

Evidence-based Practice Center Systematic Review Protocol

Project Title: Systematic Review of Calcineurin Inhibitors for Kidney Transplant

I. Background and Objectives for the Systematic Review

Approximately 17,000 kidney transplants occur each year in the United States, accounting for almost 60% of all organ transplants.¹ Kidney transplantation is the treatment of choice for end stage renal disease. Causes of renal failure are varied, including diabetes, hypertension, glomerular and cystic kidney diseases, and autoimmune disorders. Kidney transplantation offers a better quality of life and a survival benefit for most patients. The 2012 Organ Procurement and Transplantation Network (OPTN) and Scientific Registry of Transplant Recipients (SRTR) annual report showed that the conditional graft half-life (defined as the time to when half of the grafts surviving at least one year are still functioning) was 12.5 years for deceased donor transplants and 15.3 years for living donor transplants in 2009-2010.² Survival rates continue to improve; a recent analysis of more than 250,000 kidney transplant recipients demonstrated that death-censored graft half-life for all deceased donor transplants increased from 10.2 years in 1989 to 14.3 years in 2005, and remained approximately 16.5 years for living donor transplants during the same time period.³

Calcineurin inhibitors (CNIs) are the cornerstone of immunosuppression for kidney transplantation. Cyclosporine and tacrolimus are the most commonly used CNIs in renal transplant recipients during the past 20 years. Cyclosporine was initially approved in 1983 by the U.S. Food and Drug Administration (FDA) for immunosuppression following organ transplantation, and in 1995 a microemulsion formulation of cyclosporine (associated with better bioavailability and more consistent absorption) was approved. Cyclosporine formulations are usually administered twice daily. Tacrolimus received FDA approval in 1994 for liver transplant recipients, and in 1997 for kidney transplants. Tacrolimus is usually administered twice daily, but recently became available as an extended release once-daily formulation. FDA-approved generic equivalents are available for tacrolimus immediate release formulations, as well as modified and unmodified cyclosporine.

Tacrolimus-based regimens are currently the mainstay at most kidney transplant programs in the United States. Over 85% of kidney transplant recipients are discharged from their transplant admission on tacrolimus as part of their maintenance immunosuppressive regimen.² This is largely because tacrolimus is more potent and is associated with less rejection and nephrotoxicity than cyclosporine.⁴ However, tacrolimus is also associated with increased neurotoxicity and gastrointestinal side effects compared to cyclosporine.⁵ It has also been associated with an increased incidence of new onset diabetes and the development of metabolic syndrome, which are significant concerns because the main cause of death among kidney transplant recipients is cardiovascular disease.^{6,7}

CNIs are effective immunosuppressants but they have extensive toxicity profiles. Tacrolimus and cyclosporine both require careful management to ensure sufficient dosing for therapeutic effectiveness while avoiding toxicity. Two primary strategies have been employed to balance

efficacy while limiting side effects: routine monitoring of CNI drug levels to guide dosing adjustments, and minimization of CNI use to the lowest therapeutic levels. Alternatively, CNI use may be withdrawn or avoided entirely in favor of other immunosuppressant therapies.

CNI Monitoring

The primary commercial assays used for monitoring CNI drug levels are mass spectrometry and immunoassays. Cyclosporine is measured with high performance liquid chromatography (HPLC), fluorescence polarization immunoassay (FPIA), enzyme-multiplied-immunoassay techniques (EMIT), or liquid chromatography-tandem mass spectrometry (LC-MS/MS).

Tacrolimus can be monitored with LC-MS/MS, enzyme-linked immunosorbent assay (ELISA), or microparticle enzyme immunoassay (MEIA). Compared with the immunoassays, HPLC and LC-MS/MS offer more precise measures of the parent compound while minimizing measurement of metabolites, but they can also be more expensive, time-consuming, labor-intensive techniques and less standardized making them provider dependent. It is also unclear whether long-term health outcomes vary with each assay methodology.

The ability to accurately measure low-range CNI concentrations with methods such as LC-MS/MS are important as CNI target therapeutic ranges have decreased over time.⁸ The Report of the European Consensus Conference recommended that assays achieve a limit of quantification (LOQ) of 1 ng/mL. However, randomized trials demonstrating the value of CNI monitoring itself are lacking. Moreover, although LC-MS/MS is one of the most popular methods for currently measuring tacrolimus, there is no standardization between laboratories.

Selection of the appropriate timing and target values for measuring CNI drug levels is another important component of clinical care. It is recommended that tacrolimus be monitored at trough levels (usually just prior to morning dose administration) as this timepoint is thought to correlate well with concentration of the drug in circulation. However, a recent publication reported that pooled data from three large randomized controlled trials was unable to find any significant correlations between tacrolimus trough levels at five time points (day 3, 10 and 14, and months 1 and 6 post-transplant) and the incidence of biopsy proven rejection in kidney transplant recipients.⁹

Trough monitoring of cyclosporine (C_T) is also common, but recent research has suggested that monitoring cyclosporine at 2 hours after dosing (C_2) yields effective monitoring of cyclosporine while enabling lower doses and less risk of toxicity.^{10,11} However, C_2 level monitoring is not practical because it needs to be measured within an interval of 2 hours \pm 15 minutes in order to avoid large shift in concentrations, while the trough measurement can be done within 10-14 hours. The question of whether trough monitoring should be replaced with monitoring at C_2 is unresolved, and determining the optimal timepoint can lead to more efficient, safer, and higher value care.

CNI Management and Minimization Strategies

Immunosuppressant regimens designed to reduce or eliminate exposure to CNI toxicity risks have been investigated in recent years.¹² Four alternative approaches to full dose CNI therapy have emerged: 1) CNI minimization, which reduces the amount of the drug administered. This strategy may be undertaken from the time of transplant (de novo), or later post-transplant

(elective) as a result of an adverse event such as nephrotoxicity or BK viral infection; 2) CNI conversion, which tapers CNI dosing at any time post-transplant until achieving full replacement with alternative immunosuppressants. This strategy may be undertaken at any time post-transplant and is usually a result of an unacceptable CNI related adverse event; 3) CNI withdrawal, which slowly eliminates the amount of drug administered early or late post-transplant; 4) CNI avoidance, which substitutes other drugs such as sirolimus or belatacept for immunosuppression. All of these strategies also involve the use of concurrent immunosuppressant agents in standard or low doses, and may also include induction agents to maintain sufficient therapeutic effectiveness. No clear consensus exists about the comparative efficacy and safety of these alternatives to full dose CNI regimens.

Table 1. Alternatives to Full Dose CNI Regimens

Strategy	Definition	Timing
Minimization	Lower dosage of CNI	Planned de novo, or result of adverse event
Conversion	Tapering of CNI dose until eliminated and replaced with other immunosuppressant	Usually result of adverse event
Withdrawal	Tapering of CNI dose until eliminated; may be replaced with other immunosuppressant	Planned de novo or result of adverse event
Avoidance	No CNI given; other immunosuppressant used	Planned de novo

Another important consideration is the treatment of high-risk populations. Advances in immunosuppression and improved transplant outcomes have led to liberalized criteria for donors and recipients (e.g., HIV is no longer a universal contraindication to transplantation.) These patients present special challenges as there are drug-drug interactions between CNIs and protease inhibitors.^{13,14} Additionally, as the volume of patients seeking retransplant grows, the number of highly sensitized patients has increased as has the popularity of desensitization protocols employing high-dose induction and maintenance immunosuppression.¹⁵ As more potent, tacrolimus-based immunosuppression has become the clinical standard, opportunistic infections such as cytomegalovirus (CMV), EBV and BK viremia and nephropathy have emerged as complications, and data suggest these are more common with tacrolimus than with cyclosporine.^{16,17} Immunosuppressant regimens that minimize or avoid CNIs may play an important role in the care of these patients.

II. The Key Questions

During the topic refinement and public comment period, the key questions were revised for clarity and cohesiveness. The key questions were further revised by the Evidence Based Practice (EPC) team and are presented below. The former KQ3 (which compared alternative regimens to full dose CNI use) and KQ4 (which compared alternative regimens to each other) were streamlined into KQ3a. This modification is intended to align the KQ more closely with typical clinical decision making, where an immunosuppressive regimen may be chosen de novo from among multiple strategies; and to reflect the available literature, often comprised of studies that simultaneously assess multiple types of immunosuppressive drug regimens.

KQ1. Monitoring assays for calcineurin inhibitors

KQ1a. In adult renal transplants, how do liquid chromatographic and mass spectrometric analytical techniques compare with immunoassay analysis for therapeutic monitoring of full dosing regimens of the calcineurin inhibitors (CNIs), cyclosporine and tacrolimus?

KQ1b. In adult renal transplants, how do liquid chromatographic and mass spectrometric analytical techniques compare with immunoassay analysis for therapeutic monitoring of lower CNI doses used in minimization, conversion, or withdrawal strategies?

KQ2. Cyclosporine monitoring timepoints

In adult renal transplants, how does two-hour post-administration cyclosporine monitoring (C_2) compare with trough monitoring (C_T) for health outcomes?

KQ3. Management of alternatives to full dose CNI regimens

KQ3a. In adult renal transplants, how do immunosuppressive regimens designed to reduce or eliminate exposure to CNI toxicity compare with each other and with full dose CNI regimens for health outcomes?

KQ3b. How does the type of induction agent (including when no induction is used,) and the use of concurrent immunosuppressive agents, impact outcomes of regimens that reduce or eliminate CNI exposure?

PICOTS

The *population* for all key questions includes adult kidney transplant recipients treated with full dose CNI immunosuppression or an alternative immunosuppressive regimen as specified above. Recipients from all kidney donor types will be included. Retransplant kidney patients will be included, but recipients of multi-organ transplants will be excluded. Populations at increased risk for graft rejection or other adverse outcomes will be included for all key questions but analyzed as subgroups, as described below in the **Methods**, under *Data Synthesis*. With regard to alternative regimens (KQ3), population will include patients switched from standard dosing because of adverse events during therapy for the “intent to treat analysis” in order to determine which patients would be best on alternate regimens.

Specific interventions, comparators, and outcomes are defined for each key question below.

KQ1a and 1b. Monitoring assays for calcineurin inhibitors

- **Interventions:**
 - High performance liquid chromatography (HPLC)
 - Liquid chromatography-tandem mass spectrometry (LC-MS/MS)
- **Comparators:**
 - Fluorescence polarization immunoassay (FPIA)
 - Enzyme-multiplied-immunoassay techniques (EMIT)
 - Enzyme-linked immunosorbent assay (ELISA)
 - Microparticle enzyme immunoassay (MEIA)
- **Outcomes:**
 - **Analytical validity outcomes**

- Analytic accuracy (analytic sensitivity and specificity)
- Analytic precision (e.g., intra-assay agreement, inter-assay agreement, and measurement reproducibility)
- Limit of quantification
- Inter-laboratory comparisons (e.g., inter-laboratory agreement, measurement reproducibility)

Intermediate-term clinical outcomes

- Organ survival
- Acute cellular and/or antibody mediated rejection
 - as defined by study (e.g. ascertained by “for cause” versus “per protocol” biopsies)
 - as defined by Banff criteria used in study
- Chronic allograft injury (e.g. rejection or dysfunction, as defined by study)
- Glomerular filtration rate (GFR), as measured by study
- Serum creatinine
- Immunosuppression regimen changed due to adverse events
- Patient adherence

Long-term clinical outcomes

- All-cause mortality
- Quality of life
- Healthcare utilization
- Impact on provider workflow (if measured in surveys of providers or lab staff, evaluated in time and motion studies, or discussed in grey literature)
- Patient preferences (if measured in surveys, reported in studies of patient adherence, or discussed in grey literature)

Adverse events

- Acute and/or chronic nephrotoxicity
 - include GFR threshold and how measured
- New onset diabetes after transplant
- Major adverse cardiac events
- Malignancy
- Infections
 - timing of infections
 - clinical impact of infections on patients
- Other adverse outcomes (e.g., hyperkalemia, hypomagnesaemia, hyperuricemia, gastrointestinal complications, post-transplant hypertension or hyperlipidemia, proteinuria, hematologic side effects, neurologic complications, hair loss/gain)

KQ2. Cyclosporine monitoring timepoints

- **Intervention:**
 - Two-hour post-administration monitoring of cyclosporine
- **Comparator:**
 - Trough monitoring of cyclosporine

Source: www.effectivehealthcare.ahrq.gov

Published online: October 8, 2014

- **Outcomes:**
 - All outcomes as described for KQ1

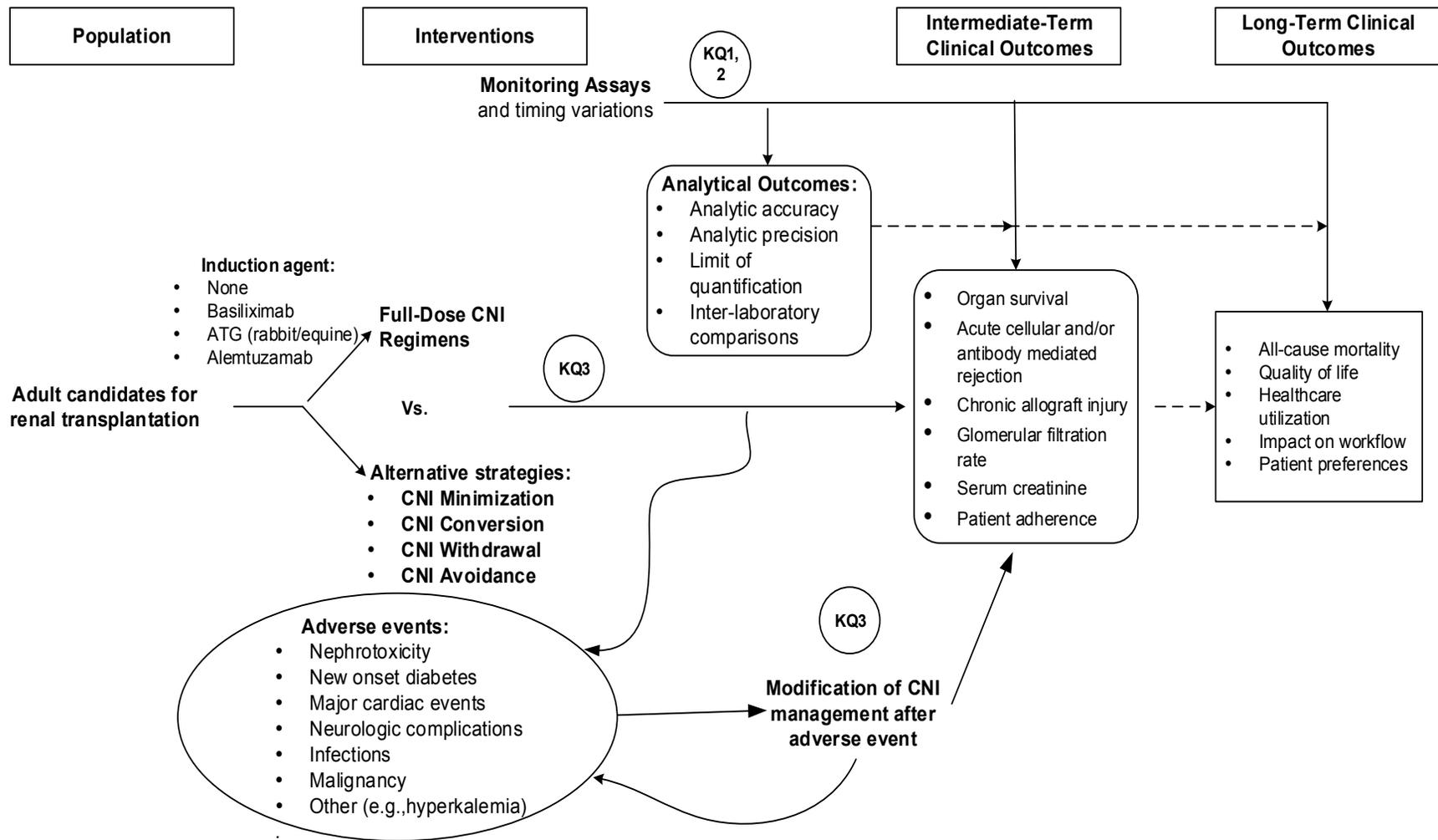
KQ3a and 3b. Management of alternatives to full dose CNI regimens

- **Interventions:**
 - CNI minimization strategies
 - CNI conversion strategies
 - CNI withdrawal strategies
 - CNI avoidance strategies
- **Comparators:**
 - Full dose CNIs
 - CNI minimization/conversion/withdrawal/avoidance strategies compared to each other
- **Outcomes:**
 - All intermediate-term and long-term clinical outcomes, and adverse events, as described for KQ1

The *timing* of patient follow up and drug management for all key questions will include the immediate post-transplant period (through 6 months); short-term follow up (7 months through 1 year), and long-term follow up (more than 1 year).

The *setting* for all key questions will include all settings where immunosuppressive therapy for transplant recipients is administered or monitored.

III. Figure 1. Analytic Framework for Calcineurin Inhibitors for Kidney Transplant



IV. Methods

Criteria for Study Inclusion and Exclusion

As suggested in the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Comparative Effectiveness Reviews, the inclusion criteria are listed below in separate categories pertaining to publication type, study design, patient characteristics, test characteristics, and reported data.¹⁸

Publication Criteria

1. Full-length articles. The article must be published as a full-length, peer-reviewed study. Abstracts and meeting presentations will not be included because they do not include sufficient details about experimental methods to permit an evaluation of study design and conduct; they may also only contain a subset of measured outcomes.^{19,20} Additionally, it is not uncommon for abstracts that are published as part of conference proceedings to have inconsistencies when compared with the final study publication or to describe studies that are never published as full articles.²¹⁻²⁵
2. Redundancy. To avoid double-counting of patients, when several reports of the same or overlapping groups of patients are available, only outcome data from the report with the largest number of patients will be included. We will make an exception and include data from a smaller study when it reports data on an outcome that was not provided by the largest report or reports longer follow-up data for an outcome.
3. English language. Moher et al.²⁶ have demonstrated that exclusion of non-English language studies from meta-analyses has little impact on the conclusions drawn. Juni et al.²⁷ found that non-English studies typically were of lower risk of bias and that excluding them had little effect on effect size estimates in the majority of meta-analyses they examined. Although we recognize that in some situations exclusion of non-English studies could lead to bias, we believe that the few instances in which this may occur do not justify the time and cost typically necessary for translating studies.
4. Publication date. To capture the most relevant data, we will include studies published on or after January 1, 1994. This date was chosen as it reflects the timeframe in which the commonly used forms of CNIs received FDA approval. Tacrolimus received approval in 1994 for use in liver transplants and in 1997 for use in kidney transplants, and the modified formulation of cyclosporine received approval in 1995. Studies published prior to this date are likely to use formulations of CNIs that are no longer in common use.

Study Design Criteria

1. Therapeutic monitoring (KQ1a, KQ1b, KQ2). Studies of any design—randomized, cross-sectional, case-control, or cohort—will be considered for inclusion. Both retrospective and prospective studies will be considered for inclusion, but retrospective studies must use consecutive enrollment or enrollment of a random sample of participants.

- a. For KQ1a, studies must compare mass spectrometry to a commercially available immunoassay, such as enzyme multiplied immunoassay technique (EMIT) for cyclosporine or microparticle enzyme immunoassay (MEIA) for tacrolimus, and report on at least one analytical or patient-oriented clinical outcome (see description of PICOTS in the background section for a description of outcomes).
 - b. For KQ1b, we will look for studies that directly compare different analytical methods (e.g., immunoassay or mass spectrometry) for monitoring lower levels of CNIs. In the absence of such studies, we will include as indirect evidence studies that report measures of analytical validity for individual technologies when used for monitoring lower CNI levels.
 - c. For KQ2, study must include a comparison of C2 monitoring to trough monitoring. Other timepoints may be included in the study and considered in this review.
2. CNI strategies (KQ3a, KQ3b). Comparative studies (randomized or nonrandomized) that compare one CNI strategy to another will be considered for inclusion. CNI strategies of interest include full dose CNIs, reduced dose CNI (CNI-minimization), conversion from CNI to another immunosuppression agent following an adverse event (CNI-conversion), planned withdrawal following initial use of CNI (CNI-withdrawal), or complete avoidance of CNI (CNI-avoidance). Studies will be excluded if they compare only: full dose cyclosporine to full dose tacrolimus; once daily dosing to twice daily dosing of full dose CNIs; or full dose generic equivalents to full dose branded CNIs. Studies that only evaluate drug pharmacokinetics will also be excluded. Finally, studies that examine genotyping related to immunosuppressant metabolism or organ transplantation will be excluded.
 3. Harms. The adverse events and harms reported by any study included to address any of the questions will be used to assess harms and adverse events related to therapeutic monitoring of CNIs and CNI treatment strategies. In addition to these studies, we will also consider other resources, such as patient registries, studies specifically conducted to measure harms, and gray literature sources.

Patient Criteria

1. Type of patient. To be included, the study must have reported data obtained from patients who were recipients of a first or subsequent living (related or non-related) or deceased donor renal transplant (brain dead donors, donors after cardiac death, and expanded criteria donors.) In cases where a study mixes patients receiving different types of solid organ transplants (e.g., mixes renal transplant recipients and liver transplant recipients), the study must report outcomes separately for each organ recipient group.

Studies that include populations of special interest to clinicians and researchers, such as patients at risk for recurrent glomerular disease, patients with diabetes, patients with BK viremia, and patients with hepatitis C virus (HCV) or human

immunodeficiency virus (HIV) will be included. Data for these patient populations will be analyzed separately through subgroup analyses.

This review will not consider studies in which 15 percent or more of patients receive another solid organ in addition to a kidney transplant (e.g., kidney with pancreas), unless findings are reported separately for those with just kidney transplants.

2. Adults. At least 85 percent of patients must have been age 18 or older, or data must have been reported separately for patients age 18 or older.

Intervention Criteria

1. Monitoring method. For KQ1a, KQ1b, and KQ2, the monitoring tests will include commercially available immunoassay tests appropriate to the specific CNI drug. These include high performance liquid chromatography (HPLC), fluorescence polarization immunoassay (FPIA), enzyme multiplied-immunoassay techniques (EMIT), and enzyme-linked immunosorbent assay (ELISA). These assay techniques will be compared to mass spectrometry, including liquid chromatography-tandem mass spectrometry (LC-MS/MS).
2. CNI strategy. Studies will be included if they examine tacrolimus or modified cyclosporine formulations (including FDA approved extended release tacrolimus formulations), with steroid maintenance, withdrawal, or avoidance, and any of the following concomitant maintenance immunosuppressive agents: sirolimus, everolimus, mycophenolate mofetil, enteric coated mycophenolate sodium, azathioprine, and belatacept. This review will not include studies of investigational immunosuppressive agents that are not FDA approved, or studies using non-modified cyclosporine formulations (or formulations not commercially available).
3. Induction agent. Induction agents frequently play an important role in immunosuppressive therapy. For this review, induction agents will be viewed as effect modifiers within the broader immunosuppressive regimens, and treated as covariates in our analysis. Studies that are designed to examine the effectiveness of an induction agent as a primary intervention will not be included.

Studies will be included if they use one or more of the following agents for induction therapy: basiliximab, antithymocyte globulin preparations (rabbit and equine), and alemtuzumab. Studies will also be included if they do not use an induction agent. Studies using daclizumab or muromonab OKT3 or other agents no longer commercially available as induction therapy will not be included.

Data Criteria

1. The study must report data pertaining to one of the outcomes of interest (see the Key Questions section for a list).
2. We will include data from timepoints and outcomes reported from studies with at least 20 patients (or 10 patients per study group) with the condition of interest who represent at least 50% of the patients originally enrolled in the study.

Searching for the Evidence: Literature Search Strategies for Identifying Relevant Studies to Answer the Key Questions

Literature searches will be performed by Medical Librarians within the Evidence-based Practice Center (EPC) Information Center, and will follow established systematic review protocols. We will search the following databases using controlled vocabulary and text words: EMBASE, MEDLINE, PubMed, and The Cochrane Library. Searches will cover the literature published from January 1, 1994 through 2014. Search dates may be adjusted based on the quantity and quality of the available literature.

The following gray literature sources will be searched using text words: ClinicalTrials.gov, Centers for Medicare and Medicaid (CMS) Medicare Coverage Database, Centers for Disease Control and Prevention (CDC), ECRI Health Devices, Healthcare Standards, Internet, Medscape, National Guideline Clearinghouse™ (NGC), the U.S. Food and Drug Administration (FDA), and the websites of relevant organizations (e.g. American Society of Transplantation, American Society of Transplant Surgeons, American Transplant Congress, World Transplant Congress, Centre for Evidence Based Transplantation, Organ Procurement and Transplantation Network, Scientific Registry of Transplant Recipients, Infectious Diseases Society of America, National Kidney Foundation.) An example search strategy is shown in Appendix A.

Literature screening will be performed in duplicate using the database Distiller SR (Evidence Partners, Ottawa, Canada). Literature search results will initially be screened for relevancy. Relevant abstracts will be screened against the inclusion and exclusion criteria in duplicate. Studies that appear to meet the inclusion criteria will be retrieved in full and screened again in duplicate against the inclusion and exclusion criteria. All disagreements will be resolved by consensus discussion among the two original screeners. The literature searches will be updated during the Peer Review process, before finalization of the review.

Data Abstraction and Data Management

Data will be abstracted using Microsoft Word or Excel. Duplicate abstraction on a 10-percent random sample will be used to ensure accuracy. All discrepancies will be resolved by consensus discussion among the two original abstracters and an additional third person as needed. Elements to be abstracted include general study characteristics (e.g., country, setting, study design, enrolled number of patients), patient characteristics (e.g., age, sex, and comorbidities), details of CNI monitoring method (e.g., type of test used, timepoint for monitoring), CNI treatment strategy (e.g., minimization strategy, control strategy), risk of bias items, and outcomes data.

Assessment of Methodological Risk of Bias of Individual Studies

For studies addressing clinical outcomes, we will use items from an internal validity item bank for comparative studies to assess the risk of bias of each individual study. This item bank was developed by ECRI Institute²⁸ and informed by empirical studies of the impact of study design on bias in comparative studies and is consistent with the guidance in the AHRQ Methods Guide for Comparative Effectiveness Reviews.²⁹ Each item chosen will address an

aspect of study design or conduct that can help to protect against bias, such as randomization of group assignment, or blinding outcome assessors to patient group assignment. Each item is phrased as a question that can be answered “Yes,” “No,” or “Not Reported,” and each is phrased such that an answer of “Yes” indicates that the study reported a protection against bias on that aspect. A list of potential items is shown in Appendix B.

Studies will be rated as “Low,” “Medium,” or “High” risk of bias. For a controlled/comparative study to be rated as Low risk of bias, the following questions must all be answered “Yes”: items 1, 2, and 4 (appropriately randomized or used methods to enhance group comparability) and items 6 and 7 (group comparability), and at least 10 of the other questions must be answered “Yes” (see the ECRI item list in Appendix B). The trial will be rated as High risk of bias if all five of the critical items above are answered “No.” The trial will be rated as Moderate risk of bias if it does not meet the criteria for either Low or High.

For studies that address the test performance characteristics of the analytical methods used to monitor CNI levels, we will base our assessment of methodological quality on items selected from the checklist recently proposed by Sun et al.³⁰ This checklist includes items that assess internal validity, reporting adequacy, validity of statistical analysis, and external validity. Like the instrument used to assess comparative studies, each item can be answered “Yes,” “No,” or “Not reported” and each is phrased such that an answer of “Yes” indicates that the study reported a protection against bias on that aspect. The list of items is shown in Appendix B.

Data Synthesis

For studies reporting on patient-oriented clinical outcomes, we plan to perform meta-analysis when appropriate and possible. Decisions about whether meta-analysis is appropriate will depend on the judged clinical homogeneity of the different study populations, monitoring methods, CNI protocols, and outcomes. When meta-analysis is not possible (due to limitations of reported data) or is judged to be inappropriate, the data will be synthesized using a descriptive, narrative review approach.

We will compute effect sizes and measures of variance using standard methods and will perform random-effects meta-analysis using the Hartung-Knapp method.^{31,32} Meta-regression and subgroup analysis will be used to explore possible causes of heterogeneity. Potential covariates include population descriptors, including baseline immunological risk factors for rejection, such as age, race, and transplant type; type of induction agent; concomitant immunosuppressive agents; and type of CNI strategy (in studies using multiple strategies).

Subgroup analyses will be performed to isolate effects potentially associated with specific populations. Subgroups will be identified according to the following criteria: kidney donor type (living donor or deceased donor; for deceased donors: expanded criteria donors, donation after cardiac death, standard donor or CDC high risk donors); patient age; patient ethnicity; retransplants; patients who are immunologically sensitized by a calculated panel reactive antibody (CPRA) or PRA $\geq 20\%$; patients receiving a deceased donor kidney transplant with cold ischemic time (CIT) > 12 hours; patients experiencing delayed graft function defined as requiring dialysis in the first seven days post-transplant; patients who

experience CNI-related adverse events; patients at higher risk for infections; patients with diabetes, HIV, HCV, BK nephropathy or BK viremia; or patients at higher risk for these or other severe comorbidities.

A descriptive, narrative review approach will be used to synthesize data from studies reporting on the analytical accuracy of mass spectrometry and immunoassays to monitor low dose CNIs (KQ1b). This approach will be taken largely due to the complex and heterogeneous nature of this type of data.

Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes

We will use a formal grading system that conforms with the CER Methods Guide Manual recommendations on grading the strength of evidence.^{18,33,34} The primary domains assessed in this system include risk of bias, directness, consistency, precision, and publication bias. Additional domains may be used when appropriate. These domains include dose-response association, all plausible confounders would increase the effect, and strength of association. The output is a rating of the strength of evidence: high, moderate, low, or insufficient. This rating is made separately for each outcome of each comparison of each KQ.

If the evidence is sufficient to permit a conclusion, then the rating is deemed high, moderate, or low. The rating will be provided by two independent raters, and discrepancies will be resolved by consensus. A rating of insufficient will be given when the evidence does not permit a conclusion for the outcome of interest for that KQ. Below, we discuss the primary domains and how they will be considered as input to the rating:

Risk of bias (see the section Assessment of Methodological Risk of Bias of Individual Studies above). Study limitations concern the degree to which the included studies for a given outcome have a high likelihood of adequate protection against bias (i.e., have good internal validity). If the evidence permits a conclusion, and all else being equal, a set of studies at low risk of bias yield a higher strength of evidence rating than a set of studies at moderate or high risk of bias.

Directness. Directness relates to (a) whether evidence links interventions directly to a health outcome of specific importance for the review, and (b) for comparative studies, whether the comparisons are based on head-to-head studies.

Consistency. Consistency is the degree to which included studies find either the same direction or similar magnitude of effect.

Precision. Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome, based on the sufficiency of sample size, number of events and width of confidence intervals relative to a clinically important effect estimate.

Reporting bias. This will be addressed by noting the presence of abstracts or ClinicalTrials.gov entries describing studies that did not subsequently appear as full published articles. If many such studies exist, this will decrease the strength of the evidence.

Assessing Applicability

Several *a priori* factors may limit the applicability of findings. Small sample size will be an important limitation in many studies, and addressing this through meta-analysis may be challenging if there is substantial heterogeneity in study design, intervention, and reporting of outcomes. Imprecision in laboratory results, between and within labs, will also present a challenge to applying our findings. Patient adherence to prescribed CNI regimens is another important factor that may limit the findings, and, similarly, imperfect fidelity to monitoring protocols (i.e., variation in when clinical staff actually collect samples for laboratory testing) represents an inherent limitation.

Several patient subgroup factors may cause or explain heterogeneity of treatment effect. These include donor type, retransplants, co-morbid medical conditions, and highly sensitized patient populations. Variation in combinations of immunosuppressive and inductive agents will also present an important limitation to generalizing the findings.

V. References

1. U.S. Organ Procurement and Transplantation Network (OPTN). Organ Procurement and Transplantation Network home page. [internet]. Washington (D.C.): U.S. Organ Procurement and Transplantation Network (OPTN), U.S. Department of Health and Human Services (HHS); [accessed 2013 Sep 05]. [2 p]. Available: <http://optn.transplant.hrsa.gov/>.
2. Matas AJ, Smith JM, Skeans MA, et al. OPTN/SRTR 2012 Annual Data Report: kidney. *Am J Transplant*. 2014 Jan;14 Suppl 1:11-44. Also available: <http://dx.doi.org/10.1111/ajt.12579>. PMID: 24373166
3. Lamb KE, Lodhi L, Meier-Kriesche HU. Long term renal allograft survival in the United States: a critical reappraisal. *Am J Transplant*. 2011;11:450-62.
4. Ekberg H, TedescoSilva H, Demirbas A, et al. Reduced exposure to calcineurin inhibitors in renal transplantation. *New Eng J Med*. 2007 Dec 20;357(25):2562-75. Also available: <http://dx.doi.org/10.1056/NEJMoa067411>. PMID: 18094377
5. Webster AI, Woodroffe RC, Taylor RS, et al. Tacrolimus versus cyclosporine as primary immunosuppression for kidney transplant recipients. *Cochrane Database Syst Rev*. 2005 Oct 19;4:CD003961.
6. Claes K, Meier-Kriesche HU, Schold JD, et al. Effect of different immunosuppressive regimens on the evolution of distinct metabolic parameters: evidence from the Symphony study. *Nephrol Dial Transplant*. 2012;27:850-7.
7. Almeida CC, Silveira MR, de Araujo VE, et al. Safety of immunosuppressive drugs used as maintenance therapy in kidney transplantation: a systematic review and meta-analysis. *Pharmaceuticals*. 2013;6(10):1170-94. Also available: <http://dx.doi.org/10.3390/ph6101170>. PMID: 24275847
8. Wallemacq P, Armstrong VW, Brunet M, et al. Opportunities to optimize tacrolimus therapy in solid organ transplantation: report of the European Consensus Conference. *Ther Drug Monit*. 2009;31:139-52.
9. Bouamar R, Shuker N, Hesselink DA, et al. Tacrolimus predose concentrations do not predict the risk of acute rejection after renal transplantation: a pooled analysis from three randomized-controlled clinical trials. *Am J Transplant*. 2013 May;13(5):1253-61.
10. Knight SR, Morris PJ. The clinical benefits of cyclosporine C2-level monitoring: a systematic review. *Transplantation*. 2007 Jun 27;83(12):1525-35. Also available: <http://dx.doi.org/10.1097/01.tp.0000268306.41196.2c>.
11. Kim SJ, Prasad GV, Huang M, et al. A comparison of the effects of C2-cyclosporine and C0-tacrolimus on renal function and cardiovascular risk factors in kidney transplant recipients. *Transplantation*. 2006;82:924.
12. Moore J, Middleton L, Cockwell P, et al. Calcineurin inhibitor sparing with mycophenolate in kidney transplantation: a systematic review and meta-analysis. *Transplantation*. 2009;87:591-605.
13. van Maarseveen EM, Rogers CC, TrofeClark J, et al. Drug-drug interactions between antiretroviral and immunosuppressive agents in HIV-infected patients after solid organ transplantation: a review. *AIDS Patient Care STDS*. 2012 Oct;26(10):568-81. Also available: <http://dx.doi.org/10.1089/apc.2012.0169>. PMID: 23025916
14. Blumberg EA, Rogers CC, AST Infectious Diseases Community of Practice. Human immunodeficiency virus in solid organ transplantation. *Am J Transplant*. 2013 Mar;13 Suppl 4:169-78. Also available: <http://dx.doi.org/10.1111/ajt.12109>. PMID: 23465009
15. Montgomery RA, Lonze BE, King KE, et al. Desensitization in HLA-incompatible kidney recipients and survival. *New Eng J Med*. 28 Jul 2011;365(4):318-26. Also available: <http://dx.doi.org/10.1056/NEJMoa1012376>. PMID: 21793744

16. Hirsch HH, Vincenti F, Friman S, et al. Polyomavirus BK replication in de novo kidney transplant patients receiving tacrolimus or cyclosporine: a prospective, randomized, multicenter study. *Am J Transplant*. 2013 Jan;13(1):136-45. Also available: <http://dx.doi.org/10.1111/j.1600-6143.2012.04320.x>. PMID: 23137180
17. Dharnidharka VR, Cherikh WS, Abbott KC. An OPTN analysis of national registry data on treatment of BK virus allograft nephropathy in the United States. *Transplantation*. 2009 Apr 15;87(7):1019-26. Also available: <http://dx.doi.org/10.1097/TP.0b013e31819cc383>. PMID: 19352121
18. Owens D, Lohr K, Atkins D, et al. Grading the strength of a body of evidence when comparing medical interventions. In: Agency for Healthcare Research and Quality (AHRQ). *Methods guide for comparative effectiveness reviews*. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2009 Jul. Also available: http://effectivehealthcare.ahrq.gov/ehc/products/122/328/2009_0805_grading.pdf.
19. Chalmers I, Adams M, Dickersin K, et al. A cohort study of summary reports of controlled trials. *JAMA*. 1990 Mar 9;263(10):1401-5. PMID: 2304219
20. Neinstein LS. A review of Society for Adolescent Medicine abstracts and Journal of Adolescent Health Care articles. *J Adolesc Health Care*. 1987 Mar;8(2):198-203. PMID: 3818406
21. Dunder Y, Dodd S, Williamson P, et al. Case study of the comparison of data from conference abstracts and full-text articles in health technology assessment of rapidly evolving technologies: does it make a difference? *Int J Technol Assess Health Care*. 2006 Jul;22(3):288-94.
22. De Bellefeuille C, Morrison CA, Tannock IF. The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. *Ann Oncol*. 1992 Mar;3(3):187-91. PMID: 1586615
23. Scherer RW, Langenberg P. Full publication of results initially presented in abstracts. In: *Cochrane Library [Cochrane methodology review]*. Issue 2. Oxford: Update Software; 2001 Jan 01 [accessed 2001 Apr 23]. [35 p]. Available: <http://www.cochrane.org/index.htm>.
24. Yentis SM, Campbell FA, Lerman J. Publication of abstracts presented at anaesthesia meetings. *Can J Anaesth*. 1993 Jul;40(7):632-4. PMID: 8403137
25. Marx WF, Cloft HJ, Do HM, et al. The fate of neuroradiologic abstracts presented at national meetings in 1993: rate of subsequent publication in peer-reviewed, indexed journals. *AJNR Am J Neuroradiol*. 1999 Jun-Jul;20(6):1173-7. PMID: 10445467
26. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses? *J Clin Epidemiol*. 2000 Sep;53(9):964-72. PMID: 11004423
27. Juni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol*. 2002 Feb;31(1):115-23. PMID: 11914306
28. Treadwell JR, Tregear SJ, Reston JT, et al. A system for rating the stability and strength of medical evidence. *BMC Med Res Methodol*. 2006;6:52. Also available: <http://www.biomedcentral.com/1471-2288/6/52>. PMID: 17052350
29. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. [AHRQ Publication No. 12-EGC047-EF]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2012 Mar. 33 p. (Methods guide for comparative effectiveness reviews; Also available: <http://www.effectivehealthcare.ahrq.gov>.)

30. Sun F, Bruening W, Erinoff E, et al. Addressing challenges in genetic test evaluation. Evaluation frameworks and assessment of analytic validity Methods research report. (Prepared by the ECRI Institute Evidence-based Practice Center under Contract No. HHS 290-2007-10063-I.) AHRQ publication no. 11-EHC048-EF. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2011 Jun. 107 p. Also available: <http://www.ncbi.nlm.nih.gov/books/NBK56750/pdf/TOC.pdf>.
31. Hartung J, Knapp G. A refined method for the meta-analysis of controlled clinical trials with binary outcome. *Stat Med*. 2001 Dec 30;20(24):3875-89. PMID: 21640820
32. Cornell JE, Mulrow CD, Localio R, et al. Random-effects meta-analysis of inconsistent effects: a time for change. *Ann Intern Med*. 2014 Feb 18;160(4):267-71. Also available: <http://dx.doi.org/10.7326/M13-2886>. PMID: 24727843
33. Owens DK, Lohr KN, Atkins D, et al. Grading the strength of a body of evidence when comparing medical interventions-Agency for Healthcare Research and Quality and the Effective Health Care Program. *J Clin Epidemiol*. 2010 May;63(5):513-23. PMID: 19595577
34. Agency for Healthcare Research and Quality (AHRQ). Methods guide for medical test reviews [AHRQ Publication No. 12-EHC017]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2012 Jun. 188 p. Also available: http://effectivehealthcare.ahrq.gov/ehc/products/246/558/Methods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf.

VI. Definition of Terms

Assay: a laboratory test used to detect the level of a substance, such as a therapeutic drug, in a biologic (usually blood or urine) sample

Calcineurin: a protein phosphatase that activates T cells, causing an immunological response

Calcineurin inhibitor (CNI): type of drug that suppresses the immune system by blocking calcineurin; cyclosporine and tacrolimus are calcineurin inhibitors frequently used as therapy following organ transplant

Graft half life: the estimate of the median survival time of an organ graft after transplant

Immunoassay: an assay that relies on antibodies binding to specific antigens; frequently used to measure levels of therapeutic drugs, such as immunosuppressants

Immunosuppression: reducing the ability of the immune system to fight infection; can be intentionally induced to prevent rejection of a transplanted organ

Induction agent: a drug used to augment the effectiveness of immunosuppressive therapy, usually in the immediate post-transplant phase when risk of organ rejection is greatest

Mass spectrometry: a laboratory test that uses ionization to analyze molecules; effective in detecting therapeutic levels of drugs in a biologic sample

Nephrotoxicity: exposure of the kidneys to poisonous (toxic) chemicals; many drugs, including CNIs, can cause nephrotoxicity and result in kidney damage

Neurotoxicity: exposure of the nervous system to toxic substances, potentially resulting in damage to the brain or nervous system

Trough monitoring: measurement of a drug when it is present at the lowest level in the body; this is usually immediately prior to administration of a new dose

VII. Summary of Protocol Amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol. Example table below:

Table 2. Protocol Amendments Example Table

Date	Section	Original Protocol	Revised Protocol	Rationale
This should be the effective date of the change in protocol.	Specify where the change would be found in the protocol.	Describe the language of the original protocol.	Describe the change in protocol.	Justify why the change will improve the report. If necessary, describe why the change does not introduce bias. Do not use justification as "because the AE/TOO/TEP/Peer reviewer told us to" but explain what the change hopes to accomplish.

VIII. Review of Key Questions

AHRQ posted the key questions on the Effective Health Care Website for public comment. The EPC refined and finalized the key questions after review of the public comments, and input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the key questions are specific and relevant.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than \$10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than \$10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than \$1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than \$1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder

This project was funded under Contract No. HHS290201200011I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Appendix A. Sample of Search Strategy

Table A-1. EMBASE/MEDLINE (Presented in EMBASE.com syntax)

Set Number	Concept	Search Statement
1	Kidney transplantation	'kidney graft'/exp OR 'kidney graft' OR 'kidney transplantation' OR 'renal graft dysfunction'/exp OR 'renal graft dysfunction' OR (kidney OR renal) NEAR/2 (allograft* OR alograft* OR transplant* OR homograft* OR graft*)
2	Immunosuppressive drugs	'tacrolimus'/exp OR tacrolimus OR 'cyclosporin'/exp OR cyclosporin OR 'cyclosporine'/exp OR cyclosporine OR 'ciclosporine'/exp OR ciclosporine OR 'mustopic oint'/exp OR 'mustopic oint' OR 'tsukubaenolide'/exp OR 'tsukubaenolide' OR 'cipol'/exp OR cipol OR 'cyclokat'/exp OR cyclokat OR 'deximune'/exp OR deximune OR 'implanta'/exp OR implanta OR 'immunosporin'/exp OR immunosporin OR imusporin OR 'vekacia'/exp OR vekacia OR 'prograf'/exp OR prograf OR 'advagraf'/exp OR advagraf OR 'hecoria'/exp OR hecoria OR 'neoral'/exp OR 'gengraf'/exp OR gengraf OR 'astagraf'/exp OR astagraf OR 'ol-27-400' OR 'CSA-neoral' OR 'cya-nof' OR neoral
3	CNI	'calcineurin inhibitor'/exp OR calcineurin NEAR/2 inhibit* OR 'cni'
4	Cyclosporin only (for KQ2)	Cyclosporin'/exp OR Cyclosporine OR cyclosporin OR cipol OR cyclokat OR deximune OR implanta OR imusporin OR vekacia OR ciclosporin OR CsA-Neoral OR CyA-NOF OR Neoral OR OL 27-400
5	Combine sets (CNIs)	2 or 3
6	Combine sets (kidney transplant and CNI)	1 and 5
7	Immunoassay/Mass Spectrometry	'immunoassay'/exp OR immunoassay* OR 'mass spectrometry'/exp OR 'mass spectrometry' OR 'high performance liquid chromatography'/exp OR (mass NEAR/1 spectrometr*) OR 'ms' OR 'gc-ms' OR 'hplc-ms' OR 'high performance liquid chromatography' OR 'hplc' OR (fluorescence NEAR/1 polarization) OR 'fpia' OR 'enzyme multiplied immunoassay' OR 'emit' OR 'enzyme linked immunosorbent assay' OR 'elisa' OR 'microparticle enzyme immunoassay' OR 'meia' OR ('liquid chromatography' NEAR/2 'mass spectrometry') OR 'loc-ms' OR 'antibody conjugated magnetic immunoassay' OR ACMIA
8	Low dose CNI's/CNI minimization	'low drug dose'/exp OR 'dosage schedule comparison'/exp OR 'treatment withdrawal'/exp OR 'drug withdrawal'/exp OR ((low OR lower* OR reduce OR reduction OR minimize OR minimization OR minimal OR withdraw* OR avoid* OR eliminate* OR taper* OR alternative OR conversion) NEAR/4 (dose* OR dosing OR dosage* OR drug* OR calcineurin OR tacrolimus OR cyclosporine* OR 'CNI' OR strategy OR strategies OR regimen*))
9	Cyclosporine monitoring timepoints	((('2' OR 'two') NEAR/1 hour*) OR trough OR ((time OR 'time point' OR timepoint* OR timing OR duration) AND (cyclospor* NEAR/2 level*)) OR "c1" OR "c0" OR "c2" OR ('area under' NEXT/1 curve) OR time NEAR/1 series
10	Drug monitoring terms	'drug monitoring'/exp OR 'drug monitoring' OR ((drug OR therapy OR therapeutic) AND (monitor* OR measure* OR surveillance)) OR 'pharmacodynamics'/exp OR 'area under the curve'/exp OR 'pharmacokinetics'/exp OR bioequivalence OR (drug NEAR/3 (clearance OR activation OR adsorp* OR absorp* OR bioavailabilit* OR distribution)) OR ('area under' NEXT/4 curve) OR (limit NEXT/3 quantification) OR 'loq'

Set Number	Concept	Search Statement
11	Diagnostic test filter	('diagnostic accuracy'/exp OR 'diagnosis':lnk OR 'receiver operating characteristic':de OR 'roc curve'/exp OR 'roc curve' OR 'sensitivity and specificity':de OR 'sensitivity' OR 'specificity' OR 'accuracy':de OR 'precision'/exp OR 'precision':de OR 'prediction and forecasting'/exp OR 'prediction and forecasting' OR 'diagnostic error'/exp OR 'diagnostic error' OR 'maximum likelihood method':de OR 'test retest reliability'/exp OR (test NEXT/3 reliability) OR 'reliability'/exp OR 'validity'/exp OR 'measurement repeatability'/exp OR 'likelihood' OR 'predictive value'/exp OR 'predictive value' OR 'ppv' OR ((false OR true) NEAR/1 (positive OR negative)) OR ('area under' NEXT/4 curve) OR (limit NEXT/3 quantification) OR 'loq' OR (('inter assay' OR 'inter-assay' OR 'inter laboratory' OR 'inter-laboratory') NEAR/2 (agreement OR measurement OR reproducibility))
12	RCT or controlled study filter	('randomized controlled trial'/exp OR 'randomized controlled trial' OR 'randomization'/exp OR 'randomization' OR 'double blind procedure'/exp OR 'double blind procedure' OR 'single blind procedure'/exp OR 'single blind procedure' OR 'placebo'/exp OR 'placebo' OR 'latin square design'/exp OR 'latin square design' OR 'crossover procedure'/exp OR 'crossover procedure' OR 'triple blind procedure'/exp OR 'triple blind procedure' OR 'controlled study'/exp OR 'controlled study' OR 'clinical trial'/exp OR 'clinical trial' OR 'comparative study'/exp OR 'comparative study' OR 'cohort analysis'/exp OR 'cohort analysis' OR 'follow up'/exp OR 'follow up' OR 'intermethod comparison'/exp OR 'intermethod comparison' OR 'parallel design'/exp OR 'parallel design' OR 'control group'/exp OR 'control group' OR 'prospective study'/exp OR 'prospective study' OR 'retrospective study'/exp OR 'retrospective study' OR 'case control study'/exp OR 'case control study' OR 'major clinical study'/exp OR 'major clinical study' OR 'evaluation study'/exp OR 'evaluation study' OR random*:de OR random*:ti OR placebo* OR (singl* OR doubl* OR tripl* OR trebl* AND (dummy OR 'blind'/exp OR blind OR sham)) OR 'latin square' OR isrctn* OR actrn* OR (nct* NOT nct))
13	Systematic Review/Meta-analysis filter	('research synthesis' OR pooled OR 'systematic review'/exp OR 'systematic review' OR 'meta analysis'/exp OR 'meta analysis' OR (('evidence base' OR 'evidence based'/exp OR 'evidence based' OR methodol* OR systematic OR quantitative* OR studies OR search*)) AND ('review'/exp OR 'review' OR 'review'/it))
14	Combine sets (KQ1a and KQ1b)	6 AND 7 AND (10 OR 11 OR 12 OR 13)
15	Combine sets (KQ2)	1 AND 4 AND 9 AND 10 AND (11 OR 12 OR 13)
16	Combine sets (KQ3a)	6 AND 8 AND (12 OR 13)
17	Remove unwanted publication types	(14 OR 15 OR 16) NOT ('book'/exp OR 'book' OR 'case report'/exp OR 'case report' OR 'conference paper'/exp OR 'conference paper' OR 'editorial'/exp OR 'editorial' OR 'letter'/exp OR 'letter' OR 'note'/exp OR 'note' OR book:it,pt OR 'edited book':it,pt OR 'case report':it,pt OR 'case reports':it,pt OR comment:it,pt OR conference:it,pt OR editorial:it,pt OR letter:it,pt OR news:it,pt OR note:it,pt OR proceeding:it,pt)

EMBASE.com Syntax:

* = truncation character (wildcard)

NEAR/*n* = search terms within a specified number (*n*) of words from each other in any order

NEXT/*n* = search terms within a specified number (*n*) of words from each other in the order specified

- / = search as a subject heading
- exp = “explodes” controlled vocabulary term (e.g., expands search to all more specific related terms in the vocabulary’s hierarchy)
- mj = denotes a term that has been searched as a major subject heading
- :de = search in the descriptors field (controlled terms and keywords)
- :lnk = floating subheading
- :it,pt. = source item or publication type
- :ti. = limit to title
- :ti,ab. = limit to title and abstract fields

Appendix B. Sample Risk of Bias Items

ECRI Risk of Bias Item List for Controlled or Comparative studies

1. Were patients randomly assigned to the study's groups?
2. Did the study use appropriate randomization methods?
3. Was there concealment of group allocation?
4. For nonrandomized trials, did the study employ any other methods to enhance group comparability?
5. Was the process of assigning patients to groups made independently from physician and patient preference?
6. Did the patients in different study groups have similar levels of performance on the outcome of interest at the time they were assigned to groups?
7. Were the study groups comparable for all other important factors at the time they were assigned to groups?
8. Did the study enroll all suitable patients or consecutive suitable patients?
9. Was the comparison of interest prospectively planned?
10. If the patients received ancillary treatment(s), was there a $\leq 5\%$ difference between groups in the proportion of patients receiving each specific ancillary treatment?
11. Were the two groups treated concurrently?
12. Was compliance with treatment $\geq 85\%$ in both of the study's groups?
13. Were patients blinded to the treatment they received?
14. Was the healthcare provider blinded to the groups to which the patients were assigned?
15. Were those who assessed the patient's outcomes blinded to the group to which the patients were assigned?
16. Was the integrity of blinding of patients, physicians, or outcome assessors tested and found to be preserved?
17. Was the outcome measure of interest objective and was it objectively measured?
18. Was a standard instrument used to measure the outcome?
19. Was there $\leq 15\%$ difference in the length of follow-up for the two groups?
20. Did $\geq 85\%$ of the patients complete the study?
21. Was there a $\leq 15\%$ difference in completion rates in the study's groups?
22. Was the funding for this study derived from a source that would not benefit financially from results in a particular direction?

Risk of Bias Items for Analytic Validity Studies

1. Reporting adequacy: Was the execution of the index test described in sufficient detail to permit replication of the test?
2. Internal validity and reporting adequacy: Are both positive and negative control samples tested in the study?
3. Internal validity and reporting adequacy: Are positive control samples used in the study appropriately verified as “positive”?
4. Internal validity and reporting adequacy: Are negative control materials used in the study appropriately verified/known to be “negative”?
5. Internal validity and reporting adequacy: Are negative control materials used in the study from the same type of tissue, and collected, stored, and processed in the same way that positive control sample materials used clinically for testing will be?
6. Internal validity and reporting adequacy: Were the tests performed with positive or negative control samples being blinded to the testers?
7. Internal validity and reporting adequacy: Were the testing results interpreted with positive or negative control samples being blinded to the interpreters?
8. Internal validity and reporting adequacy: Were criteria for determining a testing result as positive, negative, indeterminate, or uninterpretable appropriate and set a priori?
9. Internal validity and reporting adequacy: For measuring the limit of detection of the test, has the absolute amount of the positive control samples been appropriately measured?
10. Internal validity and reporting adequacy: Has the assay linearity range been established?
11. Internal validity and reporting adequacy: Has the issue of cross-reactivity been thoroughly evaluated?
12. Internal validity and reporting adequacy: Has the reproducibility of the test when performed multiple times on a single specimen been established?
13. External validity and reporting adequacy: Has the reproducibility of the test been adequately established, namely has the reproducibility been assayed across different operators, different instruments, different reagent lots, different days of the week, different laboratories?
14. Internal validity and reporting adequacy: Was the rate of yield of useable results of the test assayed?
15. Validity of statistical analysis and reporting adequacy: Was the statistical analysis performed appropriately?
16. External validity and reporting adequacy: Were the study data from a multisite collaborative, proficiency testing, or interlaboratory exchange programs?

17. External validity and reporting adequacy: Did the testing performed in the study represent routine laboratory testing in preanalytic, analytic and postanalytic aspects?