The efficacy of loop diuretics in acute renal failure: Assessment using Bayesian evidence synthesis techniques

Sriram Sampath, MD (Gen Med); John L. Moran, FRACP, FJFICM, MD; Petra L. Graham, PhD; Sue Rockliff, BA, Grad Dip Lib; Andrew D. Bersten, MD, FANZCA, FJFICM; Keith R. Abrams, PhD

Objective: **To quantify the therapeutic efficacy of loop diuretics in acute renal failure using Bayesian evidence synthesis, because despite widespread use, the role of diuretics is controversial.**

Data Source: **Randomized controlled trials or nonrandomized studies, 1966 to January 2007, identified from MEDLINE and EMBASE databases and manual bibliographic search.**

Study Selection: **Studies with assessable predefined end points, exclusive of those pertaining to acute renal failure prophylaxis or chronic renal failure.**

Data Extraction: **Data extraction was performed jointly by the first two authors; independent study assessment was via standard checklist, unblinded.**

Data Synthesis: **The primary outcome was mortality; secondary outcomes were time to renal function normalization and total number of dialyses. Bayesian hierarchical random effects estimates of treatment effects were determined as risk ratio for mortality, incidence rate ratio for dialysis number, and mean difference for continuous measures. Bayesian outcome probabilities were calculated as probability (*P*) that risk ratio or incidence rate ratio of loop diuretics >1 and probability that mean difference >0. Five randomized controlled trials and eight nonrandomized studies were identified. Loop diuretics were not associated with decreased mortality in either randomized controlled trials or nonrandomized studies: overall risk ratio 1.10; 95% credible interval 0.85, 1.42; *P* (risk ratio >1) = 83.8%. The oliguric period was decreased by loop diuretics: overall mean difference −7.70 days; 95% credible interval −12.51, −2.08; *P* (mean difference >0) = 0.7%. Although the dialysis rate credible interval, loop diuretics vs. control, spanned unity (incidence rate ratio 0.71; 95% credible interval 0.47, 1.06), the probability that the incidence rate ratio exceeded unity indicated a substantial benefit: *P* (incidence rate ratio >1 = 4.1%. Uremic duration was not substantially different, loop diuretics vs. control: overall mean difference −1.54 days; 95% credible interval −5.62, 2.46; *P* [mean difference >0] = 17.8%).**

Conclusions: **Loop diuretics were not associated with improved survival benefit in acute renal failure, despite reduction in oliguric period and high probability of a significant reduction in dialysis numbers. Further studies to clarify this dichotomy appear mandated. (Crit Care Med 2007; 35:2516–2524)**

Key Words: sodium potassium chloride symporter inhibitors; kidney failure, acute; meta-analysis; Bayes theorem; review; evidence-based medicine

Acute renal failure (ARF) is associated with a mortality rate of 45% to 70% (1). Loop diuretics have been used in the nondialytic management of ARF for >30 yrs, but appraisals (2, 3) of the therapeutic options in ARF have not supported their routine use, and no decrease in patient mortality has been shown. A more recent (2006) evaluation (4), investigating the role of frusemide in the prevention or treatment of ARF, noted the paucity of randomized controlled trials (RCTs) studying the effect of frusemide in established ARF and concluded that frusemide was "not associated with any significant clinical benefits." The lack of a well-accepted definition of ARF (3, 5) also confounds the interpretation of these studies.

The lack of evidence from RCTs has increasingly prompted analysis of evidence from observational data and a "broad perspective on meta-analysis" (6) may be of advantage as hypothesis generating. Similarly, Norris and Atkins (7) suggested that the difficulty of both conducting RCTs and applying the obtained results to the "general" population mandates the inclusion of nonrandomized studies in systematic reviews. Our purpose in this study was a) to perform an extended search of the literature to identify RCTs and nonrandomized studies of loop diuretic use in ARF; and b) to assess the efficacy of this intervention in ARF by incorporating evidence from both types of studies into a final pooled effect estimate using Bayesian evidence synthesis techniques (8, 9). Bayesian evidence synthesis is a cutting-edge meta-analytic technique allowing combination of estimates from different sources, such as study types. The hierarchical approach of evidence synthesis and the particular assumptions underlying it allow each study to borrow strength from the other studies result-

From the Intensive Care Unit, St John's Medical College Hospital, Bangalore, India (SS); Department of Intensive Care Medicine, The Queen Elizabeth Hospital, Woodville, SA, Australia (JLM); Department of Statistics, Division of Economic and Financial Studies, Macquarie University, Sydney NSW, Australia (PLG); Department of Library Services, The Queen Elizabeth Hospital, Woodville, SA, Australia (SR); Department of Critical Care Medicine, Flinders Medical Centre and School of Medicine, Flinders University, South Australia (ADB); and Centre for Biostatistics & Genetic Epidemiology, Department of Health Sciences, University of Leicester, Leicester, UK (KRA).

Supported by departmental funds.

The authors have not disclosed any potential conflicts of interest.

Address requests for reprints to: John L. Moran, FRACP, FJFICM, MD, Department of Intensive Care Medicine, The Queen Elizabeth Hospital, 28 Woodville Road, Woodville, SA 5011, Australia.

DOI: 10.1097/01.CCM.0000284503.88148.6F

ing in shrinkage of each estimate toward the overall mean. Standard meta-analytic methods are typically only suitable for pooling evidence from similar studies because of the difficultly incorporating heterogeneity between the studies and study types into overall effect estimates. Bayesian methods have no such difficulty and aim to obtain the best estimate given all of the evidence, effectively incorporating all of the heterogeneity. Unlike standard methods, it is also easy to obtain quantities of interest, such as the probability that the true risk ratio of mortality under diuretics is greater than the null (≡1). The advantages of employing Bayesian methods in such a setting were set forth by Spiegelhalter et al (8).

METHODS

Definitions. A recent systematic review of 28 studies of postoperative ARF found that no two studies used the same criteria; the operative definition used in this study was "a syndrome of diverse origins characterized by the abrupt reduction of renal function" (10). A consensus definition of ARF has only recently been proposed (11), and in the absence of a definition of ARF that would potentially encompass a 30-yr history of clinical trials of varying quality, ARF was defined inclusively as an acute worsening of renal function determined by either blood or urinary variables (12).

Data Sources. A preliminary search using OVID (limited to English language) was carried out for the period 1966 to January 2007 using the search terms "Diuretics (explode) and Kidney failure, acute (explode)" in the MEDLINE and EMBASE databases. These search terms were then combined in a more detailed search (Appendix 1) using the "highly sensitive search strategy" described by Robinson and Dickersin (13). A similar strategy was used in the CINAHL, Cochrane database of systemic reviews, and ACP journal club databases using the search software OVID. In addition, the same technique was used in PubMed to see if additional articles could be identified. MEDLINE from 1958 to 1965 was also searched for relevant articles. The Australian Digital Thesis program and parts of the National Digital Thesis and Dissertations were searched, and the bibliographies of the abstracted review articles and trials were reviewed for suitable publications.

Study Selection and Data Extraction. All RCTs and nonrandomized studies in humans were included provided that the primary therapeutic end point, (hospital) mortality, was reported. The selection and inclusion of nonrandomized studies were consistent with the principles outlined by the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) Group in 2000 (14). Appropriate secondary end point specification was somewhat difficult due to the lack of comparator studies, and the following were used: Those likely to reflect a beneficial economic outcome (3, 15, 16)

1. Time taken (in days) for return of spontaneous urine output >1500 mL
2. Total number of dialyses needed.

And those likely to reflect a physiologically plausible outcome (17–19)

3. Time taken (in days) to spontaneously normalize plasma creatinine to <2.2 mg/dL (199 μmol/L) or urea <40 mg% (14 mmol/L)

Studies of chronic renal failure (CRF), or where the patient population was predominantly CRF (however defined) or pediatric, and studies in which diuretics were administered in a prophylactic manner were excluded. Data extraction was done jointly by two authors (SS and JLM). In studies where subgroups of patients were reported, the results were pooled for analysis. As there were both randomized and nonrandomized studies, the checklist suggested by Downs and Black (20) (score ranging from 0 [worst] to 31) was used for assessment of study quality; the studies were independently assessed, unblinded, by two authors (SS and JLM).

Data Synthesis. For the purpose of both presentation and analysis, trials were categorized as either RCTs or nonrandomized after Norris and Atkins (7), so that heterogeneity arising from different study types in the meta-

Table 1. Characteristics of studies included for meta-analysis

Lead Author (Reference No.)	Country Year	Study Type	Study Quality	Etiology	Mean Age	Gender Ratio, M/F	Study Features	Definition of ARF
Beroniade (26)	Rumania 1969	NR	3	Mixed	34	ND	Escalating drug dose	Not described
Cantarovich (28)	Argentina 1971	RCT	7	Mixed	ND	ND	Two subgroups in treatment arms combined for analysis	Urine output <400 mL/24 hrs
Cantarovich (29)	Argentina 1973	NR	5	Mixed	ND	ND	High dose of frusemide	Mannitol test
Chandra (31)	India 1975	NR	12	Mixed	41	ND	Adult and pediatric study, only adult data studied	Urine output <400 mL/24 hrs
Kleinknecht (32)	France 1976	RCT	11	Mixed	ND	31/35	Escalating up to 1200 mg/day of frusemide	Criteria defined
Minuth (35)	US 1976	NR	5	Mixed	55	76/28		Urine/blood variables
Borirakchanyavat (27)	Thailand 1978	NR	7	Leptospirosis	41	13/01	No dialysis required in any patient	Undefined
Brown (25)	UK 1981	RCT	14	Mixed	52	31/25	Initial 1 g of frusemide to both groups	Criteria defined
Lumlertgul (33)	Thailand 1989	NR	10	Malaria	24	8/0	Only two of five subgroups analyzable	Urinary indexes
Shilliday (36)	UK 1997	RCT	22	Mixed	58	42/34	Dopamine and mannitol given in both groups	Creatinine >180 mmol/L
Mehta (34)	US 2002	NR	18	Mixed	5	72/28	Preexisting renal dysfunction in 25% of patients	Criteria defined
Uchino (37)	Australia 2004	NR	17	Mixed	67	64/36	Preexisting renal dysfunction in 29% of patients	Criteria defined
Cantarovich (30)	France 2004	RCT	24	Mixed	58	67/33	Initial frusemide given to both groups	Criteria defined

M/F, male/female; ARF, acute renal failure; NR, nonrandomized study; ND, not described; RCT, randomized controlled trial; Mixed, mixed etiologies of ARF.

analysis could be incorporated. To estimate an overall treatment effect, the treatment effects and standard errors were initially estimated for each category of study, and then the study category estimates were combined into an overall estimate. This is known in the Bayesian setting as evidence synthesis. The methods described by Prevost et al. (21) and Spiegelhalter et al. (8) were employed with priors chosen to reflect the plausible range of variable values. The methods and priors are described in detail in Appendix 2.

Risk ratios (RR) were used for the binary data, mean differences (MD) for the continuous data, and incidence rate ratios (IRR) for the count data. Bayesian methods also allowed estimation of the probability of an outcome from the posterior distribution. For each of these meta-analyses we calculated the probability that the RR or IRR of loop diuretics was >1 and the probability that the MD was >0. Results are presented as the median estimate with 95% credible intervals (CI). Heterogeneity is presented as the variance between or within study categories. As a sensitivity analysis, the following covariates—quality score, average age of study cohort, proportion of males, control arm risk, and study completion year—were included (one at a time) as a common effect in the evidence synthesis model to determine whether each acted as a treatment effect modifier. While additional variables, such as baseline creatinine and time to dialysis, may have modified the effects of the treatment, only the specified covariates were available across all of the included studies. Control arm risk was used as a surrogate for patient severity of illness (22). Year of termination of the particular study was used, rather than the year of publication, to avoid publication lead-time bias. Publication bias was of interest but could not be formally assessed as there were fewer than ten studies in each of the RCT and nonrandomized study groups (23).

RESULTS

A preliminary search identified 1,690 articles, 558 in MEDLINE and 1,132 in EMBASE. The highly sensitive search strategy adopted in PubMed format and searches in other databases did not reveal any additional studies. Thirty-six articles were identified as clinical trials, of which 24 were excluded; 17 of these articles were case reports or noncomparative studies, six of these trials involved the prophylaxis of ARF, and one study (24) was a preliminary report whose results were included in a later trial (25). A further trial (26) was identified from a hand search of review articles. These 13 (25–37) satisfied criteria for inclusion and analysis and were divided into RCTs and nonrandomized studies (Table 1).

The studies spanned >35 yrs from 1969 to 2004. Total patient number was 3,111; 2,520 patients were reported in eight nonrandomized studies (26, 27, 29, 31, 33–35, 37) and 591 patients in 5 RCTs (25, 28, 30, 32, 36). The mean (SD) age was 48 (15) years, female proportion 30%, and mean mortality rate 54%. Mortality was the only outcome reported in all 13 studies and was described as hospital mortality in four studies (31, 34, 35, 37) or mortality at 21 days (36) and 30 days (30) after enrollment in 2; no deaths were recorded in two studies (27, 33). Overall median (range) quality score was 11 (3–24) and, for RCTs and nonrandomized groups, median score was 14 (7–24) and 9 (3–18), respectively. Missing patient descriptive variables were also noted in a number of studies (Table 1).

The etiology of ARF in the two trials (27, 33) from Thailand was malaria and leptospirosis, while in the other trials there were varying etiologies (Table 1). There was no single consistent definition of ARF in the trials analyzed, and

Table 2. Patient and therapy characteristics in trials studied

Lead Author (Reference No.)	Total No. of Patients	Control Deaths, No. (%)	Treatment Deaths, No. (%)	Control Survivors, No.	Treatment Survivors, No.	Dosage of Diuretic	Delivery Technique	Duration of Therapy	Deafness Incidence
Beroniade (26)	24	6 (50)	3 (25)	6	9	Frusemide 60–480 mg	Not described	Until onset of diuresis	ND
Cantarovich (28)	47	7 (54)	15 (44)	6	19	Frusemide 600–3200 mg fixed/geometric progression	IV infusion 30 mins to 10 hrs	Until onset of diuresis	ND
Cantarovich (29)	58	11 (58)	18 (46)	8	21	Frusemide 2000 mg/day	IV infusion	Until onset of diuresis	ND
Chandra (31)	17	3 (60)	5 (42)	2	7	Frusemide 200–2000 mg/day	IV infusion	Until onset of diuresis	2/16
Kleinknecht (32)	66	12 (36)	13 (39)	2	20	Frusemide 150–1200 mg	Intermittent IV infusion	Until onset of diuresis	ND
Minuth (35)	104	12 (48)	47 (59)	13	22	Frusemide 40–500 mg	Intravenous	Undefined	ND
Borirakchanyavat (27)	14	0	0	8	6	Frusemide 500 mg/day	IV	7–8 days	ND
Brown (25)	56	16 (57)	18 (64)	12	10	Frusemide 2 mg/min or 1 g tid	IV or oral	Defined biochemical/urinary criteria	2/56
Lumlertgul (33)	8	0	0	4	4	Frusemide 200 mg 6 hourly	IV	4 days	ND
Shilliday (36)	92	15 (50)	41 (68)	15	20	Frusemide or torasemide 3 mg/kg	IV bolus 6 hourly	21 days	1/92
Mehta (34)	552	110 (48)	184 (56)	116	142	Frusemide (median 80 mg)	Not known	Undefined	ND
Uchino (37)	1743	357 (57)	697 (62)	269	420	Frusemide (mean 240 mg/24 hrs)	Not known	Undefined	ND
Cantarovich (30)	330	50 (30)	59 (35)	114	107	Frusemide 25–35 mg/kg/day	IV or oral	Until renal recovery	3/166

ND, not defined; IV; intravenous.

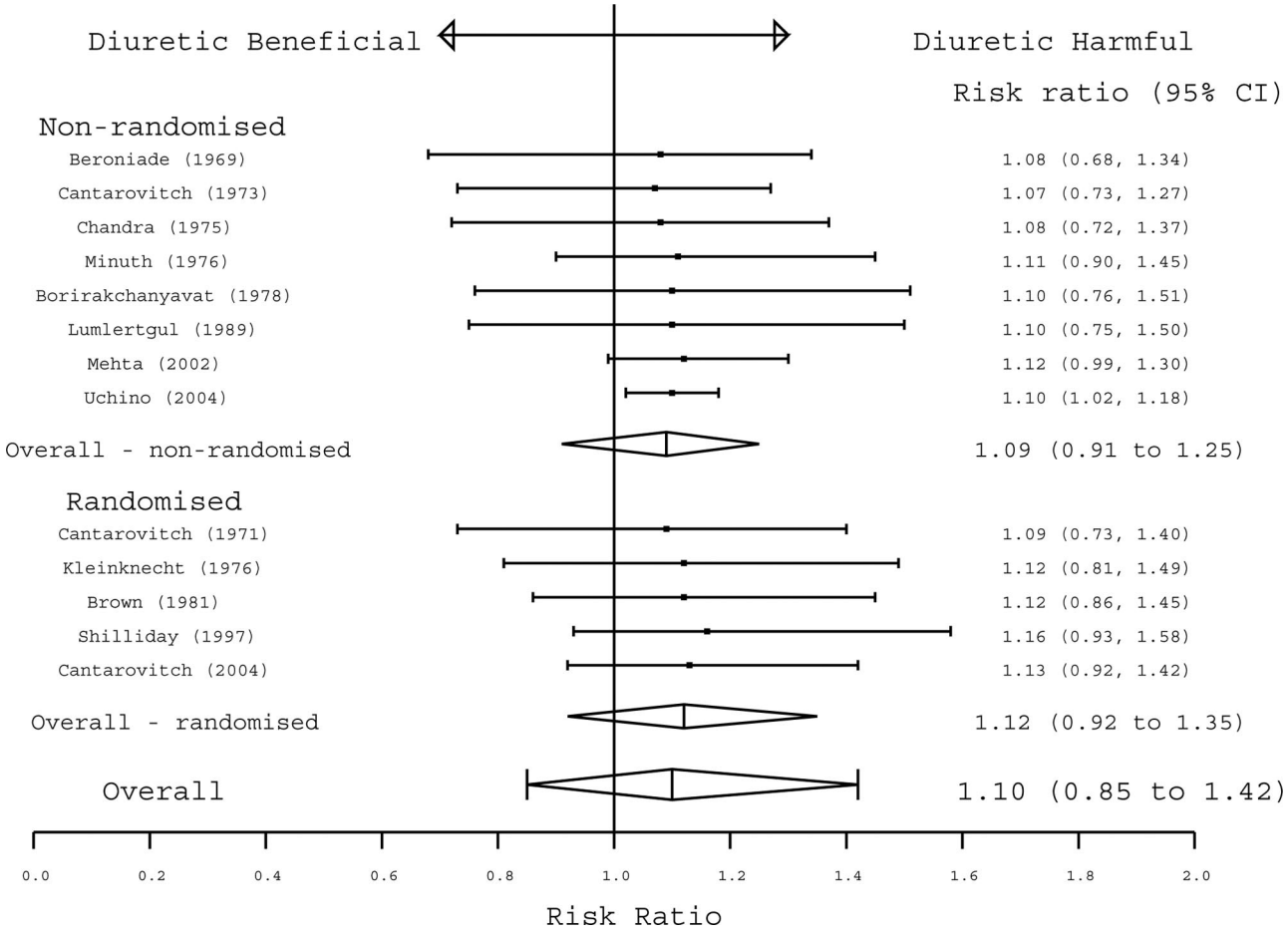

Figure 1. Forest plot showing effect of randomized and nonrandomized studies on mortality treatment effect as risk ratio. *Small solid squares*, study estimates; *vertically capped horizontal lines*, 95% credible intervals (*CI*); *vertical lines within vertically capped diamond-shaped boxes*, subgroup and overall point estimates and 95% CI; *vertical straight line*, the null effect.

in two trials (26, 27) ARF was not formally defined. Exclusion criteria were defined in only six studies (25, 30, 32, 34, 36, 37), and there were no consistent criteria for patient exclusion. The duration, dosage, and technique of loop diuretic administration in the studies were variable (Table 2). In one trial (28) there were two subgroups with different techniques of diuretic administration; these subgroups were combined together for analysis. Frusemide was the loop diuretic in all RCTs except one (36) in which frusemide and torasemide were used; for the purposes of analysis these subgroups were combined (Table 2). Of the five RCTs, two were assessed as formally placebo-controlled (28, 32) and in three (25, 30, 36) additional active renal therapies were initially used in both control and treatment arms. In the two recent large nonrandomized studies (34, 37), thiazide diuretics were used in addition to loop diuretics, and the combined results were used in analysis. In both these studies, preexisting non-dialysis-dependent renal dysfunction was reported equally in treatment and control arms (Table 1).

Influence of Loop Diuretics on Mortality. Thirteen studies had information on mortality (Fig. 1). While the credible interval contained one, both group and overall point estimates indicated an adverse effect of loop diuretics (overall RR 1.10; 95% CI 0.85, 1.42). Furthermore, there was an 83.8% probability that the RR >1 (i.e., a high probability that loop diuretics were associated with harm). Heterogeneity between and within the study types was small (between types 0.01, nonrandomized 0.01, RCTs 0.01).

Influence of Loop Diuretics on the Time Taken to Normalize Creatinine/Urea in ARF. In the eight assessable studies, the time taken (in days) to normalize creatinine or urea was decreased for both groups and overall; however, the difference was not significant (overall pooled MD -1.54 days; 95% CI -5.62, 2.46) (Fig. 2). There was a 17.8% probability that loop diuretics were associated with longer time to normalization of creatinine. Little heterogeneity between or within the study types (between types 2.10, nonrandomized studies 6.22, RCTs 1.60) was observed.

Influence of Loop Diuretics on the Time Taken to Diuresis of ≥1500 mL/day in ARF. The use of loop diuretics significantly decreased the oliguric period in six assessable trials for both groups and overall (overall MD -7.70 days; 95% CI -12.51, -2.08). There was a 99.3% probability that diuretics were associated with shorter times to diuresis compared with control. Heterogeneity of treatment effect was observed within the RCT group (25.17) and to a lesser extent within the nonrandomized group (5.69) and overall (2.54) (Fig. 3).

Influence of Loop Diuretics on the Number of Dialyses. The credible interval for the incidence rate of dialyses needed suggested no significant decrease with

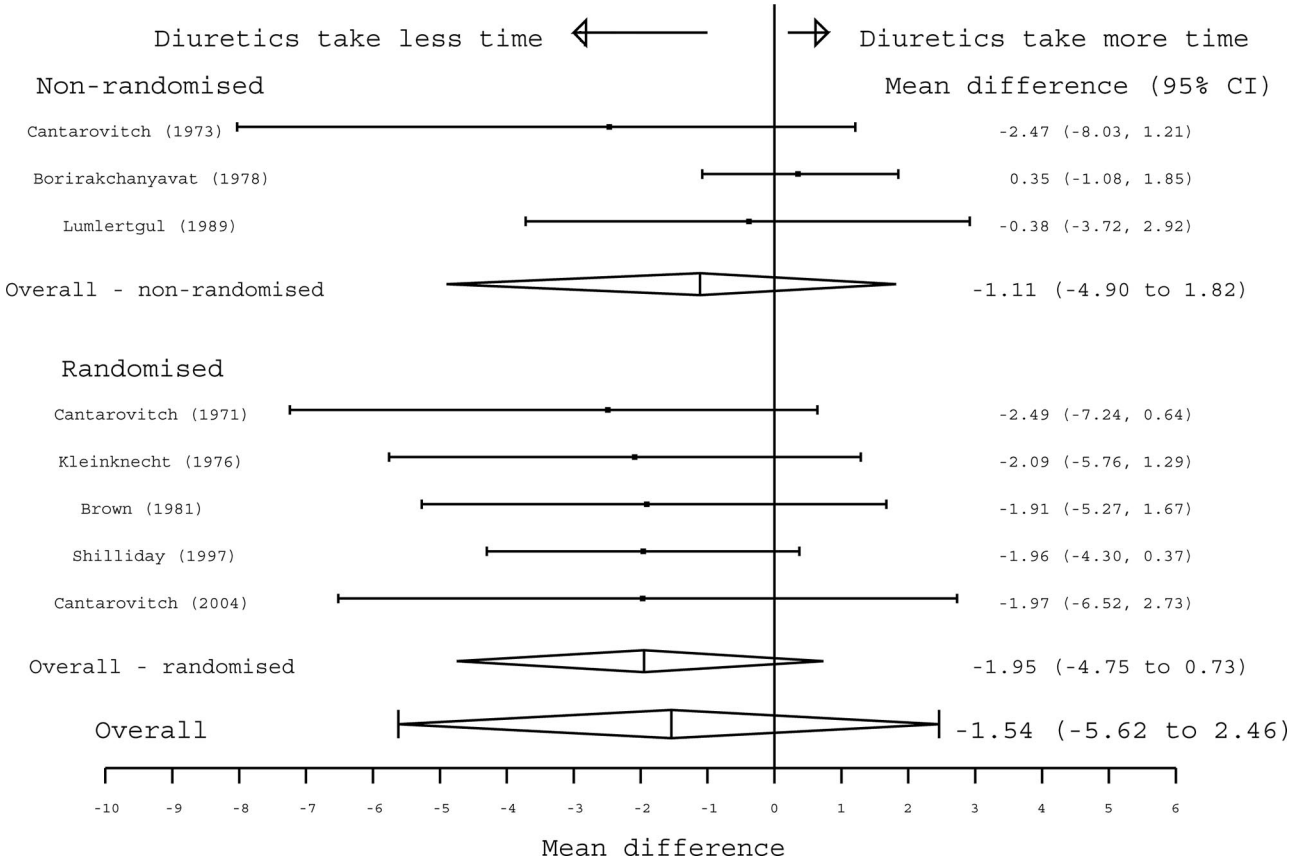

Figure 2. Forest plot showing effect of randomized and nonrandomized studies on time taken to normalize creatinine/urea, as mean difference (days). *Small solid squares*, study estimates; *vertically capped horizontal lines*, 95% credible intervals (*CI*); *vertical lines within vertically capped diamond-shaped boxes*, subgroup and overall point estimates and 95% CI; *vertical straight line*, the null effect.

the use of loop diuretics trials in the six assessable trials overall (IRR 0.71; 95% CI 0.47, 1.06). However, the overall probability that the incidence rate of dialyses was decreased under loop diuretics compared with placebo was 95.9%, that is, $1 - P(\text{IRR} > 1) = 95.9\%$. Little heterogeneity was noted overall (0.02) or for the nonrandomized studies (0.07) or RCTs (0.07) (Fig. 4).

Sensitivity Analysis: Influence of Quality Score, Average Age, Proportion of Males, Control Arm Risk, and Year of Study Termination on Estimates. No substantial modifying effect of average patient age, proportion of males, or underlying control arm risk on RR of mortality was seen using each covariate as a common effect in the evidence synthesis model (Table 3). While 95% credible intervals on the coefficients for quality score and study termination date included 0, each had ≥90% probability that the slope was >0.

DISCUSSION

The paucity of RCTs of loop diuretics in ARF (38) stands in contradistinction to their undoubted widespread use, as attested to by recent nonrandomized studies (34, 37) and editorial reviews (39, 40). In terms of the Bayesian methods used in this meta-analysis, the probability of adverse mortality effect (RR >1) of loop diuretics was substantial at 84%.

This appears somewhat unexpected, given the theoretical basis for the use of loop diuretics in ARF (39). As proposed by Majumdar and Kjellstrand (16), human ARF is not due to a single physiologic insult and there appear to be multiple occasions in the course of ARF when loop diuretics could have an advantageous effect (38). The potential detrimental effects of diuretic use in ARF were first emphasized by the nonrandomized study of Mehta et al. (34) and were related to direct adverse effects on renal physiology, exacerbated by excessive preload reduction (38, 40) and/or indirect effects consequent on underestimation of severity of illness with delay in recognition of ARF and institution of dialytic therapy (30, 34, 39). With respect to formally addressing the adverse effects of diuretics postulated by Mehta et al. (34), no effect of underlying risk (as a surrogate for severity of illness) on treatment effect was demonstrated and no assessment could be made, via sensitivity analysis, of the effect of either baseline creatinine or time to first dialysis from ARF due to selective bias in reporting (41). Point effect estimates (as odds ratio) for both recent large nonrandomized studies were adverse; the odds ratio was 1.37 in the Mehta et al. (34) study and approximately 1.2 for all models in the Uchino et al. (37) study, as reiterated by editorial commentary (39, 40). Similarly, the point estimate for the largest RCT (30) was an odds ratio of 1.26 (95% CI 0.79, 1.99). These individual estimates were remarkably consistent with the overall mortality estimate of the current study (in the odds ratio metric, 1.24; 95% CI 0.89, 1.74).

In the current study, the use of loop diuretics was found to significantly decrease the oliguric period by a mean of 7.7 days (Fig. 4), with a probability of shorter time to normalization of renal indexes of 82.2%, that is, $1 - P(\text{MD} > 0)$, although the credible interval was wide at −5.62 to 2.46 days for this secondary end

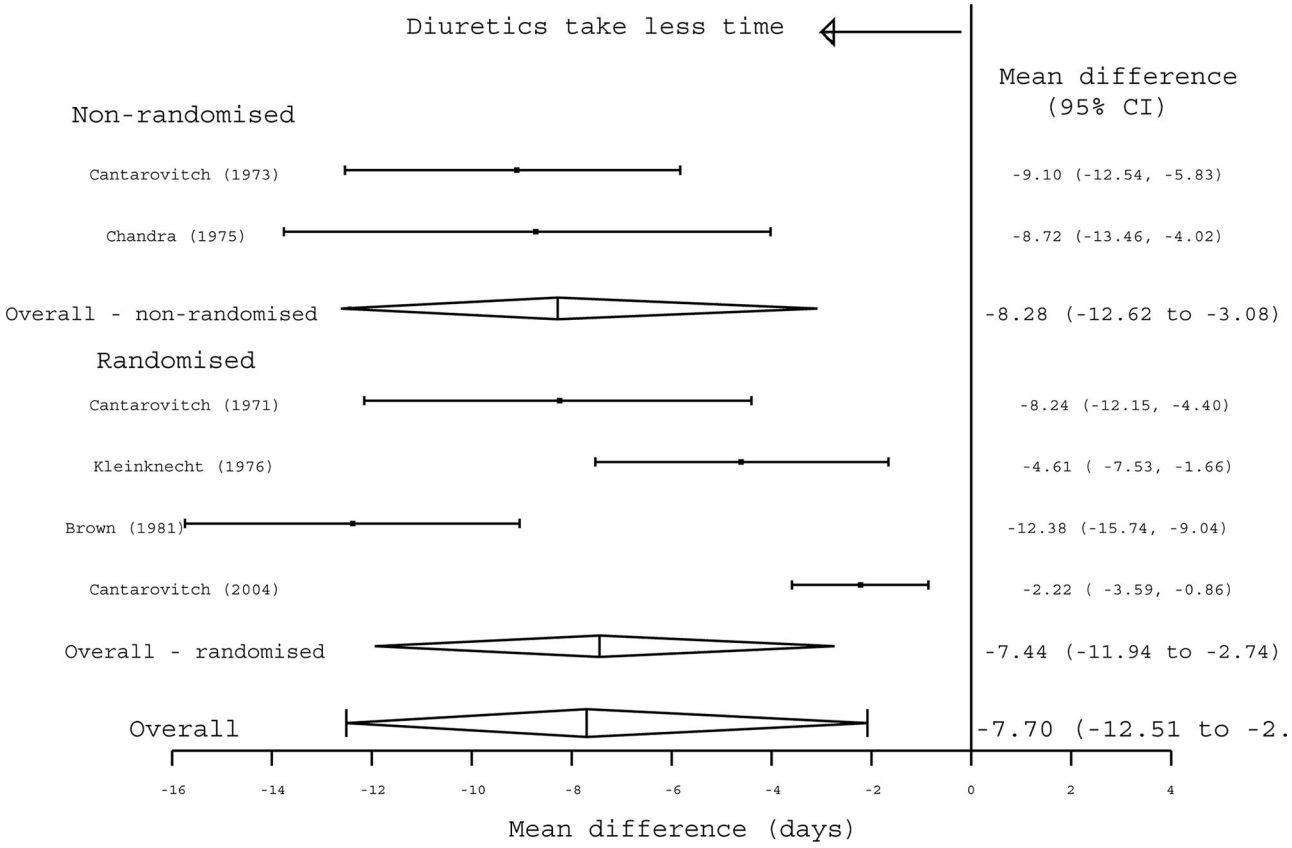

Figure 3. Forest plot showing effect of randomized and nonrandomized studies on time taken to diuresis of ≥1500 mL/day, as mean difference (days). *Small solid squares*, study estimates; *vertically capped horizontal lines*, 95% credible intervals (*CI*); *vertical lines within vertically capped diamond-shaped boxes*, subgroup and overall point estimates and 95% CI; *vertical straight line*, the null effect.

point (Fig. 2). The need for dialysis has also been considered a quantifiable cost factor (15), and although the credible interval indicated no substantial decrease in the rate of dialyses under loop diuretics compared with control (IRR 0.71; 95% CI 0.47–1.06), again, the probability that the IRR was greater than unity indicated a *realizable* benefit of loop diuretics: $P(\text{IRR}) > 1 = 4.1\%$. Thus, the "beneficial" and, presumably, not unexpected effects of loop diuretics were increment of urine volume and potential for decrease in need for dialysis. Of note, there was considerable selective bias in reporting these secondary end points (41).

The tendency for loop diuretics to increase mortality despite improved uremic indexes appears counterintuitive when ARF is viewed as a discrete organ failure. However, few cases of ARF are due to a disease process, such as glomerulonephritis, that is relatively isolated to the kidney; most represent the renal manifestation of a generalized insult, such as severe infection (1, 37). Consequently, factors unrelated to kidney function may determine outcome, and frusemide does have a number of nonrenal effects that could influence mortality. In chronic heart failure, 1 mg/kg intravenous frusemide results in reduced stroke volume and elevated blood pressure, peripheral resistance, pulmonary artery occlusion pressure, and plasma renin, norepinephrine, and vasopressin levels (42). Similar effects in patients with ARF could contribute to organ dysfunction without necessarily impairing renal recovery, and adverse circulatory effects may inadvertently occur if inappropriate diuresis results in reduced ventricular preload and cardiac output. Frusemide has also been shown to have immunosuppressive effects on peripheral blood mononuclear cells similar to equimolar concentrations, hence equivalent doses, of hydrocortisone (43). Given that the dose of frusemide administered in the studies of ARF can easily exceed 300 mg/day (Table 2), it is likely that significant immunosuppression, with its attendant risks, may have contributed to nonrenal causes of death. Thus, initial beneficial effects with respect to secondary outcomes may not be manifest in primary outcomes; indeed, the recently reported improvement in secondary outcomes, but increase in mortality (in certain subsets), with steroids in acute lung injury is consistent with this notion (44). The divergence between the adverse survival and "beneficial" secondary end point effects may have also been due to a selection effect of diuretics upon the sickest patients (a "frailty") (45), although, in the absence of individual time-to-event patient data (46), these questions cannot be definitively addressed.

Both quality score and trial termination date, considered as modifying covariates, had positive coefficients, with the high probability of being greater than zero indicating that the more recent the study or the better the study quality, the greater the association with an adverse mortality effect (see the sensitivity analysis in the Results section; also Table 3). Five of the six studies with a quality score <11 (median value) (26–29, 35) were performed and/or published during the period 1969–1978, and a majority (10 of 13) were published >20 yrs ago (25–29, 31–33, 35, 36). Similarly, 87.3% of total patient number was contributed by the

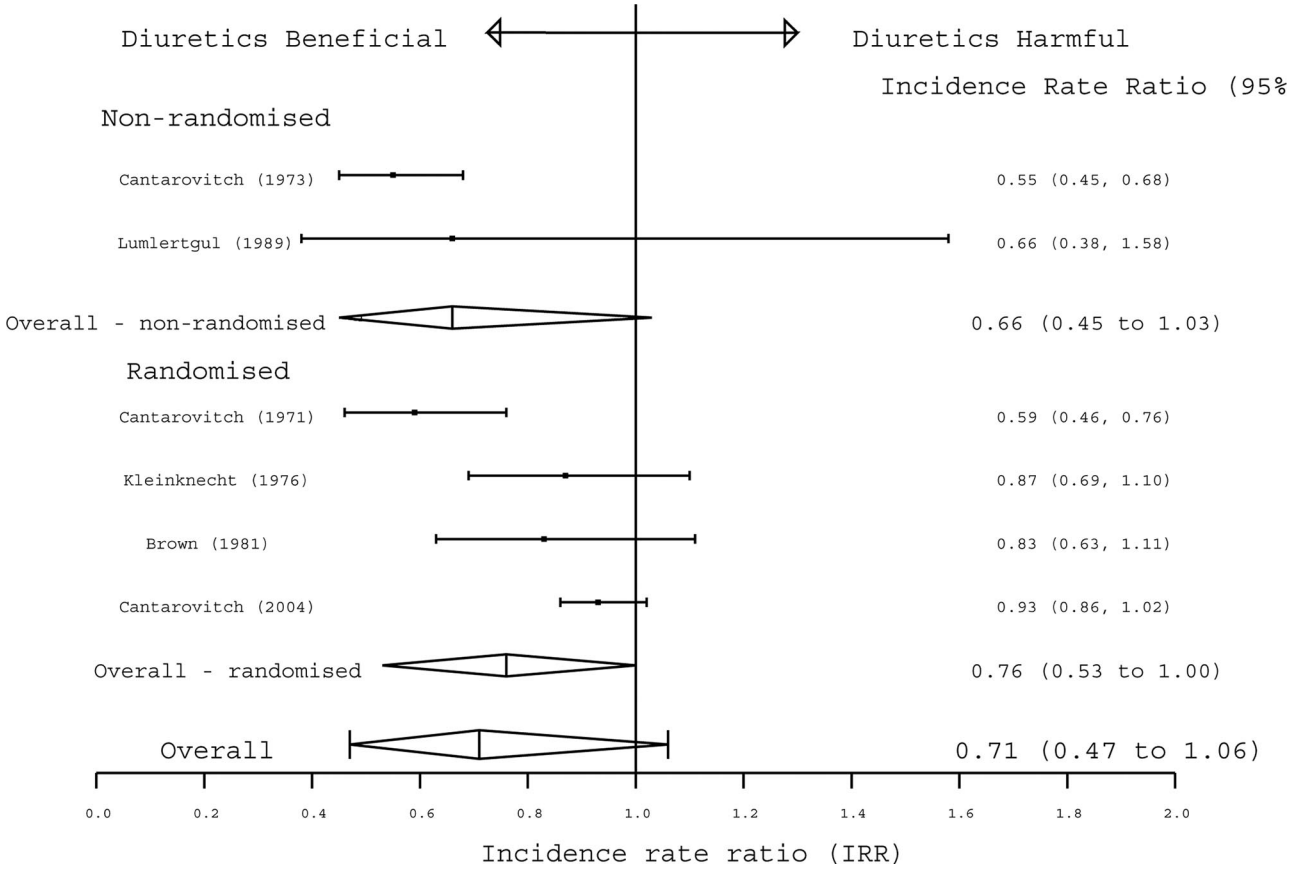

Figure 4. Forest plot showing effect of randomized and nonrandomized studies on dialysis rate, as incidence rate ratio. *Small solid squares*, study estimates; *vertically capped horizontal lines*, 95% credible intervals (*CI*); *vertical lines within vertically capped diamond-shaped boxes*, subgroup and overall point estimates and 95% CI; *vertical straight line*, the null effect.

Table 3. Summary results of including covariates as a common effect in the hierarchical model on the risk ratio of mortality under loop diuretics vs. control

Covariate	No. of Included Studies,[a] Nonrandomized/RCTs	Median Slope[b]	95% CI[c]	p (slope >0), %[d]
Quality score	8/5	0.02	−0.01, 0.04	91.2
Average age	6/3	0.01	−0.01, 0.04	73.0
Proportion males	5/4	0.29	−1.89, 2.39	61.2
Control arm risk	8/5	−0.18	−0.84, 0.48	28.8
Study completion date	8/5	0.01	−0.00, 0.02	90.0

RCT, randomized controlled trial; CI, credible interval.

[a]Number of studies (nonrandomized or RCTs) in which the covariate was reported and thus suitable for analysis; [b]median empirical estimate(of the true posterior variables) of the effect magnitude of the covariate (≡ "coefficient") acting as a common effect in the hierarchical model. Note that allowing covariates to have a different effect for each study type gave very similar results (data not presented); [c]95% Bayesian credible intervals; [d]probability that the "slope" coefficient was >0 in magnitude.

three most recent studies (30, 34, 37), which were notable in that a significant percent of intensive care patients were included, suggesting a determining effect of calendar year-dependent case-mix differences.

The studies of Ho and Sheridan (4) and Bagshaw et al. (47), using frequentist analysis, found no clinical benefit for the use of frusemide in ARF but were unable to effectively quantify the consequent risk of benefit or harm. The current Bayesian analysis generated overall point estimates and intervals for the RCT group that were comparable with these studies (4, 47). However, the inclusion of the larger recent observational studies and the particular insight afforded by Bayesian evidence synthesis have allowed a more adequate assessment of the efficacy of diuretics in ARF. The Bayesian methods used here resulted in point estimates of individual studies borrowing strength from each other with shrinkage toward the overall effect estimate, resulting in an apparent "difference" from raw effect estimates (Table 2). The current analysis thus complements recent studies (4, 47) by providing additional information from all of the available studies of diuretic use in ARF. Furthermore, probability statements quantifying the risk of harm or benefit were easily accessed; in this study, the probability that diuretics were associated with higher mortality compared with control was 84%. To determine the effect of the priors on the results, a sensitivity analysis (not presented) was undertaken. Here, priors were made progressively less informative, resulting in very similar point estimates to those reported but with wider credible intervals. The wider intervals reflected an increasing level of uncertainty; however, the overall conclusions remain unchanged.

Importantly, for the overall gestalt of loop diuretic effect, estimates obtained from nonrandomized studies and controlled trials were consistent, although nonrandomized studies may have been subject to patient and treatment selection bias. Given that loop diuretics are still widely used in ARF, the current evidence, suggesting a mortality detriment, would appear to mandate further large-scale trials to more accurately define their role.

REFERENCES

1. Thadhani R, Pascual M, Bonventre JV: Acute renal failure. *N Engl J Med* 1996; 334: 1448–1460
2. Dishart MK, Kellum JA: An evaluation of pharmacological strategies for the prevention and treatment of acute renal failure. *Drugs* 2000; 59:79–91
3. Star RA: Treatment of acute renal failure. *Kidney Int* 1998; 54:1817–1831
4. Ho KM, Sheridan DJ: Meta-analysis of frusemide to prevent or treat acute renal failure. *BMJ* 2006; 333:420
5. Bellomo R, Kellum JA, Ronco C: Defining acute renal failure: Physiological principles. *Intensive Care Med* 2004; 30:33–37
6. Gotzsche PC: Why we need a broad perspective on meta-analysis. It may be crucially important for patients. *BMJ* 2000; 321:585–586
7. Norris SL, Atkins D: Challenges in using nonrandomized studies in systematic reviews of treatment interventions. *Ann Intern Med* 2005; 142:1112–1119
8. Spiegelhalter DJ, Abrams KR, Myles JP: Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester, West Sussex: Wiley, 2004
9. Sutton AJ, Abrams KR: Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res* 2001; 10:277–303
10. Novis BK, Roizen MF, Aronson S, et al: Association of preoperative risk factors with postoperative acute renal failure. *Anesth Analg* 1994; 78:143–149
11. Bellomo R, Ronco C, Kellum JA, et al: ADQI workgroup. Acute renal failure—Definition, outcome measures, animal models, fluid therapy and information technology needs: The Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 2005; 8:R204–R212
12. Albright RC Jr: Acute renal failure: A practical update. *Mayo Clin Proc* 2001; 76:67–74
13. Robinson KA, Dickersin K: Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002; 31:150–153
14. Stroup DF, Berlin JA, Morton SC, et al: Meta-analysis of observational studies in epidemiology: A proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; 283:2008–2012
15. Conger JD: Interventions in clinical acute renal failure: What are the data? *Am J Kidney Dis* 1995; 26:565–576
16. Majumdar S, Kjellstrand CM: Why do we use diuretics in acute renal failure? *Semin Dial* 1996; 9:454–459
17. Brezis M, Rosen SN, Epstein FH: The pathophysiological implications of medullary hypoxia. *Am J Kidney Dis* 1989; 13:253–258
18. Lameire NH, De Vriese AS, Vanholder R: Prevention and nondialytic treatment of acute renal failure. *Curr Opin Crit Care* 2003; 9:481–490
19. Shilliday I, Allison ME: Diuretics in acute renal failure. *Ren Fail* 1994; 16:3–17
20. Downs SH, Black N: The feasibility of creating a checklist for the assessment of the methodological quality both of randomized and non-randomized studies of health care interventions. *J Epidemiol Community Health* 1998; 52:377–384
21. Prevost TC, Abrams KR, Jones DR: Hierarchical models in generalized synthesis of evidence: An example based on studies of breast cancer screening. *Stat Med* 2000; 19:3359–3376
22. Moran J, Solomon P, Warn D: Methodology in meta-analysis: A study from critical care meta-analytic practice. *Health Serv Outcomes Res Methodol* 2004; 5:207–226
23. Sterne JA, Gavaghan D, Egger M: Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *J Clin Epidemiol* 2000; 53:1119–1129
24. Brown CB, Ogg CS, Cameron JS, et al: High dose frusemide in acute reversible intrinsic renal failure: A preliminary communication. *Scott Med J* 1974; 19:35–40
25. Brown CB, Ogg CS, Cameron JS: High dose frusemide in acute renal failure: A controlled trial. *Clin Nephrol* 1981; 15:90–96
26. Beroniade V: Furosemide in acute and chronic renal failure. *Dialysis Renal Transplant* 1969; 5:241–245
27. Borirakchanyavat V, Vongsthongsri M, Sitprija V: Furosemide and acute renal failure. *Postgrad Med J* 1978; 54:30–32
28. Cantarovich F, Fernandez JC, Locatelli A, et al: Frusemide in high doses in the treatment of acute renal failure. *Postgrad Med J* 1971; 47(Suppl):13–17
29. Cantarovich F, Galli C, Benedetti L, et al: High dose frusemide in established acute renal failure. *BMJ* 1973; 4:449–450
30. Cantarovich F, Rangoonwala B, Lorenz H, et al: High-dose furosemide for established ARF: A prospective, randomized, double-blind, placebo-controlled, multicenter trial. *Am J Kidney Dis* 2004; 44:402–409
31. Chandra M, Agarwal SS, Mitra MK, et al: Some observations in the management of acute renal failure by massive intravenous frusemide therapy. *J Assoc Physicians India* 1975; 23:415–422
32. Kleinknecht D, Ganeval D, Gonzalez-Duque LA, et al: Furosemide in acute oliguric renal failure: A controlled trial. *Nephron* 1976; 17: 51–58
33. Lumlertgul D, Keoplung M, Sitprija V, et al: Furosemide and dopamine in malarial acute renal failure. *Nephron* 1989; 52:40–44
34. Mehta RL, Pascual MT, Soroko S, et al: PICARD Study Group. Diuretics, mortality, and nonrecovery of renal function in acute renal failure. *JAMA* 2002; 288:2547–2553
35. Minuth AN, Terrell JB Jr, Suki WN: Acute renal failure: A study of the course and prognosis of 104 patients and of the role of furosemide. *Am J Med Sci* 1976; 271:317–324
36. Shilliday IR, Quinn KJ, Allison ME: Loop diuretics in the management of acute renal failure: A prospective, double-blind, placebo-controlled, randomized study. *Nephrol Dial Transplant* 1997; 12:2592–2596
37. Uchino SM, Doig GSP, Bellomo RM, et al: Diuretics and mortality in acute renal failure. *Crit Care Med* 2004; 32:1669–1677
38. Schetz M: Diuretics in acute renal failure? *Contrib Nephrol* 2004; 144:166–181
39. Lameire N, Vanholder R, Van Biesen W: Loop diuretics for patients with acute renal failure: Helpful or harmful? *JAMA* 2002; 288: 2599–2601
40. Noble DW: Acute renal failure and diuretics: Propensity, equipoise, and the need for a clinical trial. *Crit Care Med* 2004; 32: 1794–1795
41. Williamson PR, Gamble C: Identification and impact of outcome selection bias in meta-analysis. *Stat Med* 2005; 24:1547–1561
42. Francis GS, Siegel RM, Goldsmith SR, et al: Acute vasoconstrictor response to intravenous furosemide in patients with chronic congestive heart failure: Activation of the neurohumoral axis. *Ann Intern Med* 1985; 103:1–6
43. Yuengsrigul A, Chin TW, Nussbaum E: Immunosuppressive and cytotoxic effects of furosemide on human peripheral blood mononuclear cells. *Ann Allergy Asthma Immunol* 1999; 83:559–566
44. The National Heart Lung and Blood Institute Acute Respiratory Distress Syndrome Clinical Trials Network: Efficacy and safety of corticosteroids for persistent acute respiratory distress syndrome. *N Engl J Med* 2006; 354: 1671–1684
45. Vaupel JW, Yashin AI: Heterogeneity's Ruses: Some surprising effects of selection on population dynamics. *Am Stat* 1985; 39:176–185
46. Stewart LA, Parmar MK: Meta-analysis of the literature or of individual patient data: Is there a difference? *Lancet* 1993; 341: 418–422
47. Bagshaw SM, Haase M, Ghali WA, et al: Loop diuretics in the management of acute renal failure: A systematic review and meta-analysis. *Crit Care Resusc* 2007; 9:60–68
48. Speigelhalter D, Thomas A, Best N: WINBUGS. http://www.mrc-bsu.cam.ac.uk/bugs. Accessed December 21, 2004
49. Smith B: Bayesian Output Analysis Program (BOA): Version 1.1.5 for R. http://www.public-health.uiowa.edu/boa. Accessed November 6, 2005

APPENDIX 1

Highly Sensitive Search Strategy Used in Search Software OVID

exp review/ or exp meta-analysis/ or exp cohort studies/ or (exp randomized controlled trials/ or randomized controlled trial.pt.) or controlled clinical trial.pt. or exp random allocation/ or exp double-blind method/ or exp single-blind method/ or clinical trial.pt. or exp clinical trials/ or clinical trial.mp. or ((single or double or treb$ or triple) and (mask$ or blind$)).mp. or (latin square or placebo$ or random$ or control$ or prospective$ or volunteer$ or placebos).mp. or (exp research design/ or exp placebos/ or comparative study/ or exp evaluation studies/ or exp follow-up studies/ or exp prospective studies/ or exp cross-over studies/ or "Review Literature"/ or review articles.mp.) and (exp kidney failure, acute/ or acute renal failure.mp. or exp kidney tubular necrosis, acute/) and (exp diuretics/ or bumetanide/ or exp furosemide/ or exp sodium-potassium-chloride symporters/ or frusemide.mp. or furosemide/ or lasix.mp. or (piretanide or torasemide or ethacrynic acid).mp. or loop diuretics.mp.).

APPENDIX 2

Each of these Bayesian hierarchical models used WinBUGS software (48). Three chains each with differing starting values were employed. A burn-in of 100,000 iterations was run, and a further 100,000 iterations were obtained for calculating estimates and credible intervals. Convergence was checked using the Brooks, Gelman, and Rubin and Geweke criteria available in the Bayesian Output Analysis Program (BOA) (49) and deemed reasonable in all cases.

The statistical model used here closely follows that of Prevost et al. (21). Let y_{ij} be the observed effect of the ith study of type j with associated sample variance s_{ij}^2. For the mortality data, the observed effect is the log relative risk, for the count data the observed effect is the log incidence rate ratio, and for the continuous outcomes the observed effect is the mean difference. Then we assume a three-level hierarchical model of the form.

$$y_{ij} \sim N(\delta_{ij}, s_{ij}^2)$$
$$\delta_{ij} \sim N(\phi_j, \sigma_j^2)$$
$$\phi_j \sim N(\mu, \tau^2) \quad [1]$$

Here μ represents the overall treatment effect (in the scale described previously) with between study-type variability τ^2. The ϕ_j represents the treatment effect with associated variance σ_j^2 in the nonrandomized and RCT groups ($j = 1,2$), respectively, and δ_{ij} represents the underlying effect in the ith study of type j. In the Bayesian framework, prior distributions are required for μ, τ^2, and σ_j^2. Specification of these prior distributions depends on the scale of the observed effect and should be locally noninformative in the absence of pertinent external evidence, yet clinically plausible. However, for variance variables, sensitivity of the results to any prior distribution used should be assessed (8).

For log risk ratio, the same arguments as those of Prevost et al. (21) are made for choosing the priors as they cover the plausible range of variable values in this study also. For τ we assume that it is very unlikely that the risk ratio for a given study type will be more than twice or less than half the underlying population risk ratio so that the log risk ratio in the jth study type should be within $\mu \pm \log(2)$. Assuming a normal prior distribution, this means that it should be unlikely that τ will exceed $\log(2)/1.96 = 0.35$. A half-normal (HN) distribution (8) (i.e., a normal distribution truncated below zero) with mean 0 and variance 1/30 has approximately this property. Similarly, priors for each σ_j may be chosen by assuming that the between study variability within each type j is likely to be greater than the variability in risk ratio between study types. Suppose then that the log risk ratio for each study δ_{ij} of type j is highly likely to be within $\phi_j \pm \log(4)$. Again, assuming a normal prior distribution, σ_j should be unlikely to exceed $\log(4)/1.96 = 0.71$. An HN distribution (8) with mean 0 and variance 1/8 has approximately this property. To choose a prior distribution for μ, we assume that the overall risk ratio is not likely to exceed 500 in favor of either the treatment or the control, that is, $\log(500)/1.96 = 3.17$. As such, an N(0,10) distribution would be a reasonable choice as a prior for μ. The same arguments may be made when choosing the prior distribution for the log incidence rate ratio.

The analysis of the continuous end points used the mean difference and had a different range of plausible values. As such, the choice of prior distributions for τ, σ_j, and μ needed to be reconsidered. First we determined a suitably diffuse prior distribution for τ by initially noting that between the study types a median range of approximately 3 days was clinically plausible. The median of a half normal distribution is 1.09 SD, so that $SD = 3/1.09 = 2.75$. With an SD of 2.75, it would be extremely unlikely that a difference of more than about 8.25 (since $3 \cdot SD = 3 \cdot 2.75 = 8.25$) days will occur between study types. So we take an $HN(0,2.75^2)$ distribution as our clinically plausible prior distribution for τ. To determine reasonable prior distributions for σ_j, we note that a median range of 7 days between any two studies is clinically plausible. This gives $SD = 7/1.09 = 6.42$. Again, it is clinically reasonable to believe that a difference in the level of variability about 0 is highly unlikely to exceed 19 days (since $3 \cdot SD = 3 \cdot 6.42 = 19$), and we take an $HN(0,6.4^2)$ distribution as our prior distributions for each σ_j. Finally, we assume that the overall mean difference, μ, is not likely to be >30 days so that we get $SD = 30/3 = 10$ and we take an $N(0,10^2)$ distribution as a reasonable prior for μ. For all end points, sensitivity analyses were conducted in which the corresponding SDs for the HN prior distributions were increased four-fold to represent more diffuse clinical ranges of plausibility. In each case, the results obtained, although quantitatively different, were similar point estimates with increased levels of uncertainty, and the conclusions drawn were thus qualitatively the same.

Note that for the sensitivity analyses, in which a covariate was included (see Prevost (21) for details) in the previously mentioned model, a vague Normal distribution with mean 0 and variance 1000 was given to the coefficient of that covariate.