Review and Practical Excursus of the Propensity Score: Low Protein Diet Compared to Mediterranean Diet in Patients With Chronic Kidney Disease

In depth review

Vincenzo Calabrese¹, Guido Gembillo¹, Martina Buda², Valeria Cernaro¹, Elisa Longhitano¹, Antonello Pontoriero³, Matteo Polizzi⁴, Daniela Metro¹, Domenico Santoro¹

1 Unit of Nephrology and Dialysis, Department of Clinical and Experimental Medicine, A.O.U. ‘G. Martino’, University of Messina, Messina, Italy
2 General and Oncologic Surgery Unit, Faculty of Medicine, University of Messina, G. Martino University Hospital, Messina, Italy
3 Unit of Nephrology, Department of medicine ASP 5 Messina, P.O. Milazzo
4 University of Messina, Messina, Italy

Corresponding author:
Vincenzo Calabrese
Tel: 328736905
calabrese@outlook.it

ABSTRACT

Although Randomized clinical trials (RCT) represent the gold standard to compare two or more treatments, the impact of observational studies cannot be ignored. Obviously, these latter are performed on unbalanced sample, and differences among the compared groups could be detected. These differences could have an impact on the estimated association between our allocation and our outcome. To avoid it, some methods should be applied in the analysis of observational cohort. Propensity score (PS) can be considered as a value which sums up and balances the known variables. It aims to adjust or balance the probability of receiving a specific allocation group, and could be used to match, stratify, weight, and perform a covariate adjustment. PS is calculated with a logistic regression, using allocation groups as the outcome. Thanks to PS, we compute the probability of being allocated to one group and we can match patients obtaining two balanced groups. It avoids computing analysis in unbalanced groups.

We compared low protein diet (LPD) and the Mediterranean diet in CKD patients and analysed them using the PS methods. Nutritional therapy is fundamental for the prevention, progression and treatment of Chronic Kidney Disease (CKD) and its complications. An individualized, stepwise approach is essential to guarantee high adherence to nutritional patterns and to reach therapeutic goals. The best dietary regimen is still a matter of discussion. In our example, unbalanced analysis showed a significant renal function preservation in LPD, but this correlation was denied after the PS analysis. In conclusion, although unmatched analysis showed differences between the two diets, after propensity analysis no differences were detected. If RCT cannot be performed, balancing the PS score allows to balance the sample and avoids biased results.

KEYWORDS: Chronic Kidney Disease, Low Protein Diet, Matching, Mediterranean Diet, Nutritional Therapy, Propensity Score, Randomized Clinical Trials
Introduction

Clinical investigations are mainly categorized in observational and interventional studies, the latter including randomized controlled trials (RCT) [1]. Comparative effectiveness studies belong to the family of observational studies and aim to compare two active treatments to identify which one is more efficient in improving the time course of a disease or reducing the risk of a given condition in real life (i.e., in a context different from an RCT) [2]. From this perspective, this type of study design differs from RCTs because the latter specifically contemplate ‘no intervention’ (i.e., the placebo arm).

Treatments are candidates to be investigated by a study of comparative effectiveness only when the same treatment was proved to be effective versus a control in an RCT. The main reason why these studies are considered with caution by the scientific community is the lack of randomization, which implies that the results of these studies are prone to a peculiar type of bias called ‘confounding by indication’ [3]. In a given treatment-outcome pathway, a confounder is a variable that is associated with the treatment (i.e., it differs between the study arms). It is not an effect of the treatment, does not lie in the causal pathway between the treatment and the outcome, and represents a risk factor for the outcome. In real life, a confounder can increase, reduce, or definitely obscure the true effect of treatment on an outcome. Despite these challenges, observational studies of effectiveness offer opportunities to examine questions impossible to be investigated by RCTs [4]. First, they can be used to examine the effectiveness of medication that has already obtained marketing authorization and for which funding for further trials may be limited. Second, they can allow the examination of effectiveness for rare treatment indications. Third, a large observational study can be more representative of a clinical population and less prone to selection bias than a trial.

In observational studies of effectiveness, common methods used to adjust to confounding are multiple regression models [5], the use of instrumental variables [6], and the propensity score (PS) [7]. Briefly, multiple regression analyses are performed by including in the model all variables that meet the criteria to be considered as confounders. An instrument is a variable that predicts exposure, but conditional on exposure shows no independent association with the study outcome. As an example, we can consider an observational multicenter study that evaluates how different treatments can affect a clinical outcome. The facility allocation can be considered as the result of a ‘natural experiment’ by simulating a randomization. In this manuscript, we describe an efficient statistical technique used by researchers to mitigate the problem of confounding in observational studies of effectiveness.

Propensity score

The propensity score (PS) was described in 1983. This method allows adjusting or balancing for the probability to receive a specific allocation group, an estimation of the likelihood of being in one or in another group in relation to a set of covariates [8]. PS could be used to match, stratify, weight, and perform a covariate adjustment. If the outcome is a binary variable, matching has less bias than stratification or covariate adjustment, as in a time-dependent outcome both matching and Inverse Probability of Treatment Weighting (IPTW) are less biased than stratification or covariate adjustment. PS is calculated with a logistic regression, using allocation groups as outcome. Thanks to this method, we can compute the probability of confounder variables to be allocated in one group. Since PS has no limits of variables, it can be used in small samples and for rare diseases [9], unlike multivariate regression.
Matching
Matching with PS methods allows us to compare one or more patients with the same allocation probability, so it follows that matched patients have similar features, decreasing bias. This method consists of matching cases of two or more groups on the basis of similar predicted PS, thus allowing the comparison of groups with an equal distribution of confounders (covariate balance) [10, 11]. Imaging having two groups of patients, at first, we need to compute PS, corresponding to the probability of receiving allocation in group A, for each one of them [12]. By doing this, a binomial logistic regression is performed to select, among the study variables, those associated with the allocation variable. Patients with the same PS value are thus compared. Minimizing the differences between patients, and comparing homogeneous groups, confounding is reduced.

Stratification
The stratification by PS follows the matching methods. Strata will be created between subjects with similar PS of treatments. The Stratification method removes about 90% of bias due to covariate imbalance [13]. Formally, stratification by PS can be resumed as follows:

- choosing variables included in the PS model among personal data, comorbidities, laboratory data, and variables clinically related to outcome
- estimating PS value for each subject, with logistic regression using allocation as the dependent variable
- calculating the Cumulative Distribution Function for each subject, able to define the distribution also in a discrete and binomial variable
- ranking population based on PS value, dividing the whole sample into quartiles, tertiles, deciles, etc., based on PS values
- assessing balance for each of the K (K is the indicator of the treatment group), analyzing the baseline features
- retaining the PS value ordering that creates strata with the best covariate balance and conducting a stratified outcome analysis to estimate ATE or ATT [14].

The number of strata can be evaluated based on the number of covariates (2×covariates – 1) with groups of more than ten subjects [15]. In a large observational study, Cernaro V. et al. [16], on behalf of the Workgroup of the Sicilian Registry of Nephrology, analyzed the impact of convective dialysis on mortality and cardiovascular mortality. They performed Cox Regression analysis with incremental multivariate models but, although the independent impact of convective dialysis on mortality, many other variables were related to the outcome. Thus, as highlighted in their methods section, PS stratification was computed to perform a sensitivity analysis [17]. PS was computed through a multivariate logistic regression model including age, gender, ethnicity, arterial hypertension, diabetes mellitus, and cardiac diseases. Then, the whole sample was divided into quartiles (based on PS value) and survival analyses computed in the whole sample were repeated. These latter results confirmed the independent impact of the treatment, but in subsamples that are theoretically more homogenous because PS value was computed on the bases of the possible confounding.

Inverse Probability of Treatment Weighting (IPTW) Estimation
IPTW analyses aimed to create a weighted sample in which the distribution of each confounding variable was the same between the compared groups [18]. Patients will be allocated the reciprocal
of the PS value: each patient of the treated group receives the weight of 1/PS and each patient of the untreated group receives the weight of 1/(1-PS). A treated patient with a low PS value enters in the analysis with a high weighting because he is considered likely an untreated patient in terms of comorbidities, so a valid comparison can be made between the two [19]. Practically, in the analysis, each patient is evaluated as many times as their IPTW is. A treated patient with a PS of 0.1 will weigh 1/0.1=10 and will be considered in the analysis ten times. Similarly, a treated patient with high PS, for example 0.8, will weigh 1/0.8=1.25 and it will be considered in the analysis 1.25 times. Moreover, IPTW was at the basis of the Marginal Structured Models, a multistep estimation procedure designed to control confounding variables at different time points in longitudinal studies [20]. IPTW method is not robust against the outliers.

**Covariate adjustment**

This method uses the PS values as a covariate in a linear regression analysis. Even if there is no significant association between the covariates used to compute PS value and the outcome, the use of PS value as a covariate allows us to approximate the effect of each of the aforementioned covariates [21].

**Practical example**

To explain these methods, we will use a dataset containing 75 non-randomized patients with CKD stage III-IV. All the remaining patients gave written consent to data processing for research purposes in respect of privacy. Ethical approval was not necessary according to National Code on Clinical Trials declaration and according to Italian ministerial rules of September 6, 2002 n°6, because our observation derives from a real-life retrospective study.

Patients were followed up for one year. 40 patients followed an LPD, defined by a protein intake of 0.6 g/kg/day (Group A), and 35 patients were subjected to the Mediterranean diet (Group B). The allocation, according to the real-life observation design, was based on dietician suggestions, patient’ habits, and adherence abilities, which were evaluated during the baseline visit. Supplementary Table 1 and Table 2 summarized the details about the quantity and the nature of both diet regimens. Laboratory data were collected at the baseline visit (T0) and the annual follow-up (T1), as follows: serum urea (mg/dl), serum creatinine(mg/dl), serum phosphorous (mg/dl), serum sodium (mmol/l), serum potassium (mmol/l), white blood cells (WBC) (cc/mmc). The groups had significant differences in BMI (28.7 [25.0, 34.7] vs 26.4 [24.0, 28.0], p=0.02), age (68 ± 9 vs 74 + 13, p=0.04), and basal creatinine clearance (33 [25, 44] vs 27 [21, 36], p=0.03). Baseline features were summed up in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Whole sample</th>
<th>Group A (n= 40)</th>
<th>Group B (n= 35)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>71 ± 11</td>
<td>68 ± 9</td>
<td>74 ± 13</td>
<td>0.04</td>
</tr>
<tr>
<td>Sex (M/F)</td>
<td>45/55</td>
<td>40/60</td>
<td>49/51</td>
<td>0.32</td>
</tr>
<tr>
<td>BMI (kg/mq)</td>
<td>27.4 [24.2 – 30]</td>
<td>28.7 [25.0 – 34.7]</td>
<td>26.4 [24.0 – 28.0]</td>
<td>0.02</td>
</tr>
<tr>
<td>Clearance (ml/min)</td>
<td>31 [23 – 41]</td>
<td>33 [25 – 44]</td>
<td>27 [21 - 36]</td>
<td>0.03</td>
</tr>
<tr>
<td>Serum Urea (mg/dl)</td>
<td>73 [64 -102]</td>
<td>75 [65 -99]</td>
<td>73 [60 -121]</td>
<td>0.84</td>
</tr>
<tr>
<td>Serum creatinine (mg/dl)</td>
<td>1.8 [1.5 – 2.5]</td>
<td>1.7 [1.4 – 2.4]</td>
<td>2.0 [1.6 – 2.7]</td>
<td>0.03</td>
</tr>
<tr>
<td>Serum sodium (mmol/L)</td>
<td>141 ±3.3</td>
<td>4.7 [4.5 – 5.0]</td>
<td>4.4 [4.9 – 5.2]</td>
<td>0.40</td>
</tr>
<tr>
<td>Serum Potassium (mmol/l)</td>
<td>4.74 ± 0.58</td>
<td>4.72 ± 0.53</td>
<td>4.76 ± 0.64</td>
<td>0.68</td>
</tr>
<tr>
<td>Serum phosphorous (mg/dl)</td>
<td>3.8 [3.6 – 4.1]</td>
<td>3.7 [3.5 – 4.0]</td>
<td>3.8 [3.7 – 4.3]</td>
<td>0.35</td>
</tr>
<tr>
<td>WBC (cc/mmc)</td>
<td>7744 ± 1824</td>
<td>7575 ± 1947</td>
<td>7932±1683</td>
<td>0.46</td>
</tr>
<tr>
<td>Delta_Clearance</td>
<td>-3.50/ 0.00/ 4.00</td>
<td>-0.25/ 1.00/ 7.25</td>
<td>-5.50/2.00/ 2.00</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 1. Baseline features of whole sample and into the two groups. Body mass index (BMI); White blood cells (WBC).
An unadjusted model with GLM for repeated measures showed a significant effect on creatinine clearance of the Mediterranean diet compared to LPD, with an estimate marginal mean of -9.98 ml/min [95% CI, 15.6/, 4.3]. Adjusted model for age, BMI and sex (Table 2) appeared to confirm this significance in the between-group mean in the joint mean difference (–9.34, 95%CI –15.44/ –3.24) (Table 2).

<table>
<thead>
<tr>
<th>Variable</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mediterranean diet vs low protein diet</td>
<td>–9.34</td>
<td>0.003</td>
<td>0.119</td>
</tr>
<tr>
<td>Sex (Male vs female)</td>
<td>2.71</td>
<td>0.104</td>
<td>0.038</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.08</td>
<td>0.780</td>
<td>0.001</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>0.04</td>
<td>0.947</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 2. Between-group mean in the joint mean differences: Adjusted GLM model for repeated measures. Body mass index (BMI).

Due to the non-randomized study design and the unbalanced groups, we decided to implement the analysis with the PS matching. We computed PS value using the treatment as dependent variable of the logistic regression, and graphically evaluated it (Figure 1). The PS values were not equally distributed between the two groups. Carrying on with the matching, choosing a caliper of 0.2, 20 patients from group A were paired with 20 patients from group B (Table 3). Unmatched patients are excluded from the analysis, reducing sample’s size. This reduction of the patients admitted in the analysis is one of the major limitations of the matching.

Analyzing the standardized means of the baseline features before and after the matching, a better balance between the two groups could be shown (Figures 2a and 2b).

GLM for repeated measures performed in the matched sample did not show significant differences between the two groups (2.737, 95%CI –4.328/9.803). Also using the covariate adjustment, that uses the whole sample, the not significant relationship between the two treatments and the clearance progression was confirmed in the GLM for repeated measures including treatment and ps-value (–3.314, 95%CI –8.524/1.897).

![Figure 1. Propensity score distribution before the matching.](image)
Table 3. Groups composition based on Propensity Score Matching.

<table>
<thead>
<tr>
<th>Group A</th>
<th>Group B</th>
<th>PS value group A</th>
<th>PS value group B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>0.5728</td>
<td>0.5990</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>0.5029</td>
<td>0.4885</td>
</tr>
<tr>
<td>3</td>
<td>43</td>
<td>0.7979</td>
<td>0.8133</td>
</tr>
<tr>
<td>4</td>
<td>53</td>
<td>0.2244</td>
<td>0.2236</td>
</tr>
<tr>
<td>5</td>
<td>41</td>
<td>0.8256</td>
<td>0.8244</td>
</tr>
<tr>
<td>6</td>
<td>65</td>
<td>0.2370</td>
<td>0.2436</td>
</tr>
<tr>
<td>8</td>
<td>49</td>
<td>0.7872</td>
<td>0.7496</td>
</tr>
<tr>
<td>9</td>
<td>47</td>
<td>0.7313</td>
<td>0.7068</td>
</tr>
<tr>
<td>10</td>
<td>52</td>
<td>0.2709</td>
<td>0.2670</td>
</tr>
<tr>
<td>11</td>
<td>66</td>
<td>0.5662</td>
<td>0.5339</td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>0.6588</td>
<td>0.6768</td>
</tr>
<tr>
<td>14</td>
<td>71</td>
<td>0.6731</td>
<td>0.6888</td>
</tr>
<tr>
<td>15</td>
<td>75</td>
<td>0.1971</td>
<td>0.2084</td>
</tr>
<tr>
<td>16</td>
<td>39</td>
<td>0.6640</td>
<td>0.6990</td>
</tr>
<tr>
<td>18</td>
<td>63</td>
<td>0.3849</td>
<td>0.3833</td>
</tr>
<tr>
<td>19</td>
<td>55</td>
<td>0.4595</td>
<td>0.6256</td>
</tr>
<tr>
<td>21</td>
<td>67</td>
<td>0.6014</td>
<td>0.6256</td>
</tr>
<tr>
<td>26</td>
<td>45</td>
<td>0.4350</td>
<td>0.4386</td>
</tr>
<tr>
<td>27</td>
<td>42</td>
<td>0.2674</td>
<td>0.2947</td>
</tr>
<tr>
<td>31</td>
<td>60</td>
<td>0.4544</td>
<td>0.4280</td>
</tr>
</tbody>
</table>

Figure 2a. Balance of the covariate before and after the Matching.

Figure 2b. Propensity score distribution after the Matching.
Usefulness of propensity score

A few RCTs were conducted on ERSD patients due to high costs and their difficult organization. In these cases, a well-structured comparative effectiveness study could be done to generate hypothesis or to add results to existing RCT. For Example, Chan KE et al. conducted a large observational study including more than 10000 patients, the study’s population and structure were modeled on 4D study’s methods, using the same eligibility criteria, endpoints, and similar timeline. To reduce bias caused by known and unknown variables, patients were initially matched in statin-group and control-group based on similar lipid profiles and years of dialytic treatment. Subsequently, a logistic model was performed to compute the probability of receiving the therapy, also all Cox analyses were weighted using the IPTW methods. Differently from the unmatched baseline analysis, the baseline characteristic computed after propensity scoring showed two well-balanced groups. At the outcome analysis, all HRs computed in this observational study were compared with the HRs showed in the 4D Study, and no significant differences were found between these two studies (Figure 1). Furthermore, RCTs are often smaller than observational studies, due to the stronger inclusion criteria and the higher costs than observational design. As shown in Figure 1, PS methods computed in a big sample, allowed to find a smaller confidence interval compared to 4D RCT, without significant differences in anyone outcome.

Through these comparisons, although RCTs were the lowest-biased studies, we can speculate about the effective validity of observational comparative studies using PS methods to reduce biases.

Limitation of propensity score methods

PS is applicable when the treatment assignment is neglectable, with unknown and unmeasured confounders. Furthermore, PS value > 0 is necessary. According to G. et Lepeyre-Mestre M. [22], propensity score methods is not very able to reduce selection bias, information bias and instrumental bias. Despite PS reducing inhomogeneity between groups, some unconsidered variables can exist, hence residual bias should be taken into account in the interpretation of results and in the critical appraisal of the study [23]. Leisman D.E. et al., resumed ten “Pearls and Pitfalls” about the use of matching method [24]. They highlighted problems regarding the reduction of sample size: the number of cases does not represent the whole sample because every unpaired subject is excluded from the analysis. This can impair the external validity of the study, reducing its applicability. Consequently, the power of the study should be computed on the balanced sample, excluding the unmatched patients. Indeed, the analysis reflects the matched sample, losing information about the excluded cases. However, no patients were excluded by the analysis using the covariate adjustment and the IPTW. We highlighted that, similarly to our sample, no significant differences between matching and covariate adjustment were found. However, can be useful performing more PS methods, to compare the results. Furthermore, machine learning methods can be used to compute PS, and they reduce the variability of the PS. Last but not least, a limitation of these methods is the inability to detect interaction variables. In correlated subgroup effects, these variables could indeed invalidate the PS model and should be excluded from it [25].

Discussion

Our analysis seemed to show a slow CKD progression in patients treated with LPD compared to patients treated with Mediterranean diet. However, the unbalanced covariate distribution between the two groups must be highlighted. Conversely to classic analysis, our result showed no difference between the two groups in matched sample, where the two groups were well balanced.
Healthy dietary habits are essential to contrast the progression of chronic diseases such as CKD and the risk factors related to its development. A tailored diet that follows patients’ eating habits can enhance compliance with nutritional therapy, improving the conservative management of CKD patients.

In patients with renal impairment, optimal eating is crucial, representing a high-impact modifiable lifestyle factor for the primary prevention of CKD progression [26], and it avoids the dysregulation of fluid status, pH, electrolytes [27, 29], chronic metabolic acidosis [30], all factors that should be corrected by an adequate dietary regimen and balanced supplementation of the missing nutrients.

Nutritional therapy can be useful to slow CKD progression and delay ESRD with a consistent improvement of the patient’s quality of life [31]. LPD should be started from GFR <30 ml/min, with a protein intake below 0.8 mg/kg/die, and it shown slower CKD progression and reduction of the mortality [32]. Rhee et al. (2018) [33] in their meta-analysis of randomized controlled trials (RCTs) found that the risk of progression to ESRD was significantly lower in patients with LPD regimens than those with higher-protein diets, with serum bicarbonate augmentation. Notwithstanding its restrictions, LPD does not seem to impair the quality of life of CKD patients. The study of Piccoli et al. (2020) [34] on 422 CKD patients with stages III-V demonstrated that moderately protein-restricted diets (0.6 g/kg/day) guaranteed good compliance to therapy, with a median dietary satisfaction of 4 on a 1-5 scale with a minimal dropout.

The Mediterranean diet is a nutritious regimen first proposed by Keys in the mid-1980s that has been demonstrated to exert a favourable action on inflammation, CKD, cardiovascular health, and overall mortality [35, 37]. Different studies demonstrated a tight link between CKD prevention and Mediterranean diet regimen [38, 39]. How the Mediterranean diet exerts kidney protection is still under debate, and the anti-inflammatory and antioxidant effects were suggested [40, 41]. Moreover, tighter adherence to a healthy plant-based diet was associated with a slower eGFR decrease [42].

Asghari et al. (2017) [43] showed, in a six-year follow-up study, that adherence to the Mediterranean diet is associated with a reduced risk of 50% of incident CKD. These results are in line with the ones from the Northern Manhattan Study. In this cohort of patients, the patients with relatively preserved renal function and high adherence to the Mediterranean diet experienced an approximate 50% decreased odds for incidence of eGFR<60 ml/min/1.73m².

The effectiveness of LPD compared to the Mediterranean diet is still a matter of debate. Mediterranean diet is characterized by free fat, abundant vegetables, legumes, fresh fruits, cereals, moderate wine consumption, low milk and milk products, low meat/animal products, and frequent fish. Moreover, both the Mediterranean diet and LPD are effective in the modulation of gut microbiota, reducing protein-bound uremic toxins levels, especially in patients suffering from moderate to advanced CKD.

Davis et al. (2015) [44] tried to define nutrient content and range of servings for the Mediterranean diet, analysing the variations in the quantity of this diet components in recent literature. The Mediterranean diet’s positive effects are not only limited to metabolic influence, but the conviviality, culinary and physical activity exerts a beneficial effect on mental health, ameliorating body homeostasis and reactivity to the chronic disease [45].

A diet regimen feasible in different settings is essential for adherence to nutritional therapy. Different dietetic strategies have been investigated over the years, but which is the best nutritional regimen remains controversial. Kim et al. analysed the data of 4343 incident CKD patients, during a median follow-up of 24 years and showed that higher adherence to a balanced diet was linked to a lower risk of CKD progression.
In conclusion, although our previous analysis showed differences between the two diets, after propensity match no differences were detected, as well as after the covariate adjustment methods. In the study of Hu et al. (2021) [46] adherence to healthy nutritional patterns was associated with lower risk for renal impairment progression and all-cause mortality in CKD patients. Thus, based on our results and according to the literature, the Mediterranean diet should be a good choice for patients who are not compliant with a low-protein diet, without a significant increase of CKD progression risk [47].
BIBLIOGRAFIA


