



[Home](#) » [Editorial](#) » [Policies](#)

Grading Guide

Abstract

While grading strength of recommendations and quality of underlying evidence enhances the usefulness of clinical guidelines, the profusion of guideline grading systems undermines the value of the grading exercise. The international GRADE group has suggested an approach that may be useful for many groups developing guidelines; UpToDate has adopted the GRADE approach. The grading scheme classifies recommendations as strong (Grade 1) or weak (Grade 2), according to the balance between benefits, risks, burden, and cost, and the degree of confidence in estimates of benefits, risks, and burden. The system classifies quality of evidence (as reflected in confidence in estimates of effects) as high (Grade A), moderate (Grade B), or low (Grade C) according to factors that include the risk of bias, precision of estimates, the consistency of the results, and the directness of the evidence.

Introduction

Treatment decisions involve a trade off between benefits on the one hand, and risks, burden, and costs on the other. UpToDate provides recommendations for management of typical patients. To integrate these recommendations with their own clinical judgement, and with individual patient's values and preferences, clinicians need to understand the basis for the recommendations that expert guidelines offer. A systematic approach to grading the strength of management recommendations can minimize bias and aid interpretation. Indeed, most guideline groups have accepted the necessity for some sort of grading scheme.

While grading of recommendations represents a positive development for guideline development and interpretation, the proliferation of grading systems has proved an unfortunate consequence. Methodologists and guideline developers have given much thought and effort to considering criteria and approaches to an optimal grading system. An international group of guideline developers and methodologists (the GRADE group) has developed a system that remedies some of the disadvantages of prior systems, and that a number of organizations and groups are finding useful [1].

Strength of recommendation

UpToDate should make recommendations to administer, or not administer an intervention on the basis of tradeoffs between benefits on the one hand, and risks, burden, and costs on the other. If desirable consequences outweigh undesirable consequences, experts will recommend that clinicians

offer a treatment to appropriately chosen patients. The uncertainty associated with the trade-off between desirable and undesirable consequences will determine the strength of recommendations.

The UpToDate approach classifies recommendations into two levels, strong and weak ([see table 1](#)). A two-level grading system has the merit of simplicity. Two levels also facilitate a clear interpretation of the implications of strong and weak recommendations by clinicians [2]. We offer three ways editors and clinicians can interpret strong and weak recommendations.

If editors are confident that, on the basis of the existing evidence, most or all patients will be best served by a particular management strategy, they will make a strong recommendation: Grade 1. This confidence can arise in a number of ways, the most common of which is that high quality evidence may provide precise estimates of both benefits and risk, and the balance may be clear (recommendation for statins in patients with known atherosclerotic disease). Less frequently, a strong recommendation may follow when high quality evidence suggests that two therapies share equivalent benefits, and low quality evidence points to appreciably more harm in one than the other (recommendation for acetaminophen over aspirin in children with chickenpox). If they believe that benefits and risks and burdens are finely balanced, or appreciable uncertainty exists about the magnitude of both benefits and risks, they offer a weak (Grade 2) recommendation.

Clinicians are becoming increasingly aware of the importance of patient values and preferences in clinical decision making [3]. A second way to interpret strong and weak recommendations is in relation to patient values and preferences. For decisions in which it is clear that benefits far outweigh risks, or risks far outweigh benefits, virtually all patients will make the same choice (see Box 1). In such instances, editors can offer a strong (Grade 1) recommendation. In contrast, there are other choices in which patient values and preferences will play a crucial role and in which patients will, as a result, make different choices (see Box 2). When, across the range of patient values, fully informed patients are liable to make different choices, editors should offer weak (Grade 2) recommendations.

Box 1: Short-term aspirin reduces the relative risk of death after myocardial infarction by approximately 25 percent. Aspirin has minimal side effects and very low cost. Peoples' values and preferences are such that virtually all patients suffering a myocardial infarction would, if they understood the choice they were making, opt to receive aspirin. UpToDate can thus offer a strong recommendation for aspirin administration in this setting.

Box 2: A systematic review of randomized trials suggests that in 1000 patients with ST elevation myocardial infarction who are receiving thrombolytic therapy and aspirin and who are treated with heparin (versus no treatment with heparin), 5 fewer will die, 3 fewer will have reinfarction, and 1 fewer will have a pulmonary embolus, while 3 more will have major bleeds [4]. Further, these estimates are not precise and the advantage in decreased infarctions may be lost after 6 months. The small, imprecise, and possibly transient benefit leaves us less confident about any recommendation to use heparin in this situation. Hence, the recommendation is likely to be weak.

Following closely from this reasoning, a third way for clinicians to interpret strong recommendations is, for typical patients, just do it. On the other hand, when clinicians face weak recommendations, or when they face patients with very atypical circumstances or values, they should carefully consider the benefits, risks, and burden in the context of the individual patient before them.

How to individualize decision-making in weak recommendations remains a challenge. For weak recommendations, clinicians should have a detailed conversation with the patient to ensure that the ultimate decision is consistent with the patient's values or even use a decision aid that presents patients with both benefits and down sides of therapy [5]. For strong recommendations, using a decision aid is likely, for most patients, to constitute a poor use of time and energy.

Factors that influence the strength of a recommendation

Editors must consider a number of factors in grading recommendations ([see table 2](#)). One issue is their confidence in the best estimates of benefit and harm.

Prevention of outcomes with high patient-importance should, in general, lead to stronger recommendations than prevention of outcomes of lesser patient importance. For instance, one needs to expose 4 patients to a respiratory rehabilitation program for 1 patient to gain a small but important improvement in dyspnea in daily life [6]. In low risk patients who have suffered a myocardial infarction one might need to treat 100 patients with agents such as aspirin, beta blockers, ACE inhibitors, or statins to extend one life. Despite the much higher number needed to treat (NNT), since we value prolongation of life more highly than relieving dyspnea, the latter intervention may warrant a stronger recommendation.

The choice of adjusted dose warfarin versus aspirin for prevention of stroke in patients with atrial fibrillation illustrates a number of the factors that will influence the strength of a recommendation. A systematic review and meta-analysis found a relative risk reduction (RRR) of 46 percent in all strokes with warfarin versus aspirin. This large effect supports a strong recommendation for warfarin. Furthermore, the relatively narrow 95 percent confidence interval (RRR 29 to 57 percent) suggests that warfarin provides a RRR of at least 29 percent, and further supports a strong recommendation. At the same time, warfarin is associated with an inevitable burden of keeping dietary intake of vitamin K constant, monitoring the intensity of anticoagulation with blood tests, and living with the increased risk of both minor and major bleeding. Most patients, however, are much more stroke averse than they are bleeding averse. As a result, almost all patients with high risk of stroke would choose warfarin, suggesting the appropriateness of a strong recommendation.

This last point emphasizes the importance of the patient's baseline risk (sometimes called control event rate) of the adverse outcome that treatment is designed to avoid. Consider a 65 year-old patient with atrial fibrillation and no other risk factors for stroke. This individual's risk for stroke in the next year is approximately 2 percent. Considering the relative risk reduction and this baseline risk, one can derive the absolute magnitude of an effect ([see table 2](#)). Dose-adjusted warfarin can, relative to aspirin, reduce the risk to approximately 1 percent for an absolute risk reduction of 1 percent (= 2 percent - 1 percent). Some patients who are very stroke averse may consider the down sides of taking warfarin well worth it. Given the relative narrow range that follows from the confidence interval around the relative risk reduction one could make a strong recommendation to use warfarin if all patients were equally stroke adverse. Other patients, however, are likely to consider the benefit not worth the risks and inconvenience [7]. When, across the range of patient values, fully informed patients are liable to make different choices, editors should offer weak (Grade 2) recommendations.

When UpToDate editors make recommendations, they assume a particular set of values as they weigh the possible beneficial and detrimental outcomes. When value or preference judgements are particularly salient, editors should describe the key values attached to these outcomes and that influenced the direction of a recommendation or its grade. The limited literature regarding what

average patient values and preferences actually are, and the range of preferences, emphasizes the importance of making explicit the key values and preference judgements that drive their recommendations.

Wording of recommendations

It is desirable to provide clinicians with as many indicators as possible in interpreting strength of recommendations. UpToDate editors, when they are making a strong recommendation, should use the terminology "We recommend...". When they make a weak recommendation, they should use less definitive wording, such as, "We suggest...".

Confidence in estimates of magnitude of benefits, risks, burden, and costs

Early systems of grading evidence quality relied primarily on the basic study design (ie, randomized control trials (RCTs), or observational studies). The fundamental study design remains critically important in determining our confidence in estimates of beneficial and detrimental treatment effects. Because of prognostic differences between groups, and lack of safeguards such as blinding that can avoid biased ascertainment of outcomes, evidence based on observational studies will, in general, be appreciably weaker than evidence from RCTs. Recent years have seen, however, an increased awareness of a number of other factors that influence our confidence in our estimates of risk and benefit ([see table 3](#)).

UpToDate has chosen a three-category system of quality of evidence (as reflected in confidence in estimates of effects): high (Grade A), moderate (Grade B), and low quality (Grade C) ([see table 1](#)). The strongest evidence comes from systematic reviews that summarize one or more well-designed and well-executed randomized control trials (RCTs) yielding consistent directly applicable results. Strong evidence can also come, under unusual circumstances, from observational studies yielding very large effects.

The moderate strength category is populated by randomized trials with important limitations and by exceptionally strong observational studies. Observational studies, and on occasion RCTs with multiple serious limitations, will fill the low quality evidence category. This categorization follows the principle that all relevant clinical studies provide evidence, the strength of which varies.

Factors that modify the quality of evidence

When RCTs have addressed the impact of alternative management strategies (both benefits and harms) on all relevant outcomes they will yield high quality evidence unless they suffer from one of a number of limitations [8]. The following limitations may decrease the quality of evidence supporting a recommendation ([see table 3](#)).

1) Our confidence in recommendations decreases if the available RCTs suffer from major deficiencies that are likely to result in a biased assessment of the treatment effect [9]. These methodological limitations include a very large loss to follow-up, or an unblinded study with subjective outcomes highly susceptible to bias. How lack of blinding can influence the grading is exemplified by a recommendation to treat heparin-induced thrombocytopenia (HIT) complicated by thrombosis with danaparoid sodium. The randomized trial evidence for danaproid use in HIT comes from an unblinded trial in which the outcome was the clinicians' assessment of when the thromboembolism had resolved, a subjective judgement. As a result, an ACCP guideline panel rated the quality of the evidence as moderate rather than strong [10].

2) When several RCTs yield widely differing estimates of treatment effect (heterogeneity or variability in results) investigators look for explanations for that heterogeneity. For instance, drugs may have larger relative effects in sicker, or in less sick, populations. When heterogeneity exists but investigators fail to identify a plausible explanation, the confidence in estimates of effect from even rigorous RCTs is lower [11]. For example, RCTs of pentoxifylline in patients with intermittent claudication have shown conflicting results that so far defy explanation. Acknowledging the unexplained heterogeneity, a guideline panel rated the quality of the evidence for pentoxifylline as moderate, rather than high [12].

3) Investigators may have undertaken RCTs in populations similar, but not identical, to those under consideration. Editors should consider this indirect evidence and, to the extent they are uncertain about the applicability to their relevant population, rate down the quality of the evidence [13]. For instance, while graduated compression stockings have proved of benefit in a variety of populations at risk of venous thrombosis, they have never been tested directly in trauma patients. An ACCP guideline panel judged the available RCTs relevant to trauma patients in whom administration of low molecular weight heparin is contraindicated, but because of concern about generalizing from other populations (that is, concern about the indirectness of the evidence), rated the quality of the evidence as moderate [14].

Indirectness may also apply to the intervention (RCTs of similar but not identical interventions, different doses and formulations, for instance) and outcomes (RCTs measuring laboratory exercise capacity, for instance, when a panel is really interested in quality of life improvement). Consider sigmoidoscopic screening for colon cancer. The relevant evidence includes not only direct but weak evidence from observational studies, but also stronger but indirect evidence from randomized trials of fecal occult blood screening.

4) Investigators may have conducted RCTs but included very few patients, and observed very few events [15]. For instance, a well-designed and rigorously conducted RCT addressed the use of nadroparin, a low molecular weight heparin, in patients with cerebral venous sinus thrombosis. Of 30 treated patients, 3 had a poor outcome, as did 6 of 29 patients in the control group. The investigators' analysis suggests a 38 percent reduction in relative risk of a poor outcome, but the result was not statistically significant [16]. Because of the small number of patients, and small number of events, a guideline panel judged the quality of the evidence for anticoagulation in cerebral sinus thrombosis as moderate rather than high [17].

• **Observational studies can provide moderate or high quality evidence** — While observational studies will generally yield only low quality evidence, there may be unusual circumstances in which editors will classify such evidence as of moderate, or even high quality [18].

1) On the rare occasions when they yield extremely large and consistent estimates of the magnitude of a treatment effect, we may be confident about the results of observational studies. For example, oral anticoagulation in mechanical heart valves has not been compared to placebo in an RCT. However, evidence from observational studies suggests that the probability of suffering thromboembolic events without anticoagulation is 12.3 percent annually in bileaflet prosthetic aortic valves and higher for other valve types, and estimates of the relative risk reduction with oral anticoagulation are in the range of 80 percent. While the observational studies are likely to overestimate the true effect, the weak study design is very unlikely to explain the entire benefit. Thus, an ACCP guideline panel concluded that these data, despite the absence of randomized trials, constituted high quality evidence of the effectiveness of anticoagulation in bileaflet aortic prosthetic valves [19].

Anticoagulation among patients with AF for more than 48 hours undergoing cardioversion provides another example. In a large observational study, patients who had presented with AF and were already receiving anticoagulation were continued on warfarin; in contrast, those not already on anticoagulation underwent cardioversion without warfarin [19]. The incidence of embolization following DC electroversion was much lower in the warfarin group (0.8 versus 5.3 percent). A review of observational studies among patients undergoing cardioversion found rates of thromboembolism of 2 percent in patients who were not anticoagulated and 0.33 percent among those who received anticoagulation [20]. The magnitude of the association in these trials constitutes moderate or even high quality evidence.

2) On other occasions, all plausible biases from observational studies may be working to underestimate an apparent treatment effect. In other words, the actual treatment effect is very likely to be larger than what the data suggests. For instance, a rigorous systematic review of observational studies including a total of 38 million patients compared private for-profit versus private not-for-profit hospital care. The meta-analysis demonstrated higher death rates in the private for-profit hospitals [21].

The investigators postulated two likely sources of bias. The first was residual confounding with disease severity. It is likely that, if anything, patients in the not-for-profit hospitals were sicker than those in the for-profit hospitals. Thus, to the extent that residual confounding existed, it would bias results against the not-for-profit hospitals.

The second likely bias was the possibility that higher numbers of patients with excellent private insurance coverage could lead to a hospital having more resources and a "spill-over" effect that would benefit those without such coverage. Since for-profit hospitals are likely to admit a larger proportion of such well-insured patients than not-for-profit hospitals, the bias is once again against the not-for-profit hospitals. Because the plausible biases would all diminish the demonstrated treatment effect, one might consider the evidence from these observational studies as moderate rather than low quality.

What to do when strength of evidence differs across outcomes?

UpToDate will provide a single rating of quality of evidence for every recommendation. Recommendations, however, depend on evidence regarding a variety of outcomes. Thus, it may occasionally be necessary to report a single evidence grade when the quality of evidence differs across important outcomes. Consider, for instance, administration of clopidogrel versus aspirin for threatened stroke. A very large well-conducted RCT has shown a small incremental benefit of clopidogrel over aspirin in reducing vascular events. The confidence interval, however, included an effect very near null, suggesting a rating of moderate quality evidence [22]. In deciding on whether to recommend clopidogrel over aspirin, however, one must also consider toxicity. Case reports have suggested the clopidogrel may, on rare occasions, cause thrombotic thrombocytopenic purpura [23]. The quality of this evidence is low. Should the overall quality of evidence for clopidogrel versus aspirin, therefore, be considered moderate, or low?

In such instances, we suggest that editors should consider whether toxicity endpoints are critical to the decision regarding the optimal management strategy. If they are, one must rate the overall quality of the evidence according to the studies that address toxicity. If not, the overall rating of the evidence is based on the evidence regarding benefit. For example, if one considers thrombotic thrombocytopenic purpura to be a critical outcome, then one would rate the overall quality of the

evidence regarding clopidogrel versus aspirin as low. If, on the other hand, one considered that the outcome was so rare that it should not be considered critical, the overall evidence rating would remain moderate.

The process of grading: a checklist and an example

UpToDate editors may benefit, in developing recommendations and grading them, from reference to a checklist ([see table 4](#)). The following examples (see box 4 and 5) from the management of Wegener's granulomatosis show how editors might work through the issues.

Box 4

Question Definition

Patients: Wegener's granulomatosis not requiring immediate dialysis

Intervention: a prednisone, cyclophosphamide combination versus no drug treatment

Outcomes: mortality, respiratory tract and renal morbidity, cyclophosphamide, and steroid toxicity

Evidence Summary

Observational studies show an 8-year survival of 80 percent with treatment. Historical observational studies show a 10 percent 2-year survival without treatment. Observational studies suggest cyclophosphamide toxicity of leucopenia, gastrointestinal upset, and increased risk of malignancy. Observational studies suggest steroid toxicity depends on dose and duration, and includes aseptic necrosis of the hip, infection, osteoporosis, and Cushing's syndrome.

Quality of Evidence

The critical outcome is mortality. While studies are observational, the magnitude of the treatment effect is extremely large, and the evidence therefore high quality (Grade A).

Best estimates

Large mortality reduction, toxicity, and burden variable depending on response to treatment and treatment requirements, cost uncertain.

Judgement of benefits versus risks, burden, and cost

Benefit in mortality reduction greater than all down sides, recommend treatment.

Grade of recommendation

Magnitude of mortality reduction, and pre-eminent importance of mortality to almost all patients, dictate strong recommendation (Grade 1).

Box 5

Question Definition

Patients: Wegener's granulomatosis not requiring immediate dialysis

Intervention: initial treatment with pulse versus continuous cyclophosphamide

Outcomes: mortality, remissions, relapses, leucopenia, infections, gastrointestinal upset, hemorrhagic cystitis, late malignancies

Evidence Summary

Systematic review of 11 observational studies and 3 randomized trials. Observational studies of 202 patients suggest high rates of remission with pulse therapy and low rates of toxicity. Randomized trials or 143 patients showed statistically significant greater remission, lower infection, and lower leucopenia

rates with pulse therapy, but a trend toward more frequent relapse.

Quality of Evidence

Randomized trials without serious limitations provide direct and consistent evidence but total number of patients are few and confidence intervals wide and thus evidence moderate in quality (Grade B).

Best estimates

Greater remission and lower infection with pulse therapy, increase relapses, and no information about many outcomes.

Judgement of benefits versus risks, burden, and cost

Information available suggests benefits of pulse therapy outweigh down sides.

Grade of recommendation

Quality of evidence only moderate for outcomes available and minimal evidence for some outcomes leaves considerable uncertainty about magnitude of benefits and down sides and thus dictates a weak recommendation (Grade 2).

Summary

In the UpToDate grading system, the strength of any recommendation depends on two factors: the tradeoff between benefits and risks and burden, and the quality of the evidence regarding treatment effect. We grade the tradeoff between benefits and risks and burden in two categories; 1, in which the tradeoff is clear enough that most patients, despite differences in values, would make the same choice, leading to a strong recommendation; and 2, in which the tradeoff is less clear, and individual patients values will likely lead to different choices, leading to a weak recommendation. We grade methodological quality in three categories: randomized trials that show consistent results, or observational studies with very strong treatment effects; randomized trials with limitations, or observational studies with exceptional strengths; and observational studies without exceptional strengths or randomized trials with major weaknesses. The framework summarized in Table 1 generates recommendations from the very strong (benefit/risk tradeoff unequivocal, high quality evidence, 1A) to the very weak (benefit/risk questionable, low quality evidence, 2C).

Table 1: Grading Recommendations

Grade of Recommendation	Clarity of risk/benefit	Quality of supporting evidence	Implications
1A. Strong recommendation, high quality evidence	Benefits clearly outweigh risk and burdens, or vice versa.	Consistent evidence from well performed randomized, controlled trials or overwhelming evidence of some other form. Further research is unlikely to change our confidence in the estimate of benefit and risk.	Strong recommendations, can apply to most patients in most circumstances without reservation. Clinicians should follow a strong

			recommendation unless a clear and compelling rationale for an alternative approach is present.
1B. Strong recommendation, moderate quality evidence	Benefits clearly outweigh risk and burdens, or vice versa.	Evidence from randomized, controlled trials with important limitations (inconsistent results, methodologic flaws, indirect or imprecise), or very strong evidence of some other research design. Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.	Strong recommendation and applies to most patients. Clinicians should follow a strong recommendation unless a clear and compelling rationale for an alternative approach is present.
1C. Strong recommendation, low quality evidence	Benefits appear to outweigh risk and burdens, or vice versa.	Evidence from observational studies, unsystematic clinical experience, or from randomized, controlled trials with serious flaws. Any estimate of effect is uncertain.	Strong recommendation, and applies to most patients. Some of the evidence base supporting the recommendation is, however, of low quality.
2A. Weak recommendation, high quality evidence	Benefits closely balanced with risks and burdens.	Consistent evidence from well performed randomized, controlled trials or overwhelming evidence of some other form. Further research is unlikely to change our confidence in the estimate of benefit and risk.	Weak recommendation, best action may differ depending on circumstances or patients or societal values.
2B. Weak recommendation,	Benefits closely balanced with risks and burdens, some	Evidence from randomized, controlled trials with important limitations	Weak recommendation, alternative

moderate quality evidence	uncertainty in the estimates of benefits, risks and burdens.	(inconsistent results, methodologic flaws, indirect or imprecise), or very strong evidence of some other research design. Further research (if performed) is likely to have an impact on our confidence in the estimate of benefit and risk and may change the estimate.	approaches likely to be better for some patients under some circumstances.
2C. Weak recommendation, low quality evidence	Uncertainty in the estimates of benefits, risks, and burdens; benefits may be closely balanced with risks and burdens.	Evidence from observational studies, unsystematic clinical experience, or from randomized, controlled trials with serious flaws. Any estimate of effect is uncertain.	Very weak recommendation; other alternatives may be equally reasonable.

Table 2. Factors panels should consider in deciding on a strong or weak recommendation

Issue/what should be considered	Recommended process	Examples
Quality of evidence	Strong recommendations usually require at least moderate-quality evidence for all the critical outcomes. The lower the quality of evidence, the less likely it becomes a strong recommendation.	Many high quality randomized trials have demonstrated the benefit of inhaled steroids in asthma while only case series have examined the utility of pleurodesis in pneumothorax
Relative importance of the outcomes (benefits of therapy, harm of treatment, burdens of therapy, cost)	Authors and editors consider the relative values and preferences that patients and other stakeholders place on outcomes and the variability in values and preferences across	Preventing post-phlebotic syndrome with thrombolytic therapy in DVT in contrast to preventing death from PE. Most young, healthy people will put a high value on prolonging their lives (and thus incur suffering to do so); the elderly and infirm are

	<p>patients. If values and preferences vary widely a strong recommendation becomes less likely.</p>	<p>likely to vary in the value they place on prolonging their lives (and may vary in the suffering they are ready to experience to do so).</p>
<p>Baseline risks of outcomes (benefits of therapy, harm of treatments, burdens of therapy)</p>	<p>The higher the baseline risk of an adverse outcome, the greater the magnitude of benefit from a treatment, and the more likely a strong recommendation. If the baseline risk is very different in two subpopulations then UpToDate may make separate recommendations for these different populations.</p>	<p>a. Some surgical patients are at very low risk of post-operative DVT and PE while others surgical patients have considerably higher rates of DVT and PE b. ASA and clopidogrel in acute coronary syndromes anticoagulation have a higher risk for bleeding than ASA alone c. Taking adjusted-dose warfarin is associated with a higher burden than taking aspirin; warfarin requires monitoring the intensity of anticoagulation and a relatively constant dietary vitamin K intake.</p>
<p>Magnitude of relative risk including benefits (reduction in RR), harms (increase in RR) and burden (increase in RR)</p>	<p>Larger relative risk reductions with treatment make a strong recommendation for treatment more likely, while larger increases in the relative risk of harms make a strong recommendation for treatment less likely.</p>	<p>Clopidogrel versus aspirin leads to a smaller stroke reduction in TIA (8.7 percent percent RRR [21]) than anticoagulation versus placebo in AF (68 percent RRR).</p>
<p>Absolute magnitude of the effect (benefits. harms and burden)</p>	<p>The larger the absolute benefits with treatment, the greater the likelihood of a strong recommendation in favor of treatment. The larger the absolute increase in harms, the less likely a strong recommendation in favor of treatment.</p>	<p>The absolute reduction in stroke risk in atrial fibrillation patients at yearly stroke risk is 8 percent and in the lowest risk patients less then 1 percent.</p>

Precision of the estimates of the effects (benefits of therapy, harms of treatments and burdens of therapy)	The greater the precision the more likely is a strong recommendation.	ASA versus placebo in AF has a wider confidence interval than ASA for stroke prevention in patients with TIA.
Costs	The higher the cost of treatment, the less likely a strong recommendation.	Clopidogrel has much higher cost than aspirin as prophylaxis against stroke in patients with TIA.

Table 3. Factors panels should consider in deciding on their confidence in estimates of benefits, risks, burden, and costs

<p>Factors that may decrease the strength of evidence based on randomized control trials (RCTs):</p> <ol style="list-style-type: none"> 1. Poor quality of planning and implementation of the available RCTs suggesting high likelihood of bias 2. Inconsistency of results 3. Indirectness of evidence 4. Imprecision 5. High likelihood of publication bias
<p>Factors that may increase the strength of evidence:</p> <ol style="list-style-type: none"> 1. Large magnitude of effect 2. All plausible confounding would reduce a demonstrated effect 3. Dose-response gradient

Table 4. A checklist for developing and grading recommendations

Define the population, intervention and alternative, and the relevant outcomes
Summarize the relevant evidence (relying on systematic reviews, if possible)
<p>If randomized trials, start by assuming high quality, but then check for</p> <ul style="list-style-type: none"> • serious methodologic limitations (lack of blinding, high loss to follow-up, stopped early) • indirectness in population, intervention, or outcome (use of surrogates) • inconsistency in results • imprecision in estimates

- high likelihood of publication bias

Grade down from high to moderate or even low depending on limitations

If no randomized trials (including indirectly relevant trials), start by assuming low quality, but then check for

- large or very large treatment effect
- all plausible confounders would diminish effect of intervention
- dose-response gradient

Grade up to moderate or even high depending on special strengths

Decide on best estimates of benefits, risks, burden, and costs for relevant population

Decide on whether the benefits are, overall, worth the risks, burden, and costs for relevant population

Decide on grade of recommendation, weak or strong, bearing in mind factors in Table 2, and the following advice

- low quality evidence will seldom warrant strong recommendations
- it's hard to go wrong making a weak recommendation. If in doubt, weak recommendation will almost always be the way to go

References

1. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, for the GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008 Apr 26;336(7650):924-926.
2. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, Schünemann HJ, GRADE Working Group. Going from evidence to recommendations. *BMJ* 2008 May 17;336 (7652):1049-1051.
3. Montori V, Devereaux PJ, Straus S, Haynes RB, Guyatt G. Decision making and the patient In: Guyatt G, Rennie D, Meade M, Cook D. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. McGraw-Hill, 2008.
4. Collins R, MacMahon S, Flather M, et al. Clinical effects of anticoagulant therapy in suspected acute myocardial infarction: systematic overview of randomised trials. *Bmj* 1996; 313:652-659.
5. O'Connor AM, Stacey D, Entwistle V, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2003:CD001431.

6. Goldstein RG, EH. Gort, Guyatt, GH. Feeny, D. Economic Analysis of Respiratory Rehabilitation. *Chest* 1997; 112:370-379.
7. Devereaux PA, DR. Gardner, MJ. Putnam, W. Flowerdew, GJ. Brownell, BF. Nagpal, S. Cox, JL. Differences between perspectives of physicians and patients on anticoagulation in patients with atrial fibrillation: observational study. *Bmj* 2001; 323:1218-1222.
8. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter T, Schünemann HJ, GRADE Working Group. What is 'quality of evidence' and why is it important to clinicians? *BMJ* 2008 May 3;336(7651):995-8.
9. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW Jr, Atkins D, Meerpohl J, Schünemann HJ. GRADE Guidelines: 4. Rating the quality of evidence-study limitations (risk of bias). *J Clin Epidemiol*. 2011 Apr;64(4):407-15.
10. Warkentin TE, Greinacher A. Heparin-induced thrombocytopenia: recognition, treatment, and prevention: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004; 126:311S-337S.
11. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schunemann HJ, The GRADE Working Group. GRADE guidelines: 7. Rating the quality of evidence-inconsistency. *J Clin Epidemiol* 2011 Jul 30 (epub).
12. Clagett GP, Sobel M, Jackson MR, et al. Antithrombotic therapy in peripheral arterial occlusive disease: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004; 126:609S-626S.
13. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schunemann HJ; The GRADE Working Group. GRADE guidelines: 8. Rating the quality of evidence-indirectness. *J Clin Epidemiol* 2011 Jul 29 (epub).
14. Geerts WH, Pineo GF, Heit JA, et al. Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004; 126:338S-400S.
15. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux P, Montori V, Freyschuss B, Vist G, Jaeschke R, Williams JW, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schunemann HJ, The GRADE Working Group. GRADE guidelines: 7. Rating the quality of evidence- imprecision. *J Clin Epidemiol* 2011 Jul 30 (epub).
16. de Bruijn SF, Stam J. Randomized, placebo-controlled trial of anticoagulant treatment with low-molecular-weight heparin for cerebral sinus thrombosis. *Stroke* 1999; 30:484-488.
17. Salem DN, Stein PD, Al-Ahmad A, et al. Antithrombotic therapy in valvular heart disease--native and prosthetic: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004; 126:457S-482S.
18. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Murad MH, Schunemann HJ, The GRADE Working Group. GRADE guidelines: 9, Rating up the quality of evidence. *J Clin Epidemiol* 2011 Jul 29 (epub).

19. Bjerkelund CJ, Orning OM. The efficacy of anticoagulant therapy in preventing embolism related to D.C. electrical conversion of atrial fibrillation. *Am J Cardiol* 1969; 23:208-216.
20. Moreyra E, Finkelhor RS, Cebul RD. Limitations of transesophageal echocardiography in the risk assessment of patients before nonanticoagulated cardioversion from atrial fibrillation and flutter: an analysis of pooled trials. *Am Heart J* 1995; 129:71-75.
21. Devereaux PJ, Choi PT, Lacchetti C, et al. A systematic review and meta-analysis of studies comparing mortality rates of private for-profit and private not-for-profit hospitals. *Cmaj* 2002; 166:1399-1406.
22. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. *Lancet* 1996; 348:1329-1339.
23. Bennett CC, JM. Carwile, JM. Moake, JL. Bell, WR. Tarantolo, SR. McCarthy, LJ. Sarode, R. Hatfield, AJ. Feldman, MD. Davidson, CJ. Tsai, HM. Thrombotic thrombocytopenic purpura associated with clopidogrel. *New England Journal of Medicine*. 2000; 342:1773-1777.

[Company](#)

[News & Events](#)

[Resources](#)

[Follow Us](#)

[About Us](#)

[What's New ↗](#)

[UpToDate Login ↗](#)



[Editorial Policy](#)

[Clinical Podcasts](#)

[CME/CE/CPD](#)



[Testimonials](#)

[Press
Announcements](#)

[Mobile Apps](#)

[Wolters Kluwer ↗](#)

[In the News](#)

[Webinars](#)

Sign up today to receive the latest news and updates from UpToDate.

[Careers ↗](#)

[Events](#)

[EHR Integration](#)

[Support](#)

[Health Industry
Podcasts](#)

[Contact Us](#)

[Sign Up](#)

[Help & Training](#)

[Citing Our Content](#)

When you have to be right

[Privacy Policy](#) [Trademarks](#) [Subscription & License Agreement](#)

© 2017 UpToDate, Inc. and/or its affiliates. All rights reserved.

Licensed to: Akademiska sjukhuset Licensed to: Akademiska sjukhuset Support Tag: [0504 -
192.121.175.201 - 56F5F70EED - EI1286210 - UPT - 20171002-21:44:26GMT+00:00] Support Tag: [0504 -
192.121.175.201 - 56F5F70EED - EI1286210 - UPT - 20171002-21:44:26GMT+00:00]