability.4 The large day-to-day variability in albuminuria may hamper the accuracy and precision of drug effect estimates if only two pre-determined time points are used.

To our knowledge no study has systematically assessed the effect of the number of albuminuria measurements on antialbuminuric drug effect estimates. We therefore first questioned whether multiple consecutive albuminuria collections per visit at baseline and end-of-treatment alter the average drug effect. Secondly, we questioned whether using the albuminuria data of multiple follow-up visits alters the average drug effect estimates. Finally, we assessed whether the precision of the drug effect changes when the frequency of urine collections is increased during follow-up.

Material and Methods

Patients and clinical trials
Data from the Aliskiren combined with losartan in type 2 diabetes and nephropathy (AVOID), Selective vitamin D receptor activation for albuminuria lowering (VITAL), and Residual albuminuria lowering with endothelin antagonist atrasentan (RADAR) trials were used for this study.1, 2, 5 These trials all enrolled patients with type 2 diabetes and macroalbuminuria.

In the AVOID trial, 599 patients with type 2 diabetes and a urinary albumin/creatinine ratio (UACR) > 300mg/g (or UACR > 200 mg/g for patients on medication blocking the Renin-Angiotensin-Aldosterone-System (RAAS)) and an estimated glomerular filtration rate (eGFR) ≥ 30 mL/min were enrolled.2 After a run-in period of 3 months in which every patient received losartan 100 mg/day, patients were randomly allocated to either placebo or active treatment with
aliskiren 150mg/day with a dose increase to 300mg/day after 12 weeks. During the 24 weeks follow-up period first morning voided (FMV) urine was collected at randomization and at week 4, 8, 12, 16, and 24.

In the VITAL trial 281 patients on stable RAAS medication with type 2 diabetes and nephropathy defined as urinary albumin to creatinine ratio (UACR) 100 – 3000 mg/g and eGFR between 15 and 90 mL/min were randomly allocated to 24 weeks of treatment with either placebo or active treatment with paricalcitol 1µg/day or 2 µg/day. FMV urine collections were performed at randomization and 4, 8, 12, 16, 20 and 24 weeks thereafter. The study demonstrated a dose-dependent effect of paricalcitol on albuminuria. Paricalcitol 1µg/day did not decrease albuminuria whereas paricalcitol 2µg/day did when all study visits over time were taken into account. Because there was no effect of paricalcitol 1µg/day on albuminuria, we excluded this treatment arm for the purpose of this study.

The RADAR trial enrolled 211 patients taking the maximum recommended antihypertensive dose of RAAS medication who were diagnosed with type 2 diabetes and UACR 300 – 3500 mg/g and eGFR 30 - 75 mL/min. After a 4 to 12 weeks run in period patients were randomly allocated to placebo or 12 weeks treatment with either 0.75mg/day or 1.25 mg/day atrasentan. FMV urine collections were performed at randomization and 2, 4, 6, 8, 10 and 12 weeks thereafter.

Urine collections and albuminuria measurement:
In all three trials patients collected three consecutive FMV urine samples for albuminuria assessment before each scheduled follow-up visit. Patients started collecting urine samples two days before each study visit and then collected the third FMV urine sample on the morning of the study visit. In all trials, urinary albumin (mg/L) and urinary creatinine (g/L) were measured and subsequently expressed as the UACR (mg/g). All assessments of urinary albumin and creatinine were performed by a same company (CRL Medinet in AVOID, Quintiles in VITAL, and Quest in RADAR) in a central laboratory located in the United States, Europe or Japan. The urinary albumin concentration was determined by immunoassays.

Statistical analyses:
Baseline characteristics are presented as mean and standard deviation for continuous variables and counts and proportions for discrete variables. Since the distribution of albuminuria is skewed, albuminu-
Number and frequency of albuminuria measurements in clinical trials

Statistical analysis of albuminuria data were conducted using log-transformed UACR values, taking the non-normal distribution of albuminuria into account. All patients who collected three consecutive urine samples at all study visits were included in this study. The treatment effect was calculated based on single, the average of two, or the average of three urine samples for albuminuria measurements collected before each follow-up visit. The albuminuria measurement derived from the urine collected on the morning of the study visit was used to calculate the treatment effect for a single measurement. The urine on the day before to the study visit and at the day of the study visit was used to calculate the treatment effect for a duplicate measurement and all three urine collections were used to assess the impact of three averaged albuminuria measurements.

The drug effect and its precision at baseline to end of treatment were assessed by analyses of covariance (ANCOVA). Change in albuminuria was analyzed with treatment as classification variable and baseline UACR as covariate. The drug effect is reverse transformed to the natural scale and presented as percent change. The precision of the log drug effect is reflected by its standard error (SE).

Longitudinal analyses (across all follow-up visits) were based on a generalized estimating equation. The model contains treatment as a classification variable and visit, albuminuria level at baseline and a treatment times visit interaction as covariates. The corrected Akaike information criterion (AICc) was used to compare the model fit taking different covariance structures into account. We compared models that take into account the decreasing within patient correlation between measurements over time. We compared model fit of models with Ante dependence-, Auto Regressive- and Toeplitz covariance structures using the corrected (AICc). Based on the model fit, a full Toeplitz model was used in which equal SDs for log (UACR) were assumed at each follow-up visit, and correlations between albuminuria measurements at different visits were assumed to depend on the difference between the respective visit numbers (Supplemental Tables 1 and 2 provide details on the models used).

To determine how the standard error would change if more than three albuminuria measurements were considered at each time-point, we performed simulations using input from a patient-specific linear mixed model that was applied to the data. In this simulation, we used the estimated correlation of UACR measurements between time points within patients. The correlation between consecutive
measurements collected the days before a study visit in a participant was set to 0, as this would maximize the simulated gain in precision given by each additional measurement collected prior to a follow-up visit. In order to present the estimated standard errors with their 95% confidence intervals, we performed 1000 simulations.

Analyses were performed using SAS 9.2 with a two sided p-value < 0.05 considered statistically significant. The additional data simulation was conducted using R 3.1.0 and power calculations were performed using G*Power 3.1.9.

Results

Baseline characteristics A total of 464 (77.4%) participants from AVOID, 129 (68.8%) participants from VITAL, and 180 (85.3%) participants from RADAR had three albuminuria measurements at each study visit available and were included in this study. Key baseline characteristics of included trial participants are shown in table 1. Baseline characteristics were well balanced among randomized treatment groups. In all trials, average age was 60 to 65 years; the majority of participants were men (66% - 81%) and were Caucasian (43% - 87%). Baseline albuminuria was similar across the three trials.
Table 1. Baseline characteristics stratified by trial and treatment allocation.

<table>
<thead>
<tr>
<th>Case (n)</th>
<th>Placebo 350 - 1550</th>
<th>Placebo 140 - 1725</th>
<th>Placebo 473 - 972</th>
<th>Atrasentan 0.75 mg/day</th>
<th>Atrasentan 1.25 mg/day</th>
<th>Paricalcitol 2 µg/day</th>
<th>Aliskiren (150mg/day - 300 mg/day)</th>
<th>Placebo</th>
<th>Atrasentan 1.25 mg/day</th>
<th>PARIS</th>
<th>RADAR</th>
<th>VITAL</th>
<th>AVOID</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>172.4 (9.2)</td>
<td>172.7 (9.3)</td>
<td>175.1 (9.3)</td>
<td>174.7 (9.3)</td>
<td>174.4 (9.3)</td>
<td>172.7 (9.3)</td>
<td>172.6 (9.3)</td>
<td>173.5</td>
<td>173.4 (9.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>174.2 (9.5)</td>
<td>174.5 (9.6)</td>
<td>177.1 (9.6)</td>
<td>176.6 (9.6)</td>
<td>176.5 (9.6)</td>
<td>174.7 (9.6)</td>
<td>174.5 (9.6)</td>
<td>175.2</td>
<td>175.2 (9.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>175.1 (9.1)</td>
<td>175.2 (9.2)</td>
<td>178.0 (9.2)</td>
<td>176.8 (9.2)</td>
<td>176.8 (9.2)</td>
<td>175.1 (9.2)</td>
<td>175.0 (9.2)</td>
<td>175.8</td>
<td>176.0 (9.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>177.1 (9.9)</td>
<td>177.2 (9.9)</td>
<td>179.9 (9.9)</td>
<td>178.7 (9.9)</td>
<td>178.9 (9.9)</td>
<td>177.3 (9.9)</td>
<td>177.1 (9.9)</td>
<td>178.6</td>
<td>178.8 (9.9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4

Treatment effect between baseline and end-of-treatment

The treatment effects of aliskiren, paricalcitol, and atrasentan estimated from the baseline and end of treatment measurements are summarized in Figure 1. Relative to placebo, active treatment decreased albuminuria by approximately 20% in AVOID and VITAL, 35% in RADAR low dose and 40% in RADAR high dose trial. We observed minimal variation in the mean albuminuria lowering treatment effect if the effect was calculated from single, double, or triple consecutive urine collections at baseline and end-of-treatment (Figure 1). For example, the treatment effect ranged between -17 and -18% in VITAL and was not statistically significant (p=0.089 for three consecutive urine collections at baseline and end-of treatment) and ranged between -39 and -41% in RADAR (1.25 mg/d; p<0.01).

The standard error of the treatment effect estimate decreased (increased precision) if albuminuria measurements from two consecutive urine collections at baseline and end of treatment were averaged (Figure 1 bottom panel) compared to using single urine collections. For example, the standard error decreased from 0.1301 to 0.1135 in VITAL. There was a small further decrease in standard error if the treatment effect was estimated from three consecutive urine collections in AVOID but not in VITAL or RADAR. Additional simulation studies showed no appreciable decrease in standard error beyond three urine collections [Supplement Figure 1].
Number and frequency of albuminuria measurements in clinical trials

Figure 1. Albuminuria change between baseline and end of treatment.

Top panel: Albuminuria reduction (% and 95% CI) according to number of consecutive collected urine samples considering the change from baseline and the end-of-study visit.

Bottom panel: Standard error of the treatment effect (Δ Log albuminuria) according to number of consecutive collected urine samples considering the change from baseline.
Figure 2. Albuminuria lowering treatment effect (top panel) and standard error of the treatment effect (log scale; bottom panel) as function of number of follow-up visits. The x-axis shows the number of follow-up visits that were used to calculate the treatment effect and the time-interval during which these follow-up visits were performed in VITAL (at week 4, 8, 12, 16, 20, 24) and RADAR-JPN (at week 2, 4, 6, 8, 10, 12).

The y-axis shows the treatment effect (%), and the standard error of the treatment effect (top panel) and the standard error (log scale; bottom panel) as function of number of follow-up visits.
Treatment effect considering all follow-up visits

The albuminuria lowering treatment effects of aliskiren, paricalcitol, and atrasentan estimated from albuminuria measurements collected at follow-up visits during the treatment phase are summarized in figure 2 (top panels). Across the study, urine was collected at five follow-up visits in AVOID and six follow-up visits in VITAL and RADAR trials. The albuminuria lowering treatment effect was consistent regardless of, first, the number of consecutive urine collections at each follow-up visit and, second, the number of follow-up visits used to calculate the albuminuria lowering effect (all P-values >0.1).

Figure 2 (bottom panels) shows that the standard error of the treatment effect decreases with increasing number of visits at which urine was collected during follow-up without reaching a plateau below which the standard error did not further decrease. In AVOID and VITAL the smallest standard error (highest precision) was observed when three consecutive urine collections were performed at all follow-up visits. In RADAR, the standard error was not different between double or triple urine collections at all follow-up visits. As a result of the increase in precision the treatment effect of paricalcitol in the VITAL study became statistically significant (p=0.047 for using three consecutive urine collections at all follow-up visits).

Table 2 shows the impact of the gain in precision on sample size requirements when all albuminuria measurements during follow-up were selected versus using only the baseline and end of treatment measurements. The sample sizes only modestly decreased when triple, compared to double or single consecutive urine collections, were performed at each visit. However, there was a clear decrease in the sample size when all albuminuria measurements at all follow-up visits were considered. The sample size requirements to detect a 30% reduction in albuminuria decreased 4- to 6-fold compared with only using the baseline and end-of-treatment measurements (Table 2).

Standard errors (standard deviation) for sample size calculations are derived from our models to assessing the treatment effect between baseline and end of treatment; and from baseline considering all follow-up measurements (five follow-up visits were scheduled in AVOID while six follow-up visits were scheduled in VITAL and RADAR).
Chapter 4

Discussion

This study was conducted to assess the impact of the number of urine collections at one visit and the number of study visits on albuminuria lowering drug effects. We observed that increasing the frequency of urine collections at two predefined time-points or at multiple follow-up visits did not alter the estimate of the mean albuminuria lowering drug effect. However, the precision of the drug effect increased with an increased frequency of urine collections at baseline and end-of-treatment, and markedly increased when more albuminuria measurements were used during follow-up. The increase in precision in the absence of a change in the average drug effect led to a marked increase in statistical power, which translated into a 4- to 6-fold lower sample size requirements for trials designed to test the efficacy of an albuminuria-lowering drug. Future studies should use consecutive albuminuria measurements per visit and at multiple follow-up visits.

The largest effect on the precision of the drug effect was observed when the frequency of follow-up visits was increased. As a result of the increase in precision the albuminuria lowering effect of paricalcitol 2 µg/day in the VITAL study became statistically significant whereas it was not when the baseline and end-of treatment values alone were analyzed. We did not identify a threshold above which the precision

Table 2. Sample size requirements for clinical trials with a type 1 error rate of 5% and 20% power to detect a 30% decrease in albuminuria

<table>
<thead>
<tr>
<th></th>
<th>AVOID</th>
<th>VITAL</th>
<th>RADAR (0.75mg/d)</th>
<th>RADAR (1.25mg/d)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline and end of treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 urine sample at each visit</td>
<td>288</td>
<td>362</td>
<td>202</td>
<td>208</td>
</tr>
<tr>
<td>2 urine samples at each visit</td>
<td>278</td>
<td>356</td>
<td>200</td>
<td>206</td>
</tr>
<tr>
<td>3 urine samples at each visit</td>
<td>282</td>
<td>352</td>
<td>198</td>
<td>204</td>
</tr>
<tr>
<td><strong>Baseline and all follow-up visits</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 urine sample at each visit</td>
<td>94</td>
<td>76</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>2 urine samples at each visit</td>
<td>80</td>
<td>62</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>3 urine samples at each visit</td>
<td>76</td>
<td>58</td>
<td>30</td>
<td>30</td>
</tr>
</tbody>
</table>
Number and frequency of albuminuria measurements in clinical trials

did not further increase suggesting that the larger the number of urine collections during a trial the greater the precision and gain in statistical power. However, increasing the number of urine collections goes at the expense of additional burden on the patient and increases costs and complexity of the trial. There are a couple of additional considerations that should be taken into account. First, when using multiple albuminuria measurements over time one should also take into account the time-course of the anti-albuminuric drug effect. Importantly, the drugs included in this analysis are known to have relatively rapid onsets of action. The time course of other albuminuria lowering drugs may be different. One should be cautious about analyzing albuminuria measurements obtained directly after randomization if the drug does not have a direct effect on albuminuria since including such cases may dilute the overall drug effect and decrease statistical power. Second, the albuminuria lowering drug effect needs to be stable over time in order to justify using all albuminuria measurements over time. If this is not the case statistical power can be significantly compromised. Third, the correlation between albuminuria measurements likely depends on the time interval between visits. Therefore, the effect of using multiple albuminuria measurements over time on statistical power may depend on the time interval between visits. We suggest that the decision on the timing and frequency of urine collections in a clinical trial should be based on individual grounds taking into account drug characteristics, pharmacodynamics, patient population, aims of the trials, and operational aspects. We recommend however that multiple collections be used rather than baseline and end-of-treatment alone.

Increasing the number of urine collections at each study visit also increased the precision of the treatment effect. In the AVOID and VITAL study, the highest precision was observed when three urine samples at each study visit were collected whereas in RADAR the precision reached a maximum with two urine samples. Since no more than three consecutive urine collections were performed in each trial, we performed a simulation analysis to estimate the standard errors of the treatment effect for more than three urine samples. Those results confirmed that no appreciable gain in precision was observed when collecting more than three consecutive urine samples. Previous studies on the intra-individual variability in albuminuria also concluded that three consecutive first morning void urine samples and measurement of the albumin:creatinine ratio should be performed to quantify albuminuria. These conclusions are reflected in current
clinical practice guidelines advocating collection of three consecutive urine samples for diagnostic and prognostic purposes.\textsuperscript{8,9} Because we found that the optimal standard error was achieved with three consecutive urine collections and because three urine collections are recommended to diagnose and monitor albuminuria in clinical practice, we recommend that albuminuria assessment in clinical trials at each study visit should be based on the average of three consecutive FMV collections.

Various operational aspects of albuminuria measurements have been studied in the past. Prior studies investigated whether first morning void or daytime urine samples can replace the gold standard of 24-hour urine samples for prediction of renal and cardiovascular disease.\textsuperscript{10,11} In addition, the intraindividual variability in albuminuria over time and the stability of albuminuria after prolonged frozen storage have been thoroughly investigated.\textsuperscript{4,12} However, to the best of our knowledge, this is the first systematic analysis on the impact of the frequency of urine collections during a clinical trial to determine drug efficacy.

Our study has strengths and limitations. The strengths are the individual data of different clinical trials resulting in a well characterized cohort of patients with diabetes and nephropathy on an appropriate modern therapeutic background. Hence, our results and conclusions are mostly useful for the assessment of treatment effects on albuminuria as a surrogate outcome in trials of diabetic nephropathy. In addition, the use of different drug classes increases the generalizability of the results. The limitations are that this is a post hoc analysis and the trials were not designed and powered to address our research question. We were also unable to compare our results with gold-standard measurements as twenty-four hour urine samples were not collected at each follow-up visit in the included trials. Finally, baseline albuminuria was defined as the average albuminuria level collected at three consecutive days prior to the randomization visit. Unfortunately albuminuria was not assessed at multiple visits prior to randomization. We were therefore unable to assess the impact of the frequency of urine collections to define an optimal baseline measurement.

In conclusion, increasing the number of urine collections per study visit and the number of visits over time does not change the average drug effect estimate but markedly increases the precision of the estimate thereby enhancing statistical power of clinical trials. Since many trials already collect multiple urine samples for albuminuria measurements during follow-up, considering these data in the final

80
efficacy analysis is a simple and practical way to increase statistical power. Hence, current clinical trial designs in diabetic nephropathy trials using albuminuria as an endpoint can be significantly improved leading to smaller sample sizes, and less complex and cheaper trials.

**Acknowledgements**

This research was presented in part at the American Society of Nephrology Annual Meeting, held November 5–10, 2013, in Atlanta, Georgia. For this study no funding or other financial support was received.

**Conflicts of interest**

D.d.Z. is a consultant for and received honoraria (to employer) from AbbVie, Astellas, AstraZeneca, Chemocentryx, Johnson & Johnson, HemoCue, Novartis, Reata, Takeda, and Vitae. D.L.A. is employed by AbbVie and owns AbbVie stock. F.P. is employed by StenoDiabetes Center, a nonprofit institution owned by Novo Nordisk; in addition, F.P. has received lecture fees from Novo Nordisk, Novartis, Eli Lilly, Boehringer Ingelheim, as well as advisory honoraria from Bristol-Myers Squibb and AstraZeneca. H.H.P. reports having equity in Merck and Novo Nordisk and receiving consulting and lecture fees from AstraZeneca, Abbott, Novartis, and Reata. H.J.L.H. is a consultant for and received honoraria (to employer) from AbbVie, Astellas, Janssen Pharmaceuticals, Reata, and Vitae.
Chapter 4

References


Supplement

**Supplementary Table 1.** AICC stratified by included trial and tested covariance structure. Longitudinal models covering all follow-up visits with albuminuria measurements averaged over three consecutive measurements were considered for model testing.

<table>
<thead>
<tr>
<th>(l,j) th element</th>
<th>AVOID</th>
<th>VITAL</th>
<th>RADAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANTE(1) Ante-dependence</td>
<td>$\sigma_l \sigma_j \prod_{i=l}^{j-1} \rho_i$</td>
<td>3381.6</td>
<td>990.7</td>
</tr>
<tr>
<td>(AR1) Autoregressive 1</td>
<td>$\sigma^2 \rho^{j-i}$</td>
<td>3460.4</td>
<td>1001.8</td>
</tr>
<tr>
<td>(TOEP) Toeplitz</td>
<td>$\sigma_{j-i-j+1}$</td>
<td>3368.9</td>
<td>990.7</td>
</tr>
</tbody>
</table>
Supplementary Figure 1. Standard error (and 95% CI of the SE) considering 6 (simulated) consecutive urine samples at a single follow-up visit. The X axis shows the number of consecutive urine samples prior to the study visit. The Y axis shows the simulated standard errors and their 95% confidence interval.
Number and frequency of albuminuria measurements in clinical trials

**Supplement 2.** Listing of trial specific covariances for repeated UACR measurements considering all follow-up visits and averaged UACR measurements from three consecutive urine collections

**AVOID**
Covariance parameters for follow-up measurements

<table>
<thead>
<tr>
<th>Covariance parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEP(2)</td>
<td>0.2758</td>
<td>0.01843</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(3)</td>
<td>0.2444</td>
<td>0.01831</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(4)</td>
<td>0.1922</td>
<td>0.01838</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(5)</td>
<td>0.1716</td>
<td>0.01982</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.4114</td>
<td>0.01865</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

**VITAL**
Overall covariance parameters for follow-up measurements

<table>
<thead>
<tr>
<th>Covariance parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEP(2)</td>
<td>0.2153</td>
<td>0.02524</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(3)</td>
<td>0.1682</td>
<td>0.02499</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(4)</td>
<td>0.1499</td>
<td>0.02516</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(5)</td>
<td>0.1333</td>
<td>0.02677</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(6)</td>
<td>0.09726</td>
<td>0.03096</td>
<td>&lt;.0017</td>
</tr>
<tr>
<td>Residual</td>
<td>0.3258</td>
<td>0.02584</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

**RADAR**
Overall covariance parameters for follow-up measurements

<table>
<thead>
<tr>
<th>Covariance parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOEP(2)</td>
<td>0.1378</td>
<td>0.01501</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(3)</td>
<td>0.1311</td>
<td>0.01498</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(4)</td>
<td>0.1247</td>
<td>0.01517</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(5)</td>
<td>0.1178</td>
<td>0.01587</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TOEP(6)</td>
<td>0.1058</td>
<td>0.01731</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>0.1999</td>
<td>0.01516</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>