Pro: The value of randomized controlled studies in dialysis methods

Ajay K. Singh¹ and Jameela A. Kari²

¹Renal Division, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA and ²King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Correspondence and offprint requests to: Ajay K. Singh; Email: asingh@rics.bwh.harvard.edu

INTRODUCTION

Willis R. Whitney, a chemist, who founded the General Electric Company laboratory once said ‘Necessity is not the mother of invention. Knowledge and experiment are its parents’ [1]. Twardowski and Misra reject the primacy of experimentation. They wish us to abandon the necessity to demonstrate by rigorous randomized trials that longer duration and/or more frequent is superior to intermittent dialysis. They argue that randomized controlled studies (RCTs) are inadequate as a research methodology in evaluating a dialysis method citing limitations with each of the four RCTs—the National Cooperative Dialysis Study [2], the HEMO study [3], the Frequent Hemodialysis Network (FHN) study [4] and the FHN Nocturnal trial [5]. Twardowski and Misra suggest that ‘all progress in dialysis methods was made in research presented in case reports, case control studies and other observational studies. ‘and that’ four RCTs in hemodialysis did not provide any useful data …. observational studies allow clinical research to represent the full breadth of treated patients and offer tremendous power…’.

The approach that Twardowski and Misra is questionable. We can hardly blame rigorous experimentation because a handful of studies that have been performed so far do not support a belief, however strongly held the belief is. After all any given study may have asked the wrong question, or inadvertently reached the wrong conclusion, or been null because of lack of power, or may have problems with study quality (for example, imbalances between the randomized groups or an excessive drop-out rate). Surely, this does not mean that the whole methodological approach embodied in RCTs is wrong? RCTs have value in the investigation of interventions like the choice of dialysis methods, just as they are important to other forms of human investigation.

Twardowski and Misra acknowledge that RCTs represent the most rigorous method to determine whether a cause-and-effect relation exists between any treatment and an outcome. However, they understate the true advantage of RCTs over observational studies. In the study of interventions, RCTs are superior to observational studies because RCTs are based on random allocation of subjects to two or more intervention groups. The possibility of a systematic error is reduced because patient-related confounding factors are balanced across different interventions. While skilled biostatisticians and epidemiologists might try to attenuate the effect of confounding, these attempts are generally imperfect and residual confounding almost always exists. Besides, in RCTs, the intervention groups are treated identically, except for the experimental treatment. Subjects are analyzed within the group to which they were allocated, irrespective of whether they experience the intended intervention or not (intention to treat analysis) further reducing the chance of bias.

Knowledge and experimentation are the founding principles of evidence-based medicine. The randomized trial represents the most rigorous method to get to the truth of whether an intervention causes an outcome. We should embrace RCTs not abandon them. RCTs are essential, because as Chertow has written [6] ‘Wishing Don’t Make it So’.

Evaluating non-randomized studies on dialysis methods: all that glitters is not gold

Twardowski and Misra emphasize the importance of non-randomized studies in determining the optimal duration and frequency of dialysis. These non-randomized studies cover a broad spectrum: case reports, case series, historical controlled trials and registry data. Collectively, the observational experience has laid the foundation for performing randomized trials but is not a substitute.

Consider one of the studies cited by Twardowski and Misra in their article to support the superiority of more frequent and/or longer dialysis. This was a study performed by Twardowski between 1969 and 1973 and exemplified a ‘before and after design’ [7]. Fourteen dialysis patients were enrolled and after a
mean follow-up of ~6 months, improvements in a variety of laboratory and two clinical parameters are reported. However, while this study was pioneering it also had some important flaws. For example, Twardowski provides no information on whether the subjects enrolled in this study were incident or prevalent patients on dialysis, or how patients were selected, or whether the study recruited patients from a single or multiple centers, or whether all of the patients adhered to the study treatments. Indeed, inadequate information on patient demographics is provided. The interventions were also rather nebulous: Twardowski writes that the frequency of dialysis was increased one weekly, from 2 to 3 and from 3 to 4 dialysis treatments each week, and the ‘increase in dialysis duration without an increase in frequency of dialysis was on average 17.5% Twardowski’s study does not provide proof of causality that more dialysis is better. Likewise, Charra’s experience in Tassin [8, 9] or the Canadian experience [10], among many reported experiences, represents additional foundation stones in building the case for more rigorous investigation. Observational studies are of great value because like Twardowski’s study they provide the basis for hypothesis-testing in a randomized trial. Furthermore, as Jan Vandenbroucke has pointed out that sometimes RCTs are not feasible. Indeed, for some rarer adverse effects observational or case-control studies might provide additional information on an RCT because an RCT is either too small or has insufficient follow-up [11, 12].

Twardowski and Misra also cite data from both European and Japanese registries that ‘indicated much better survival in association with longer weekly time on dialysis in Europe and Japan. They state ‘during a conference in Dallas in 1989, several presentations indicated that the difference in mortality was not related to the patient mix or other factors but shorter dialysis time in the US.’ (Also 13) Robinson and Port revisit the conclusions from this conference in a paper in 2009 by [14]. They raise the possibility that higher survival in Japanese patients may have alternative explanations. The Japanese have a very low rate of transplantation, so that dialysis patients in Japan are healthier than those in the USA. Furthermore, they raise the possibility that differences in the dialysis mortality may also reflect differences in background mortality in the general population. Finally, case-mix adjustments when comparing data from different registries may be imperfect. For example, it is possible that US patients may be sicker than Japanese or European patients because patients with shorter life expectancy could have received shorter dialysis sessions. In essence, one cannot rule out the possibility that an association might be caused by a third factor linked to both the intervention and outcome.

Randomized trials do not always provide the answer, but perfection should not be the enemy of good

In a parody of the obsession some of us have with RCTs, Smith and Pell [15] use the example of parachutes to question whether RCTs are always necessary. Parachutes have never been tested in an RCT and a trial would be unethical. Common sense should prevail and they make the compelling argument that RCTs are neither always necessary or always provide an answer. Aside from the circumstance when an RCT is not ethically possible, a randomized trial may be difficult to complete because of difficulties with randomization or recruitment. Consider the FHN Nocturnal study [5] as a prime example of difficulty with recruitment.

In the FHN Nocturnal trial, 87 patients were randomized to either three times per week conventional hemodialysis or to nocturnal hemodialysis six times per week, all with single-use high-flux dialyzers. The investigators did not find any significant effect of nocturnal hemodialysis for either of the two co-primary outcomes (death or left ventricular mass (measured by MRI) or of death or RAND Physical Health Composite. Twardowski and Misra argue that the trial represents a type 2 error, i.e. the investigators failed to reject a false null hypothesis. The null hypothesis in the FHN nocturnal trial was that there is no difference in outcomes between frequent nocturnal home hemodialysis six times per week versus conventional three times per week hemodialysis. Whether the null hypothesis is truly false or not remains to be seen. That aside, the trial did have some major limitations that the authors acknowledge. The most important were the relatively small sample size and the lower adherence to the dialysis prescription in the frequent nocturnal arm, both of which reduced the power of the study. As well, only 87.3% of the patients who were randomized completed 12 months of follow-up and had measures of both co-primary outcomes. Moreover, ~25% of patients in the frequent nocturnal arm performed less than five hemodialysis treatments per week. Twardowski and Misra use the FHN Nocturnal trial to bolster their dissatisfaction with RCTs.

Twardowski and Misra criticize both the NCDS [2] and HEMO [3] studies. Their unhappiness with the NCDS study centered on the decision by the investigators to not consider the duration of the dialysis treatment as a part of their research question. And, the HEMO study because it asked the wrong question. However, these studies had other limitations that are extensively discussed elsewhere [16, 17]. Chief among these was the highly selected patient sample in NCDS: diabetic, elderly and hypertensive patients were excluded. This criticism with external validity could also be directed at the HEMO study. In the HEMO study, there were a disproportionately higher number of African Americans and less heavy patients than the US dialysis population. This was because subjects were recruited mostly from urban centers (63% of HEMO patients were African-American versus 41% in the US general dialysis population) and because the heaviest patients were excluded to assure that the high-dose goal could be achieved (~3% of patients randomized in the HEMO trial exceeded 100 kg, when compared with an estimated 10% of hemodialysis patients in the US). Both NCDS and HEMO provided key information. NCDS was valuable in establishing the minimum amount of dialysis dose necessary for survival and the HEMO study taught us that increasing dose by blood flow or by using a high flux dialyzer is insufficient to improve outcomes. Just because NCDS and HEMO studies did not answer all of the questions related to dialysis dose does not mean that they were not useful.

Advocating the superiority of a treatment based on an anecdote and experience rather than rigorous trials can be
Conflicts of interest statement

None declared.

References

7. Twardowski Z. Effect of long-term increase in the frequency and/or prolongation of dialysis duration on certain clinical manifestations and results of laboratory investigations in patients with chronic renal failure. Hemodial Int 2004; 8: 30–38
11. Vandebroucke JP. When are observational studies as credible as randomised trials? Lancet 2004; 363: 1728–1731

Opponent’s comments

Singh accuses us that we reject the primacy of experimentation. On the contrary, we support all experimentation for hypothesis testing as long as it is appropriately planned and conducted. We believe that we have put forth a strong argument to prove that all randomized controlled trials (RCTs) in dialysis failed owing to either inadequate planning or due to misplaced belief that they would be easy to conduct and complete. All ITT (Traumatic Intercocular Test) results were the outcome of meticulously planned and carefully conducted experiments. A careful review of the articles cited in Table 2 of our paper highlights the rigorous experimentation that led to the conclusions now forming the basis of our current knowledge about dialysis. The before-and-after studies constitute an established way of experimental design. What is in fact wrong-headed is the thought that only RCTs can bring forth the correct answer to all questions. The fact that all the RCTs in dialysis failed in some way or other is adequate proof that RCTs are not only fallible but also not applicable universally. It underscores our point that RCTs, as a means to test dialysis methods have failed.

Singh states that “RCTs are superior to observational studies because in RCTs intention to treat (ITT) analysis reduces the chance of bias. We cannot agree with this statement. Intention to treat analysis is the major weakness of RCTs. How can one deny such a bias if a patient allocated to intervention group undergoes a treatment meant for the control group but such patient’s data are included in the intervention group? This was one of the major weaknesses of the FHN Nocturnal trial [1]. Patients in the more frequent arm switched to less frequent and shorter HD (28%); those in the control arm (10%) switched to longer and more frequent HD. Obviously, calculation according to the ITT principle does not make any sense in such situation.

Singh criticizes a before-and-after study performed by Twardowski about 40 years ago. He states that this study did not provide sufficient information to judge the results. Singh

value of RCTs

835
did not review the study carefully. This study was reported in three parts, two in 1974 [2 and 3] and the third in 1975 [4] later republished in 2004 [4]. Only the last paper was cited in our review. The reference to the first paper was given on page 32 of the paper published in 1975 and on page 31 in the paper republished in 2004. Detailed data on the patients including their sex, ages, primary renal diseases, start of dialysis, uremic symptoms etc. are provided in the first two papers. All patients were dialyzed on Ultra Flo 145 dialyzers with dialyzer blood flow of 200 ml/min. All patients were observed from the start of treatment and throughout the therapy up to 4 years. Thus this group was a combination of incident and prevalent patients. All patients were recruited from one center, there were no withdrawals and all patients were compliant as was typical of patients treated 40 years ago. The patients were gradually losing residual renal function. Therefore the improvement after change to longer or more frequent dialyses was definitely related to dialysis and not to preserved/improving function of native kidneys or changes in dialyzers or dialyzer blood flow.

Singh criticizes us that we considered the FHN Nocturnal trial, as “representing a type 2 error, i.e., that the investigators failed to reject a false null hypothesis.” He argues that “The null hypothesis in the FHN nocturnal trial was that there is difference in outcomes between frequent nocturnal home hemodialysis six times per week versus conventional three times per week hemodialysis.” This argument is absolutely inconsistent with the understanding of the null hypothesis. It is common knowledge that the null hypothesis states that if the difference between the means of two groups is small then the two groups are not different and they belong to the same population. Simplistically stated, a researcher has to ‘nullify’ the null hypothesis to prove that the difference between the two groups is significant and two groups belong to two populations.

Singh blames observational studies as creating bias which makes “clinicians unwilling to enroll subjects into an “experiment” when they believe it could be harmful to their patient.” The problem of recruiting patients for RCT is not only related to the fact that doctors tell their patients that one method is better, but also to other patient related factors that Singh fails to mention. For instance, as we understand, in the FHN study, patients already on NxStage machines did not want to stop this treatment since they were satisfied with the results on the NxStage machine; other patients did not want to travel to the center for dialyzing so frequently. Patients on AKSYS PHD® machine felt so well that they did not want any other treatment.

We did “before and after study” on PHD® for FDA in four patients. Three patients were anuric and very experienced in HD (treated previously for many years). The difference in feeling on daily dialysis was so great that these patients never wanted to be dialyzed by any other dialysis method. One patient, who was not anuric, did not notice significant difference and was not interested in frequent dialysis; his chemistries showed substantial improvement but they meant nothing to him. An example of this non anuric patient tells us a lot about who benefits most from more frequent and longer dialysis. If for FHN nocturnal study, patients with urine output over 1000 ml/day were recruited, it is not surprising that they did not benefit too much or they decided to switch to routine, thrice weekly dialysis. We suspect that those 28% of patients may have had high urine output.

We have only criticized the RCTs in dialysis that have been performed thus far. We have absolutely nothing against properly conducted RCTs, if there are such to be planned and conducted in the future. The study of shorter and longer dialysis may be also performed in RCT. As a matter of fact we planned such a study which we called: “Kt/V requiem”. The comparison was supposed to be between dialysis of approximately 190 min duration and dial flow of 350-500 ml/min to achieve spKt/V urea of 1.3 with a dialysis of 4.5-6 hours duration and a blood flow of 200 ml/min to provide spKt/V urea of 1.3. We were unable to do this study for various reasons. However, such a RCT would really be the first worthwhile RCT in dialysis methods. We are confident that, if such a study is done on patients with negligible urine output (<100 ml/24hr), longer dialysis with lower blood flow will be associated with better control of blood pressure, better control of anemia with lower erythropoietin use, better nutrition, better control of phosphorus, no ‘hangover’ after dialysis, and better subjective feeling of the patients.

Zbylut J. Twardowski, MD, PhD, FACP
Professor Emeritus of Medicine
Madhukar Misra, MD, FRCP (UK)
Professor of Medicine
Division of Nephrology
Department of Medicine
University of Missouri
5 Hospital Drive, Room CE 421
Columbia, MO 65212
E-mail: twardowskiz@health.missouri.edu

References