The HEMO Study: applicability and generalizability

Michael V. Rocco¹, Alfred K. Cheung², Tom Greene³ and Garabed Eknoyan⁴, for the Hemodialysis (HEMO) Study Group⁵

¹Wake Forest University School of Medicine, Winston-Salem, North Carolina, ²Veterans Affairs Salt Lake City Health Care System and University of Utah, Salt Lake City, Utah, ³Cleveland Clinic Foundation, Cleveland, Ohio, ⁴Baylor College of Medicine, Houston, Texas and ⁵National Institutes of Diabetes, Digestive and Kidney Diseases, Bethesda, Maryland, USA

Keywords: clinical trial; dialysis dose; dialysis flux; haemodialysis; hospitalizations; mortality

Introduction

The 2002 Lasker Clinical Medical Research Award honoured Willem Kolff and Belding Scribner, whose seminal work changed kidney failure from a fatal to a treatable disease. What began as exploratory efforts to sustain life and relieve uraemic symptoms in selected patients now provides chronic life-saving replacement therapy to millions of people worldwide. Continued improvements in clinical expertise, accrued scientific information and technical advances have improved dialytic therapy and outcomes. Unfortunately, despite this encouraging trend, there remain significant differences in clinical practice and less than optimal patient outcomes. The annual mortality rate of dialysed patients, albeit variable among different countries, remains unacceptably high. Although medical comorbidities and late referral of patients to nephrologists may contribute to this high mortality rate, the dialytic prescription can also influence mortality. To address this latter concern, national guidelines for dialysis therapy have been developed, but are based, to a great extent, on observational studies and registry data.

There have been two randomized clinical trials of the dialysis prescription on outcomes in haemodialysis patients. The first, the National Cooperative Dialysis Study (NCDS), completed >20 years ago, established the importance of dialysis dose in affecting patient outcomes at dose levels substantially below current standards [1,2]. Subsequently, a large number of reports, from observational data sets, suggested that a dose of dialysis substantially higher than that provided in the NCDS trial is associated with a lower mortality rate [3–6]. In response to these observational studies, national standards for dialysis dose were developed in the United States advocating a minimum single pool Kt/V urea (spKt/V) of at least 1.2 [7]. The second randomized clinical trial, the Hemodialysis (HEMO) Study, begun in 1995, was designed to determine whether further increases in dialysis dose above current standards or the use of high-flux membranes would improve patient outcomes. The main results of the HEMO Study, published in December 2002 [8], showed no statistically significant effects by the two study interventions examined, dialysis dose and membrane flux, on mortality rates or on intermediate composite outcomes defined by time to either death or selected first cause-specific hospitalizations. Since the publication of this report, several articles have appeared which discuss the implications of the HEMO Study results for clinical practice, including generalizability, statistical power and design [9,10]. The goal of this report is to address the major questions that have been raised pending the publication of additional analyses of HEMO Study data.

Study design and conduct

The details of the HEMO study design have been described elsewhere [8,11,12]. In this article, only salient and pertinent features of the study are reviewed. The trial was conceived in 1992, a pilot study was conducted in 1993 and 1994 and the full-scale trial was conducted between March 1995 and December 2001. A total of 1846 patients were randomized to either a high dose [target urea equilibrated Kt/V (eKt/V) of 1.45, which approximates a spKt/V of 1.65] or a standard dose...
(target eKt/V of 1.05, which approximates a spKt/V of 1.25) of dialysis and to a high-flux (mean $\beta_2$-microglobulin clearance: $>20\text{ ml/min}$) or a low-flux (mean $\beta_2$-microglobulin clearance: $<10\text{ ml/min}$) membrane. Approximately 900 patients were randomized in the first 18 months of the study. From that time until the final year of the trial, patients who died, terminated haemodialysis due to kidney transplantation, switched to another type of dialysis or transferred to a dialysis facility not participating in the study were replaced with newly randomized patients. Under this recruit-to-replace strategy, more than 5200 patient-years of follow-up were accrued, which is $\sim 60\%$ greater than would have been possible using the same resources without study subject replacement. As expected, the high death rate of chronic haemodialysis patients resulted in a decrease in the length of mean patient follow-up. Thus, while the potential mean follow-up time was 4.48 years (range: 0.9–6.6 years, depending on time of randomization), the actual mean follow-up time was 2.84 years, which reflects an annual mortality rate of 16.6% in the HEMO Study.

Key entry criteria included a residual kidney urea clearance of $\leq 1.5\text{ ml/min per 35 l of urea distribution volume}$, achievement of an eKt/V $>1.3$ within two of three consecutive monitored dialysis sessions in which the high-dose goal was targeted and serum albumin $>2.6\text{ g/dl}$ by nephelometry. Centralized dialysis prescriptions and monthly monitoring of the delivered dialysis dose led to a clear separation between the high- and standard-dose groups (Figure 1). The mean difference between the two dose groups was 0.37 eKt/V units, which was 92.5% of the study goal separation of 0.40 eKt/V units. The flux intervention was monitored by dialytic $\beta_2$-microglobulin clearance every 2 months in the high-flux arm and every 6 months in the low-flux arm. This monitoring in the initial stages of the study revealed that, in certain situations, reprocessing with Renalin without bleach substantially reduced dialyser $\beta_2$-microglobulin clearance over successive reuses. To compensate for this decrease, a maximum limit of four reuses was imposed for dialysers using this reprocessing method. As shown in Figure 2, this dialyser reuse limit resulted in a clear separation in $\beta_2$-microglobulin clearance between the low-flux ($3.4 \pm 7.2\text{ ml/min}$) and high-flux groups ($33.8 \pm 11.4\text{ ml/min}$).

**Generalizability**

The design of randomized trials requires achieving a balance between internal consistency, which is the ability to determine the effects of the interventions within the trial, and external consistency, which is the ability to generalize the results of the trial to the target population. In order to maintain internal consistency, randomized trials usually exclude patient subgroups deemed likely to compromise the protocol (e.g. patients with factors that prevent accurate evaluation of the independent or dependent variables), to be unable to

**Fig. 1.** (Left) Separation of eKt/V values in the standard- and high-dose groups of the HEMO Study. (Right) Separation in urea reduction ratio (URR as %) between the standard- and high-dose groups. Box plots show 10th, 25th, 50th, 75th and 90th percentiles and mean (+).
follow the interventions or who are believed to have such severe comorbidity that early death would occur irrespective of treatment. In the HEMO study, the main exclusions were related to residual kidney function and the inability to achieve a high dose of dialysis, the rationale for which is discussed below. As a result of the need to balance internal and external consistency, patients in trials are often not exact representations of their target patient populations. This is considered an acceptable compromise based on the argument that the results of a well-conducted trial are likely to be generalizable provided the spectrum of patients included in the trial is sufficiently broad and provides substantial representation of the relevant subgroups in the target population [13–15].

On the other hand, some trials have been criticized for using entry criteria so restrictive as to grossly limit generalizability. The exclusion of diabetic, elderly and hypertensive patients from the NCDS is one such example. The HEMO Study investigators, therefore, strove to employ the broadest possible entry criteria without sacrificing the ability to conduct the study. Diabetic and hypertensive patients were included and well represented, while the exclusion criteria for age >80 years and a serum albumin level of <2.6 g/dl encompassed <10% of prevalent chronic haemodialysis patients in the US [16]. Further, patient baseline characteristics were monitored continuously to assure representation from subgroups with high medical comorbidity [11,12]. The distribution of patients in the HEMO trial did differ from the US haemodialysis population in two respects: (i) due to a preponderance of urban centres, 63% of HEMO patients were African-American vs 41% in the US general dialysis population and (ii) the heaviest patients had to be excluded to assure that the high-dose goal could be achieved. Indeed, it has been argued that the results of the trial may not be generalizable due to either over-representation of healthier and African-American patients or the exclusion of elderly and large patients. The suggestion that the HEMO Study over-recruited healthy patients is not substantiated by analysis of patient characteristics. For example, 44.6% of HEMO Study patients had diabetes mellitus, which is similar to the 43% in the US dialysis population [16,17], but higher than the rate of 5–21% reported in European registries and some observational studies over the past decade [18–21]. In addition, in spite of the upper age limit of 80 years, the mean age at randomization in the HEMO Study was 57.6 years, compared with a mean age of 61.4 years in the Centers for Medicare and Medicaid (CMS) Clinical Performance Measures (CPM) project [22] and similar to the mean age of 54–60 years in multicentre haemodialysis studies in European patients [18–21]. While differences in classification systems complicate strict comparisons among various studies, the percentage of patients with individual medical comorbidities in the HEMO Study appears similar to that reported by the United States Renal Data System (USRDS) (Table 1) [16]. Moreover, the projected death rate of US haemodialysis patients within the age limits and racial distribution of the HEMO Study was ~17.6%, which is similar to the 16.6% mortality rate observed in the HEMO Study [8]. Thus, the comorbidity of haemodialysis patients in the HEMO Study was well within the range encountered in the US dialysis population, but higher than that typically reported from Europe [18–21]. However, the percentage of older and diabetic

Fig. 2. Separation of β₂-microglobulin clearance values in the high- and low-flux groups of the HEMO Study. Box plots show 10th, 25th, 50th, 75th and 90th percentiles and mean (+).
patients is increasing in Europe, suggesting that in the future, the characteristics and comorbidities of European dialysis patients will more closely resemble those of current haemodialysis patients in the US.

Although the HEMO trial included a higher percentage of African-Americans than in the US haemodialysis population, non-African-Americans were nonetheless well represented (n = 690). Moreover, relative risks (RR) of mortality for high vs standard dose [(RR: 1.11; 95% confidence interval (CI): 0.89–1.37] and high vs low flux (RR: 1.04; 95% CI: 0.84–1.26) exceeded 1.0 among non-African-Americans [23], indicating that the high-dose or high-flux interventions did not have larger beneficial effects in non-African-Americans than in African-Americans. Thus, it is improbable that an increase in non-African-American enrolment into the study would have increased the likelihood of detecting benefits of the interventions.

Concerning the implications of excluding heavier patients, ~3% of patients randomized into the HEMO trial exceeded 100 kg, as compared with an estimated 10% of haemodialysis patients in the US [22, 24]. We believe that exclusion of a limited number of heavier patients is unlikely to have suppressed the ability of the trial to detect beneficial effects of the interventions for two reasons. First, since larger and heavier patients tend to have a lower overall mortality rate than lighter and smaller patients in the haemodialysis population [3–6], inclusion of a greater proportion of heavier patients is likely to have reduced the number of deaths observed in the trial, thus, reducing the statistical power to detect a given intervention effect. Second, it has been hypothesized that smaller and lighter patients are more likely to benefit from a higher dose of dialysis than larger and heavier patients [25]. Thus, the increased proportion of smaller and lighter patients in the HEMO trial should have increased, rather than decreased, the sensitivity of the trial to the effects of increasing the dose of dialysis.

The dialysis parameters of the HEMO Study were within the range of those encountered in current US practice. In the HEMO Study, the average time on dialysis was 190 ± 23 min in the low-dose group and 219 ± 23 min in the high-dose group and the average blood flow rates were 311 ± 51 and 375 ± 32 ml/min, respectively. These data are similar to those reported in the most recent CMS-CPM project. In that report, the average treatment time was 217 min and the average blood flow rate was 410 ml/min for patients with native arteriovenous fistula, 418 ml/min for patients with synthetic grafts and 345 ml/min in patients with catheters [22].

Finally, although a history of hypertension was prevalent in the HEMO Study cohort, blood pressure was controlled throughout the HEMO Study follow-up. The respective mean systolic and diastolic blood pressures during follow-up were 151 ± 18 and 81 ± 11 mmHg pre-dialysis and 137 ± 17 and 74 ± 10 mmHg post-dialysis. This compares with mean systolic and diastolic blood pressures of 152/80 mmHg pre-dialysis and 137/74 mmHg post-dialysis reported in the USRDS Waves 3 and 4 Study [26].

### Inclusion of prevalent patients

The question has also been raised as to whether the ability of the HEMO Study to detect beneficial effects of the interventions may have been impaired by its enrolment of prevalent, rather than only incident, haemodialysis patients. Patients with residual kidney function would result in higher and more variable levels of total urea and β₂-microglobulin clearance. The end result of these higher total clearances would result in masking the effects of the dialysis treatment and, thus, a reduced power to detect differences in the interventions due to a decrease in the proportional separation between the treatment groups for both total urea and β₂-microglobulin clearance. Thus, the inclusion of prevalent patients likely would have reduced the statistical power to detect differences in the primary and secondary outcomes of the trial.

Although patients with <3 months of haemodialysis were excluded from enrolment into the HEMO Study, 490 patients were randomized within 1 year of starting dialysis. Subgroup analyses examining the effect of the high-flux intervention in relation to the number of prior years of dialysis indicated that there was a potential benefit of high-flux dialysis for patients with >3.7 years of prior dialysis, but no benefit of high flux in patients with fewer years on dialysis. Thus, it is unlikely that a beneficial effect of high-flux dialysis would have been detected had the study been restricted to incident patients. It has also been suggested that the effects of the randomized treatment assignment might have been obscured by prior dialysis therapy in prevalent patients. In the context of a large randomized trial, however, patient characteristics at entry are assumed to be approximately balanced between the treatment groups due to randomization. In the HEMO Study, the baseline characteristics among the treatment groups, including the prior dialysis therapy, were indeed

### Table 1. Baseline comorbidities in HEMO Study patients and in the USRDS cohort

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>HEMO Study randomized patients (% present)</th>
<th>USRDS (% present)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerebrovascular disease</td>
<td>15.8</td>
<td>9.5</td>
</tr>
<tr>
<td>Ischaemic heart disease</td>
<td>27.5</td>
<td>24.5</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>28.7</td>
<td>32.1</td>
</tr>
<tr>
<td>Cardiac arrhythmia</td>
<td>17.3</td>
<td>6.9</td>
</tr>
<tr>
<td>Peripheral vascular disease</td>
<td>20.0</td>
<td>14.3</td>
</tr>
<tr>
<td>Hypertension</td>
<td>93.7</td>
<td>78.4</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>11.6</td>
<td>7.6</td>
</tr>
<tr>
<td>Positive HIV</td>
<td>0.2</td>
<td>0.7</td>
</tr>
</tbody>
</table>
well-matched [8]. The percentages of patients on high-flux dialyser membranes prior to study entry were 59.0% and 61.3% among those randomized to the low- and high-flux groups, respectively, and were 59.5% and 60.9% for patients randomized to the standard- and high-dose groups, respectively. Further, seven pre-specified baseline factors, including years of prior dialysis therapy, were included as covariates in the analyses of mortality and other secondary outcomes in order to increase the power of the comparisons of the dose and flux groups [12]. The RRs for the effects of dose and flux on mortality were also unchanged after addition of dialyser flux at entry to the seven pre-specified covariates. The possibility that treatment effects may be reduced early in follow-up due to the carry-over effects of prior treatments exists for both prevalent and incident dialysis patients and was incorporated into the study’s power calculations [12]. Unpublished data from the HEMO Study, however, indicate no evidence for a smaller treatment effect early in follow-up as compared with later in follow-up.

Implications for other interventions

The HEMO trial was designed to investigate the effects of a higher dialysis dose or the use of high-flux membranes within the context of haemodialysis as it is practised in the US. The absence of significant benefits of these particular interventions does not detract from the possibility that there may be benefits of different interventions, such as the use of very long or nocturnal treatments, more frequent haemodialysis or haemodiafiltration techniques. In fact, the absence of major benefits of the high-dose or high-flux interventions in the context of current dialysis practice gives additional impetus to studies of alternative therapies. These alternative therapies were only in the initial stages of development in the US during the design phase of the HEMO Study and are still not widely available to US patients with end-stage renal disease.

Dose by flux interaction

As indicated in our primary outcomes report [8], there was no significant interaction between the dose and flux interventions for the primary outcome of mortality ($P = 0.30$), which indicates that there is no evidence that the effect of either the dose or flux interventions varied depending on the level of the other intervention. Thus, in accordance with the study’s analysis plan and standard practice for the analysis of factorial designs [27], the primary analysis considered the effect of the dose intervention, combining patients in both flux groups, and, similarly, the effect of the flux intervention, combining patients in both dose groups. Figure 3 shows the post-hoc comparisons of the four individual dose–flux combinations in the study design on mortality. None of the six possible pair-wise comparisons of these four treatment groups approached statistical significance. The RR for the high dose/high flux combination, compared with the standard dose/low
Statistical power

Finally, when interpreting any randomized trial in which the treatment comparisons are not statistically significant, it is important to recall that the null hypothesis of no difference can never be proven completely. Under these circumstances, it is important to consider the statistical power of the trial as reflected by the confidence limits for the treatment effects. Under highly conservative assumptions, the HEMO Study was designed a priori to achieve 80% power to detect a 25% risk reduction for both interventions [12]. The actual power was greater because enrollment and event rates were higher than projected. The HEMO Study dose comparison for mortality, with a risk ratio of 0.96 for the high dose compared with the standard dose, was close to the null hypothesis of no effect and had a P-value of 0.53. The 95% CI extended from 0.84 (indicating a 16% risk reduction) to 1.09 (indicating a 9% risk increase). Analyses of the main secondary composite outcomes gave more narrow confidence limits, with maximum bounds of 9–14% (Table 2). Thus, it was possible to detect smaller treatment effects on the secondary outcomes than on the primary outcome of mortality.

Similar considerations apply to the flux intervention, where the RR of mortality was 0.92, corresponding to a risk reduction of 8% for the high-flux intervention. The 95% CI for this primary outcome was 0.81–1.05, with a P-value of 0.23, far from statistical significance. The 95% CIs for the high-flux intervention on the pre-specified main secondary outcomes are shown in Table 2. These upper limits to the size of possible risk reductions for either the dose and flux intervention were smaller than those reported by a number of recent observational studies [3–6], but do allow for the possibility of beneficial effects that were too small to be detected by the sample size (1846) in the HEMO Study. It is for this reason that the HEMO Study primary findings were interpreted as showing ‘no major overall beneficial effects’ of the interventions, rather than as establishing a complete absence of a benefit [8].

Additionally, these confidence limits refer to the overall treatment effects in analyses that include all patients and do not rule out the possibility that beneficial or detrimental effects may have occurred in certain subgroups of patients. In particular, we have reported that subgroup analyses have suggested possible beneficial effects of the high-dose intervention in women and of the high-flux intervention in patients with a greater number of years of prior dialysis on all-cause mortality [8]. Because of the high risk of ‘false positive’ findings when multiple subgroup analyses are performed, these subgroup results are considered intriguing and hypothesis-generating and require further investigation.

Conclusions

The HEMO Study cohort is representative of the majority of patients who receive chronic haemodialysis therapy in both the US and Europe. The study population was sufficiently heterogeneous for the results to be applicable to patients who receive 3.0–4.5 h of maintenance haemodialysis three times per week. There was clear separation between the high- and standard-dose arms and between the high- and low-flux arms of the study. Analysis of the dose and flux interventions does not show a major benefit of either intervention on mortality or the pre-specified main secondary outcomes. Testing the effects of ultrafiltration, blood pressure control and other potentially important interventions on clinical outcomes were not part of the study design and require separate trials. We do not recommend a decrease in the dose of dialysis for patients who have a dose of dialysis well in excess of the urea spKt/V of 1.20 but are doing well clinically. We also do not condone the maintenance of the dose of dialysis at a level below a spKt/V urea of 1.20, in accordance with US and European guidelines, although the HEMO Study did not specifically address dialysis dose below this level. Finally, the data collected in the HEMO Study should provide important insights into many aspects of care for chronic haemodialysis patients and valuable guidance for the design of new clinical trials, such as haemodialysis or haemofiltration with very high middle molecule clearances and daily or long dialytic therapies.

Conflict of interest statement. None declared.

References


24. RPA/ASN position paper on the frequency of hemodialysis in outpatients, 2002

25. Lowrie EG, Chertow GM, Lew N, Lazarus JM, Owen W. The urea clearance \( \times \) dialysis time \( (K_t) \) as an outcome-based measure of hemodialysis dose. *Kidney Int* 1999; 56:729–737
