# Come interpretare le raccomandazioni delle linee guida: il metodo GRADE

Giorgio Costantino
IRCCS Fondazione Ca' Granda, Ospedale Maggiore Policlinico
Milano
**Gr**uppo di **A**utoformazione **M**etodologica

# Gianni e Giulio, gemelli

Gianni

- 48 anni, sovrappeso,
- Si reca in PS al Policlinico di Milano alle 18.30 per la comparsa di senso di oppressione precordiale dalle 18, della durata di 15 minuti

Giulio

- 48 anni, sovrappeso,
- Si reca in PS al Policlinico di Milano alle 21.30 per la comparsa di senso di oppressione precordiale dalle 21, della durata di 15 minuti

All'ECG RS, 82 bpm, alterazioni aspecifiche della ripolarizzazione

# Al Policlinico, alle ore 20 cambia il medico di guardia

## Gianni

- Sospetta Sindrome coronarica acuta
- Esegue coronarografia di urgenza

## Giulio

- Sospetta sindrome ansiosa
- Dimesso senza passare dal via

# Linea guida

- Obiettivo: rendere più omogeneo il trattamento, (sperando di renderlo più omogeneo rispetto a ciò che dovrebbe essere giusto fare…)

ovvero

- Raccomandazioni di comportamento clinico, prodotte attraverso un processo sistematico, allo scopo di assistere medici e pazienti nel decidere le modalità di assistenza più appropriate in specifiche circostanze cliniche

# Problemi delle linee guida: Confronto LG fibrillazione atriale in PS

Analizzate linee guida ESC, AHA, Canadian

- Su 19 quesiti specifici:
  - 4 quesiti nessuna concordanza tra linee guida
  - 8 concordanza solo parziale (alcuni aspetti non trattati da linee guida o differenze di tipo di farmaco)
  - 7 concordi

# Fibrillazione Atriale

## Linee guida AHA 2014

- In patients with pre-excitation and AF, digoxin, nondihydropyridine calcium channel antagonists, or intravenous amiodarone should not be administered as they may increase the ventricular response and may result in ventricular fibrillation (274). (Level of Evidence: B)

## Linee guida ESC 2012

- In pre-excitation, preferred drugs are class I antiarrhythmic drugs or amiodarone. (I C)

## Conflitto di interesse


## Perché?

## Metodologia

OPEN ACCESS Freely available online
PLOS MEDICINE

# How Evidence-Based Are the Recommendations in Evidence-Based Guidelines?

**Finlay A. McAlister[1*], Sean van Diepen[1], Rajdeep S. Padwal[1], Jeffrey A. Johnson[2,3], Sumit R. Majumdar[1]**

1 The Division of General Internal Medicine, University of Alberta, Edmonton, Canada, 2 The Institute of Health Economics, University of Alberta, Edmonton, Canada, 3 The School of Public Health, University of Alberta, Edmonton, Canada

**Funding:** This study was completed without external funding support, and none of the salary-support funders for any of the authors had input into the design or conduct of the study; collection, management, analysis, or interpretation of the data; nor preparation, review, approval, or decision to submit the manuscript for publication.

**Competing Interests:** FAM is co-chair of the Central Review Committee
Hypertensi
and RSP is
Canadian H
Education
Commitee.

ABSTRACT

**Background**

Treatment recommendations for the same condition from different guideline bodies often disagree, even when the same randomized controlled trial (RCT) evidence is cited. Guideline appraisal tools focus on methodology and quality of reporting, but not on the nature of the supporting evidence. This study was done to evaluate the quality of the evidence (based on consideration of its internal validity, clinical relevance, and applicability) underlying therapy
al practice guidelines.

iscular risk management recommendations was
diabetes mellitus, dyslipidemia, and hypertension)
rom the United States, Canada, and Europe). Of the
nine guidelines, 231 (68%) cited RCT evidence but
mendations were based on high-quality evidence.
most often because of reservations about the
is specified in the guideline recommendation (64/
ted surrogate outcomes (59/126 cases, 47%).

## Idee diverse


## autorevolezza

# Come vengono definite le raccomandazioni?

**Variant 1:** **Older than age 40.**

| lure | Rating | Comments | RRL* |
|---|---|---|---|
| | 8 | | ☢ |
| | 4 | | ☢☢☢ |
| | 3 | | ☢☢☢ |
| trast | 1 | | ☢☢☢ |

opriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate — *Relative Radiation Level

**Table 1** Brief description of the generic levels of evidence and guideline statement grades used

| Evidence level | Definition | Guideline statement grade |
|---|---|---|
| Ia | A good recent systematic review of studies designed to answer the question of interest | A+ |
| Ib | One or more rigorous studies designed to answer the question, but not formally combined | A– |

| Classes of recommendations | Definition | Suggested wording to use |
|---|---|---|
| Class I | Evidence and/or general agreement that a given treatment or procedure is beneficial, useful, effective. | Is recommended/is indicated |
| Class II | Conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of the given treatment or procedure. | |
| *Class IIa* | *Weight of evidence/opinion is in favour of usefulness/efficacy.* | Should be considered |
| *Class IIb* | *Usefulness/efficacy is less well established by evidence/opinion.* | May be considered |
| Class III | Evidence or general agreement that the given treatment or procedure is not useful/effective, and in some cases may be harmful. | Is not recommended |

| | |
|---|---|
| Level of evidence A | Data derived from multiple randomized clinical trials or meta-analyses. |
| Level of evidence B | Data derived from a single randomized clinical trial or large non-randomized studies. |
| Level of evidence C | Consensus of opinion of the experts and/ or small studies, retrospective studies, registries. |

- Pz di 85 anni viene per edemi declivi e dispnea, PA 95/60, marcati edemi declivi e turgore giugulare, FE 25%.
- Che farmaco iniziereste?

Diuretico nello scompenso cardiaco

Ace inibitore nello scompenso cardiaco

# GRADE

## (**Grading** of Recommendations Assessment, Development and Evaluation)

- Obiettivo: fornire una metodologia condivisa e universale per proporre raccomandazioni
- In particolare:
  - Dividere qualità delle evidenze da forza delle raccomandazioni
  - Spiegare il motivo alla base di questi giudizi in modo trasparente (allegare il “grade evidence profile”) proponendo un metodo sistematico e riproducibile per proporre la forza delle raccomandazioni

# Metodo

Approccio strutturato in 3 fasi principali:

- Formulazione di un quesito clinico con scelta e valutazione formale degli outcome ad esso correlati; valutazione sistematica della letteratura scientifica e della qualità delle prove reperite
- Valutazione di benefici e rischi associati all'intervento considerando le preferenze dei pazienti, la fattibilità e l'impiego di risorse necessario
- Definizione formale della forza della raccomandazione

| Factors that can strengthen a recommendation | Comment |
| --- | --- |
| **Quality of the evidence** | The higher the quality of evidence, the more likely is a strong recommendation. |
| **Balance between desirable and undesirable effects** | The larger the difference between the desirable and undesirable consequences, the more likely a strong recommendation warranted. The smaller the net benefit and the lower certainty for that benefit, the more likely weak recommendation warranted. |
| **Values and preferences** | The greater the variability in values and preferences, or uncertainty in values and preferences, the more likely weak recommendation warranted. |
| **Costs (resource allocation)** | The higher the costs of an intervention – that is, the more resources consumed – the less likely is a strong recommendation warranted |

- Therefore, unlike many other grading systems, the GRADE system emphasizes that weak recommendations in the face of high quality evidence are common because of factors other than the quality of evidence influencing the strength of a recommendation. For the same reason it allows for strong recommendations based on the evidence from observational studies.

# Forza della raccomandazione

**Box 2 Examples of implications of strong and weak recommendations**

**Strong recommendation for intervention**

*For patients*—Most people in this situation would want the recommended course of action and only a small proportion would not

*For clinicians*—Most people should receive the intervention

*For quality monitors*—Adherence to this recommendation could be used as a quality criterion or performance indicator. If clinicians choose not to follow such a recommendation, they should document their rationale

**Weak recommendation for intervention**

*For patients*—Most people in this situation would want the suggested course of action, but many would not

*For clinicians*—Examine the evidence or a summary of the evidence yourself and be prepared to discuss that evidence with patients, as well as their values and preferences

*For quality monitors*—Clinicians' discussion or consideration of the pros and cons of the intervention, and their documentation of the discussion, could be used as a quality criterion.

**No specific recommendation**

The advantages and disadvantages are equivalent

The target population has not been identified

Insufficient evidence on which to formulate a recommendation

- Strong: applicalo a quasi tutti
- Weak: pensaci e discutine

# Qualità dell'evidenza:

- High (Further research is very unlikely to change our confidence in the estimate of effect);
- Moderate (Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate);
- Low (Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate);
- Very low (Any estimate of effect is very uncertain).

# Possibilità di ridurre o aumentare la qualità dell'evidenza

- Studi a rischio di bias
- Inconsistenza dei risultati (eterogeneità non spiegabile tra gli studi)
- Confronti indiretti (tra interventi o per outcome)
- Imprecisione (intervalli di confidenza larghi)
- Bias di pubblicazione (pubblicazione più facile di studi con risultato positivo)

| Study Design | Quality of Evidence | Lower if | Higher if |
|---|---|---|---|
| Randomised trial → | High | Risk of bias<br>-1 Serious<br>-2 Very serious<br><br>Inconsistency<br>-1 Serious<br>-2 Very serious<br><br>Indirectness<br>-1 Serious<br>-2 Very serious<br><br>Imprecision<br>-1 Serious<br>-2 Very serious<br><br>Publication bias<br>-1 Likely<br>-2 Very likely | Large effect<br>+1 Large<br>+2 Very large<br><br>Dose response<br>+1 Evidence of a gradient<br><br>All plausible confounding<br>+1 Would reduce a demonstrated effect or<br><br>+1 Would suggest a spurious effect when results show no effect |
| | Moderate | | |
| Observational study → | Low | | |
| | Very low | | |

FIGURE 2. GRADE approach to rating quality of evidence. See Figure 1 legend for expansion of abbreviation.

# GRADE evidence profile

A GRADE evidence profile allows presentation of key information about all relevant outcomes for a given health care question. It presents information about the body of evidence (*e.g.* number of studies), the judgments about the underlying quality of evidence, key statistical results, and a grade for the quality of evidence for each outcome.

Hide table **Summary of findings for the main comparison. Implantable cardioverter defibrillators compared to β-blocker therapy for prevention of sudden cardiac death in people with cardiac channelopathies**

**Implantable cardioverter defibrillators compared to β-blocker therapy for prevention of sudden cardiac death in people with cardiac channelopathies**

**Patient or population:** people with cardiac channelopathies
**Settings:** Thailand
**Intervention:** implantable cardioverter defibrillators
**Comparison:** β-blocker therapy

| Outcomes | Illustrative comparative risks* (95% CI) | | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|---|---|
| | Risk with β-blocker therapy | Risk with ICD | | | | |
| All-cause mortality | Study population | | RR 0.11 (0.01 to 0.83) | 86 (2 RCTs) | ⊕⊕⊖⊖ Low [1,2] | All included study participants experienced a previously documented or clinically suspected episode of terminated sudden cardiac arrest due to Brugada syndrome |
| | 18 per 100 | **2 per 100** (1 to 15 per 100) | | | | |
| Fatal and non-fatal cardiovascular events | Study population | | RR 1.49 (0.66 to 3.34) | 86 (2 RCTs) | ⊕⊕⊖⊖ Low [1,2] | - |
| | 18 per 100 | **27 per 100** (12 to 61 per | | | | |

# Take home message

- Forza della raccomandazione:
  - Forte (di regola applica la raccomandazione)
  - Debole (di regola discuti la raccomandazione)

- Livello di evidenza
  - Alto (non servono ulteriori studi)
  - Moderato (ulteriori studi potrebbero cambiare l'effetto)
  - Basso (ulteriori studi verosimilmente cambieranno l'effetto)

# A favore del GRADE


- Omogeneizzazione
- Tentativo di rendere sistematico l'approccio alle raccomandazioni
- Inserire le preferenze del paziente e il costo beneficio
- semplicità

ORIGINAL ARTICLE

# GRADE guidelines system is reproducible when instructions are clearly operationalized even among the guidelines panel members with limited

The interrater agreement for the domain of quality of evidence was good (kappa value: 0.68; 95% CIs: 0.54, 0.84), and fair for balance of benefit and harms (kappa value: 0.4; 95% CIs: 0.25, 0.57) and use of resources (kappa value: 0.28: 95% CIs: 0.12, 0.42). The interrater agreement was moderate for the GRADE domain of patients' values and preferences (kappa value: 0.44; 95% CI: 0.31, 0.56). The interrater agreement for making a for/against recommendation was good (kappa value: 0.74; 95% CIs: 0.33, 0.91) and fair for strong/weak recommendation (kappa value: 0.39; 95% CIs: 0.18, 0.68).

ment for the domain of quality of evidence was good (kappa value: 0.68; 95% CIs: 0.54, 0.84), and fair for balance of benefit and harms (kappa value: 0.4; 95% CIs: 0.25, 0.57) and use of resources (kappa value: 0.28: 95% CIs: 0.12, 0.42). The interrater agreement was moderate for the GRADE domain of patients' values and preferences (kappa value: 0.44; 95% CI: 0.31, 0.56). The interrater agreement for making a for/against recommendation was good (kappa value: 0.74; 95% CIs: 0.33, 0.91) and fair for strong/weak recommendation (kappa value: 0.39; 95% CIs: 0.18, 0.68).

**Conclusions:** Although not all elements of GRADE system had good agreement, the interrater agreement for assessing the quality of evidence and issuing a recommendation of for vs. against among panel members who had limited exposure to GRADE methodology was good. This is probably because GRADE has operationalized these two areas in more detail than other domains. Further operationalization of all GRADE domains such as with the GRADE evidence to decision frameworks would likely improve its reproducibility. 


# CONTRO (2)


## RECOMMENDATION

We recommend that ventricular rate be assessed at rest in all patients with persistent and permanent AF or AFL (Strong Recommendation, Moderate-Quality Evidence).

We recommend that heart rate during exercise be assessed in patients with persistent or permanent AF or AFL and associated exertional symptoms (Strong Recommendation, Moderate-Quality Evidence).

We recommend that treatment for rate control of persistent or permanent AF or AFL should aim for a resting heart rate of $<100$ bpm (Strong Recommendation, High-Quality Evidence).

**Values and preferences.** These recommendations place a high value on the randomized clinical trials and other clinical studies demonstrating that ventricular rate control of AF is an effective treatment approach for many patients with AF.


UNO NASCE, E POI MUORE.
IL RESTO SONO CHIACCHIERE.
ALTAN. TERAPIA
Un distillato di Altan come sollievo alla vita quotidiana
Salani Editore

- Pz di 85 anni viene per edemi declivi e dispnea, PA 95/60, marcati edemi declivi e turgore giugulare, FE 25%.
- Che farmaco iniziereste?

Diuretico nello scompenso cardiaco

Ace inibitore nello scompenso cardiaco