Appendix: Methods for Guideline Development

Aim

The overall aim of the project was to create a clinical practice guideline with recommendations for the care of the KTRs using an evidence-based approach. After topics and relevant clinical questions were identified, the pertinent scientific literature on those topics was systematically searched and summarized.

Overview of process

The development of the guideline included sequential and concurrent steps:

- Appoint the Work Group and Evidence Review Team (ERT), which were responsible for different aspects of the process.
- Confer to discuss process, methods and results.
- Develop and refine topics.
- Define specific populations, interventions or predictors and outcomes of interest.
- Create and standardize quality assessment methods.
- Create data-extraction forms.
- Develop literature search strategies and run searches.
- Screen abstracts and retrieve full articles based on predetermined eligibility criteria.
- Extract data and perform critical appraisal of the literature.
- Grade quality of the outcomes of each study.
- Tabulate data from articles into summary tables.
- Grade the quality of evidence for each outcome and assess the overall quality and findings of bodies of evidence with the aid of evidence profiles.
- Write recommendations and supporting rationale statements.
- Grade the strength of the recommendations based on the quality and strength of the evidence and other considerations.
- Peer review by KDIGO Board of Directors in December 2008 and the public (March 2009), with subsequent revisions.

The Work Group, KDIGO Co-Chairs, ERT, liaisons and KDIGO support staff met for four 2-day meetings for training in the guideline development process, topic discussion and consensus development.

Creation of groups

The KDIGO Co-Chairs appointed the Co-Chairs of the Work Group, who then assembled the Work Group to be responsible for the development of the guideline. The Work Group consisted of domain experts, including individuals with expertise in adult and pediatric nephrology, transplant surgery and medicine, critical-care medicine, cardiology, infectious diseases, oncology and epidemiology, along with a patient advocate. Tufts Center for Kidney Disease Guideline Development and Implementation at Tufts Medical Center in Boston, MA, USA, was contracted to provide expertise in guideline development methodology and systematic evidence review. The ERT consisted of physician–methodologists with expertise in nephrology and internal medicine, and research associates and assistants. The ERT instructed and advised Work Group members in all steps of literature review, critical literature appraisal and guideline development. The Work Group and the ERT collaborated closely throughout the project. The ERT also included methodological input and assistance with literature searches from methodology experts at the Cochrane Renal Group in Sydney, Australia.

Systematic Review: General Process

The first task of the Work Group was to define the overall topics and goals for the guideline. The Work Group Co-Chairs drafted a preliminary list of topics. The Work Group identified the key clinical questions. The Work Group and ERT further developed and refined each topic, specified screening criteria, literature search strategies and data-extraction forms.

The ERT performed literature searches, and organized screening of abstracts and articles. The ERT also coordinated the methodological and analytic processes of the report, and defined and standardized the methodology of performing literature searches, data extraction and summarizing the evidence. Throughout the project, the ERT offered suggestions for guideline development, led discussions on systematic review, literature searches, data extraction, assessment of quality and applicability of articles, evidence synthesis, grading of evidence and recommendations and consensus development. With input from the Work Group, the ERT finalized eligible studies, performed all data extraction and summarized data into summary tables. They also created preliminary evidence profiles (described below), which were completed by the Work Group members. The Work Group members reviewed all included articles, data-extraction forms and summary tables for accuracy and completeness. The Work Group took the primary role of writing the recommendations and rationale statements, and retained final responsibility for the content of the recommendation statements and the accompanying narrative.

For questions of treatments in the KTRs, systematic reviews of the eligible RCTs were undertaken (Table 32). For these topics, the ERT created detailed data-extraction forms, and extracted information on baseline data for the populations, interventions, study design, results and provided an assessment of quality of evidence. The ERT then tabulated studies in summary tables, and assigned grades for the quality of the evidence in consultation with the Work Group.

For nontreatment questions, that is questions related to prevalence, evaluation and risk relationships, the ERT conducted systematic searches, screened the yield for relevance and provided lists of citations to the Work Group. The ERT created summary tables of selected observational incidence and predictor studies. The Work Group took primary responsibility for reviewing and summarizing this literature in a narrative format. The ERT also searched online databases for estimates of incidence rates of different cancers among larger countries representative of different regions. The primary database used was Cancer Mondial (http://www-dep.iarc.fr). SIRs for cancer in solid-organ transplant recipients were taken from a meta-analysis by Grulich et al. (623).

For topics on which previous or ongoing KDIGO or KDOQI guidelines have provided recommendations for KTRs, new systematic reviews were not performed. These include anemia, hepatitis C, mineral and bone disorders and pediatric nutrition. For these topics, the relevant recommendations and rationale text were excerpted and refined as necessary. The Work Group Chairs and selected members conferred with Co-Chairs of the concurrent KDIGO mineral and bone disorder guideline and KDOQI pediatric nutrition guideline on transplant bone disease (Chapter 21), and growth and development (Chapter 24).

Refinement of Topics

The Work Group Co-Chairs prepared the first draft of the scope of work document as a series of mock (preliminary) recommendations to be considered by Work Group members. At their first 2-day meeting, members added further mock guideline topics until the initial working document included all topics of interest to the Work Group. The inclusive, combined set of questions formed the basis for the deliberation and discussion that followed. The Work Group strove to ensure that all topics deemed clinically relevant and worthy of review were identified and addressed. The four major topic areas of interest for the care of KTRs included immunosuppression, infections, CVD and cancer. In addition, there were several miscellaneous topics.

At the initiation of the guideline development process, it was agreed that this guideline would focus on patients who have had kidney transplantations. Thus, with few exceptions (e.g. the timing of vaccinations), all topics, systematic reviews and study eligibility criteria were restricted to patients with existing kidney transplantations. The guideline does not address management issues regarding choosing patients for kidney transplantation, pretransplant care, intraoperative care (except for the timing of initiating immunosuppression) or management of patients who have lost their grafts. In addition, in regards to care of comorbidities and complications after kidney transplantation (e.g. infections, cancer and CVD), this guideline focuses primarily on monitoring and prevention of the conditions, as opposed to treatment of the conditions (with some exceptions, e.g. for infectious diseases). However, where the recommended treatment of conditions differed from the general population (e.g. due to drug interactions with immunosuppression agents), standard treatment recommendations are offered.

Based on the list of topics, the Work Group and ERT developed a list of specific research questions for which systematic review would be performed (Table 32). For each systematic review topic, the Work Group Co-Chairs and the ERT formulated well-defined systematic review research questions using a well-established system (931). For each question, explicit criteria were agreed on for the population, intervention or predictor, comparator, outcomes of interest and study design features. A list of outcomes of interest was generated. The Work Group ranked patient-centered clinical outcomes (such as death, graft loss or infections) as more important than intermediate outcomes (such as cholesterol level or hypertension). The outcomes were further categorized as being of critical, high or moderate importance to KTRs. Outcomes of low importance were not considered for the purpose of systematic review and evidence synthesis. The specific criteria used for each topic are described below in the description of the review topics. In general, eligibility criteria were determined based on clinical value, relevance to the guideline and clinical practice, determination whether a set of studies would affect recommendations or the strength of evidence and practical issues such as available time and resources.

Literature Searches and Article Selection

The MEDLINE, Cochrane Central Registry for trials, and Cochrane database of systematic reviews were searched from 1985 through January 2007 by the ERT to capture all citations relevant to the topic of kidney transplantation, including original articles, systematic reviews and previous guidelines. The Cochrane Renal Group ran parallel searches in their Renal Registry database and these supplemented the primary ERT searches. The search was updated through February 2008 and supplemented by articles identified by Work Group members through November 2008.

Appendix

Table 32: Systematic review topics and screening criteria

Chapter 1: Induction Therapy	
Population	KTRs in the first 24 h after transplant
Predictor, reference standard	IL2 (mab) vs. no induction, antithymoglobulin vs. no induction, antithymoglobulin vs. IL2
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, infection, cancer, NODAT, fracture, BMD, erythrocytosis, neutropenia, quality of life, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 50
Chapter 2: Initial Maintenance Immunosuppressive Medications	
Population	KTRs
Intervention, reference standard	Tac vs. CsA (CsA or CsA-ME) (with AZA, MMF, Sirolimus, Everolimus), CNI vs. non-CNI regimens, MMF vs. AZA, MMF formulation vs. other MMF formulation, CNI-sparing (withdrawal), CNI-free, steroid withdrawal, steroid avoidance
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, infection, cancer, NODAT, erythrocytosis, neutropenia, fracture, BMD, hypertension, hyperuricemia, hyperlipidemia, quality of life, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 100
Chapter 3: Long-Term Maintenance Immunosuppressive Medications	
Population	KTRs
Intervention	Tac vs. CsA (CsA or CsA-ME) (with AZA, MMF, sirolimus, everolimus), CNI vs. non-CNI regimens, MMF vs. AZA, MMF formulation vs. other MMF formulation, CNI-sparing (withdrawal), CNI-free, steroid withdrawal, steroid avoidance
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, infection, cancer, NODAT, erythrocytosis, neutropenia, fracture, BMD, hypertension, hyperuricemia, hyperlipidemia, quality of life, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 100
Chapter 4: Strategies to Reduce Drug Costs	
Population	KTRs
Intervention	CsA-ME generics, other generic medications
Outcomes	Pharmacokinetics, pharmacodynamics
Study design	RCT
Minimum number of subjects	N $\geq$ 20
Chapter 5: Monitoring Immunosuppressive Medications	
Population	KTRs
Intervention	MMF fixed dose vs. AUC-adjusted doses, C_0 vs. C_2 CsA to determine dosing, anti-HLA antibodies
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, infection, cancer, NODAT, erythrocytosis, neutropenia, fracture, BMD, hypertension, hyperuricemia, hyperlipidemia, quality of life, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 10
Chapter 6: Treatment of Acute Rejection	
Population	KTRs with biopsy-proven acute rejection
Predictor	Adding induction agents or other (intravenous immunoglobulin, plasma exchange), change of maintenance regimen
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, infection, cancer, NODAT, fracture, BMD, erythrocytosis, neutropenia, quality of life, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 100
Chapter 7: Treatment of Chronic Allograft Injury	
Population	KTRs with CAN or biopsy-proven CNI toxicity
Intervention, predictor	Reduction in CNI, change in maintenance immunosuppression, adding ancillary treatments (ACE-I, ARB, etc.), CNI dose reduction, CNI withdrawal, replacement of CNI with another immunosuppression agent, comparisons with placebo or other treatments
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, biopsy changes, infection, cancer, NODAT, erythrocytosis, neutropenia, fracture, BMD, hypertension, hyperuricemia, hyperlipidemia, quality of life, adverse events

Continued.

Table 32: Continued

Chapter 7: Treatment of Chronic Allograft Injury	
Study design	RCT
Minimum number of subjects	N ≥ 100
Chapter 8: Monitoring Kidney Allograft Function	
Population	KTRs
Intervention	Protocol monitoring vs. no protocol, different frequencies of monitoring
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN
Study design	RCT; minimum follow-up time: ≥6 months
Minimum number of subjects	N ≥ 10
Chapter 9: Kidney Allograft Biopsy	
Population	KTRs
Intervention	Protocol biopsy vs. not, different protocols, treatment of 'borderline' rejection based on protocol biopsy vs. no biopsy
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, recurrent disease
Study design	RCT; minimum follow-up time: ≥6 months
Minimum number of subjects	N ≥ 10
Chapter 10: Recurrent Kidney Disease	
Population	KTRs with biopsy-proven recurrent disease
Intervention	Any
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, recurrent disease, GFR/SCr or eGFR, biopsy changes, serious adverse events
Study design	RCT
Minimum number of subjects	N ≥ 20
Chapter 11: Preventing, Detecting, and Treating Nonadherence	
Population	KTRs
Intervention	Any
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, use of immunosuppressive medications as prescribed
Study design	RCT
Minimum number of subjects	N ≥ 20
Chapter 12: Vaccination	
Population	KTRs (for PCP: any solid-organ recipient)
Intervention	Polyoma virus/BKV nephropathy: biopsies, urine NAT, urine decoy cells EBV: acyclovir/ganciclovir, change immunosuppression agent, intravenous immunoglobulin, anti-CD20 antibody HSV/VZV: acyclovir/ganciclovir PCP: sulfamethoxazole–trimethoprim vs. dapsone vs. pentamidine, prophylaxis vs. no prophylaxis, different protocols HBV: monitoring, drug prophylaxis UTI: antibiotic prophylaxis TB: PPD, QuantiFERON screening Fungal: screening, prophylaxis
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, relevant disease, adverse events
Study design	RCT
Minimum number of subjects	N ≥ 10
Chapter 13: Viral Diseases	
Population	KTRs
Intervention	Polyoma virus/BKV nephropathy: reduce immunosuppression, cidofovir, leflunomide CMV: reduce immunosuppression, gancyclovir, valgancyclovir, intravenous immunoglobulin, acyclovir EBV: acyclovir, gancyclovir, reduce immunosuppression, intravenous immunoglobulin, anti-CD20 antibody HBV: interferon (timing), pegylated interferon, lamivudine, adefovir, entecavir
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, relevant disease, BKV nephropathy, change in management, HBV, liver disease progression (by biopsy), hepatocellular carcinoma, adverse events

Continued.

Appendix

Table 32: Continued

Chapter 13: Viral Diseases	
Study design	RCT, cohort
Minimum number of subjects	N ≥ 20 for RCT; N ≥ 100 for cohort
Chapter 14: Other Infections	
Population	KTRs
Intervention	Antibiotic prophylaxis, PPD, Quantiferon screening, screening and prophylaxis for fungal infections
Outcomes	UTI, active TB, fungal disease, mortality, acute rejection, graft loss, kidney function, DGF, CAN, adverse events
Study design	RCT, cohort
Minimum number of subjects	N ≥ 20 for RCT; N ≥ 100 for cohort
Chapter 15: Diabetes Mellitus	
Population	KTRs with NODAT
Intervention	Change in immunosuppressive medications
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, CVD events, fasting glucose
Study design	RCT, cohort; minimum follow-up time: ≥6 months
Minimum number of subjects	N ≥ 100 for RCT; N ≥ 500 for cohort
Chapter 16: Hypertension, Dyslipidemias, Tobacco Use, and Obesity	
Population	KTRs with CVD risk factors
Intervention	Smoking cessation, obesity: weight loss
Outcomes	Reduction in risk factor, all-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, CVD
Study design	RCT; minimum follow-up time: ≥6 months
Minimum number of subjects	N ≥ 20
	Systematic reviews were not performed for hypertension or dyslipidemia
	Referred to KDOQI Guidelines for hypertension and dyslipidemia
Chapter 17: Cardiovascular Disease Management	
Population	KTRs with CVD
Intervention	Aspirin, dipyridamole, ticlopidine, clopidogrel, cilostazol, pentoxyifylline
Outcomes	Bleeding
Study Design	RCT
Minimum number of subjects	N ≥ 20
Chapter 18: Cancer of the Skin and Lip	
Population	KTRs
Intervention	Not applicable
Outcomes	Incidence
Study Design	Registry data
Minimum number of subjects	N ≥ 1000
Chapter 19: Non-Skin Malignancies	
Population	KTRs
Intervention	Not applicable
Outcomes	Incidence
Study design	Registry data or systematic review
Minimum number of subjects	N ≥ 1000
Chapter 20: Managing Cancer with Reduction of Immunosuppressive Medication	
Population	KTRs with cancer
Intervention	Change in immunosuppressive regimens
Outcomes	Mortality, acute rejection, graft loss, kidney function, DGF, CAN, adverse events
Study design	RCT
Minimum number of subjects	N ≥ 10
Chapter 21: Transplant Bone Disease	
	Systematic review not performed
	Referred to KDIGO CKD–MBD Guideline
Chapter 22: Hematological Complications	
Population	KTRs with anemia, erythrocytosis or neutropenia
Intervention	Erythrocyte stimulation therapies, changes in immunosuppressive medications, granulocyte CSF, other treatments

Continued.

Table 32: Continued

Chapter 22: Hematological Complications	
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, quality of life, CVD, infections, hemoglobin or hematocrit, neutropenia duration, adverse events
Study design	RCT
Minimum number of subjects	N $\geq$ 10
Chapter 23: Hyperuricemia and Gout	
Population	KTRs with hyperuricemia
Intervention	Changes in immnosuppressive medications, allopurinol, serum uric acid
Outcomes	Gout, all-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, CVD events
Study Design	RCT, cohort
Minimum number of subjects	N $\geq$ 20 for RCT; N $\geq$ 100 for cohort
Chapter 24: Growth and Development	
Population	Adult and pediatric KTRs
Intervention	Growth hormone, immunosuppressive regimens
Outcomes	Growth, growth retardation, development
Study design	RCT, cohort, systematic review
Minimum number of subjects	N $\geq$ 20 for RCT; N $\geq$ 100 for cohort
Chapter 25: Sexual Function and Fertility	
Population	Kidney transplant patients with sexual dysfunction, mothers who are pregnant, have a transplant during pregnancy, are lactating, or fathers who has a transplant at conception
Intervention	Erectile dysfunction medications
Outcomes	All-cause mortality, DGF, slow graft function, acute rejection, graft failure/survival, kidney function, CAN, erectile dysfunction, pregnancy outcomes, pregnancy complications, immunosuppression medication levels in milk
Study design	RCT, cohort
Minimum number of subjects	N $\geq$ 10 for RCT; N $\geq$ 1 for mothers, fathers or N $\geq$ 50 for pregnancy in cohort
Chapter 26: Lifestyle	
	Systematic review not performed
Chapter 27: Mental Health	
Population	Kidney transplant patients with depression

ACE-I, angiotensin-converting enzyme inhibitor; Anti-HLA, anti-human leukocyte antigen; ARB, angiotensin II receptor blocker; AUC, area under the concentration-time curve; AZA, azathioprine; BKV, BK polyoma virus; BMD, bone mineral density; CAN, chronic allograft nephropathy; CKD–MBD, chronic kidney disease–mineral and bone disorder; CMV, cytomegalovirus; CNI, calcineurin inhibitor; CsA, cyclosporine A; CsA-ME, cyclosporine A microemulsion; CSF, colony-stimulating factor; CVD, cardiovascular disease; DGF, delayed graft function; EBV, Epstein-Barr virus; eGFR, estimated glomerular filtration rate; GFR, glomerular filtration rate; HBV, hepatitis B virus; HSV, herpes simplex virus; IL2 (mab), interleukin-2 (monoclonal antibody); KDIGO, Kidney Disease: Improving Global Outcomes; KDOQI, Kidney Disease Outcomes Quality Initiative; KTRs, kidney transplant recipients; MMF, mycophenolate mofetil; N, number of subjects; NAT, nucleic acid testing; NODAT, new-onset diabetes after transplantation; PCP, *pneumocystis jirovecii* pneumonia; PPD, purified protein derivative; RCT, randomized controlled trial; SCr, serum creatinine; Tac, tacrolimus; TB, tuberculosis; UTI, urinary tract infection; VZV, varicella zoster virus.

During citation screening, journal articles reporting original data were reviewed. Editorials, letters, stand-alone abstracts, unpublished reports and articles published in non–peer-reviewed journals were excluded. The Work Group also decided to exclude publications from journal supplements and Transplantation Proceedings journal because of potential differences in the process of how they get solicited, selected, reviewed and edited compared to peer-reviewed publications.

Potentially relevant existing systematic reviews were examined. If these reviews were deemed to adequately address topics of interest (even if only selected outcomes were reviewed), *de novo* searches on these topics were limited to the time period since the end of the literature search within the systematic reviews.

The MEDLINE and Cochrane search results were screened by the ERT for relevance using predefined eligibility criteria (Table 32). Restrictions by sample size and duration of follow-up were based on methodological and clinical considerations. Generally, it was deemed that trials with fewer than 100 people would be unlikely to have sufficient power to find significant differences in patient-centered clinical outcomes in KTRs. However, for specific topics where sparse data were available, lower sample-size thresholds were used to provide some information for descriptive purposes.

Appendix

Table 33: Literature search yield of RCTs

Topic	Abstracts identified[a]	RCTs retrieved	RCTs accepted	RCTs data-extracted	RCTs included in summary tables[b]	Systematic reviews in evidence profiles
Immunosuppression	15 094	134	93	87	84	7
Monitoring and infections		24	23	17	17	5
CVD		11	2	2	0	0
Malignancy		0	0	0	0	1
Miscellaneous		18	18	13	13	2

CVD, cardiovascular disease; RCT, randomized controlled trial.
[a]All topics and all study designs combined.
[b]Available at www.kdigo.org.

For most topics, the minimum mean duration of follow-up of 6 months was chosen based on clinical reasoning. For the treatments of interest, the proposed effects on patient-centered clinical outcomes require long-term exposure and, typically, would not be expected to become evident before several months of follow-up. For all treatment topics, all RCTs in children with five or more individuals per arm were included.

From the onset of the guideline development process, it was known that for numerous topics of interest (e.g. care of comorbidities and complications after kidney transplantation) very few or no RCTs of KTRs exist. In addition, several topics required data on predictors of outcomes as opposed to treatment efficacy. Therefore, for selected topics, large observational studies were reviewed. As described below, in general, associations from only multivariable regression analyses were considered. The observational studies were not graded for quality. For these topics, the ERT completed its search in December 2007 and did not update the search.

Literature Yield for Systematic Review Topics

Table 33 and Figure 2 summarize the numbers of abstracts screened, articles retrieved and data extracted and included in summary tables.

Data Extraction

The ERT designed data-extraction forms to capture information on various aspects of the primary studies. Data fields for all topics included study setting, patient demographics, eligibility criteria, kidney transplantation details, numbers of subjects randomized, study design, study funding source, descriptions of interventions (or predictors), description of outcomes, statistical methods used, results, quality of outcomes (as described below), limitations to generalizability and free-text fields for comments and assessment of biases.

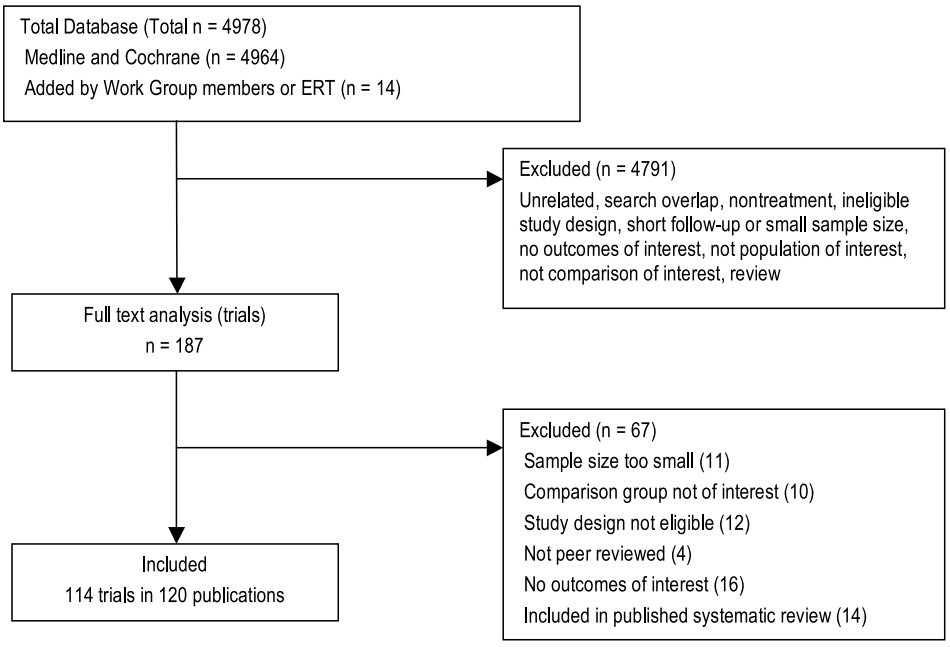

Figure 2: Literature search diagram for systematically reviewed RCTs. ERT, Evidence Review Team; RCT, randomized controlled trial.

Table 34: Hierarchy of outcomes

Hierarchy[a]	Outcomes[b]
Critical importance	Mortality, graft loss, cardiovascular events, malignancy (except skin)
High importance	Acute rejection, CAN, skin cancer, NODAT, infection (disease), bone fracture, quality of life
Moderate importance[c]	DGF, kidney function, proteinuria, lipids, blood pressure, BMD, bone marrow suppression, diarrhea, infection (marker)

BMD, bone mineral density; CAN, chronic allograft nephropathy; DGF, delayed graft function; NODAT, new-onset diabetes after transplantation.

[a]Outcomes of lesser importance are excluded from review.

[b]This categorization was the consensus of the Work Group for the purposes of this transplant guideline only. The lists are not meant to reflect outcome ranking for other areas of CKD management. The Work Group acknowledges that not all clinicians, patients or families or societies would rank all outcomes the same.

[c]All surrogate (intermediate) outcomes that were evaluated were classified as moderate.

Table 35: Classification of study quality

Good quality: Low risk of bias and no obvious reporting errors, complete reporting of data. Must be prospective. If study of intervention: Must be RCT.
Fair quality: Moderate risk of bias, but problems with study/paper are unlikely to cause major bias. If study of study/intervention: Must be prospective.
Poor quality: High risk of bias or cannot exclude possible significant biases. Poor methods, incomplete data, reporting errors. Prospective or retrospective.

Summary Tables

Summary tables were developed to tabulate the data from studies pertinent to each question of intervention (see Supporting Tables 2 and 4 as examples at http://www3.interscience.wiley.com/journal/118499698/toc). Each summary table contains a brief description of the outcome, baseline characteristics of the population, intervention, results and methodological quality. Baseline characteristics include a description of the study size, country of residence, age, percentage of deceased donors and dates of transplant. Intervention and concomitant therapies and the results were all captured. The final column was assigned for a grade for methodological quality. The studies were listed by outcome within the table based on the hierarchy of important outcomes (Table 34). Categorical and continuous outcomes were summarized in separate sets of tables. Work Group members were asked to proof all data in summary tables on RCTs. Separate sets of summary tables were created for nonrandomized studies of incidence and predictors of outcomes.

Due to the large number of recommendations included here and the large volume of literature reviewed, summary tables are not published with this report. They are available at http://www3.interscience.wiley.com/journal/118499698/toc.

Evaluation of Individual Studies

Study size and duration

The study (sample) size is used as a measure of the weight of the evidence. In general, large studies provide more precise estimates. Similarly, longer-duration studies may be of better quality and more applicable, depending on other factors.

Methodological quality

Methodological quality (internal validity) refers to the design, conduct and reporting of the outcomes of a clinical study. A three-level classification of study quality was used (Table 35). Given the potential differences in quality of a study for its primary and other outcomes, the study quality was assessed for each outcome. Variations of this system have been used in most KDOQI and all KDIGO guidelines, and have been recommended by the US Agency for Healthcare Research and Quality Evidence-Based Practice Center program (http://effectivehealthcare.ahrq.gov/repFiles/2007_10DraftMethodsGuide.pdf; last accessed March 30, 2009).

Each study was given an overall quality grade. Each reported outcome was then evaluated and given an individual grade depending on the quality of reporting and methodological issues specific to that outcome. However, the quality grade of an individual outcome could not exceed the quality grade for the overall study.

Results

The type of results used from a study was determined by the study design, the purpose of the study and the Work Group's question(s) of interest for which the results were used. Decisions were based on the screening criteria and outcomes of interest.

Grading the quality of evidence and the strength of a recommendation

A structured approach, based on GRADE (932–934) and facilitated by the use of Evidence Profiles (see Table 36 for an example), was employed in order to grade the quality of the overall evidence and the strength of recommendations. For each topic, the discussion on grading of the quality of the evidence was led by the ERT, and the discussion regarding the strength of the recommendations was led by the Work Group Chairs. The 'strength of a recommendation' indicates the extent to which one can be confident that adherence to the recommendation will do more good than harm. The 'quality of a body of evidence' refers to the

Table 36: Example of an evidence profile[a]

							Summary of findings		
Outcome	No. of studies and study design	Total N (N on study drug)	Methodological quality of studies per outcome	Consistency across studies	Directness of the evidence generalizability/ applicability	Other considerations	Quality of evidence for outcome	Qualitative and quantitative description of effect	Importance of outcome
Mortality	4 RCTs (High)	1584 (799)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Not adequately powered (−1)	Low	No difference between CsA low dose vs. CsA standard dose	Critical
Graft loss	3 RCTs (High)	1473 (746)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Not adequately powered (−1)	Low	No difference between CsA low dose vs. CsA standard dose	Critical
CVD events	0 RCTs								Critical
Cancer	3 RCTs (High)	1256 (635)	Some limitations (−)	No important inconsistencies (0)	No uncertainty (0)	None (0)	Moderate	No difference between CsA low dose vs. CsA standard dose	Critical
Acute rejection	4 RCTs (High)	1584 (799)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	None (0)	Moderate	No difference between CsA low dose vs. CsA standard dose	High
CAN	1 RCT (High)	111 (53)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Sparse data (−1)	Low	No difference between CsA low dose vs. CsA standard dose	High
NODAT	1 RCT (High)	789 (399)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Sparse data (−1)	Low	No difference between CsA low dose vs. CsA standard dose	High
Infection (disease)	4 RCTs (High)	1584 (799)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	None (0)	Moderate	No difference between CsA low dose vs. CsA standard dose	High
Delayed graft function	3 RCTs (High)	1473 (746)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	None (0)	Moderate	No difference between CsA low dose vs. CsA standard dose	Moderate

Continued.

Table 36: Continued

Outcome	No. of studies and study design	Total N (N on study drug)	Methodological quality of studies per outcome	Consistency across studies	Directness of the evidence generalizability/ applicability	Other considerations	Summary of findings: Quality of evidence for outcome	Summary of findings: Qualitative and quantitative description of effect	Summary of findings: Importance of outcome
Kidney function	4 RCTs (High)	1584 (799)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Sparse data (−1)	Low	Kidney function is better in the CsA low dose than CsA standard dose in two of three trials and statistically significantly better in two trials.	Moderate
Proteinuria	1 RCT (High)	789 (399)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Sparse data (−1)	Low	No difference between CsA low dose vs. CsA standard dose	Moderate
Lipids	2 RCTs (High)	1145 (582)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	None (0)	Moderate	No difference between CsA low dose vs. CsA standard dose	Moderate
Blood pressure	2 RCTs (High)	1145 (582)	Some limitations (−1)	No important inconsistencies (0)	No uncertainty (0)	Sparse data (−1)	Low	No difference between CsA low dose vs. CsA standard dose	Moderate
Adverse events	1 RCT (High)	328 (164)						There is no evidence for differences in adverse event profiles between CsA low dose vs. CsA standard dose	Depends on outcome
Balance of potential benefits and harm: No net benefit or harm							**Quality of overall evidence:** Low		

CAN, chronic allograft nephropathy; CsA, cyclosporine A; CVD, cardiovascular disease; N, number; NODAT, new onset diabetes after transplant; RCT, randomized controlled trials.
[a] Topic 3.1: CsA low dose vs. CsA standard dose.

Appendix

Table 37: GRADE system for grading quality of evidence

Step 1: Starting grade for quality of evidence based on study design	Step 2: Reduce grade	Step 3: Raise grade	Final grade for quality of evidence for an outcome
Randomized trials = High Observational study = Low Any other evidence = Very Low	*Study quality* −1 level if serious limitations −2 levels if very serious limitations *Consistency* −1 level if important inconsistency *Directness* −1 level if some uncertainty −2 levels if major uncertainty *Other* −1 level if sparse or imprecise data −1 level if high probability of reporting bias	*Strength of association* +1 level is strong,[a] no plausible confounders +2 levels if very strong,[b] no major threats to validity *Other* +1 level if evidence of a dose response gradient +1 level if all residual plausible confounders would have reduced the observed effect	High Moderate Low Very Low

Modified with permission (933); adapted from (932,935).
[a]Strong evidence of association is defined as 'significant RR of >2 (<0.5)' based on consistent evidence from two or more observational studies, with no plausible confounders.
[b]Very strong evidence of association is defined as 'significant RR of >5 (<0.2)' based on direct evidence with no major threats to validity.

extent to which our confidence in an estimate of effect is sufficient to support a particular recommendation (934).

Grading the quality of evidence for each outcome

Following GRADE, the quality of a body of evidence pertaining to a particular outcome of interest was initially categorized based on study design. For questions of interventions, the initial quality grade was 'High' when the body of evidence consisted of RCTs. In theory, the initial grade would have been 'Low' if the evidence consisted of observational studies or 'Very Low' if it consisted of studies of other study designs; however, the quality of bodies of evidence was formally determined only for topics where we performed systematic reviews of RCTs. The grade for the quality of evidence for each intervention/outcome pair was decreased if there were serious limitations to the methodological quality of the aggregate of studies, if there were important inconsistencies in the results across studies, if there was uncertainty about the directness of evidence including limited applicability of the findings to the population of interest, if the data were imprecise or sparse, or if there was thought to be a high likelihood of bias. The final grade for the quality of the evidence for an intervention/outcome pair could be one of the following four grades: 'High,' 'Moderate,' 'Low' or 'Very Low' (Table 37).

Grading the overall quality of evidence

Each clinical outcome was ranked by the Work Group as to its level of clinical importance to the patient. The quality of the overall body of evidence was then determined based on the quality grades for all outcomes of interest, taking into account explicit judgments about the relative importance of each outcome. The resulting four final categories for the quality of overall evidence were: 'A,' 'B,' 'C' or 'D' (Table 38) (932). This evidence grade is indicated within each recommendation.

Assessment of the net health benefit across all important clinical outcomes

The net health benefit was determined based on the anticipated balance of benefits and harm across all clinically important outcomes. The assessment of net medical benefit was affected by the judgment of the Work Group and the ERT. The assessment of net health benefit is summarized in Table 39.

Grading the strength of the recommendations

The strength of a recommendation is graded as Level 1 or Level 2. Table 40 shows the KDIGO nomenclature for grading the strength of a recommendation and the implications of each level for patients, clinicians and policy makers. Recommendations can be for or against doing something. Table 41 shows that the strength of a recommendation is determined not just by the quality of the evidence, but also by other, often complex, judgments regarding the size of

Table 38: Final grade for overall quality of evidence

A: High quality of evidence. We are confident that the true effect lies close to that of the estimate of the effect.
B: Moderate quality of evidence. The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
C: Low quality of evidence. The true effect may be substantially different from the estimate of the effect.
D: Very low quality of evidence. The estimate of effect is very uncertain, and often will be far from the truth.

Table 39: Balance of benefits and harm

When there was evidence to determine the balance of medical benefits and harm of an intervention to a patient, conclusions were categorized as follows:
- Net benefits = the intervention clearly does more good than harm.
- Tradeoffs = there are important tradeoffs between the benefits and harm.
- Uncertain = it is not clear whether the intervention does more good than harm.
- No net benefits = the intervention clearly does not do more good than harm.

the net medical benefit, values and preferences and costs. Formal decision analyses, including cost analysis, were not conducted.

Ungraded statements

The KDIGO consensus statement on grading (933) had recommended a category for a 'consensus-based statement.' This category was designated for guidance by the Work Group based predominantly on expert opinion in areas of low- or very low-quality evidence. However, it became clear that 'consensus-based' was not a distinguishing feature, since all recommendations are supported by Work Group consensus. Still, it was felt that having a category that allows the Work Group to issue general advice would be useful. Typically, an ungraded statement meets the following criteria: it provides guidance based on common sense; it provides reminders of the obvious; it is not sufficiently specific to allow application of evidence to the issue, and therefore it is not based on systematic evidence review. Common examples include recommendations about frequency of testing, referral to specialists and routine medical care. We strove to minimize the use of ungraded recommendations.

This grading scheme with two levels for the strength of a recommendation together with four levels of grading the quality of the evidence, and the option of an ungraded statement for general guidance, was adopted by the KDIGO Board in December 2008.

Format for Guideline Recommendations

Each section contains one or more specific recommendations. Within each recommendation, the strength of recommendation is indicated as level 1 or level 2 and the quality of the supporting evidence is shown as A, B, C or D. These are followed by a brief background with relevant definitions of terms, then the rationale starting with a 'chain of logic,' which consists of declarative sentences summarizing the key points of the evidence base and the judgments supporting the recommendation. This is followed by a narrative in support of the rationale. In relevant sections, research recommendations suggest future research to resolve current uncertainties.

Limitations of Approach

While the literature searches were intended to be comprehensive, they were not exhaustive. MEDLINE and various Cochrane databases were the only databases searched. Hand searches of journals were not performed, and review articles and textbook chapters were not systematically searched. However, important studies known to the domain experts that were missed by the electronic literature searches were added to retrieved articles and reviewed by the Work Group. Not all topics and subtopics covered by this guideline could be thoroughly and systematically reviewed. Decisions to restrict the topics were made to focus the systematic reviews on those topics where existing evidence was thought to be likely to provide support for the guideline. Although nonrandomized studies were reviewed, the majority of the ERT and Work

Table 40: KDIGO nomenclature and description for grading recommendations

	Implications		
Grade*	Patients	Clinicians	Policy
Level 1: 'We recommend'	Most people in your situation would want the recommended course of action and only a small proportion would not	Most patients should receive the recommended course of action	The recommendation can be adopted as a policy in most situations
Level 2: 'We suggest'	The majority of people in your situation would want the recommended course of action, but many would not	Different choices will be appropriate for different patients. Each patient needs help to arrive at a management decision consistent with her or his values and preferences	The recommendation is likely to require debate and involvement of stakeholders before policy can be determined

*The additional category 'Not Graded' was used, typically, to provide guidance based on common sense or where the topic does not allow adequate application of evidence. The most common examples include recommendations regarding monitoring intervals, counseling and referral to other clinical specialists. The ungraded recommendations are generally written as simple declarative statements, but are not meant to be interpreted as being stronger recommendations than Level 1 or 2 recommendations.

Appendix

Table 41: Determinants of strength of recommendations

Factor	Comment
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the more likely a strong recommendation is warranted. The narrower the gradient, the more likely a weak recommendation is warranted.
Quality of the evidence	The higher the quality of evidence, the more likely a strong recommendation is warranted.
Values and preferences	The more variability in values and preferences, or more uncertainty in values and preferences, the more likely a weak recommendation is warranted.
Costs (resource allocation)	The higher the costs of an intervention—that is, the more resources consumed—the less likely a strong recommendation is warranted.

Reprinted with permission (934).

Group resources were devoted to review of randomized trials, since these were deemed to be most likely to provide data to support level 1 recommendations with very high- or high-quality (A or B) evidence. Where randomized trials are lacking, it was deemed to be sufficiently unlikely that studies previously unknown to the Work Group would result in a higher-quality level 1 recommendations. A small number of supplemental sets of evidence were collected with a nonsystematic review approach. Any such evidence that is summarized is noted. Decisions to take a nonsystematic review approach for these topics were made due to time constraints and resource limitations.

Review of the Guideline Development Process

Several tools and checklists have been developed to assess the quality of the methodological process for guideline development. These include the Appraisal of Guidelines for Research and Evaluation (AGREE) criteria (936) and the Conference on Guideline Standardization (COGS) checklist (937). Supporting Table 62 shows the COGS criteria that correspond to the AGREE checklist and how each one of them is addressed in this guideline.