A key issue in the analysis of outcome trials is the adjustment for baseline covariates that influence the primary outcome. Imbalance of an important covariate between treatment groups at baseline is of considerable concern if one treatment group is favored over another with respect to the hypothesis testing outcome. With the use of the Reduction of Endpoints in NIDDM with the Angiotensin II Antagonist Losartan (RENAAL) study database as an example, the influence of baseline proteinuria on the primary composite endpoint, ESRD, and ESRD or death after adjusting for baseline proteinuria as a continuous covariate was examined. Increasing baseline proteinuria was associated with increased risk for renal events, confirming that proteinuria is an important covariate for renal outcomes. When the randomization was stratified according proteinuria <2000 mg/g or ≥2000 mg/g, within the higher proteinuria stratum (≥2000 mg/g), patients in the losartan group had a higher baseline mean proteinuria value. When the imbalance was adjusted, an increase in the magnitude and the significance of the risk reduction with losartan for each outcome was observed. No apparent interaction between treatment effect and baseline proteinuria was found, and there was no heterogeneity in the treatment response in patients with different baseline proteinuria levels. After proteinuria was adjusted as a continuous variable, greater treatment effects were observed in the RENAAAL study. This effect was due solely to the imbalance in baseline proteinuria. Considering the importance of proteinuria as a risk factor, adjustment for baseline proteinuria as a continuous covariate should be prespecified in the design and analysis of clinical trials involving renal outcomes, even when patients are stratified on the basis of level of proteinuria.

The goal of a randomized clinical trial is to achieve a statistically unbiased and efficient treatment comparison. One of the key issues in the analysis of outcome trials is adjustment for a baseline covariate that has an important influence on the primary outcome. This issue has been debated among clinical trialists, regulatory agencies, and statisticians. There is no commonly accepted standard, and few examples exist to show the consequence for such a decision.

An imbalance of an important covariate at baseline between treatment groups is of considerable concern if one treatment group is favored over another with respect to the primary (hypothesis testing) outcome. A recent survey of 50 clinical studies (1) showed that baseline covariate differences were noted in 17 (34%). In the International Conference on Harmonization 9 guidelines, investigators are advised to identify the covariates that may influence the primary outcome and to prespecify the method to be used to account for them to compensate for any imbalance between groups. A baseline covariate can be considered at two separate stages in a clinical trial: in the randomization process (using stratified randomization), or alternatively in the data analysis. Regardless of method used, it is important to identify covariates and to prespecify how they will be addressed. This important point is the subject of new guidelines released by the Committee for Proprietary Medicinal Products (The European Agency for the Evaluation of Medicinal Products) (2). The techniques used for adjusting a covariate in statistical analyses are dependent on the nature of the covariate or the outcome, the relationship between the covariate and the outcome variable, and the statistical model used in the analysis. For most time-to-event analyses, a Cox proportional hazards regression model is widely used. When a continuous covariate, such as proteinuria, is stratified at random-
ization, there are two ways to adjust for it in the analysis: Using the stratum factor to reflect the dependence of baseline hazard risk on the covariate or by including the continuous covariate in the model. Understanding the merits and limitations for each method is important for clinical trialists to plan an analysis in advance and to explain the results to clinicians.

To explore the effect of adjusting for baseline covariates in different ways, we used the Reduction in Endpoints in NIDDM with the Angiotensin II Antagonist Losartan (RENAAL) study as an example. It has been known for some time that proteinuria is a strong predictor of risk for the progression of renal disease to ESRD in patients both with and without diabetes (3–7). The RENAAL study examined the treatment effect of losartan on a composite endpoint that was composed of the doubling of the serum creatinine concentration, ESRD, or death. As such, the primary outcome and its renal components would be expected to be highly influenced by proteinuria levels. Therefore, the RENAAL study was designed to account for proteinuria as a covariate using stratified randomization (8). Patients were stratified into two groups according to their baseline proteinuria levels: <2000 mg/g and ≥2000 mg/g (determined as the ratio of milligrams of urinary albumin to grams of urinary creatinine). It was assumed that this stratification would sufficiently account for proteinuria as a covariate and would prevent any imbalance between the treatment groups.

The aim of this post hoc analysis was to explore the effect of adjusting for baseline proteinuria as a continuous variable compared with the original adjustment by proteinuria strata on the primary endpoint, the composite of ESRD or death, and ESRD alone in the RENAAL study. The potential reason for differences between the two methods was also examined.

Materials and Methods

Study Design

The RENAAL study was a multinational, randomized, double-blind trial that compared the effects of losartan versus placebo on a background of conventional antihypertensive medication in patients with type 2 diabetes and nephropathy. A total of 250 centers in 28 countries in Asia, Europe, Central America, South America, and North America participated in the study. The methods have previously been described in detail (8,9). Briefly, after a 6-wk screening phase, patients were randomized to either losartan 50 mg (titrated to 100 mg as needed) or placebo. Additional antihypertensive medications (calcium channel blockers, β-blockers, centrally acting agents, and diuretics, excluding angiotensin-converting enzyme inhibitors or other angiotensin receptor antagonists) were permitted to reach the goal BP of <140/90 mmHg (systolic/diastolic). Patients were followed for a mean of 3.4 yr (range 2.3 to 4.6). Because of a small number of patients expected to be enrolled at most investigative sites, a blocking factor of 2 was used in the generation of random numbers.

Patient Population

A total of 1513 patients with type 2 diabetes and nephropathy, aged 31 to 70 yr and of both genders, were randomized into the study. Nephropathy was defined by the presence on two occasions of a ratio of urinary albumin to urinary creatinine excretion in a morning specimen of 300 mg/g (or a rate of urinary protein excretion of at least 0.5 g/d) and serum creatinine values between 1.3 and 3.0 mg/dl, with a lower limit of 1.5 for male patients who weighed >60 kg. There was no upper limit of proteinuria required for patients who were randomized into the study.

Outcomes

The primary outcome of the RENAAL study was a composite of the doubling of serum creatinine, ESRD, or death from any cause. In addition, prespecified endpoints included the individual components (doubling of serum creatinine, ESRD, or death) as well as the composite endpoint of ESRD or death.

Statistical Analyses

The prespecified primary (original) analysis has been described elsewhere in detail (9). Briefly, intention-to-treat analysis using a stratified Cox regression model adjusted by treatment (losartan versus placebo) and region (North America, Latin America, Europe, and Asia) was performed. The baseline hazard function was stratified by baseline proteinuria levels (urinary albumin to creatinine ratio of <2000 or ≥2000 mg/g), in agreement with the randomization procedure. In addition, the interaction between treatment response and proteinuria stratum was explored.

Consistent with the prespecified analyses, post hoc analyses include all 1513 randomized participants in the RENAAL study. The relationship between baseline renal function (determined by level of proteinuria) with renal endpoints was explored. Baseline proteinuria was divided into four categories (<1000, 1000 to 2000, 2000 to 4000, and ≥4000 mg/g). The categories were chosen ad hoc, with the aim of providing a partition within each original prespecified proteinuria stratum. A multivariate Cox regression model was performed with indicators of proteinuria categories as covariates. The lowest category was used as a common reference to compute the hazard ratio and 95% confidence interval for the remainder of the categories.

To explore whether the response to treatment was homogeneous over all levels of renal impairment, we conducted two analyses: (1) The interaction of treatment effect and baseline proteinuria was examined using a multivariate Cox model with treatment group, region, continuous proteinuria, and the interaction of treatment group with proteinuria included in the model; and (2) the response to treatment effect across baseline proteinuria categories (<1000, 1000 to 2000, 2000 to 4000, and ≥4000 mg/g) was determined using a Cox model with region, proteinuria subgroup, and the interaction of treatment group with proteinuria category included in the model.

To determine whether baseline proteinuria was balanced between the treatment groups, we performed a comparison that included all patients within each randomization stratum in proteinuria. The P value was calculated using the Wilcoxon sum rank test.

To explore the effect of adjusting for baseline proteinuria as a continuous variable (referred to as the adjusted model) compared with the original adjustment by proteinuria strata (referred to as the original model), we used a Cox model; however, baseline proteinuria strata (<2000 or ≥2000 mg/g) was replaced with continuous proteinuria as an additional covariate. In contrast to the Kaplan-Meier curves without any adjustment, the event curves were also adjusted by continuous proteinuria at baseline using the Cox model–based product-limit estimate (10).

All statistical analyses were performed using SAS version 8. All tests were two-tailed. The P value for statistical significance was P < 0.05, with the exception of the primary endpoint analyses, for which statistical significance was considered at P < 0.048, which is consistent with the originally published findings for the RENAAL study.
Results

Effects of Baseline Proteinuria on Renal Outcomes

Figure 1 shows the risk associated with increasing baseline proteinuria. In patients with baseline proteinuria levels of 1000 to 2000 mg/g, the risk for developing ESRD increased 2.6-fold compared with patients with proteinuria <1000 mg/g. In patients with proteinuria ≥4000 mg/g, the effect was even more pronounced, with a 16.8-fold increased risk for developing ESRD.

In the RENAAL study, no apparent interaction between treatment effect and baseline proteinuria was found (P = 0.1412 for the primary composite endpoint, P = 0.5183 for ESRD or death, and P = 0.9151 for ESRD). To confirm that the response to treatment was consistent across all levels of baseline proteinuria, we examined the treatment effect on the development of ESRD in the four post hoc baseline proteinuria categories. Table 1 shows that losartan reduced the risk for developing ESRD in all of the baseline proteinuria categories and confirms that a treatment response was observed at all levels of renal impairment.

Effects of an Imbalance in Baseline Proteinuria

As shown in Figure 2, an imbalance was observed in the ≥2000 mg/g proteinuria stratum, in which patients in the losartan group had a median protein excretion of 3460.3 mg/g, whereas those in the placebo group had a median protein excretion of 3136.5 mg/g (P = 0.0117). No significant difference in median baseline proteinuria was observed between the treatment groups for patients in the lower proteinuria stratum (<2000 mg/g; Figure 2).

Because proteinuria is a continuous variable, the covariate-adjusted analysis for proteinuria is shown in Table 2 along with the data for the original prespecified analysis. The risk reduction with losartan as compared with placebo increased from 16.1 (P = 0.022) to 22.2% (P = 0.001) for the primary composite endpoint. Similarly, the treatment effect of losartan increased in magnitude for ESRD or death and for ESRD alone. Figure 3 shows the Kaplan-Meier plots after adjustment for ESRD by proteinuria stratum (Figure 3A) and after adjustment for proteinuria as a continuous covariate (Figure 3B). Note the greater separation between the curves after adjustment for proteinuria as a continuous covariate. To confirm that the change in treatment effect was due to the adjustment for the imbalance in proteinuria in the higher stratum, we conducted similar analyses by proteinuria stratum. Table 2 demonstrates that the adjustment resulted in a greater treatment effect with losartan only in the high proteinuria stratum, in which an imbalance in baseline proteinuria was

Table 1. Losartan treatment effects on ESRD by proteinuria subgroup

<table>
<thead>
<tr>
<th>Baseline Proteinuria (mg/g)</th>
<th>Losartan</th>
<th>Placebo</th>
<th>Hazard Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Events (Rate)</td>
<td>n</td>
</tr>
<tr>
<td>&lt;1000</td>
<td>326</td>
<td>17 (16.6)</td>
<td>330</td>
</tr>
<tr>
<td>1000 to 2000</td>
<td>175</td>
<td>22 (41.3)</td>
<td>181</td>
</tr>
<tr>
<td>2000 to 4000</td>
<td>158</td>
<td>56 (132.1)</td>
<td>180</td>
</tr>
<tr>
<td>≥4000</td>
<td>92</td>
<td>52 (266.3)</td>
<td>71</td>
</tr>
</tbody>
</table>

*The interaction between proteinuria stratum and treatment group was not significant (P = 0.5825). CI, confidence interval; rate, event rate per 1000 patient-years of follow-up.
observed between the treatment groups. Kaplan-Meier plots are shown for ESRD after adjustment by proteinuria stratum (Figure 3C) and after adjustment for proteinuria as a continuous covariate (Figure 3D). Again, the greater separation between the two treatments was observed but only in the ≥2000 mg/g stratum.
Discussion

In the RENAAL study, the prespecified analysis accounted for the impact of baseline proteinuria as a covariate by including the proteinuria strata used at randomization. Subsequent analyses of the unblinded RENAAL study database revealed that baseline proteinuria levels were unbalanced in the upper stratum. Adjustment for baseline proteinuria as a continuous covariate provided a more accurate estimate of the treatment effect observed with losartan. The original stratum-adjusted analysis (9) failed to account for the imbalance in proteinuria values between the treatment groups within the high proteinuria stratum. Therefore, the manner by which a covariate highly correlated to outcome is adjusted can have a meaningful impact on the observed treatment difference.

The aim of a randomized clinical trial is not to determine the relationship between a covariate and the outcome variable; however, the relationship between a covariate and the outcome may affect the statistical analysis and thereby the interpretation of the results. The estimated treatment effect based on a model adjusted for a significant covariate may be influenced by any of the following factors: (1) The association of a covariate with the outcome, (2) an imbalance in the covariate between the treatment groups, (3) the interaction of treatment effect and covariate, and (4) inclusion of the covariate in the model. This analysis demonstrated that significant improvement in the treatment effect in the post hoc analysis was not due to a meaningful interaction between proteinuria and treatment effect, and there was not a significant effect of heterogeneity in treatment response to losartan. The greatest contributing factor to the difference in treatment response between the two analyses was the imbalance in baseline proteinuria.

On the basis of available outcomes data at the time of its design, the RENAAL study was designed to stratify proteinuria at randomization. Despite this, a higher mean level of baseline proteinuria was observed in the losartan group. The reason for the observed imbalance is not clear. However, few patients per site and many patients with large values of proteinuria may have increased the chance of an imbalance between the two treatment groups. In addition, the lack of an upper limit for proteinuria at baseline increased the likelihood that an imbalance might occur. In contrast to proteinuria, 14 other baseline covariates, which were prespecified in the original analysis, were not significantly different between the treatment groups. Furthermore, adjusting for the 14 variables in the analysis did not result in a change in the treatment effect despite some of them being strong predictors for the outcomes.

Because of the importance of baseline proteinuria on renal outcome, simple stratification may not result in balance. We propose that it be addressed by stratification at randomization with strata that have defined upper limits or better blocking paradigms or by treating proteinuria as a continuous variable in the Cox model. Regardless of which method is used, the decision should be prespecified at the time of study design. If adjustment for proteinuria occurs, then it is important to confirm that there is no meaningful interaction between proteinuria and treatment effect and that the response to treatment is consistent across all levels of baseline proteinuria. In studies with small numbers of patients, this is most easily achieved by an interaction test, whereas in larger studies, analysis by subgroup is more appropriate. Importantly, adjustment of baseline proteinuria as a continuous covariate when an imbalance occurs at baseline may not always improve the treatment effect, although it is likely to improve the estimate of the treatment effect. For example, if the imbalance had occurred in the placebo arm, then the adjustment would have resulted in a decrease in the treatment effect observed with losartan. Of interest, the RENAAL study is not the only large, randomized, double-blind, placebo-controlled renal outcomes study that has shown an imbalance in baseline proteinuria. Of seven major renal outcomes trials (9,11–16) completed within the past 15 yr, an imbalance in baseline proteinuria was observed in two of the studies (9,14).

The intention of this article was not to correct the estimate of the treatment effect reported for the RENAAL study; the goal was to illustrate the conditions under which covariate adjustment should occur and how it can affect the estimate of the treatment effect. In future clinical studies on renal endpoints, every effort should be made to reduce and correct for the imbalance of baseline proteinuria that may occur by chance. Including proteinuria as a covariate in the analysis should be considered in the primary analysis even with stratification for proteinuria at randomization.

Acknowledgments

The data contained in this article were presented as a poster at the American Society of Nephrology Meeting, San Diego, November 12 to 17, 2003, SU-PO674 (J Am Soc Nephrol 14: 682A, 2003).

References


Access to UpToDate on-line is available for additional clinical information at http://www.jasn.org/