Methods for guideline development

Kidney International Supplements (2013) 3, 287–296; doi:10.1038/kisup.2013.39

AIM
The overall aim of this project was to develop an evidence-based CPG for the management of dyslipidemia and CKD. The guideline consists of recommendation statements, rationales, and a summary of systematically generated evidence on relevant pre-defined clinical topics. The general guideline development method is described at http://www.kdigo.org/home/guidelines/development as well as below.

OVERVIEW OF PROCESS
The development process for the KDIGO Clinical Practice Guideline for Lipid Management in Chronic Kidney Disease included the following steps:
- Appointing Work Group members and the evidence review team (ERT)
- Discussing process, methods, and results
- Developing and refining topics
- Identifying populations, interventions or predictors, and outcomes of interest
- Selecting topics for systematic evidence review
- Standardizing quality assessment methodology
- Developing and implementing literature search strategies
- Screening abstracts and retrieving full-text articles on the basis of pre-defined eligibility criteria
- Creating data extraction forms
- Extracting data and performing critical appraisal of the literature
- Grading the methodology and outcomes in individual studies
- Tabulating data from individual studies into summary tables
- Grading quality of evidence for each outcome across studies, and assessing the overall quality of evidence across outcomes with the aid of evidence profiles
- Grading the strength of recommendations on the basis of the quality of evidence and other considerations
- Finalizing guideline recommendations and supporting rationales
- Sending the guideline draft for peer review to the KDIGO Board of Directors in August 2012 and for public review in November 2012
- Editing the guideline
- Publishing the final version of the guideline

The Work Group Co-Chairs, KDIGO Co-Chairs and ERT met for a two-day meeting to go over the guideline development process, evidence review topics, and systematic review findings. Following this, the Work Group, KDIGO Co-Chairs and KDIGO support staff met held a two-day meeting to revisit the available evidence, formulate recommendation statements, deliberate on rationale for recommendations, and to develop consensus.

Commissioning of Work Group and ERT
The KDIGO Co-Chairs appointed the Work Group Co-Chairs, who then assembled the Work Group of domain experts, including individuals with expertise in internal medicine, adult and pediatric nephrology, cardiology, hypertension, pharmacology, epidemiology, and endocrinology. The Tufts Center for Kidney Disease Guideline Development and Implementation at Tufts Medical Center in Boston, Massachusetts, USA, was contracted to conduct systematic evidence review and provide expertise in guideline development methodology. The ERT consisted of physician—methodologists with expertise in nephrology and evidence-based clinical practice guideline development, a project coordinator, a research assistant, and a project manager/medical writer.

Defining Scope and Topics
The Work Group Co-Chairs and the ERT defined the overall scope and goals of the guideline (including a list of critical and important outcomes) and then drafted a preliminary list of topics and key clinical questions. They also reviewed the topics in the KDOQI guideline,1 which the ERT also had helped to develop. The Work Group and ERT further developed and refined each topic and specified screening criteria, literature search strategies, and data extraction forms (Table 6).

Establishing the Process for Guideline Development
The ERT performed systematic literature searches and organized abstract and article screening. The ERT also coordinated the methodological and analytical processes, and defined and standardized the methodology for performing literature searches, data extraction, and summarizing the evidence. The Work Group took the primary role of writing the recommendation statements and rationales and retained final responsibility for their content. The Work Group Co-Chairs and the ERT prepared the first draft of the scope of work document as a series of open-ended questions to be considered by Work Group members.

Formulating Questions of Interest
Questions of interest were formulated according to the PICODD (Population, Intervention, Comparator, Outcome, study Design and Duration of follow-up) criteria. Details of the PICODD criteria are presented in Table 6.
### Table 6 | Systematic review topics and screening criteria

| Topic | Population | Intervention | Comparator | Outcome | Study design | Minimum duration of follow-up | Minimum N of Subjects |
|-------|------------|--------------|------------|---------|--------------|-----------------------------|-----------------------|
| **Lipid-lowering agents** | Adults and children with CKD of any severity, with or without dyslipidemia and diabetes; kidney transplant recipients; CKD subgroups in large studies of the general population | ≥ 1 lipid-lowering agent (typically statin, niacin, colestipol, or cholestyramine). Excluded dietary supplements, phosphate binders, apheresis, stanols, and sterols. | Active or control | Categorical: All-cause mortality, cardiovascular mortality, clinical cardiovascular events, ESRD; graft failure, doubling of SCR; halving of GFR Continuous: changes in TC, LDL-C, or HDL-C or TGs | RCTs with parallel-group design; systematic reviews, CKD-subgroup analyses of general population RCTs if no evidence of selection bias and of sufficient size | 4 weeks for continuous lipid outcomes; 1 year for clinical outcomes; if general population study, 1 year | ≥ 100 per arm for adults, ≥ 25 per arm for children; if general population study, ≥ 500 per arm for adults or ≥ 100 per arm for children in full study |
| **Diet or lifestyle modification** | Adults and children with CKD of any severity, with or without dyslipidemia and diabetes; kidney transplant recipients; CKD subgroups in large studies of the general population | Weight loss, special diet, or exercise; also structured care vs. usual care | Different diet or lifestyle modification or agent or placebo | Categorical: All-cause mortality, cardiovascular mortality, clinical cardiovascular events, ESRD; graft failure, doubling of SCR; halving of GFR Continuous: changes in TC, LDL-C, or HDL-C or TGs | RCTs with parallel-group design; systematic reviews, CKD-subgroup analyses of general population RCTs if no evidence of selection bias and of sufficient size | 4 weeks for continuous lipid outcomes; 1 year for clinical outcomes; if general population study, 1 year | ≥ 25 per arm |
| **Drug interactions (update of Tables 32-37 in KDOQI 2003 guideline)** | General population | Any statin and any other drug | NA | Change in bioavailability of statin | Systematic reviews | NA | NA |
| **Change in LDL-C level by statin** | General population | Any statin | Other agent or placebo | Change in LDL-C | Systematic review or meta-analysis, 2006-2011 | NA | NA |
| **Adverse events from statin+fibrate therapy** | General population (typically focused on familial hypercholesterolemia or mixed dyslipidemia) | Any statin or statins + any fibrate or fibrates | Any statin or statins alone (also captured data vs. fibrate alone or placebo) | Any adverse event, any serious adverse event, discontinuation owing to drug. AKI, cancer, rhabdomyolysis, myalgia, increased creatine kinase, increased creatinine, increased ALT or AST, any other specified; in children, also measures of growth, development, cognitive function | Any | Any | NA |
| **Frequency of lipids testing** | Any regimen with variable timing of measurement: e.g., more vs. less testing, some vs. no testing | Active or placebo | Measures of compliance, cardiovascular outcomes, mortality | RCTs or systematic reviews | Any | 6 months | ≥ 50 per arm |

Abbreviations: AKI, acute kidney injury; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CKD, chronic kidney disease; ESRD, end-stage renal disease; GFR, glomerular filtration rate; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; NA, not applicable; RCT, randomized controlled trial; SCR, serum creatinine; TC, total cholesterol; TG, triglyceride.
Ranking of Outcomes

The Work Group ranked outcomes of interest on the basis of their importance for informing clinical decision making (Table 7). Mortality, cardiovascular mortality, cardiovascular or cerebrovascular events, ESRD, and graft failure were considered to be of critical importance; doubling of SCr and halving of GFR, high importance; and change in TC, LDL-C, or HDL-C or TG, moderate importance. The importance of adverse events was considered to depend on the event severity.

Literature Searches and Article Selection

Systematic search strategies were developed by the ERT with input from the Work Group Co-Chairs. Modules were created for RCTs, kidney disease, dyslipidemia, and lipid-lowering agents. For the primary search, search terms were limited to the year 2000 and later to capture trials that would affect current clinical practice and because the KDOQI dyslipidemia guideline covered through 2000. Five new topics were added to the KDOQI systematic review for studies in the general population: effect of diet or lifestyle modification; an update of drug interactions with statins and fibrates; changes in LDL-C levels associated with various statins; adverse events from statin and fibrate use; and frequency of lipid-level testing. These searches were not restricted to 2000 and later. The text words or medical subject headings (MeSH) that were included are provided in the Supplemental Appendix 1. In addition, the ERT searched for existing relevant systematic reviews. The final searches were conducted in August 2011. The ERT searched MEDLINE, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews. The ERT also relied on Work Group members to identify large, general population RCTs reporting on CKD subgroups. The search yield was also supplemented by articles provided by Work Group members through June 2013.

For selection of studies, all members of ERT independently and manually screened the abstracts using the computerized screening program Abstrackr. To establish relevance and consensus among reviewers, the entire team screened and achieved consensus on an initial batch of 500 abstracts. A total of 11,337 citations were initially screened. Journal articles reporting original data, meta-analyses, and systematic reviews were selected for evidence review, based on a priori criteria for eligible evidence. Editorials, letters, abstracts, unpublished reports, and articles published in non-peer-reviewed journals were not included. The Work Group also decided to exclude publications from journal supplements because of potential differences in the process of how they are solicited, selected, reviewed, and edited compared to peer-reviewed publications. The overall search yield along with the number of abstracts identified and articles reviewed for each topic are presented in Table 8.

Data Extraction

Data extraction was done by an ERT member. Although no duplicate extraction was independently performed, data from each study was examined by another reviewer to confirm accuracy. The ERT, in consultation with the Work Group Co-Chairs, designed forms to capture data on design, methodology, sample characteristics, interventions, comparators, outcomes, results, and limitations of individual studies. Methodology and outcomes were also systematically graded (see the section on grading below) and recorded during the data extraction process.

Summary Tables

Summary tables were developed for each comparison of interest. Summary tables contain outcomes of interest, relevant population characteristics, description of intervention and comparator, results, and quality grading for each outcome. Categorical outcomes and continuous lipid outcomes were tabulated separately. For studies not exclusively examining CKD populations, only those reporting analysis by CKD subgroups were tabulated.

Work Group members proofed all summary table data and quality assessments. Summary tables are available at http://www.kdigo.org/home/guidelines/lipids.

Table 7 | Hierarchy of outcomes

| Hierarchy | Outcome |
|-----------|---------|
| Critical importance | Mortality, cardiovascular mortality, cardiovascular or cerebrovascular events, ESRD, graft failure |
| High importance | Doubling of SCr or halving of GFR |
| Moderate importance | Change in TC, LDL-C, or HDL-C or TGs |
| Importance dependent on severity | Adverse events |

Abbreviations: ESRD, end-stage renal disease; GFR, glomerular filtration rate; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; SCr, serum creatinine; TC, total cholesterol; TG, triglyceride

Table 8 | Literature yield for RCTs*

| Intervention | Abstracts identified | Articles retrieved | Studies with data extracted | Statin vs. Placebo | Atorvastatin vs. Placebo, kidney transplant recipients | Statin vs. Lifestyle | Statin vs. Placebo, ADPKD | Atorvastatin 80 mg vs. 10 mg | Low vs. Moderate Protein Diet | Ezetimibe vs. Placebo |
|-------------|---------------------|-------------------|-----------------------------|-------------------|------------------------------------------------|-------------------|------------------------|--------------------------|-------------------------|------------------------|
| Agent or Diet/Lifestyle | 11,337 | 120 | 16 | 9 | 2 | 1 | 1 | 1 | 1 | 1 |
| Adverse Events | 11,337 | 89 | 11 | 11 | across all comparisons |

Abbreviations: ADPKD, autosomal dominant polycystic kidney disease; CKD, chronic kidney disease; RCT, randomized controlled trial.

*Counts not listed for the four other topics for various reasons: no RCTs were found for frequency of testing and existing systematic reviews were used for change in LDL-C by statin and drug interactions.
Evidence Profiles
Evidence profiles were constructed to assess the quality and record quality grades and descriptions of effect for each outcome across studies, as well as the quality of overall evidence and description of net benefits or harms of the intervention or comparator across all outcomes. These profiles aim to make the evidence synthesis process transparent. Decisions in the evidence profiles were based on data from the primary studies listed in corresponding summary tables and on judgments of the ERT and Work Group. When the body of evidence for a particular comparison of interest consisted of only one study, the summary table provided the final level of synthesis and an evidence profile was not generated. Each evidence profile was initially constructed by the ERT and then reviewed, edited, and approved by the Work Group. The work products created by the ERT for summarizing the evidence base are listed in Table 9.

Grading of Quality of Evidence for Outcomes of Individual Studies
Methodological quality (internal validity) refers to the design, conduct, and reporting of outcomes of a clinical study. A previously devised three-level classification system for quality assessment was used to grade the overall study quality and quality of all relevant outcomes in the study (Table 10). Grading of individual studies was done by one of the reviewers, then confirmed by another, and finalized in a group meeting. Variations of this system have been used in most KDOQI and all KDIGO guidelines and have been recommended by the US Agency for Healthcare Research and Quality Evidence-based Practice Center program.107

Each study was given an overall quality grade based on its design, methodology (randomization, allocation, blinding, definition of outcomes, appropriate use of statistical methods, etc.), conduct (dropout percentage, outcome assessment methodologies, etc.) and reporting (internal

Table 10 | Classification of study quality

| Study 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 or paper are unlikely to cause major bias. If study of intervention, must be prospective. |
| Poor quality  | High risk of bias or cannot rule out possible significant biases. Poor methods, incomplete data, reporting errors. Prospective or retrospective. |

Abbreviation: RCT, randomized controlled trial.

Table 9 | Work products for the guideline*

| Topic                                             | Summary table of RCTs | Evidence profile |
|---------------------------------------------------|-----------------------|-----------------|
| Lipid-lowering agents or diet/lifestyle modification |                       |                 |
| Atorvastatin vs. atorvastatin                       | +                     | - (single study) |
| Ezetimibe vs. placebo (simvastatin+ezetimibe vs. simvastatin) | +                     | - (single study) |
| Statin vs. placebo in ADPKD                         | +                     | - (single study) |
| Statin vs. placebo in CKD                           | +                     | + (8 studies)    |
| Statin vs. usual care                               | +                     | - (single study) |
| Low vs. moderate protein diet                        | +                     | - (single study) |
| Statin + ezetimibe vs. placebo                      | +                     | - (single study) |
| Statin vs. placebo in kidney transplant recipients  | +                     | + (2 studies)    |
| Statin vs. lifestyle in kidney transplant recipients| +                     | + (2 studies)    |
| Statin vs. placebo in children                      | +                     | - (single study) |
| Statin vs. placebo in CKD with dialysis             | +                     | + (2 studies)    |
| Exercise vs. control                                | +                     | - (single study) |
| Drug interactions (update of Tables 32–37 in KDOQI 2003 guideline) | +                     | - (single pre-existing systematic review) |
| Drug interactions                                   |                       |                 |
| Adverse events from statin+fibrate therapy          |                       |                 |
| Any adverse event                                   | +                     | - (no evidence profiles prepared re adverse events) |
| Serious adverse event                               | +                     | - (no evidence profiles prepared re adverse events) |
| Treatment-related adverse event                     | +                     | - (no evidence profiles prepared re adverse events) |
| Discontinuation due to adverse event                | +                     | - (no evidence profiles prepared re adverse events) |
| Increased ALT or AST                                | +                     | - (no evidence profiles prepared re adverse events) |
| Increased creatine kinase                           | +                     | - (no evidence profiles prepared re adverse events) |
| Increased Scr                                      | +                     | - (no evidence profiles prepared re adverse events) |
| Rhabdomyolysis                                     | +                     | - (no evidence profiles prepared re adverse events) |
| Other adverse event                                 | +                     | - (no evidence profiles prepared re adverse events) |
| Frequency of lipids testing                         | Frequency of lipids testing | - (0 studies) |
| Change in LDL-C by statin                           | Change in LDL-C by statin | +           |
| Change in LDL-C by statin                           | Change in LDL-C by statin | - (single pre-existing systematic review) |

Abbreviations: ADPKD, autosomal dominant polycystic kidney disease; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CKD, chronic kidney disease; LDL-C, low-density lipoprotein cholesterol; RCT, randomized controlled trial; Scr, serum creatinine.

Coding: +: work product is indicated for the topic of interest; −: work product is not indicated for the topic of interest. General population studies with a CKD subgroup are included.
consistency, clarity, thoroughness and precision, etc.). Each reported outcome was then evaluated and given an individual grade depending on the quality of reporting and methodological issues specific to that outcome. However, the quality grade of an individual outcome could not exceed the quality grade for the overall study.

**Grading the Quality of Evidence and the Strength of a Guideline Recommendation**

A structured approach, based on GRADE108–110 and facilitated by the use of evidence profiles, was used 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.109

**Grading the quality of evidence for each outcome across studies.** Following GRADE, the quality of a body of evidence pertaining to a particular outcome of interest was initially categorized on the basis of study design. For questions of interventions, the initial quality grade was ‘High’ if the body of evidence consisted of RCTs, ‘Low’ if it consisted of observational studies, and ‘Very Low’ if it consisted of studies of other study designs. For questions of interventions, the Work Group decided to use only RCTs. The grade for the quality of evidence for each intervention-outcome pair was then lowered 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 a CI spanning a range >1) or sparse (only 1 study or total N <500), or if there was thought to be a high likelihood of bias. Once consensus is reached in a group meeting, 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 11).

**CKD subgroup analyses.** The following criteria were devised to grade the quality of CKD-subgroup analyses of RCTs that were not specifically designed for or limited to individuals with CKD. CKD subgroups were graded only if they were of an acceptable size for the topic of interest (e.g., 50 per arm with CKD for lipid guideline). These criteria will be considered along with the assessment of whether the putative subgroup effect is plausible with regard to direction and size of effect.

For the complete set of subgroup grading criteria, see Figure 3. Briefly, the study quality was graded according to

| Table 11 | 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 |
| – 1 level if serious limitations | +1 level if strong⁵, no plausible confounders | Moderate = Further research is likely to have an important impact on confidence in the estimate of effect, and may change the estimate |
| – 2 levels if very serious limitations | +2 levels if very strong⁶, no major threats to validity | Low = Further research is very likely to have an important impact on confidence in the estimate, and may change the estimate |
| Consistency | Other | | |
| – 1 level if important inconsistency | +1 level if evidence of a dose-response gradient | Very Low = Any estimate of effect is very uncertain |
| Observational study = Low | Directness | +1 level if all residual plausible confounders would have reduced the observed effect | |
| – 1 level if some uncertainty | | |
| – 2 levels if major uncertainty | | |
| Any other evidence = Very Low | Other | +1 level if sparse or imprecise data⁷ | |
| – 1 level if high probability of reporting bias | | |

**Abbreviation:** GRADE, Grading of Recommendations Assessment, Development and Evaluation.

⁵Strong evidence of association is defined as ‘significant relative risk of >2 or <0.5’ based on consistent evidence from two or more observational studies, with no plausible confounders.

⁶Very strong evidence of association is defined as ‘significant relative risk of >5 or <0.2’ based on direct evidence with no major threats to validity.

⁷Sparse if there is only one study or if total N <500. Imprecise if there is a low event rate (0 or 1 event) in either arm or confidence interval spanning a range >1.

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: 2058–2065;110 accessed http://www.nature.com/ki/journal/v70/n12/pdf/5001875a.pdf
Figure 3 | Grading the quality of CKD subgroups of non-CKD trials. CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate.

whether the CKD subgroup defined by kidney function or proteinuria measured at baseline (rather than after the start of treatment), whether or not analyses of subgroups were pre-specified before randomization, whether or not intervention and comparator groups were balanced within the CKD subgroup. The directness of the CKD subgroup analysis was graded on the basis of whether or not the CKD subgroup results of trials that were not specifically designed for CKD were applicable to patients with CKD and also whether or not a test for interaction by baseline kidney function (or the level of proteinuria) was performed.
Grading the overall quality of evidence. The quality of the overall body of evidence was then determined on the basis of the quality grades for all outcomes of interest, taking into account explicit judgments about the relative importance of each outcome. The resulting four final categories for the quality of overall evidence were: ‘A’, ‘B’, ‘C’ or ‘D’ (Table 12).

Assessment of the net health benefit across all important clinical outcomes. The net health benefit was determined on the basis of the anticipated balance of benefits and harms across all clinically important outcomes (Table 13). The assessment of net benefit also involved the judgment of the Work Group and the ERT.

Developing the recommendations. Draft recommendation statements were developed by the Work Group Co-Chairs and Work Group members with input from all Work Group members. The health benefits, side effects, and risks associated with each recommendation were considered when formulating the guideline, as well as information on patient preferences when available. Recommendation statements were revised in a multi-step process during teleconferences and a face-to-face meeting, as well as subsequent drafts by email. All Work Group members provided feedback on final drafts of the recommendation. The final draft was sent for internal and external peer review, and was further revised by the Work Group Co-Chairs and members. All Work Group members approved the final version of the guideline.

Grading the strength of the recommendations. The strength of a recommendation is graded as level 1 or level 2. Table 14 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. Each recommendation includes an explicit link between the quality of the evidence and the strength of the recommendation.

| Table 12 | 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. |

When there was evidence to determine the balance of medical benefits and harms of an intervention to a patient, conclusions were categorized as follows:

- For statistically significant benefit or harm, report as ‘benefit [or harm] of drug X.’
- For non-statistically significant benefit or harm, report as ‘possible benefit [or harm] of drug X.’
- In instances where studies are inconsistent, report as ‘possible benefit [or harm] of drug X.’
- ‘No difference’ can only be reported if a study is not imprecise.
- ‘Insufficient evidence’ is reported if imprecision is a factor.

| Table 13 | Balance of benefits and harms |

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

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

| Table 14 | KDIGO nomenclature and description for grading recommendations |

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

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

| Table 15 | 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 the more uncertainty in values and preferences, the more likely a weak recommendation is warranted. Values and preferences were obtained from the literature where possible or were assessed in the judgment of the Work Group where robust evidence was not identified. |
| Costs (resource allocation) | The higher the costs of an intervention—that is, the more resources consumed—the less likely a strong recommendation is warranted. |
Table 16 | The Conference on Guideline Standardization (COGS)\textsuperscript{112} checklist for reporting clinical practice guidelines

| Topic | Description | Discussed in KDIGO Clinical Practice Guideline for Lipid Management in Chronic Kidney Disease |
|-------|-------------|------------------------------------------------------------------------------------------|
| 1. Overview material | Provide a structured abstract that includes the guideline's release date, status (original, revised, updated), and print and electronic sources. | Abstract and Methods for Guideline Development. |
| 2. Focus | Describe the primary disease/condition and intervention/service/technology that the guideline addresses. Indicate any alternative preventative, diagnostic or therapeutic interventions that were considered during development. | Management of elevated levels of TC, LDL-C, or HDL-C or TGs and lipid-lowering agents in adults and children with CKD (non-dialysis-dependent or dialysis) or a kidney transplant. |
| 3. Goal | Describe the goal that following the guideline is expected to achieve, including the rationale for development of a guideline on this topic. | This CPG is intended to assist the practitioner caring for patients with CKD and dyslipidemia and to prevent deaths, CVD events, and progression to kidney failure while optimizing patients' quality of life. |
| 4. User/setting | Describe the intended users of the guideline (e.g., provider types, patients) and the settings in which the guideline is intended to be used. | Target audience is practicing nephrologists and other healthcare providers for adults and children with CKD and dyslipidemia. |
| 5. Target population | Describe the patient population eligible for guideline recommendations and list any exclusion criteria. | Adults and children with CKD and dyslipidemia. |
| 6. Developer | Identify the organization(s) responsible for guideline development and the names/credentials/potential conflicts of interest of individuals involved in the guideline's development. | Organization: KDIGO Names/credentials/potential conflicts of interest of individuals involved in the guideline’s development are disclosed in the Biographic and Disclosure Information. |
| 7. Funding source/sponsor | Identify the funding source/sponsor and describe its role in developing and/or reporting the guideline. Disclose potential conflict of interest. | KDIGO is supported by the following consortium of sponsors: Abbott, Amgen, Bayer Schering Pharma, Belo Foundation, Bristol-Myers Squibb, Chugai Pharmaceutical, Coca-Cola Company, Dole Food Company, Fresenius Medical Care, Genzyme, Hoffmann-LaRoche, International Society of Nephrology, J.C. Penney, Kyowa Hakko Kirin, NACCO—The Organization for Transplant Professionals, NKF-Board of Directors, Novartis, Pharmacosmos, PUMC Pharmaceutical, Robert and Jane Cizik Foundation, Shire, Takeda Pharmaceutical, Transwestern Commercial Services, Vifor Pharma, and Wyeth. No funding is accepted for the development or reporting of specific guidelines. All stakeholders could participate in open review. |
| 8. Evidence collection | Describe the methods used to search the scientific literature, including the range of dates and databases searched, and criteria applied to filter the retrieved evidence. | Topics were triaged either to a) systematic review, b) systematic search followed by narrative summary, or c) narrative summary. For systematic reviews on treatment with different lipid-lowering agents or lifestyle modifications, we searched for RCTs in MEDLINE, Cochrane Central Registry for trials, and Cochrane database of systematic reviews. Screening criteria are outlined in the Methods for Guideline Development chapter. The search was updated through August 2011 and supplemented by articles identified by Work Group members through June 2013. We also searched for pertinent existing guidelines and systematic reviews. |
| 9. Recommendation grading criteria | Describe the criteria used to rate the quality of evidence that supports the recommendations and the system for describing the strength of the recommendations. Recommendation strength communicates the importance of adherence to a recommendation and is based on both the quality of the evidence and the magnitude of anticipated benefits and harms. | Quality of individual studies was graded in a three-tiered grading system (see Table 10). Quality of evidence and strength of recommendations were graded following the GRADE approach (Tables 12 and 14). The Work Group could provide general guidance in ungraded statements. |
| 10. Method for synthesizing evidence | Describe how evidence was used to create recommendations, e.g., evidence tables, meta-analysis, decision analysis. | For systematic review topics, summary tables and evidence profiles were generated. For recommendations on treatment interventions, the steps outlined by GRADE were followed. The guideline had undergone internal review by the KDIGO Board in August 2012 and external public review in November 2012. Public review comments were compiled and fed back to the Work Group, which considered comments in its revision of the guideline. |
| 11. Prerelease review | Describe how the guideline developer reviewed and/or tested the guidelines prior to release. | The requirement for an update will be assessed in five years from the publication date or earlier if important new evidence becomes available in the interim. Such evidence might, for example, lead to changes to the recommendations or may modify information provided on the balance between benefits and harms of a particular therapeutic intervention. |
| 12. Update plan | State whether or not there is a plan to update the guideline and, if applicable, an expiration date for this version of the guideline. | |
available evidence and the strength of that recommendation. However, Table 15 shows that the strength of a recommendation is determined not only by the quality of the evidence but also by other, often complex judgments regarding the size of the net medical benefit (potential risks versus 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, and it is not sufficiently specific to allow for application of evidence to the issue and therefore it is not based on systematic evidence review. Common examples include recommendations about frequency of testing, referral to specialists, and routine medical care. We strove to minimize the use of ungraded recommendations.

This grading scheme, with two levels for the strength of a recommendation together with four levels of grading the quality of the evidence, as well as the option of an ungraded statement for general guidance, was adopted by the KDIGO Board in December 2008. The Work Group took on 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 guideline recommendations.** Each chapter 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. The recommendation statements and grades are followed by the supporting evidence is shown as A, B, C, or D. The recommendation statements and grades are followed by a supporting rationale with evidence tables if available.

**Table 16 | Continued**

| Topic | Description | Discussed in KDIGO Clinical Practice Guideline for Lipid Management in Chronic Kidney Disease |
|-------|-------------|---------------------------------------------------------------------------------------------|
| 13. Definitions | Define unfamiliar terms and those critical to correct application of the guideline that might be subject to misinterpretation. | Abbreviations and Acronyms. |
| 14. Recommendations and rationale | State the recommended action precisely and the specific circumstances under which to perform it. Justify each recommendation by describing the linkage between the recommendation and its supporting evidence. Indicate the quality of evidence and the recommendation strength, based on the criteria described in Topic 9. | Each guideline chapter contains recommendations for lipid management in CKD patients. Each recommendation builds on a supporting rationale with evidence tables if available. The strength of the recommendation and the quality of evidence are provided in parenthesis within each recommendation. |
| 15. Potential benefits and harms | Describe anticipated benefits and potential risks associated with implementation of guideline recommendations. | The benefits and harm for each comparison of interventions are provided in summary tables and summarized in evidence profiles. The estimated balance between potential benefits and harm was considered when formulating the recommendations. |
| 16. Patient preferences | Describe the role of patient preferences when a recommendation involves a substantial element of personal choice or values. | Recommendations that are level 2 or “discretionary” indicate a greater need to help each patient arrive at a management decision consistent with her or his values and preferences. |
| 17. Algorithm | Provide (when appropriate) a graphical description of the stages and decisions in clinical care described by the guideline. | No overall algorithm. |
| 18. Implementation considerations | Describe anticipated barriers to application of the recommendations. Provide reference to any auxiliary documents for providers or patients that are intended to facilitate implementation. Suggest review criteria for measuring changes in care when the guideline is implemented. | These recommendations are intended for a global audience. Review criteria were not suggested because implementation with prioritization and development of review criteria have to proceed locally. Furthermore, most recommendations are discretionary, requiring substantial discussion among stakeholders before they can be adopted as review criteria. Suggested audit criteria were provided to assess impact of guideline after publication. Research recommendations were also outlined to address current gaps in the evidence base. |

Abbreviations: CKD, chronic kidney disease; CPG, clinical practice guideline; CVD, cardiovascular disease; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HDL-C, high-density lipoprotein cholesterol; KDIGO, Kidney Disease: Improving Global Outcomes; LDC-C, low-density lipoprotein cholesterol; RCT, randomized controlled trial; TC, total cholesterol; TG, triglyceride.
the only database searched. Hand searches of journals were not performed, and review articles and textbook chapters were not systematically searched. However, any important studies known to domain experts that were missed by the electronic literature searches were added to retrieved articles and reviewed by the Work Group.

Review of 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 2) 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. Table 16 and Supplemental Appendix 2 online show, respectively, the criteria which correspond to the COGS checklist and the Institute of Medicine standards, and how each one of them is addressed in this guideline.

SUPPLEMENTARY MATERIAL
Supplemental Appendix 1: Online search strategies.
Supplemental Appendix 2: 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/home/guidelines/lipids