What is the best method of proteinuria measurement in clinical trials of endothelin receptor antagonists?

Jonathan D. Towler a, Neeraj Dhaun b, Margaret MacDougall c, Vanessa Melville a, Jane Goddard b, David J. Webb a,⁎

a Clinical Pharmacology Unit, University of Edinburgh, The Queen's Medical Research Institute, 47 Little France Crescent, Edinburgh, EH16 4TJ, United Kingdom
b Department of Renal Medicine, Royal Infirmary of Edinburgh, United Kingdom
c Centre for Population Health Sciences, University of Edinburgh Medical School, Teviot Place, Edinburgh, EH8 9AG, United Kingdom

A R T I C L E  I N F O
Article history:
Received 15 November 2011
Accepted 10 February 2012

Keywords:
Protein–creatinine ratio
Albumin–creatinine ratio
Endothelin receptor antagonists
Sitaxentan
Chronic kidney disease
Hypertension

A B S T R A C T

Aims: To determine whether protein–creatinine ratio (PCR) and albumin–creatinine ratio (ACR) are comparable to 24 h urine protein in terms of agreement and repeatability, and therefore whether they are suitable for monitoring and comparing reduction in proteinuria in clinical trials of endothelin receptor antagonists.

Main methods: Using data from a recent study of sitaxentan in 27 patients with proteinuric chronic kidney disease, the assays were compared with reference to their agreement, repeatability, the number of measurements required to obtain accurate results and correlation with reduction in proteinuria at baseline.

Key findings: The median coefficient of variation was lower for PCR than 24 h urine protein (25 vs. 28%) but the range was higher (70 vs. 47%). When converted into the same units, mean difference between 24 h urine protein and both PCR (0.03 g/day), and ACR (0.10 g/day), was small. However, scatter increased with mean level of proteinuria, such that agreement fell substantially above 1.5 g/day. According to 2-factor within-subjects ANOVA, the assay used was not a significant source of variation (PCR p = 0.63, ACR p = 0.38). With 3 measurements at each time point, baseline proteinuria correlated equally well with change in proteinuria, and percentage change was detected accurately by all 3 methods.

Significance: PCR and ACR may well be suitable replacements for 24 h urine protein in the clinical trial context due to their similar accuracy and repeatability, greater convenience and lower cost. However, a randomised control trial comparing all 3 assays in a larger and more diverse population is necessary before 24 h urine protein can be replaced.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Within the last two decades it has been established that proteinuria is not only a marker of the severity of chronic kidney disease (CKD), but also contributes directly to disease progression. Protein overload of proximal tubular cells promotes an inflammatory response that ultimately results in interstitial fibrosis (Abbate et al., 2006). It has since been demonstrated in both animal models and clinical trials that lowering urine protein retards the progression of CKD and can, in some cases, lead to disease remission (Remuzzi et al., 2006; The GISEN group (1997)). Consequently, interventions which reduce the extent of urinary protein excretion are a key target for the treatment of CKD, which has a prevalence of around 12% in the European population (de Zeeuw et al., 2005).

In order to compare the antiproteinuric effects of various drugs, an accurate and repeatable assay for urine protein is required. The current gold standard for assessment of proteinuria is 24 h urine protein excretion, and alternative methods must therefore be evaluated by comparison with this assay. Whilst 24 h urine collections overcome the problem of circadian rhythmicity in urine protein excretion (Koopman et al., 1989), inconvenience and patient forgetfulness often lead to incomplete collections and consequently, inaccurate assessment of proteinuria (Price et al., 2005). Furthermore, handling and analysis of 24 h urine samples is more costly and open to error than alternative approaches. By contrast, protein–creatinine ratio (PCR) and albumin–creatinine ratio (ACR) can be calculated from spot urine samples, with urine protein or albumin concentration divided by creatinine concentration to correct for variation in urine volume. Spot samples are both cheaper and more convenient than 24 h samples, and a number of studies have demonstrated the ability of both PCR and ACR to rule out significant proteinuria (Guy et al., 2009) and predict renal outcomes and mortality (Methven et al., 2010; Heerspink et al., 2010) in patients with CKD. As a result, they have largely replaced 24 h urine protein measurements in clinical screening for proteinuria. ACR is...
universally recommended for detection of proteinuria in the diabetic population (National Collaborating Centre for Chronic Conditions, 2008; Scottish Intercollegiate Guidelines Network, 2008; National Kidney Foundation, 2002), whereas some groups favour PCR in patients without diabetes (Scottish Intercollegiate Guidelines Network, 2008). However, the characteristics which make PCR and ACR suitable alternatives for screening purposes differ from those required to monitor change in proteinuria with a drug candidate. Screening assays are primarily used to stratify a target population into broad groups, whereas in drug trials the absolute level of proteinuria is required, demanding a higher degree of accuracy from the assay. Large drug trials such as AASK have used PCR as a surrogate for 24 h urine protein, but have not compared the two assays with each other (Lea et al., 2005).

The endothelin system represents an ideal target for antiproteinuric interventions. It is now clear that upregulation of this system is not merely a consequence of increased urine protein excretion, but that it also plays a part in exacerbating the degree of proteinuria, primarily through activation of ET$_A$ receptors (Barton, 2008; Neuhofer and Pittrow, 2009). Sitaxentan is a selective ET$_A$ antagonist, the effects of which were examined in a recent study in subjects with CKD (Dhaun et al., 2011). At the end of the 6-week treatment period, proteinuria had fallen by an average of 30%. Using data from this study, we tested the hypothesis that PCR or ACR are suitable replacements for 24 h urine protein in the measurement of proteinuria in clinical trials of antiproteinuric drugs. Their suitability was assessed in terms of agreement, repeatability, the requirement for multiple measurements to obtain accurate readings and correlation with change in proteinuria at baseline.

Materials and methods

Study subjects and protocol

The clinical study was carried out in 27 subjects with stable, stage 1 to 4 non-diabetic CKD and overt proteinuria (>300 mg/day). This double-blind study involved three consecutive 6-week phases in which subjects were randomly assigned to placebo, nifedipine LA 30 mg and sitaxentan 100 mg. All 27 subjects participated in all 3 phases, each of which was separated by a 2-week washout period. Measurements were made on 3 consecutive days prior to the start of treatment, at the end of week 3, and at the end of week 6 in each phase. In accordance with written instructions, urine was collected by the subject for 24 hours on each of these days, from which urine protein was measured. PCR was determined from an aliquot of the first and second 24 h collections at each time point, whilst on the third day a true spot urine was taken at the end of the 24 h period. ACR was measured retrospectively from frozen urine samples, although only one was available for each time point. Urine protein was measured using a colorimetric method with pyrogallol red and urine albumin with an immunoturbidimetric assay.

Statistical analysis

Data were analysed using SPSS 17 (SPSS Inc.). A statistical significance level of 0.05 was used for all hypothesis testing. Spearman’s rho was used to measure the strength of the correlation between mean 24 h urine protein and PCR from an aliquot (days 1 and 2) or a true spot urine (day 3). It was also used to assess correlation between baseline proteinuria and change in proteinuria with each of the assays. To enable direct comparison with 24 h urine protein (g/day), PCR and ACR data (mg/mmol) had to be expressed in the same units. In the absence of a standardised conversion factor between the assays, the equation of the regression line derived from a scatter plot of 24 h urine protein against aliquoted PCR and ACR was applied to the raw data (24 h urine protein $=73.3 \times$PCR $+7.7$ and $59.2 \times$ACR $-15.4$). Bland-Altman plots were then constructed using these transformed data, for a visual and quantitative comparison of agreement. The limits of agreement and their 95% confidence intervals were also calculated, and assessed for clinical significance. The coefficient of repeatability and coefficient of variation were used to compare the repeatability of 24 h urine protein and PCR. It was not possible to include ACR in this analysis as only one measurement was available for each time point. Two-factor within-subjects ANOVA was performed to assess the significance of time and the assay used as sources of variance in the measurement of proteinuria. This analysis was also used to test the significance of multiple measurements on the percentage reduction in proteinuria reported by each of the assays.

Results

PCR: aliquot vs. true spot urine

The correlation between PCR and mean 24 h urine protein at baseline was calculated separately for measurements made on days 1, 2 and 3 (PCR 1, 2 and 3). Correlation of PCR 3 with mean 24 h urine protein (spot urine, rho $=0.96$) was similar to that of PCR 1 and PCR 2 (aliquots of 24 h urine collection, rho $=0.94$ and rho $=0.97$ respectively). Likewise, Spearman’s rho was comparable for all three PCR readings when week 3 and week 6 were included in the analysis (rho $=0.96$, rho $=0.97$ and rho $=0.96$ respectively). In all cases the correlation was highly significant (p<0.0001). In addition, the median coefficient of variation (CV) was consistent between PCRs measured on days 1, 2 and 3 at all 3 time points in the placebo phase (23.4%, 24.4% and 24.1% respectively).

Agreement

Agreement of PCR and ACR with 24 h urine protein was assessed using the Bland-Altman method (day 3, baseline, placebo phase) [Fig. 1a and b]. As anticipated with data transformed in the manner described above, the mean difference was small in both cases (PCR $=0.03$ g/day, ACR $=0.10$ g/day). However, with the inclusion of their 95% confidence intervals, the limits of agreement span a very large range (PCR $=-2.29$ to 2.25 g/day, ACR $=-1.58$ to 1.78 g/day). Fig. 1 also suggests that scatter increases with rising levels of proteinuria. The data remain relatively clustered below 1.5 g/day with both PCR [Fig. 1a] and ACR [Fig. 1b], but above this value greater variation is apparent. Adjustment was not made for this possible downward trend in agreement both because it was not consistent across the full range of mean concentrations and because there were insufficient data to confirm that the trend was genuine.

Repeatability

The mean coefficient of repeatability (CR) for day 3 readings at baseline was lower with ACR (2.02 g/day) than with PCR (2.38 g/day) and 24 h urine protein (2.46 g/day). The median CV for all 9 placebo readings was higher with 24 h urine protein than PCR (28% vs. 25%) [Fig. 2a]. Similarly, the median CV was higher for 24 h urine protein at all 3 time points in all 3 phases (mean $=17\%$ vs. $14\%$), with the sole exception of week 3 in the nifedipine phase. However, the range of PCR data was greater than that of 24 h urine protein at baseline in the placebo phase (span $=70\%$ vs. $47\%$). Box plots constructed from the data gathered at week 6 support this observation in the nifedipine and sitaxentan phases, although not in the placebo phase [Fig. 2b].

Neither the assay used (p = 0.80), nor time (p = 0.61), nor the interaction between assay and time (p = 0.63), were identified as significant sources of variation in the data by 2-factor within-subjects ANOVA (n = 27). These findings were reproduced with the transformed ACR data (p = 0.54, p = 0.63 and p = 0.38 respectively).
Multiple measurements

Taking an average of three 24 h urine protein or PCR measurements tended to demonstrate a slightly larger reduction in proteinuria than with only 1 or 2 measurements, but this trend was not significant for either assay. According to Two-factor within-subjects ANOVA, neither the number of repeat measurements nor number of repeat measurements–time interaction were significant for 24 h urine protein ($p=0.58$, $p=0.48$) or PCR ($p=0.85$, $p=0.53$).

Correlation with change in proteinuria at baseline

The correlation of mean baseline proteinuria with mean change in proteinuria in the sitaxentan phase was calculated for 24 h urine protein and PCR. Spearman’s rho was similar ($\rho = -0.80$ and $\rho = -0.78$ respectively) [Fig. 3a and b] and highly significant in both cases ($p<0.0001$).

Using day 3 measurements only, PCR demonstrated a stronger correlation between baseline and change in proteinuria ($\rho = -0.76$) than both ACR and 24 h urine protein ($\rho = -0.57$, $\rho = -0.55$) [Fig. 3c, d and e]. Spearman’s rho was significant in all cases ($p<0.0001$, $p=0.003$ and $p=0.007$ respectively).

Discussion

We have demonstrated that PCR is comparable to 24 h urine protein in terms of repeatability, the number of measurements required to obtain accurate results and its correlation with reduction in proteinuria at baseline. ACR is also comparable to 24 h urine protein with regards to repeatability and association with change in proteinuria. These findings indicate that both PCR and ACR may be suitable alternatives to 24 h urine protein for monitoring change in proteinuria in clinical trials of antiproteinuric drugs, such as endothelin receptor antagonists.

One of the main problems associated with 24 h urine protein is poor compliance, which often results in inaccurate collections. In this study, 24 h urine creatinine excretion did not differ significantly from placebo at week 3 or week 6 in either the nifedipine or sitaxentan phases, indicating that collection of 24 h urine samples was consistent throughout. However, the consistency of these collections might not have been as robust were it not for the comprehensive written and verbal instructions given over the course of the study. Conversely, a key characteristic favouring PCR is the ability to measure proteinuria from a single spot sample. However, in this study the first and second PCR reading at each time point was calculated from an aliquot of the 24 h collection, which could potentially mask the circadian rhythmicity of proteinuria that affects true spot samples. As convenience is the distinguishing characteristic of the PCR assay, it was important to establish whether the first two PCR measurements produced values that were comparable to the third PCR from a true spot urine. Indeed, measurements with all three PCRs
showed a consistently strong correlation with mean 24 h urine protein, and in all cases this correlation was highly significant. Furthermore, the median CV of PCR measurements was consistent on all three days, suggesting that repeatability with the aliquots and true spot urine samples was comparable.

Correlation, however, is not an indicator of agreement but rather a measure of the strength of association between two quantitative variables. Therefore, the Bland-Altman method was used to assess the degree of agreement between the assays. From Fig. 1 it appears that as mean proteinuria increases, agreement falls, as demonstrated by an increase in scatter above 1.5 g/day. The reason for this apparent decrease in agreement is not certain, but may be due to a decline in the accuracy of the assays above their linear range, where manual dilutions are performed. Indeed, urine samples with very high protein (including albumin) concentrations can be erroneously reported as low or normal with certain immunoassay techniques (Lamb et al., 2009). This is known as the prozone phenomenon and results from the presence of very high concentrations of antigen. This phenomenon could have affected the immunoturbidimetric assay for urine albumin, but is not relevant with the colorimetric urine protein assay. The second

![Fig. 3. Correlation between baseline proteinuria and change in proteinuria with (A) 24 h urine protein, (B) protein–creatinine ratio (PCR), (C) 24 h urine protein – day 3 only, (D) PCR – day 3 only, and (E) albumin–creatinine ratio (ACR) – day 3 only. Spearman’s rho is shown in the bottom left corner of each plot.](image-url)
observation to be made from Fig. 1 is the wide span of the limits of agreement in both comparisons. Differences within this range are clearly significant in the clinical context, indicating that 24 h urine protein and transformed PCR or ACR values cannot be considered interchangeable.

Median variability of PCR was lower than 24 h urine protein, both before and during active treatment, but the overall range of PCR readings was greater. The difference in inter-quartile range between assays was small, indicating that the increase in overall range was the result of one or two subjects with variation that was high, but not sufficient for them to be considered outliers. With a sample size of only 27, it is not possible to state whether measurements in these subjects were anomalous or if significant variation is a genuine issue in a small subset of the population. The mean coefficient of repeatability was similar with 24 h urine protein and PCR, and lowest with ACR. When both PCR and ACR were compared with 24 h urine protein, neither assay, nor time, nor the assay x time interaction, were significant, indicating that variance is comparable in all three methods. Therefore, taking into account the lower median CV, comparable mean CR and non-significant ANOVA findings, repeatability with PCR appears to be at least equivalent to 24 h urine protein. Likewise, considering the lower CR and non-significant ANOVA findings, repeatability with ACR seems to be comparable to that of 24 h urine protein.

In addition to cumbersome 24 h collections, the other main inconvenience related to monitoring proteinuria is the use of repeat measurements. In the clinical study, three 24 h collections were taken at baseline, week 3 and week 6; a total of 27 collections over three phases. Without great care, poor compliance is likely to be compounded by this number of repeats. Therefore, we investigated what impact taking only one or two measurements would have had on the outcome of the study. Using an average of three measurements at each time point tended to demonstrate a greater reduction in proteinuria than would have been apparent with fewer measurements. This observation was made with both 24 h urine protein and PCR, suggesting that it was not purely incidental. However, according to Two-factor within-subjects ANOVA, neither number of repeats, nor the number of repeat measurements–time interaction were statistically significant, indicating a lack of evidence that number of repeat measurements has an effect on percentage change in proteinuria or that the extent of such an effect can vary significantly according to time. This is reflected in the wide confidence intervals of percentage change in proteinuria, especially with a single reading at each time point (± 15% for 24 h urine protein, ± 17% for PCR, ± 13% for ACR). Therefore, in the majority of cases taking one or two measurements at each time point would have been sufficient to obtain a reasonable picture of average reduction in proteinuria at the group level. However, the high CV for both assays means that this could lead to grossly inaccurate assessment of proteinuria on an individual basis. As such, it is advisable to carry out three repeats to average out anomalous data and ensure reliable results in all subjects. This is true for both 24 h urine protein and PCR.

A meta-analysis of 11 studies using ACE inhibitors in non-diabetic renal disease found that patients with the highest urine protein excretion at baseline have the greatest reduction in proteinuria with treatment (Jafar et al., 2001). We compared the correlation of mean baseline proteinuria with mean change in proteinuria for each assay to see if this observation was upheld with an ET<sub>A</sub> antagonist. All three assays demonstrated a strong correlation with change in proteinuria at baseline. The strength of this association increased when three 24 h urine protein measurements were considered, and was consistently high with both one and three PCR measurements. Therefore, all three assays support the conclusion that patients with high baseline urine protein excretion experience the greatest reduction in proteinuria with sitaxentan.

This clinical study had a number of limitations which should be addressed before 24 h urine protein can be legitimately replaced by spot tests. The first is the size of the study sample. There were 27 subjects involved in the clinical trial, which provided the power necessary to meet the primary endpoint and demonstrate a significant reduction in proteinuria. However, it is possible that a larger sample size may have unmasked significance in a number of the statistical tests carried out in this study. Therefore, it is not appropriate to make generalisations from this study about the characteristics of each assay in the general population, due to the small sample size. The second limitation is the exclusion criteria. Notably, this study was in patients with non-diabetic CKD and therefore the results presented here may not apply to those with diabetes mellitus or other significant comorbidities such as heart or lung disease and peripheral vascular disease. The third limitation relates to the time at which the spot urine (PCR 3) was taken. Koopman et al. (Koopman et al., 1989) found that the greatest correlation with 24 h urine protein is observed with a PCR from an early morning specimen. Patients in this study voided the day 3 spot sample in the morning, which may have strengthened the correlation between PCR 3 and mean 24 h urine protein. The fourth limitation pertains to the storage of spare samples from which the ACR was determined. Freezing samples at −20 °C has been shown to yield erroneously low urinary albumin concentrations, the magnitude of which increases with the duration of storage (Brinkman et al., 2005). However, storage at −80 °C and assessment by immunonephelometry has been shown to prevent loss of albumin at 12 months (Brinkman et al., 2007). Spare samples used in this study were stored at −80 °C but detection was with an immunoturbidimetric method. Also, samples were stored for 4 years, the impact of which has not been investigated at this low temperature. Whilst further research into the effects of frozen storage with different detection methods may be of some use in interpreting these results, minimising the duration of storage is likely to yield a more accurate comparison of multiple assays.

Conclusion

Since repeat measurements are advisable regardless of the assay, and since spot urines are far more convenient and less open to error than 24 h urine collections, it is only necessary to demonstrate that PCR and ACR are equivalent to 24 h urine protein in order to recommend them as replacements. This study, involving an endothelin receptor antagonist, suggests that PCR and ACR are comparable to 24 h urine protein in terms of both agreement and repeatability. However, a randomised control trial comparing all three assays in a larger and more diverse population is necessary before 24 h urine protein can be appropriately substituted in the clinical trial context.

Addendum

Sitaxentan has been voluntarily withdrawn by Pfizer Ltd due to unacceptable side effects. However, the findings in this manuscript remain true for selective endothelin A receptor antagonism.

Sources of funding

This study was funded by Encysive Pharmaceuticals, Inc. In June 2008, Encysive was acquired by Pfizer, Inc. N.D. was supported by the British Heart Foundation and National Health Service endowments.

Conflict of interest statement

N.D., J.G., and D.J.W. have all received research grants from Pfizer. N.D. and J.G. have held academic research fellowships funded by educational grants from Pfizer. J.G. and D.J.W. have acted as consultants to Pfizer.
Acknowledgments

We thank the study participants for their contribution to this work, as well as Neil Davie and Simon Teal, for their continued support.

References


