*Full research paper*

# A systematic breakdown of the levels of evidence supporting the European Society of Cardiology guidelines

Wouter B van Dijk[1], Diederick E Grobbee[1],
Martine C de Vries[2], Rolf H H Groenwold[3],
Rieke van der Graaf[4] and Ewoud Schuit[1,5]

## Abstract

**Aims:** Reviews of clinical practice guidelines have repeatedly concluded that only a minority of guideline recommendations are supported by high-quality evidence from randomised controlled trials. The aim of this study is to evaluate whether these findings apply to the whole cardiovascular evidence base or specific recommendation types and actions.

**Methods:** All recommendations from current European Society of Cardiology guidelines were extracted with their class (I, treatment is beneficial; II, treatment is possibly beneficial; III, treatment is harmful) and level of evidence (A, multiple randomised controlled trials/meta-analyses; B, single randomised controlled trials/large observational studies; C, expert opinion/small studies). Recommendations were categorised by type (therapeutic, diagnostic, other) and actions (e.g. pharmaceutical intervention/non-invasive imaging/test).

**Results:** In total, 3531 recommendations (median 128, interquartile range 108–150) were extracted from 27 guidelines. Therapeutic recommendations comprised 2545 (72.1%) recommendations, 411 (16.1%) were supported by level of evidence A, 833 (32.7%) by B and 1301 (51.1%) by C. Class I/III (should/should not) recommendations on minimally invasive interventions were most supported by level of evidence A (55/183, 30.1%) (B [70/183, 38.3%], C [58/183, 31.7%]), while class I/III recommendations on open surgical interventions were least supported by level of evidence A (15/164, 9.1%) (B [34/164, 20.7%], C [115/164, 70.1%]). Of all (831, 23.5%) diagnostic recommendations, just 44/503 (8.7%) class I/III recommendations were supported by level of evidence A (B (125/503, 24.9%), C (334/503, 66.4%)).

**Conclusion:** Evidence levels supporting European Society of Cardiology guideline recommendations differ widely between recommendation types and actions. Attributing to this variability are different evidence requirements, therapeutic/diagnostic recommendations, different feasibility levels for trials (e.g. open surgical/pharmacological) and many off-topic/policy recommendations based on expert opinion.

## Introduction

Clinical practice guidelines form the crest in translating science into clinical practice. Clinicians report cardiovascular guidelines to be their main source of information for clinical decision making.[1] As such, cardiovascular guidelines influence the care provided to millions of people worldwide.[2,3]

To justify this epistemological status, guidelines should be grounded in objective, high-quality evidence. Yet, a recent comparison of cardiovascular guidelines

[1]Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands
[2]Department of Medical Ethics and Health Law, Leiden University Medical Center, The Netherlands
[3]Department of Clinical Epidemiology, Leiden University Medical Center, The Netherlands
[4]Department of Medical Humanities, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, The Netherlands
[5]Cochrane Netherlands, University Medical Center Utrecht, The Netherlands

**Corresponding author:**
Wouter B van Dijk, Julius Center for Health Sciences and Primary Care, Universitair Medisch Centrum Utrecht, Universiteitsweg 100 Utrecht, Utrecht 3584, CG Netherlands.
Email: W.B.vanDijk-7@umcutrecht.nl

from the American College of Cardiology (ACC)/ American Heart Association (AHA) and the European Society of Cardiology (ESC) published between 2008 and 2018 showed that a limited number of recommendations is supported by evidence from multiple high-quality randomised controlled trials (RCTs; ACC/AHA < 10% and ESC < 15%), and a majority by expert opinion and smaller studies.[4,5] These results fuel the criticism that guideline development lacks transparency on how recommendations are conceived,[6–8] and previous claims that the evidence base underlying cardiovascular guidelines is poor.[5] However, to know which paucities in the evidence base are problematic and where to focus improvement efforts to fill these it is necessary to identify areas of recommendations not supported by high-quality evidence and identify the underlying reasons. To reveal where gaps exist in the current cardiovascular evidence base, and allow better interpretation of the evidence underlying recommendations, this paper aims to identify which types of recommendations (e.g. therapeutic or diagnostic) and which recommended actions (e.g. pharmaceutical intervention or non-invasive imaging) are supported by which level of evidence (LoE) in the guidelines of the ESC.

## Methods

All documents referred to as clinical practice guideline were collected from the ESC website (https://www.escardio.org/Guidelines). Documents were categorised as comprehensive guidelines, focused updates, definition guidelines, position/expert consensus papers and other documents. Only current comprehensive guidelines and focused updates (short updates to comprehensive guidelines) were included for further analysis. Disease definitions and position papers were excluded from further analysis, because they were not considered representative of entire topics. The search and selection of guidelines was performed by one author (WBvD).

From the remaining guideline documents, recommendations were extracted together with their corresponding class and evidence level. To appraise recommendations and value their underlying evidence the ESC uses a system of classes of recommendations and levels of evidence to grade recommendations (see Table 1).[9] Extraction was performed by one reviewer (WBvD) using Tabula, version 1.2.1 (https://tabula.technology/) and the results were exported to Excel, version 16.21.1. Only recommendations that contained a clearly indicated class and LoE were extracted. All recommendations were manually checked to have a recommendation text, class and LoE. One author (WBvD) categorised every guideline into a cardiovascular sub-specialty area (electrophysiology, coronary artery disease, heart failure and myocardial disease, congenital and valvular heart disease, general cardiology, vascular medicine), and every recommendation by type (therapeutic, diagnostic, other (e.g. definitions)) and recommended action (main categories: pharmaceutical intervention, open surgical intervention, minimally invasive/device intervention, non-invasive test/imaging, laboratory test, invasive test). Recommendations regarding preventive interventions were categorised as therapeutic and, similarly, recommendations about screening were regarded as diagnostic. A random sample of recommendations was manually cross-validated on extraction and classification by a second author (ES). Interrater reliability was assessed using

**Table 1.** European Society of Cardiology classes of recommendations and LoEs.

| | | |
|---|---|---|
| Recommendations are divided into the following classes | | |
| Class I | Evidence and/or general agreement that a treatment or procedure is beneficial, useful, effective Use of treatment or procedure is recommended | |
| Class II | Conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a treatment or procedure | |
| | Subclass a | Weight of evidence/opinion is in favour of usefulness/efficacy. Use of treatment or procedure should be considered |
| | Subclass b | Usefulness/efficacy is less well established by evidence/opinion. Use of treatment or procedure may be considered |
| Class III | Evidence or general agreement that a treatment or procedure is not useful/effective, and in some cases may be harmful. Use of treatment or procedure is not recommended | |
| The evidence underlying these recommendations is split into one of the following levels | | |
| LoE A | Data derived from multiple randomised clinical trials or meta-analyses | |
| LoE B | Data derived from a single randomised clinical trial or large non-randomised studies | |
| LoE C | Consensus of expert opinion and/or small studies, retrospective studies, registries | |

LoE: level of evidence.

Cohen's kappa statistic and was considered sufficient (≥0.7), indicating further extraction could be performed by one author (WBvD).

The overall number of recommendations, distributions and percentages of classes, evidence levels, types and recommended actions were calculated. Because of the high variety in recommendation numbers per guideline, these were summarised by calculating medians and interquartile ranges. For clarity recommendation types were also reported per cardiovascular subspecialty area and recommended action per type.

The ESC states to have published over 100 guidelines,[10] while just 57 were published on the ESC website. We contacted the ESC to query about the remaining guidelines, and they pointed us to the website of the *European Heart Journal* (EHJ; https://academic.oup.com/eurheartj). A cross-reference of the *EHJ* website showed all current guidelines were published on the ESC website. We therefore focused only on guidelines as published on the ESC website.

## Results

On 1 May 2019, 37 published documents were identified on the ESC website as current guidelines (see Supplementary Table 1). Ten documents were excluded because they concerned disease definitions ($N = 1$), position/expert consensus papers ($N = 5$), other documents ($N = 2$) or did not contain clearly stated recommendations and evidence levels ($N = 2$); leaving 27 guidelines for analysis.

The current 27 guidelines were published between 2003 and 2018 and provided 3531 recommendations. They comprised a median of 128 recommendations per guideline (interquartile range 108–150) (see Supplementary Table 2), most recommendations were of class I (1684 recommendations; 47.7%), followed by class II (1577 recommendations; 44.7%) and III (270 recommendations; 7.6%) (Figure 1).
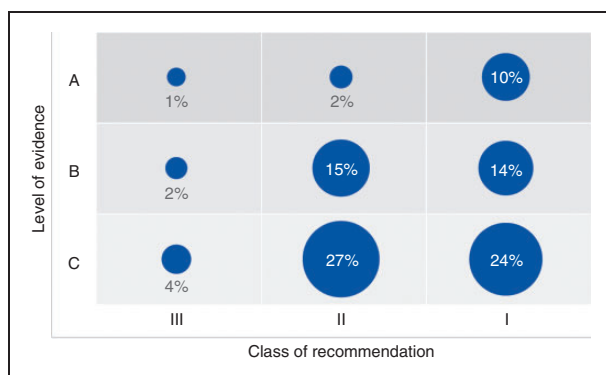


**Figure 1.** Overall proportions of recommendation classes and levels of evidence.

Of the 1684 class I recommendations 360 (21.4%) recommendations were supported by LoE A, 489 (29.0%) recommendations by LoE B, and 835 (49.6%) recommendations by LoE C; of the 1577 class II recommendations 86 (5.5%) recommendations were supported by LoE A, 535 (33.9%) by LoE B, and 956 (60.6%) by LoE C; of the 270 class III recommendations 53 (19.6%) recommendations were supported by LoE A, 79 (29.3%) by LoE B, and 138 (51.1%) by LoE C.

Notably, the number of recommendations supported by LoE C varied widely between guidelines (range 36.7–76.4%).

### Overall distributions of types and recommended actions

Overall, therapeutic recommendations comprised 2545 recommendations (72.1%), diagnostic recommendations 831 recommendations (23.5%), and other recommendations 155 (4.4%).

Among the 2545 therapeutic recommendations, 1134 (44.6%) recommendations were of class I, 1189 (46.7%) recommendations were of class II, and 222 (8.7%) recommendations were of class III (Figure 2). Of these, class I comprised 300 (26.5%) recommendations supported by LoE A, 350 (30.9%) recommendations supported by LoE B, and 484 (42.7%) recommendations supported by LoE C; class II comprised 63 (5.3%) recommendations supported by LoE A, 411 (34.6%) recommendations supported by LoE B, and 715 (60.1%) recommendations supported by LoE C; class III comprised 48 (21.6%) recommendations supported by LoE A, 72 (32.4%) recommendations supported by LoE B, and 102 (45.9%) recommendations supported by LoE C.

The three therapeutic actions recommended most were pharmaceutical interventions (1245 recommendations, 48.9%), open surgical interventions (367 recommendations, 14.4%), and minimally invasive interventions (341 recommendations, 13.4%) (Figure 3). Pharmaceutical interventions comprised 236 (19.0%) recommendations supported by LoE A, 419 (33.7%) recommendations supported by LoE B, and 592 (47.6%) recommendations supported by LoE C; open surgical interventions comprised 20 (5.4%) recommendations supported by LoE A, 78 (21.3%) recommendations supported by LoE B, and 269 (73.3%) recommendations supported by LoE C; minimally invasive interventions comprised 63 (18.4%) recommendations supported by LoE A, 141 (41.3%) recommendations supported by LoE B, and 132 (40.2%) recommendations supported by LoE C.

Among the 831 diagnostic recommendations, 456 (54.9%) recommendations were of class I, 328 (39.5%) recommendations were of class II, and 47
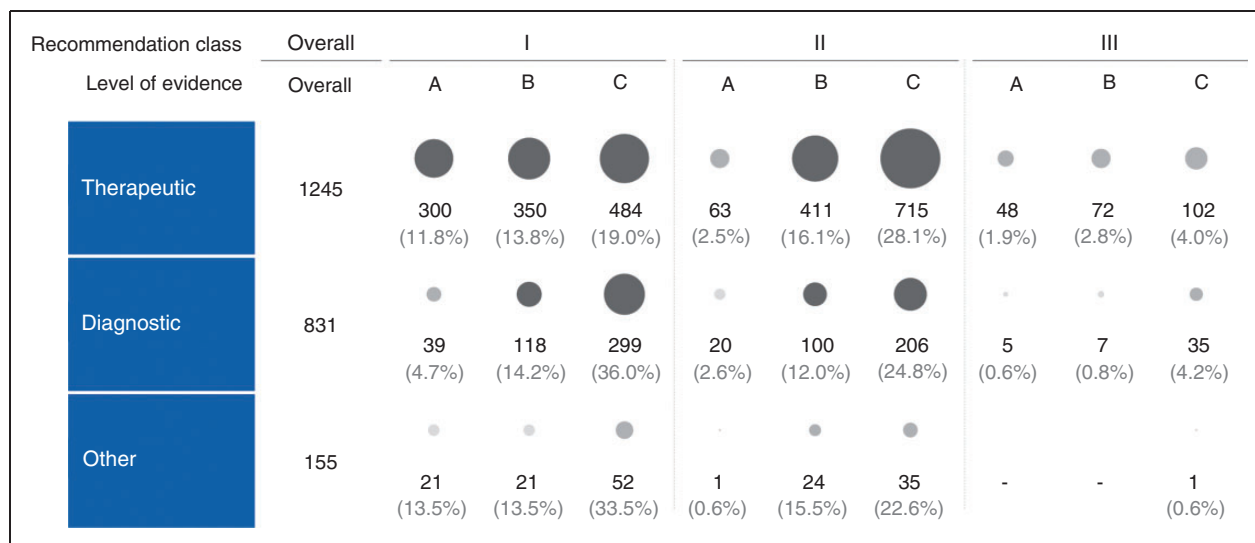
**Figure 2.** Proportions of types of recommendations by classes and evidence levels.

(5.7%) recommendations were of class III. Of these, class I comprised 39 (8.6%) recommendations supported by LoE A, 118 (25.9%) recommendations supported by LoE B, and 299 (65.6%) recommendations supported by LoE C; class II comprised 22 (6.7%) recommendations supported by LoE A, 100 (30.5%) recommendations supported by LoE B, and 206 (62.8%) recommendations supported by LoE C; class III comprised 5 (10.6%) recommendations supported by LoE A, 7 (14.9%) recommendations supported by LoE B, and 35 (74.5%) recommendations supported by LoE C.

The three diagnostic actions recommended most were non-invasive tests/imaging (378 recommendations, 45.6%), laboratory tests (156 recommendations, 18.8%), and invasive tests/imaging interventions (108 recommendations, 13.0%). Non-invasive tests/imaging comprised 14 (3.7%) recommendations supported by LoE A, 112 (29.6%) recommendations supported by LoE B, and 252 (66.7%) recommendations supported by LoE C; laboratory tests comprised 39 (25.0%) recommendations supported by LoE A, 32 (20.1%) recommendations supported by LoE B, and 85 (54.5%) recommendations supported by LoE C; invasive tests/ imaging comprised 7 (6.5%) recommendations supported by LoE A, 19 (17.6%) recommendations supported by LoE B, and 82 (75.9%) recommendations supported by LoE C.

### Distributions of types and recommended actions per subspecialty

General cardiology was the largest subspecialty with 875 (24.7%) recommendations, followed by coronary artery disease (726 recommendations, 20.6%) and congenital and valvular heart disease (677

recommendations, 19.2%) (Figure 4). The largest proportion of class I recommendations supported by LoE A was found on coronary artery disease (169 recommendations, 23.3%). Congenital and valvular heart disease comprised the most recommendation supported by LoE C (505 recommendations, 74.6%) as a result of a large number of recommendations on open surgical interventions.

## Discussion

This in-depth analysis of the ESC guidelines shows that evidence levels supporting recommendations in cardiovascular guidelines vary widely per type, recommended action and subspecialty. Overall, just 14.1% of the recommendations are supported by multiple RCTs or meta-analyses (LoE A). However, when stratified to their types and recommended actions we found that some recommendation groups are less substantiated by high-quality evidence than others. Therapeutic recommendations, in particular pharmaceutical, minimally invasive and lifestyle recommendations, appear to be supported by higher quality evidence than diagnostic recommendations. We found recommendations on open surgical interventions, non-invasive tests/imaging and invasive tests/imaging to be least supported by high-quality evidence, attributing to the low evidence levels of recommendations in the ESC guidelines in general.

In their recent review, Fanaroff et al. reported similar distributions of the overall LoEs in the cardiovascular guidelines of the ESC to those found in this review.[4] In addition, the authors reviewed the guidelines of the ACC/AHA of which they noted similar results. Yet the present study showed that these
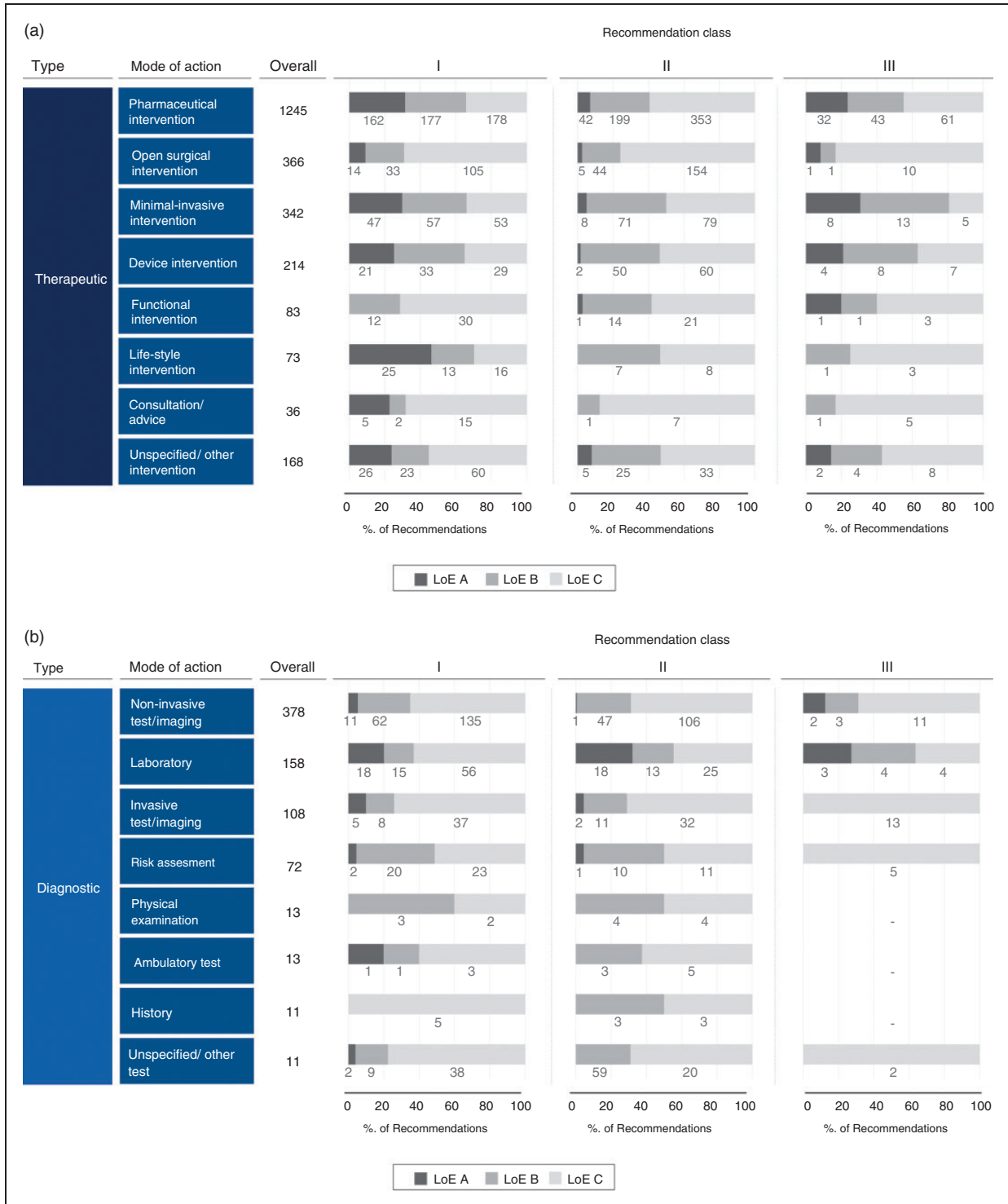
**Figure 3.** Recommendations of modes of action and evidence levels by classes: (a) therapeutic; (b) diagnostic.

numbers and their accompanying conclusions do not apply to the evidence base as a whole. Instead the quality of evidence supporting recommendations differs substantially between subspecialties, recommendation types and recommended actions. Systematic guideline analyses in other medical and surgical subspecialties have shown comparable distributions of few recommendations supported by level A evidence,[11–17] and
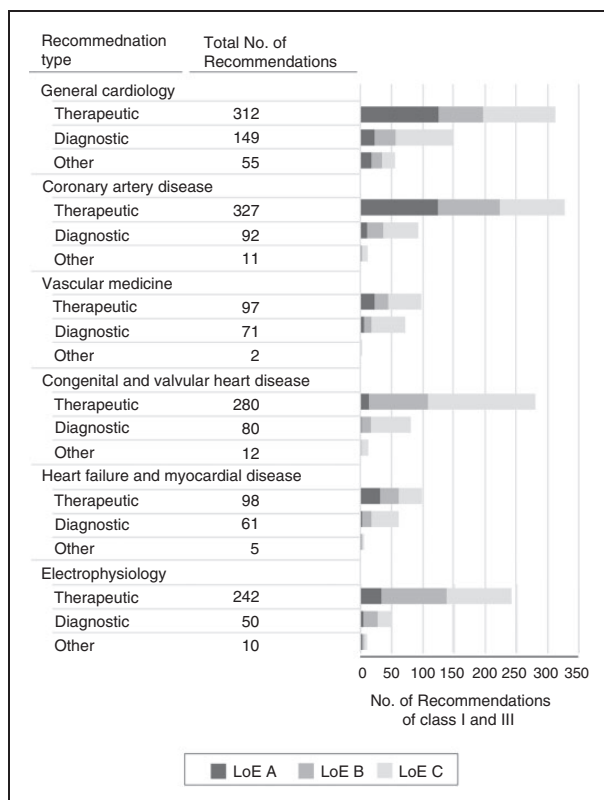
**Figure 4.** Recommendations by subspecialty, type and level of evidence.

might, consequently, need comparable distinctions on types and recommended actions.

A decade ago, Tricoci et al. already reported similar findings for the ACC/AHA guidelines. The authors identified several shortcomings in the organisation of clinical research and the guideline development process as possible explanations for this shortage in recommendations supported by high-quality evidence. These shortcomings included fragmentation of the research enterprise (a lack of common goals, vision and collaboration), missing incentives to fill evidence gaps and potential conflicts of interests. Fanaroff et al.[4] found that the distributions of LoEs of the ACC/AHA guidelines did not change between 2008 and 2018, i.e. since the review by Tricoci et al., and that the ESC guidelines exhibited similar LoE distributions and trends over time.[5]

## Implications for the cardiovascular evidence base

The findings of the present analysis provide focus for improvement efforts of the current cardiovascular evidence base.

Recommendations on pharmaceutical, minimally invasive and lifestyle interventions were found to be most supported by high-quality evidence (30–40%)

and put the low overall evidence levels (15%) in a more positive perspective as such. Yet, it is still low when used to support adequate evidence-based decision-making in practice and should be improved.

To improve the cardiovascular evidence base, more focus should be put on generating evidence for diagnostic recommendations and recommendations on open surgical interventions; two areas still mainly supported by expert opinion and small studies. Although lower evidence levels (LoE B) are understandable for research demonstrating diagnostic test accuracy, as these studies in general will have a cross-sectional design, the highest level (i.e. RCTs/meta-analyses) should be required to determine the consequences for patients of implementing a new diagnostic tool in clinical practice. Akin to this, it is important to distinct guideline recommendations on their goal, sometimes allowing lower evidence levels when only discussing test accuracies.

Hence, many recommendations on open surgical interventions also lack support from high-quality evidence. Evidence should be generated to fill these paucities to increase the reliability of recommendations on these interventions. Yielding research results to fill these evidence paucities on surgical treatments can be more difficult due to methodological challenges. Justly executed surgical intervention trials are by nature more complex than pharmaceutical intervention trials due to the increased number of variables at play[18] and difficulties in blinding patients and doctors.[19] Undesirably, RCTs are relatively less common in surgery as a consequence.[20] Despite these challenges in performing surgical trials, they cannot be condoned for not evaluating surgical treatments because of the size of their effect on the lives of patients. Execution of surgical trials could, for instance, be improved by designing trials more pragmatically or to supplement them with the increasing amounts of observational data available resulting from recent technological advancements, for instance by moving towards a learning healthcare system.[21]

Yet, it needs to be recognised that it might never be possible to support all guideline recommendations by the same (high) levels of evidence. For instance, recommendations to initiate treatment may need stronger evidence (e.g. RCTs and/or meta-analyses) than recommendations not to use specific treatments (e.g. a case series), because for the latter it will not always be feasible or ethical to require these levels of evidence.

## Implications for the development of guidelines

It is indisputable that we will always be in need of more high-quality evidence to fill paucities in the existing evidence base. If recommendations are important for clinical practice they should be included and efforts should be made to support them with evidence when evidence

is lacking. Yet, besides improving the evidence base, guidelines should also be improved, handling the evidence paucities as well as possible. In over a decade, distributions of LoEs in guidelines have barely changed.[4] This flat-line in LoEs might not only be maintained by the paucities in the cardiovascular evidence base, but could also be maintained as a consequence of the organisation of the guideline development process. Guideline committees (i.e. task forces) should reflect on the contents of the recommendations they issue in their documents.

First, guideline committees should consider the large quantity of recommendations supported by LoE C that guidelines comprise. Currently, guideline authors have a so-called wide margin of appreciation giving them substantial freedom in the contents of recommendations they include in guidelines. Whether such a margin of appreciation should be allowed in guidelines is a matter of debate.[6,8] Regardless of this, it results in a wide range of topics covered in guidelines, inside and outside the field of cardiovascular care. Current guidelines, for instance, contain recommendations stating that the diagnosis and prognosis of a disease should always be explained to patients (class I, LoE C).[22] General recommendations like these can be of importance to the homogenisation process of different European practices, but are also at risk of stating the obvious, not rising above the level of presumed textbook knowledge. Guideline committees should be aware of these risks and consciously choose whether they want to issue such recommendations. Also, deliberations should be made when issuing recommendations on adjacent specialties. Recommendations on vaginal delivery in healthy women (class I, LoE C)[23] and brain magnetic resonance imaging when neurological examination indicates Parkinsonism, ataxia or cognitive impairment (class I, LoE C)[22] might give an impression of overestimation, compromising the trust in cardiovascular guidelines as a whole.

Second, guideline committees should reflect on the goal they have with their recommendations. One-third of the recommendations of the current guideline on cardiovascular disease prevention in clinical practice concerns policy topics. These recommendations range from promoting healthy school diets (class I, LoE B) to increasing fuel taxes (class I, LoE C).[24] Other recommendations cover measures against drink-driving (class I, LoE B),[24] advice against binge drinking (class I, LoE C),[10] and requirements of resident training programmes (class I, LoE C).[25] Although the social engagement shown from these recommendations is positive, they do not add value in direct patient care. Moreover, these recommendations often contain political opinions nearly impossible to support with solid scientific evidence. The ESC describes guidelines as documents to help physicians weigh benefits and risks of diagnostic and therapeutic procedures,[26] and as these recommendations comprise neither diagnostic nor therapeutic procedures, they might be better served in separate policy guidelines.

In addition, it might be concluded that current LoEs are too crude, leaving uncertainties in the reliability of the underlying evidence. It is currently not possible to easily separate small trials from large observational studies (both LoE B) and expert opinion from small studies (both LoE C). The initiative of the ACC/AHA to indicate the origin of evidence underlying evidence levels (e.g, level C now indicates whether it is based on expert opinion or limited data)[27] should be more widely adopted to delineate the trustworthiness of LoEs supporting recommendations. Another strong initiative is the grading of recommendations assessment and evaluation (GRADE) framework commonly used in Cochrane reviews to standardise guideline recommendation classes and LoEs.[28] The GRADE framework provides guidance on advising standard key factors for recommendations, their classes and LoEs to improve the quality, consistency and reliability of recommendations.

More in general, efforts should be made to increase transparency on the guideline development process and policies. In an attempt to compare the evidence used in cardiovascular guidelines to the evidence found by the systematic literature search performed during the guideline development process, we discovered that such a comparison was impossible because the search strategies used for evidence identification were not published by the ESC. In addition, the governing policies noted that only peer-reviewed literature should be considered during the formal literature review.[26] Enforcing this policy, however, would conflict with the use of level C evidence in terms of expert opinion.

## Study limitations

Several limitations should be held into account for the present study. First, we tried to select recommendation type categories which would speak for themselves and would thus be open for as little debate as possible. Yet, recommendations were still interpreted and categorised by hand, exposing the categorisation process to possible misclassifications.

Second, the quality of evidence underlying the LoEs was not independently assessed in this study. It would therefore be possible that some LoEs falsely suggest a lower qualities of evidence than actually used. For instance, evidence categorised as LoE B could consist of (small) RCTs and of observational studies. Similarly, LoE C can contain expert opinion evidence and small studies such as case series, etc. Alternatively, evidence

standards could have been shifted in the years between different guidelines or interpreted differently by distinct guideline committees, consequently skewing the results of this study.

Third, this study only focused on the cardiovascular guidelines of the ESC. Although the review by Fanaroff et al.[4] showed similar distributions for ACC/AHA and ESC guidelines, one should be cautious in applying the salient findings of this study to studies on other medical and surgical societies. Nonetheless, the large category types and recommended actions might give an indication of the structural distribution of evidence in the current knowledge base in general.

## Conclusion

The evidence base underlying the cardiovascular guidelines of the ESC differs widely by recommendation types and recommended actions. Different reasons attribute to this high variability, including different evidence levels for therapeutic and diagnostic recommendations, different feasibility levels of trials for different interventions (e.g. open surgical vs. pharmacological interventions) and many off-topic/policy recommendations based on expert opinion. The cardiovascular research enterprise should focus on increasing evidence on diagnostic and open surgical topics by redesigning how evidence is generated, and by leveraging the increasing amounts of data available, for instance in a learning healthcare system. In addition, guideline authors should avoid issuing (off-topic) recommendations based on expert opinion/small studies as much as possible, and clinical research should focus on incentivising research on open surgical interventions.

### Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Dr van der Graaf reported being member of an independent ethical advisory committee to Sanofi. Dr Grobbee reported being a member of the committee of practice guidelines of the European Society of Cardiology. The other authors (van Dijk, Schuit, Groenwold & de Vries) reported no conflicts of interest.

### Funding

### References

1. American College of Cardiology. Real-world evidence: perceptions of cardiologists. *Cardiol Mag* 2018; 47: 28–29.
2. Vaucher J, Marques-Vidal P, Waeber G, et al. Population impact of the 2017 ACC/AHA guidelines compared with the 2013 ESH/ESC guidelines for hypertension management. *Eur J Prev Cardiol* 2018; 25: 1111–1113.
3. Pencina MJ, Navar-Boggan AM, D'Agostino RB, et al. Application of New Cholesterol Guidelines to a Population-Based Sample. *N Engl J Med* 2014; 370: 1422–1431.
4. Fanaroff AC, Califf RM, Windecker S, et al. Levels of evidence supporting American College of Cardiology/American Heart Association and European Society of Cardiology guidelines, 2008–2018. *JAMA* 2019; 27715: 1069–1080.
5. Tricoci P, Allen JM, Kramer JM, et al. Scientific evidence underlying the ACC/AHA clinical practice guidelines. *JAMA* 2009; 301: 831–841.
6. Bakris G and Sorrenntino M. Redefining hypertension – assessing the new blood-pressure guidelines. *N Engl J Med* 2018; 378: 497–499.
7. Remaley AT, Shamburek RD and Amar M. The new ACC/AHA cardiovascular risk guidelines: impact and controversies. *Clin Chem* 2014; 60: 1365–1367.
8. Shekelle PG. Clinical practice guidelines: what's next? *JAMA* 2018; 320: 757–758.
9. European Society of Cardiology, European Society of Hypertension. ESC/ESH Guidelines for the management of arterial hypertension. *Eur Heart J* 2018; 39: 3021–3104.
10. Windecker S. The European Society of Cardiology clinical practice guidelines. *Eur Heart J* 2017; 38: 2696–2697.
11. Alseiari M, Meyer KB and Wong JB. Evidence underlying KDIGO (kidney disease: improving global outcomes) guideline recommendations: a systematic review. *Am J Kidney Dis* 2016; 67: 417–422.
12. Feuerstein JD, Akbari M, Gifford AE, et al. Systematic analysis underlying the quality of the scientific evidence and conflicts of interest in interventional medicine subspecialty guidelines. *Mayo Clin Proc* 2014; 89: 16–24.
13. Feuerstein JD, Akbari M, Gifford AE, et al. Systematic review: the quality of the scientific evidence and conflicts of interest in international inflammatory bowel disease practice guidelines. *Aliment Pharmacol Ther* 2013; 37: 937–946.
14. Brito JP, Domecq JP, Murad MH, et al. The Endocrine Society guidelines: when the confidence cart goes before the evidence horse. *J Clin Endocrinol Metab* 2013; 98: 3246–3252.
15. Koh C, Zhao X, Samala N, et al. AASLD clinical practice guidelines: a critical review of scientific evidence and evolving recommendations. *Hepatology* 2013; 58: 2142–2152.
16. Wright JD, Pawar N, Gonzalez JSR, et al. Scientific evidence underlying the American College of Obstetricians and Gynecologists' practice bulletins. *Obstet Gynecol* 2011; 118: 505–512.
17. Poonacha TK and Go RS. Level of scientific evidence underlying recommendations arising from the National Comprehensive Cancer Network clinical practice guidelines. *J Clin Oncol* 2011; 29: 186–191.

18. Ergina PL, Cook JA, Blazeby JM, et al. Challenges in evaluating surgical innovation Europe PMC Funders Group. *Lancet* 2009; 374: 1097–1104.

19. Wartolowska K, Judge A, Hopewell S, et al. Use of placebo controls in the evaluation of surgery: systematic review. *BMJ* 2014; 348: 1–15.

20. Gelijns AC, Ascheim DD, Parides MK, et al. Randomized trials in surgery. *Surgery* 2009; 145: 581–587.

21. Ferguson RE and Fiore L. The Changing Face of Clinical Trials – Pragmatic trials. *N Engl J Med* 2016; 375: 454–463.

22. Brignole M, Moya A, de Lange FJ, et al. 2018 ESC Guidelines for the diagnosis and management of syncope. *Eur Heart J* 2018; 39: 1883–1948.

23. Regitz-Zagrosek V, Roos-Hesselink JW, Bauersachs J, et al. 2018 ESC Guidelines for the management of cardiovascular diseases during pregnancy. *Eur Heart J* 2018; 39: 3165–3241.

24. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J* 2016; 37: 2315–2381.

25. Neumann F-J, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J* 2018; 40: 87–165.

26. European Society of Cardiology. *Governing policies and procedures for the writing of ESC clinical practice guidelines*. European Society of Cardiology, 2017, Sophia Antipolis, Biot, France.

27. Halperin JL, Levine GN, Al-Khatib SM, et al. Further evolution of the ACC/AHA clinical practice guideline recommendation classification system: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2016; 133: 1426–1428.

28. Guyatt GH, Oxman AD, Kunz R, et al. GRADE: going from evidence to recommendations. *BMJ* 2008; 336: 1049–1051.