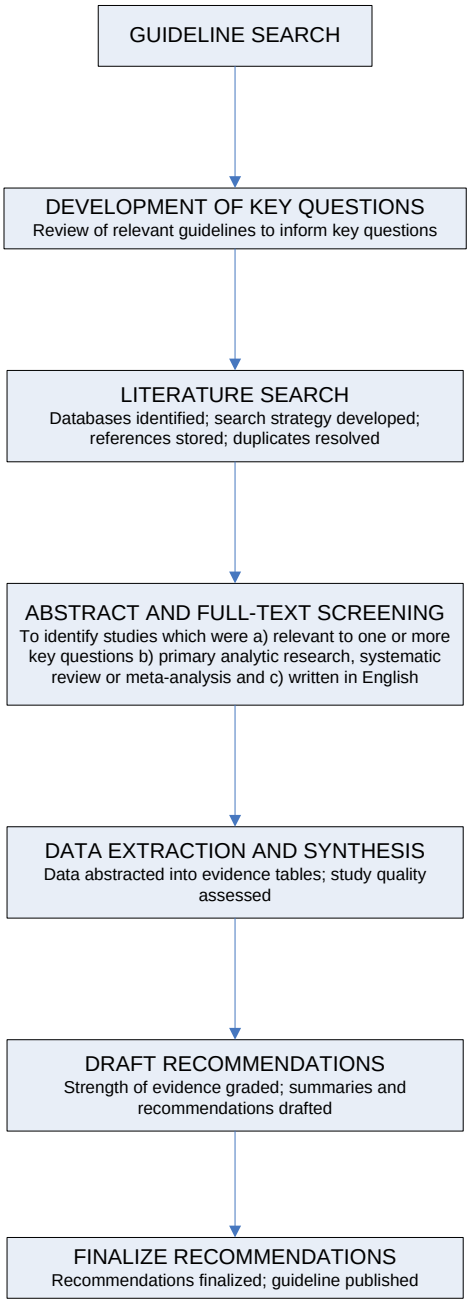
VII. Methods

This guideline was based on a targeted systematic review of the best available evidence on CAUTI prevention. We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [32-34] to provide explicit links between the available evidence and the resulting recommendations. Our guideline development process is outlined in *Figure 1.*

Figure 1. The Guideline Development Process



Development of Key Questions

We first conducted an electronic search of the National Guideline Clearinghouse® (Agency for Healthcare Research and Quality), Medline® (National Library of Medicine) using the Ovid® Platform (Ovid Technologies, Wolters Kluwer, New York, NY), the Cochrane® Health Technology Assessment Database (Cochrane Collaboration, Oxford, UK), the NIH Consensus Development Program, and the United States Preventive Services Task Force database for existing national and international guidelines relevant to CAUTI. The strategy used for the guideline search and the search results can be found in Appendix 1A. A preliminary list of key questions was developed from a review of the relevant guidelines identified in the search.[1,35,36] Key questions were finalized after vetting them with a panel of content experts and HICPAC members.

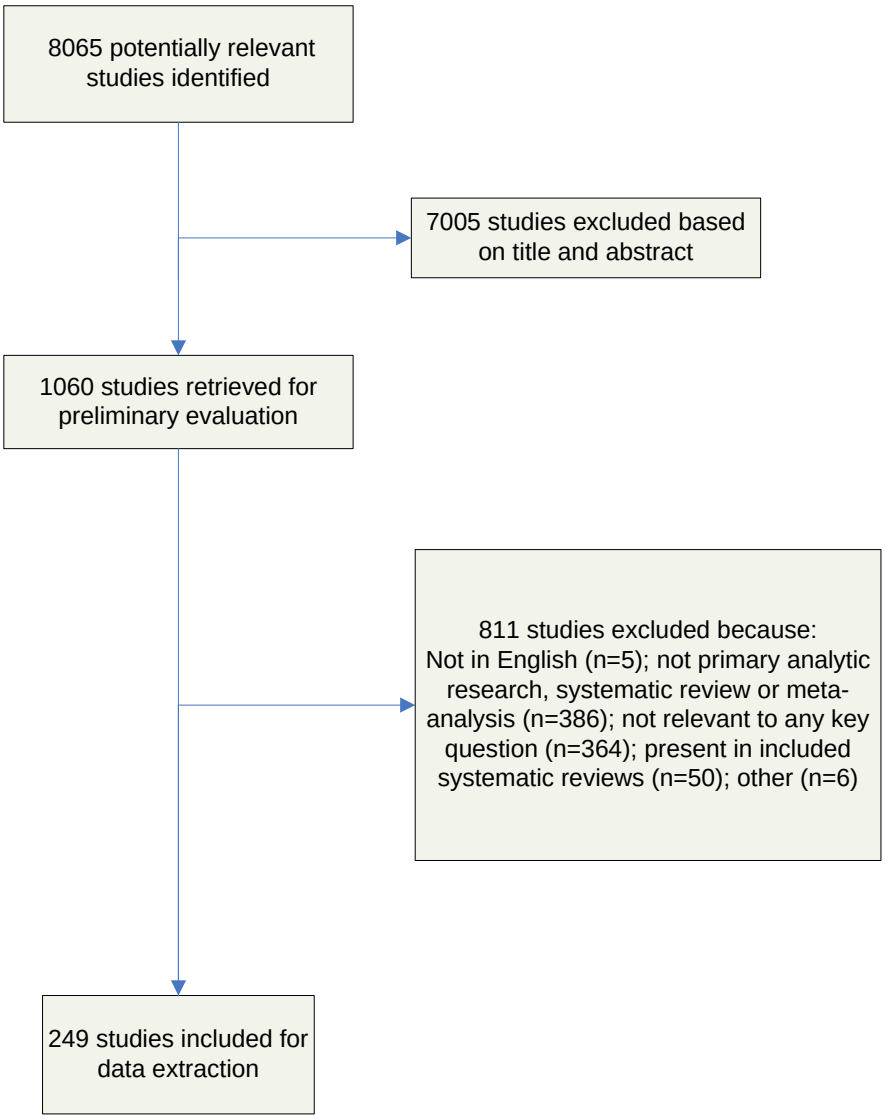
Literature Search

Following the development of the key questions, search terms were developed for identifying literature relevant to the key questions. For the purposes of quality assurance, we compared these terms to those used in relevant seminal studies and guidelines. These search terms were then incorporated into search strategies for the relevant electronic databases. Searches were performed in Medline® (National Library of Medicine) using the Ovid® Platform (Ovid Technologies, Wolters Kluwer, New York, NY), EMBASE® (Elsevier BV, Amsterdam, Netherlands), CINAHL® (Ebsco Publishing, Ipswich, MA) and Cochrane® (Cochrane Collaboration, Oxford, UK) (all databases were searched in July 2007), and the resulting references were imported into a reference manager, where duplicates were resolved. For Cochrane reviews ultimately included in our guideline, we checked for updates in July 2008. The detailed search strategy used for identifying primary literature and the results of the search can be found in Appendix 1B.

Study Selection

Titles and abstracts from references were screened by a single author (C.V.G, R.K.A., or D.A.P.) and the full text articles were retrieved if they were 1) relevant to one or more key questions, 2) primary analytic research, systematic reviews or meta-analyses, and 3) written in English. Likewise, the full-text articles were screened by a single author (C.V.G. or D.A.P.) using the same criteria, and included studies underwent a second review for inclusion by another author (R.K.A.). Disagreements were resolved by the remaining authors. The results of this process are depicted in *Figure 2.*

Figure 2: Results of the Study Selection Process



Data Extraction and Synthesis

Data on the study author, year, design, objective, population, setting, sample size, power, follow-up, and definitions and results of clinically relevant outcomes were extracted into evidence tables (Appendix 2). Three evidence tables were developed, each of which represented one of our key questions. Studies were extracted into the most relevant evidence table. Then, studies were organized by the common themes that emerged within each evidence table. Data were extracted by one author (R.K.A.) and cross-checked by another (C.V.G.). Disagreements were resolved by the remaining authors. Data and analyses were extracted as originally presented in the included studies. Meta-analyses were performed only where their use was deemed critical to a recommendation, and only in circumstances where multiple studies with sufficiently homogenous populations, interventions, and outcomes could be analyzed. Systematic reviews were included in our review. To avoid duplication of data, we excluded primary studies if they were also included in a systematic review captured by our search. The only exception to this was if the primary study also addressed a relevant question that was outside the scope of the included systematic review. Before exclusion, data from the primary studies that we originally captured were abstracted into the evidence tables and reviewed. We also excluded systematic reviews that analyzed primary studies that were fully captured in a more recent systematic review. The only exception to this was if the older systematic review also addressed a relevant question that was outside the scope of the newer systematic review. To ensure that all relevant studies were captured in the search, the bibliography was vetted by a panel of clinical experts.

Grading of Evidence

First, the quality of each study was assessed using scales adapted from existing methodology checklists, and scores were recorded in the evidence tables. Appendix 3 includes the sets of questions we used to assess the quality of each of the major study designs. Next, the quality of the evidence base was assessed using methods adapted from the GRADE Working Group.[32] Briefly, GRADE tables were developed for each of the interventions or questions addressed within the evidence tables. Included in the GRADE tables were the intervention of interest, any outcomes listed in the evidence tables that were judged to be clinically important, the quantity and type of evidence for each outcome, the relevant findings, and the GRADE of evidence for each outcome, as well as an overall GRADE of the evidence base for the given intervention or question. The initial GRADE of evidence for each outcome was deemed high if the evidence base included a randomized controlled trial (RCT) or a systematic review of RCTs, low if the evidence base included only observational studies, or very low if the evidence base consisted only of uncontrolled studies. The initial GRADE could then be modified by eight criteria.[34] Criteria which could decrease the GRADE of an evidence base included quality, consistency, directness, precision, and publication bias. Criteria that could increase the GRADE included a large magnitude of effect, a dose-response gradient, or inclusion of unmeasured confounders that would increase the magnitude of effect (Table 3). GRADE definitions are as follows:

1. High - further research is very unlikely to change confidence in the estimate of effect
2. Moderate - further research is likely to affect confidence in the estimate of effect and may change the estimate
3. Low - further research is very likely to affect confidence in the estimate of effect and is likely to change the estimate
4. Very low - any estimate of effect is very uncertain

After determining the GRADE of the evidence base for each outcome of a given intervention or question, we calculated the overall GRADE of the evidence base for that intervention or question. The overall GRADE was based on the lowest GRADE for the outcomes deemed critical to making a recommendation.

Table 3. Rating the Quality of Evidence Using the GRADE Approach

Type of Evidence	Initial Grade	Criteria to Decrease Grade	Criteria to Increase Grade	Overall Quality Grade
RCT	High	Quality Serious (-1 grade) or very serious (-2 grades) limitation to study quality Consistency Important inconsistency (-1 grade) Directness Some (-1 grade) or major (-2 grades) uncertainty about directness Precision Imprecise or sparse data (-1 grade) Publication bias High risk of bias (-1 grade)	Strong association Strong (+1 grade) or very strong evidence of association (+2 grades) Dose-response Evidence of a dose-response gradient (+1 grade) Unmeasured Confounders Inclusion of unmeasured confounders increases the magnitude of effect (+1 grade)	High Moderate Low Very low
Observational study	Low			
Any other evidence (e.g., expert opinion)	Very low			

Formulating Recommendations

Narrative evidence summaries were then drafted by the working group using the evidence and GRADE tables. One summary was written for each theme that emerged under each key question. The working group then used the narrative evidence summaries to develop guideline recommendations. Factors determining the strength of a recommendation included 1) the values and preferences used to determine which outcomes were "critical," 2) the harms and benefits that result from weighing the "critical" outcomes, and 3) the overall GRADE of the evidence base for the given intervention or question (Table 4).[33] If weighing the "critical outcomes" for a given intervention or question resulted in a "net benefit" or a "net harm," then a "Category I Recommendation" was formulated to strongly recommend for or against the given intervention respectively. If weighing the "critical outcomes" for a given intervention or question resulted in a "trade off" between benefits and harms, then a "Category II Recommendation" was formulated to recommend that providers or institutions consider the intervention when deemed appropriate. If weighing the "critical outcomes" for a given intervention or question resulted in

an "uncertain trade off" between benefits and harms, then a "No Recommendation" was formulated to reflect this uncertainty.

Table 4. Formulating Recommendations		
HICPAC Recommendation	**Weighing Benefits and Harms for Critical Outcomes**	**Quality of Evidence**
STRONG (I)	Interventions with net benefits or net harms	IA – High to Moderate IB – Low or Very Low (Accepted Practice) IC – High to Very Low (Regulatory)
WEAK (II)	Inteventions with trade offs between benefits and harms	High to Very Low
No recommendation/ unresolved issue	Uncertain trade offs between benefits and harms	Low to Very Low

For Category I recommendations, levels A and B represent the quality of the evidence underlying the recommendation, with A representing high to moderate quality evidence and B representing low quality evidence or, in the case of an established standard (e.g., aseptic technique, education and training), very low quality to no evidence based on our literature review. For IB recommendations, although there may be low to very low quality or even no available evidence directly supporting the benefits of the intervention, the theoretical benefits are clear, and the theoretical risks are marginal. Level C represents practices required by state or federal regulation, regardless of the quality of evidence. It is important to note that the strength of a Category IA recommendation is equivalent to that of a Category IB or IC recommendation; it is only the quality of the evidence underlying the IA recommendation that makes it different from a IB.

In some instances, multiple recommendations emerged from a single narrative evidence summary. The new HICPAC categorization scheme for recommendations is provided in Table 1, which is reproduced below.

Table 1. Modified HICPAC Categorization Scheme for Recommendations	
Category IA	A strong recommendation supported by high to moderate quality evidence suggesting net clinical benefits or harms
Category IB	A strong recommendation supported by low quality evidence suggesting net clinical benefits or harms or an accepted practice (e.g., aseptic technique) supported by low to very low quality evidence
Category IC	A strong recommendation required by state or federal regulation.
Category II	A weak recommendation supported by any quality evidence suggesting a trade off between clinical benefits and harms
No recommendation/ unresolved issue	Unresolved issue for which there is low to very low quality evidence with uncertain trade offs between benefits and harms

Category I recommendations are defined as strong recommendations with the following implications:

1. For patients: Most people in the patient's situation would want the recommended course of action and only a small proportion would not; request discussion if the intervention is not offered.
2. For clinicians: Most patients should receive the recommended course of action.
3. For policymakers: The recommendation may be adopted as a policy.

Category II recommendations are defined as weak recommendations with the following implications:

1. For patients: Most people in the patient's situation would want the recommended course of action, but many would not.
2. For clinicians: Different choices will be appropriate for different patients, and clinicians must help each patient to arrive at a management decision consistent with her or his values and preferences.
3. For policymakers: Policy making will require substantial debate and involvement of many stakeholders.

It should be noted that Category II recommendations are discretionary for the individual institution and are not intended to be enforced.

The wording of each recommendation was carefully selected to reflect the recommendation's strength. In most cases, we used the active voice when writing Category I recommendations - the strong recommendations. Phrases like "do" or "do not" and verbs without auxiliaries or conditionals were used to convey certainty. We used a more passive voice when writing Category II recommendations - the weak recommendations. Words like "consider" and phrases like "is preferable," "is suggested," "is not suggested," or "is not recommended" were chosen to reflect the lesser certainty of the Category II recommendations. Rather than a simple statement of fact, each recommendation is actionable, describing precisely a proposed action to take.

The category "No recommendation/unresolved issue" was most commonly applied to situations where either 1) the overall quality of the evidence base for a given intervention was low to very low and there was no consensus on the benefit of the intervention or 2) there was no published evidence on outcomes deemed critical to weighing the risks and benefits of a given intervention. If the latter was the case, those critical outcomes will be noted at the end of the relevant evidence summary.

Our evidence-based recommendations were cross-checked with those from guidelines identified in our original systematic search. Recommendations from previous guidelines for topics not directly addressed by our systematic review of the evidence were included in our "Summary of Recommendations" if they were deemed critical to the target users of this guideline. Unlike recommendations informed by our literature search, these recommendations are not linked to a key question. These recommendations were agreed upon by expert consensus and are designated either IB if they represent a strong recommendation based on accepted practices (e.g., aseptic technique) or II if they are a suggestion based on a probable net benefit despite limited evidence.
All recommendations were approved by HICPAC. Recommendations focused only on efficacy, effectiveness, and safety. The optimal use of these guidelines should include a consideration of the costs relevant to the local setting of guideline users.

Reviewing and Finalizing the Guideline

After a draft of the tables, narrative summaries, and recommendations was completed, the working group shared the draft with the expert panel for in-depth review. While the expert panel was reviewing this draft, the working group completed the remaining sections of the guideline, including the executive summary, background, scope and purpose, methods, summary of recommendations, and recommendations for guideline implementation, audit, and further research. The working group then made revisions to the draft based on feedback from members of the expert panel and presented the entire draft guideline to HICPAC for review. The guideline was then posted on the Federal Register for public comment. After a period of public comment, the guideline was revised accordingly, and the changes were reviewed and voted on by HICPAC. The final guideline was cleared internally by CDC and published and posted on the HICPAC website.

Updating the Guideline

Future revisions to this guideline will be dictated by new research and technological advancements for preventing CAUTI and will occur at the request of HICPAC.