Methods for guideline development

Kidney International Supplements (2012) 2, 243–251; doi:10.1038/kisup.2012.31

AIM
The overall aim of the project was to create a clinical practice guideline with recommendations for GN, 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.
- Assign topics to systematic review or narrative review.
- Define specific populations, interventions or predictors, and outcomes of interest for systematic review topics.
- 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.
- Incorporate existing systematic reviews and underlying studies.
- Grade quality of the outcomes of each study.
- Tabulate data from articles into summary tables.
- Update the systematic review search.
- 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 of the evidence and other considerations.

The Work Group, KDIGO Co-Chairs, ERT, and NKF support staff met for three 3-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 guidelines. The Work Group included individuals with expertise in adult and pediatric nephrology, epidemiology, and kidney pathology. For support in evidence review, expertise in methods, and guideline development, the NKF contracted with the ERT based at the Tufts Center for Kidney Disease Guideline Development and Implementation at Tufts Medical Center in Boston, Massachusetts, USA. 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.

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 and triaged topics for systematic review and narrative review. The Work Group and ERT further developed and refined each systematic review topic, specified screening criteria, literature search strategies, and data extraction forms.

The ERT performed literature searches, and conducted abstract and article screening. The ERT also coordinated the methodological and analytic processes of the report. In addition, it defined and standardized the methodology in relation to these searches and data extraction, and produced summaries of 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 GN, 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,
### Chapter 3: SSNS in Children

| PICOD criteria |
|----------------|
| **Population** | Steroid sensitive (Any definition), Children, biopsy not required, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin |
| **Intervention** | Long course or alternate day prednisone |
| **Comparator** | Short course or daily prednisone |
| **Outcomes** | Proteinuria, Complete Remission, Relapse |
| **Study design** | RCTs; No minimum follow-up |
| **Minimum N of Subjects** | No minimum N |

### FRNS in Children

| PICOD criteria |
|----------------|
| **Population** | Steroid resistance, Children, biopsy not required, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin |
| **Intervention** | Cyclosporine, Cytoxan, Chlorambucil, Tacrolimus (Prograf), Rituximab, MMF, Levamisole, Plasmapheresis, Mizoribine, AZA |
| **Comparator** | Prednisone and other comparators depending on the study |
| **Outcomes** | Proteinuria, Complete Remission, Relapse |
| **Study design** | RCTs; No minimum follow-up |
| **Nonrandomized comparative studies** | Retrospective comparative or prospective or retrospective single arm cohort |
| **Minimum duration** | 6 months |
| **Minimum N of Subjects** | No minimum N |

### SDNS in Children

| PICOD criteria |
|----------------|
| **Population** | Steroid resistance, Children, biopsy not required, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin |
| **Intervention** | Cyclosporine, Cytoxan, Chlorambucil, Tacrolimus (Prograf), Rituximab, MMF, Levamisole, Plasmapheresis, Mizoribine, AZA |
| **Comparator** | Prednisone and other comparators depending on the study |
| **Outcomes** | Proteinuria, Complete Remission, Relapse |
| **Study design** | RCTs; No minimum follow-up |
| **Nonrandomized comparative studies** | Retrospective comparative or prospective or retrospective single arm cohort |
| **Minimum duration** | 6 months |
| **Minimum N of Subjects** | No minimum N |

### Chapter 4: SRNS in Children

| PICOD criteria |
|----------------|
| **Population** | Steroid resistance (define), Children, biopsy not required, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin |
| **Intervention** | Cyclosporine, Cytoxan, Chlorambucil, Tacrolimus (Prograf), Rituximab, MMF, Levamisole, Plasmapheresis, Mizoribine, AZA |
| **Comparator** | Prednisone and other comparators depending on the study |
| **Outcomes** | Proteinuria, Complete Remission, Relapse |
| **Study design** | RCTs; No minimum follow-up |
| **Nonrandomized comparative studies** | Retrospective comparative or prospective or retrospective single arm cohort |
| **Minimum duration** | 6 months |
| **Minimum N of Subjects** | No minimum N |

### Chapter 5: MCD in Adults (biopsy proven)

| PICOD criteria |
|----------------|
| **Population** | Minimal Change Disease, biopsy-proven, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin |
| **Intervention** | Short course prednisone and Long course prednisone and Cyclosporine, Cytoxan, Chlorambucil, Tacrolimus (Prograf), Rituximab, MMF, Levamisole, Plasmapheresis, Mizoribine, AZA |
| **Comparator** | No treatment, Short course prednisone, Prednisone and other comparators depending on study |
| **Outcomes** | Change in Proteinuria, Complete Remission, Partial Remission, Relapse, GFR, SCr doubling, ESRD, Death |
| **Study design** | RCTs; No minimum follow-up |
| **Nonrandomized comparative studies** | Retrospective comparative or prospective or retrospective single arm cohort |
| **Minimum duration** | 6 months |
| **Minimum N of Subjects** | \( N \geq 10/\text{arm} \) |
**Table 32 | Continued**

| PICOD criteria |
|----------------|

### Chapter 6: FSGS in Adults

**Population**
Population FSGS, by biopsy and list FSGS subtypes, Adults, Define Nephrotic Syndrome: Urine Prot:Cr Ratio & Serum albumin

**Intervention**
Long course prednisone, Cyclosporine +/- ACE-I, MMF +/- ACE-I, Prograf +/- ACE-I, Rituximab +/- ACE-I, Lamivudine +/- ACE-I, Plasmapheresis +/- ACE-I-levamisole, Mizoribine, AZA

**Comparator**
Any treatment

**Outcomes**
Change in Proteinuria, Complete Remission, Partial Remission, Relapse, GFR, SCr doubling, ESRD, Death

**Study design**
RCTs

No minimum follow-up

**Nonrandomized comparative studies**
Retrospective comparative or prospective or retrospective single arm cohort

Minimum duration: 6 mo

**Minimum N of Subjects**
N \( \geq 10 \)/arm

### Chapter 7: MN

**Population**
Biopsy-proven MN

**Intervention**
Steroids alone (any regimen), Alkylating agent (Cyclophosphamide or Chlorambucil), CNI (Cyclosporine or Tacrolimus +/- steroids), IVIG, ACE-I or ARBs +/- steroids, AZA or Mizoribine +/- steroids, Alkylating agent, MMF +/- steroids, ACTH, Rituximab, Eculizumab, Sirolimus, Pentoxifylline, any combination

**Comparator**
Steroids, No treatment, ACE-I or ARBs, Calcineurin inhibitor (Tac, CsA), Alkylating agents

**Outcomes**
All cause mortality, ESRD, CKD 5, RRT, etc., Progression of CKD, SCr increase/GFR decrease, Change in CKD stage, Disease remission, Partial disease remission, Protocol-driven additional treatment of GN, Disease relapse, Quality of life, Proteinuria, Adverse Events: Including cancer, thromboembolic complications, pulmonary embolism, CVD especially acute MI

**Study design**
RCTs; Minimum duration \( \geq 6 \) months for remissions/AE, 5 years for survival

**Minimum N of Subjects**
N \( \geq 10 \)/arm

### Chapter 8: MPGN

**Population**
Biopsy-proven MPGN

**Intervention**
Rituximab, Eculizumab, CNI (CsA, Tac), MMF, Sirolimus, ACE-I & ARBs, Pentoxifylline, IVIG, Treatment of relapse (any), Steroid therapy (any regimen)

**Comparator**
Any

**Outcomes**
Complete & partial remission, Relapse, Categorical changes in proteinuria, Categorical changes in kidney function (Cr, GFR), ESRD, Death/survival, Adverse events

**Study design**
RCTs

Minimum follow-up 6 mo

**Nonrandomized comparative studies**
Prospective or retrospective

Minimum duration: 12 mo

**Minimum N of Subjects**
N \( \geq 20 \)

### Chapter 9: Infection-Related MN

**Population**
Patients with infection associated GN, biopsy-proven, Postinfectious GN

**Intervention**
Antiviral (lamivudine, ribavirin or interferon) for HBV, HCV, Anti-parasitic agents for malaria or other helminthic/protozoal infections. For post infectious GN: any intervention

**Comparator**
Any treatment

**Outcomes**
All cause mortality, ESRD, CKD 5, RRT, etc., Progression of CKD, SCr increase/GFR decrease, Change in CKD stage, Disease remission, Partial disease remission, Protocol-driven additional treatment of GN, Disease relapse, Quality of life, Proteinuria, AE: Including cancer, thromboembolic complications, pulmonary embolism, CVD especially acute MI

**Study design**
RCTs; No minimum duration of follow-up

**Minimum N of Subjects**
For post-infectious: N \( \geq 10 \) for RCTs, N \( \geq 20 \) for observational

### Chapter 10: IgAN

**Population**
Biopsy-proven IgAN, Primary disease only (exclude secondary disease)

**Intervention**
Any

**Comparator**
Any, regardless of ACE-I use, BP control, etc.

**Outcomes**
All cause mortality, ESRD, CKD 5, RRT, etc., Progression of CKD, SCr increase/GFR decrease, Change in CKD stage, Disease remission, Protocol-driven additional treatment of GN, Disease relapse, Proteinuria

**Study design**
RCTs

Minimum follow-up: 6 months

**Minimum N of Subjects**
N \( \geq 10 \)

---

Table 32 continued on following page
### Table 32 | Continued

| PICOD criteria |
|----------------|

**Chapter 11: HSP Nephritis**

- **Population**: Biopsy-proven HSP
- **Intervention**: Any (for RCTs and nonrandomized comparative studies)
- **Comparator**: Any (for RCTs and nonrandomized comparative studies)
- **Outcomes**: All cause mortality, ESRD, CKD 5, RRT, etc., Progression of CKD, SCr increase/GFR decrease, Change in CKD stage, Disease remission, Protocol-driven additional treatment of GN, Disease relapse, Proteinuria
- **Study design**: RCTs and nonrandomized comparative studies
- **Minimum N of Subjects**: N \( \geq 10 \)

**Chapter 12: LN Induction Therapy**

- **Population**: Biopsy-proven Lupus nephritis, class III, IV, V, (also any combination of class V + III or V + IV), Adults and pediatric
- **Intervention**: MMF, Cyclophosphamide, Rituximab, Long duration Cyclophosphamide, i.v. cyclophosphamide, Cyclosporine/Tacrolimus + MMF + Prednisone
- **Comparator**: Cyclophosphamide [p.o. or i.v.], Azathioprine, Cyclophosphamide, EURO protocol cyclophosphamide, p.o. cyclophosphamide, Cyclophosphamide. No addition of hydroxychloroquine
- **Outcomes**: Mortality, Need for RRT/renal survival, Proteinuria, Kidney function preservation in terms of SCr/eGFR such as doubling of SCr, Disease remission and Relapse, Preservation of menses (fertility), Thrombotic and thromboembolic events, Alopecia and other adverse events
- **Study design**: RCTs
- **Minimum follow-up**: 6 months
- **Nonrandomized comparative studies**
- **Prospective study design**: Minimum follow-up: 6 months
- **Minimum N of Subjects**: N \( \geq 10 \)/arm for RCTs and N \( \geq 30 \) for nonrandomized comparative studies

**Chapter 12: LN Maintenance Therapy**

- **Population**: Biopsy-proven Lupus nephritis, class III, IV, V, (also any combination of class V+III or V+IV), Both adults and pediatric
- **Intervention**: Maintenance therapy 1. MMF, 2. MMF, 3. Steroids, Hydroxychloroquine
- **Nonrandomized comparative studies**
- **Comparators**: Etanercept, TNF alpha antagonists (e.g., infliximab, etc), CTLA4-Ig and derivatives, Campath, Abetimus (LJP394)
- **Outcomes**: Mortality, Need for RRT/renal survival, Proteinuria, Kidney function preservation in terms of SCr/eGFR such as doubling of SCr, Disease remission and Relapse, Preservation of menses (fertility), Thrombotic and thromboembolic events, Alopecia and other adverse events
- **Study design**: RCTs
- **Minimum follow-up**: 12 months
- **Nonrandomized comparative studies**
- **Prospective study design**: Minimum follow-up: 12 months
- **Minimum N of Subjects**: N \( \geq 10 \)/arm for RCTs and N \( \geq 30 \) for nonrandomized comparative studies

**Chapter 13: Treatment of Pauci-immune Focal and Segmental Necrotizing GN**

- **Population**: Adults or pediatric population, ANCA Vasculitis, biopsy-proven, Positive ANCA, Wegener’s granulomatosis, microscopic polyangiitis, pauci-immune GN), Churg Strauss syndrome
- **Intervention**: RCTs: Cyclophosphamide+steroids, Cyclophosphamide+steroids+Plasmapheresis/IVIG, MMF, i.v. cyclophosphamide regimens, Pulsed cyclophosphamide, Rituximab
- **Maintenance**: Azathioprine, MMF, Cyclophosphamide, Methotrexate, Cyclosporine, Leflunomide
- **For nonrandomized comparative studies**: MMF, Rituximab, Infliximab, Campath, Abetacept, Cyclosporine, IVIG, Leflunomide
- **Plasmapheresis or immunoadsorption**
- **Comparator**: Cyclophosphamide, Cyclophosphamide+steroids, Cyclophosphamide, p.o. cyclophosphamide regimens, Continuous p.o. cyclophosphamide
- **Maintenance**: Any comparator
- **Outcomes**: Mortality, Kidney survival, Relapse, Disease free survival, Thromboembolism, Proteinuria
- **Coming off dialysis**
- **Study design**: RCTs
- **Minimum follow-up**: 6 months; For maintenance therapy trials, duration at least 1 year
- **Nonrandomized comparative studies**: Prospective or Retrospective study design
- **Minimum follow-up**: 6 months
- **Minimum N of Subjects**: Any N for RCTs and N \( \geq 30 \) for nonrandomized comparative studies

Table 32 continued on following page
refinement of topics

At the first 3-day meeting, Work Group members added comments to the scope-of-work document as prepared by the Work Group Chairs and ERT, 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 aimed to ensure that all topics deemed clinically relevant and worthy of review were identified and addressed. The major topic areas of interest for the care of GN included LN, lupus nephritis; MCD, minimal-change disease; MI, myocardial infarction; MMF, mycophenolate mofetil; MN, membranous nephropathy; mo, month; MPGN, membranoproliferative glomerulonephritis; N, number; PICOD, population, intervention, comparison, outcomes, design (study); p.o., oral; Prot, proteinuria; RCT, randomized controlled trials; RRT, renal replacement therapy; Rx, treatment; SCr, serum creatinine; SDNS, steroid-dependent nephrotic syndrome; SRNS, steroid-resistant nephrotic syndrome; SSNS, steroid-sensitive nephrotic syndrome; Tac, tacrolimus.

Table 32 | Continued

| PICOD criteria |
|----------------|
| Chapter 14: Treatment of Anti-GBM GN |
| Population | Anti-GBM disease, biopsy-proven, Anti-GBM antibody, Adults or pediatric population |
| Intervention | Prednisone+Cyclophosphamide+Plasmapheresis, Prednisone+cyclophosphamide+Immunoadsorption |
| Comparator | Prednisone+Cyclophosphamide, Prednisone+Cyclophosphamide |
| Outcomes | Mortality, Recovery of kidney function, Proteinuria |
| Study design | Any; No minimum follow-up |
| Minimum no. of Subjects | No minimum N |

ACE-I, angiotensin-converting enzyme inhibitors; ACTH, adrenocorticotropic hormone; AE, adverse events; ANCA, anti-neutrophil cytoplasmic antibody; ARB, angiotensin-receptor blocker; AZA, azathioprine; BP, blood pressure; CKD, chronic kidney disease; CNI, calcineurin inhibitors; Cr, creatinine; CsA, cyclosporine; CTLA 4-Ig, CTLA-4 Ig fusion protein; CVD, cardiovascular disease; Cyc, cyclophosphamide; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; FRNS, frequently relapsing nephrotic syndrome; FSGS, focal segmental glomerulonephritis; GFR, glomerular filtration rate; GN, glomerulonephritis; HBV, hepatitis B virus; HCV, hepatitis C virus; HSP, Henoch-Schönlein purpura; IgAN, immunoglobulin A nephropathy; i.v., intravenous; IVG, intravenous immunoglobulin; LN, lupus nephritis; MCD, minimal-change disease; MI, myocardial infarction; MMF, mycophenolate mofetil; MN, membranous nephropathy; mo, month; MPGN, membranoproliferative glomerulonephritis; N, number; PICOD, population, intervention, comparison, outcomes, design (study); p.o., oral; Prot, proteinuria; RCT, randomized controlled trials; RRT, renal replacement therapy; Rx, treatment; SCr, serum creatinine; SDNS, steroid-dependent nephrotic syndrome; SRNS, steroid-resistant nephrotic syndrome; SSNS, steroid-sensitive nephrotic syndrome; Tac, tacrolimus.

Literature Searches and Article Selection

Searches were conducted in MEDLINE and Cochrane through January 20, 2011. All searches were also supplemented by articles identified by Work Group members through November 2011. For detailed search strategies, please see Online Appendix 1.

Search results were screened by the ERT for relevance using predefined eligibility criteria, described below. For questions related to treatment, the systematic search aimed to identify RCTs as described in Table 32. For some topics, nonrandomized comparative trials were also reviewed, in addition to RCTs, to strengthen the evidence base.

For most topics, the minimum sample size was >10. For MCD and FSGS, because of sparse data, smaller studies were included.
For most topics, the minimum duration of follow-up of 6 months was chosen based on clinical reasoning. For the treatments of interest, the proposed effects on patient-important clinical outcomes require long-term exposure and, typically, would not be expected to become evident before several months of follow-up.

In addition, a search was conducted for data on predictors of kidney failure, kidney function, and remission. Only associations from multivariable regression analyses were considered. For these topics, the ERT completed its search in October 5, 2009 and did not update the search.

Included were studies of all patients with glomerular diseases, excluding those with diabetic nephropathy, thrombotic microangiopathy, amyloidosis, Alport’s and other hereditary glomerular diseases, paraproteinemia, and recurrence of GN following kidney transplantation.

Interventions of interest included all treatments for GN, including drugs, herbs, dietary supplements, tonsillectomy, infection prophylaxis, and postdiagnosis tests to determine treatment.

A list of pertinent, published systematic reviews relevant to GN guidelines was generated, organized by topic, and reviewed with the Work Group. If an existing systematic review adequately addressed a question of interest as determined by the Work Group, this was used instead of a de novo systematic review by the ERT. These systematic reviews were then used as the starting points for building the evidence base and supplemented with articles from the ERT’s own searches. 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 literature search within the systematic reviews.

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.

**Literature yield for systematic review topics.** Table 33 summarizes the numbers of abstracts screened, articles retrieved, studies data extracted, and studies included in summary tables.

**Data extraction.** The ERT designed data-extraction forms to tabulate information on various aspects of the primary studies. Data fields for all topics included study setting, patient demographics, eligibility criteria, type of GN, 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.

**Summary tables**

Summary tables were developed to tabulate the data from studies pertinent to each question of intervention. Each summary table contains a brief description of the outcome, baseline characteristics of the population, intervention, comparator results, and methodological quality of each outcome. Baseline characteristics include a description of the study size, country of residence, and baseline kidney function and proteinuria. Intervention and concomitant therapies, and the results, were all captured. 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 and non-RCTs. Separate sets of summary tables were created for predictor studies. Summary tables are available at www.kdigo.org/clinical_practice_guidelines/GN.php.

**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

---

Table 33 | Literature search yield of RCTs

| Topic       | Abstracts identified<sup>a</sup> | Studies retrieved | Studies data-extracted | No. of systematic reviews | No. of summary tables<sup>b</sup> | No. of evidence profiles<sup>b</sup> |
|-------------|----------------------------------|-------------------|------------------------|---------------------------|----------------------------------|-----------------------------------|
| Total       | 13,516                           | 418               | 94                     | 12                        | 72                               | 18                                |

RCTs, randomized controlled trials.

<sup>a</sup>All topics and all study designs combined.

<sup>b</sup>Available at: www.kdigo.org/clinical_practice_guidelines/GN.php.

---

Table 34 | Hierarchy of outcomes

| Hierarchy<sup>a</sup> | Outcomes<sup>b</sup> |
|------------------------|------------------------|
| Critical importance    | Mortality, ESRD, CKD 5, RRT |
| High importance        | Progression of CKD, Disease remission, Protocol-driven additional treatment of GN, Disease relapse, Quality of life |
| Moderate importance    | Partial disease remission, Proteinuria |

<sup>a</sup>This categorization was the consensus of the Work Group for the purposes of this GN guideline only. The lists are not meant to reflect outcome ranking for other areas of kidney disease management. The Work Group acknowledges that not all clinicians, patients or families, or societies would rank all outcomes the same.

CKD, chronic kidney disease; ESRD, end-stage renal disease; GN, glomerulonephritis; RRT, renal replacement therapy.

<sup>b</sup>Outcomes of lesser importance are excluded from review.
study quality was used (Table 35). Given the potential differences in quality of a study for its primary and other outcomes, the methodological quality was assessed for each outcome. Variations of this system have been used in most KDOQI and all KDIGO guidelines, and have been recommended for the US Agency for Healthcare Research and Quality Evidence-based Practice Center program (http://effectivehealthcare.ahrq.gov/repFiles/2007_10DraftMethods Guide.pdf). Each study was given an overall quality grade. Each reported outcome was then evaluated and given an individual quality grade depending on 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 results data for each outcome of interest were extracted including baseline values (when relevant), final values (or number of events), and net differences (between interventions). These included net change in values, RR, OR, HR, and risk difference, as reported by the studies. The CI values of the net differences and their statistical significance were also extracted. When necessary, for categorical outcomes, RR and their 95% CI were calculated based on available data. The calculated data were distinguished from the reported data in the summary tables.

Evidence profiles. Evidence profiles were constructed by the ERT and reviewed and confirmed with the Work Group members. These profiles serve to make transparent to the reader the thinking process of the Work Group in systematically combining evidence and judgments. Each evidence profile was reviewed by Work Group members. Decisions were based on facts and findings from the primary studies listed in corresponding summary tables, as well as selected existing systematic reviews, and judgments of the Work Group. Judgments about the quality, consistency, and directness of evidence were often complex, as were judgments about the importance of an outcome or the summary of effects sizes. The evidence profiles provided a structured transparent approach to grading, rather than a rigorous method of quantitatively summing up grades.

Evidence profiles were constructed for research questions addressed by at least two studies. When the body of evidence for a particular comparison of interest consisted of only one study, either an RCT or a systematic review, the summary table provides the final level of synthesis.

Grading the quality of evidence and the strength of a recommendation. A structured approach, based on GRADE,759–761 and facilitated by the use of evidence profiles, was used 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 Co-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 extent to which our confidence in an estimate of effect is sufficient to support a particular recommendation.760

Grading the quality of evidence for each outcome: Following the GRADE method, 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 (a low event rate [0 or 1 event] in either arm or CI spanning a range <0.5 to >2.0) or sparse (only one study or total N<100), 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 36). The quality of grading for topics relying on systematic reviews are based on quality items recorded in the systematic review.

Grading the overall quality of evidence: 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, weighting critical outcomes more than high or moderate. The resulting four final categories for the quality of overall evidence were: “A”, “B”, “C” or “D” (Table 37). 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. The assessment of net health benefit is summarized in Table 38.

### Table 35 | Classification of study quality

| Quality          | Description                                                                 |
|------------------|-----------------------------------------------------------------------------|
| 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 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. |

RCT, randomized controlled trial.
Grading the strength of the recommendations: The strength of a recommendation is graded as Level 1 or Level 2. Table 39 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 40 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 the net medical benefit, values, and preferences, and costs. Formal decision analyses including cost analysis were not conducted.

Ungraded statements: This category was designed to allow the Work Group to issue general advice. 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

### Table 36 | 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 and definition |
|---|---|---|---|
| Randomized trials = High | Study quality | Strength of association | High = Further research is unlikely to change confidence in the estimate of the effect |
| Observational study = Low | Consistency | Other | Moderate = Further research is likely to have an important impact on confidence in the estimate of effect, and may change the estimate |
| Any other evidence = Very Low | Directness | Other | Low = Further research is very likely to have an important impact on confidence in the estimate, and may change the estimate |

GRADE, Grading of Recommendations Assessment, Development, and Evaluation.  
*Strong evidence of association is defined as “significant relative risk of > 2 (< 0.5)” based on consistent evidence from two or more observational studies, with no plausible confounders.*  
*Very strong evidence of association is defined as “significant relative risk of > 5 (< 0.2)” based on direct evidence with no major threats to validity.*  
*Sparse if there is only one study or if total N < 100. Imprecise if there is a low event rate (0 or 1 event) in either arm or confidence interval spanning a range < 0.5 to > 2.0. Adapted by permission from Macmillan Publishers Ltd: Kidney International. Uhlig K, Macleod A, Craig J et al. Grading evidence and recommendations for clinical practice guidelines in nephrology. A position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int 2006; 70: 2058–2065761; accessed http://www.nature.com/ki/journal/v70/n12/pdf/5001875a.pdf.

### Table 37 | Final grade for overall quality of evidence

| Grade | Quality of evidence | Meaning |
|---|---|---|
| A | High | We are confident that the true effect lies close to that of the estimate of the effect. |
| B | Moderate | 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 | The true effect may be substantially different from the estimate of the effect. |
| D | Very Low | The estimate of effect is very uncertain, and often will be far from the truth. |

### Table 38 | 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
- **Trade-offs** = there are important trade-offs between the benefits and harm
- **Uncertain trade-offs** = 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

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. The Work Group took the primary role of writing the recommendations and rationale statements, and retained final responsibility for the content of the guideline statements and the accompanying narrative. The ERT reviewed draft recommendations and grades for consistency with the conclusions of the evidence review.

**Format for 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 these guidelines could be 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 guidelines. Although nonrandomized studies were reviewed, the majority of the ERT and Work Group resources were devoted to review of the randomized trials, since these were deemed to be most likely to provide data to support level 1 recommendations with very high- or high- (A or B) quality evidence. Where randomized trials were lacking, it was deemed to be sufficiently unlikely that studies previously unknown to the Work Group would result in higher-quality level 1 recommendations.

**Review of the Guideline Development Process**

Several tools and checklists have been developed to assess the quality of the methodological process for systematic review and guideline development. These include the Appraisal of Guidelines for Research and Evaluation (AGREE) criteria,

The Conference on Guideline Standardization (COGS) checklist,

and the Institute of Medicine’s recent Standards for Systematic Reviews and Clinical Practice Guidelines We Can Trust. Online Appendices 2 and 3 show the COGS criteria that correspond to the AGREE checklist and the Institute of Medicine standards, and how each one of them is addressed in this guideline.

**SUPPLEMENTARY MATERIAL**

Appendix 1: Online search strategies.
Appendix 2: The Conference on Guideline Standardization (COGS) checklist for reporting clinical practice guidelines.
Appendix 3: Concurrence with Institute of Medicine standards for systematic reviews and for guidelines.

Supplementary material is linked to the online version of the paper at http://www.kdigo.org/clinical_practice_guidelines/GN.php

---

Table 39 | KDIGO nomenclature and description for grading recommendations

| Gradea | Implications |
|--------|--------------|
| **Patients** | **Clinicians** | **Policy** |
| Level 1 | Level 2 |
| “We recommend” | “We suggest” |
| 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 evaluated as a candidate for developing a policy or a performance measure. |
| 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 substantial debate and involvement of stakeholders before policy can be determined. |

*aThe 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.

Table 40 | Determinants of strength of recommendation

| 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. |