

Hot topics and trends in cardiovascular research

Diane Gal¹, Bart Thijs², Wolfgang Glänzel^{2,3}, and Karin R. Sipido^{1*}

¹Department of Cardiovascular Sciences, KU Leuven, Campus Gasthuisberg O/N1 704, Herestraat 49, B-3000 Leuven, Belgium; ²ECOOM, Department of Managerial Economics, Strategy and Innovation, KU Leuven, Naamsestraat 61, 3000 Leuven, Belgium; and ³Department Science Policy & Scientometrics, Library of the Hungarian Academy of Sciences, Arany János u. 1, 1051 Budapest, Hungary

Received 27 July 2018; revised 20 January 2019; editorial decision 2 April 2019; accepted 17 April 2019

Aims

Comprehensive data on research undertaken in cardiovascular medicine can inform the scientific community and can support policy building. We used the publication output from 2004 to 2013 and the 2014 references to these documents, to identify research topics and trends in the field of cardiovascular disease.

Methods and results

Text fragments were extracted from the titles and abstracts of 478 000 publications using natural language processing. Through machine-learning algorithms, these text fragments combined to identify specific topics across all publications. A second method, which included cross-references, assigned each publication document to a specific cluster. Experts named the topics and document clusters based on various outputs from these semi-automatic methods. We identified and labelled 175 cardiovascular topics and 20 large document clusters, with concordance between the approaches. Overarching, strongly growing topics in clinical and population sciences are evidence-based guidance for treatment, research on outcomes, prognosis, and risk factors. 'Hot' topics include novel treatments in valve disease and in coronary artery disease, and imaging. Basic research decreases its share over time but sees substantial growth of research on stem cells and tissue engineering, as well as in translational research. Inflammation, biomarkers, metabolic syndrome, obesity, and lipids are hot topics across population, clinical and basic research, supporting integration across the cardiovascular field.

Conclusion

Growth in clinical and population research emphasizes improving patient outcomes through novel treatments, risk stratification, and prevention. Translation and innovation redefine basic research in cardiovascular disease. Medical need, funding and publishing policies, and scientific opportunities are potential drivers for these evolutions.

Keywords

Cardiovascular research • Clinical care • Outcomes • Prevention • Innovation

Introduction

Current policies for public funding of health research increasingly focus on innovation, with a final goal to improve health outcomes.¹ To support policies, roadmaps are established, for example for diabetes² and respiratory³ diseases. In the USA, the joint Academies developed a document to guide national policy in health⁴ with a dedicated document for cardiovascular medicine⁵ that includes general directions for research. In Europe, building a roadmap for cardiovascular research is one of the tasks of the ERA-CVD network.⁶ Expert

opinion guides the exercise but a macro and global-level overview of past cardiovascular research can enrich the debate and strengthen the basis for recommendations. The breadth of cardiovascular research is astounding,⁷ with research undertaken across a variety of institutions and with each piece of research having its own scope/focus or topic. It is thus challenging to review and summarize all the research that has been undertaken.

Identifying all the relevant research is the first hurdle to overcome, then classifying or identifying topics of research is the next significant hurdle. Journal classification systems offer little assistance, as they are

* Corresponding author. Tel: +32 16 330815, Fax: +32 16 3 30699, Email: Karin.Sipido@kuleuven.be

© The Author(s) 2019. Published by Oxford University Press on behalf of the European Society of Cardiology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

not granular enough to identify more specific topics within broader fields. Thesauri or medical dictionaries, such as PubMed or the International Classification of Diseases (ICD), do not offer an overview of time-dependent changes in topics or changing concepts.

Identifying key topics using semi-automatic approaches based on text analysis is an alternative solution that takes advantage of recent developments in high-level informatics. As this is not reliant on a predefined classification, it may result in different outcomes. Various methods use natural language processing (NLP) to extract topics or clusters from text. For example, the bibliometric community has compared the results when varying methods are applied to a set of astronomy publications, focusing on the importance having topic expert input throughout the process.⁸ The recent CardioScape project analysed abstracts of 2476 research projects awarded 2010–12 as published by funding bodies. The authors assigned research project to topics, based on the abstract text, using a semi-automatic process that tested and trained the data to more quickly allocate abstracts to a topic than depending solely on expert review. They produced a detailed taxonomy or classification of cardiovascular research based on the list of topics of the European Society of Cardiology, creating a hierarchical list of over 600 topics.⁹

Here, we aim to identify topics in published cardiovascular research and their evolution between 2004 and 2013, assessing whether they have appeared, disappeared, or changed over time. In a comprehensive approach, we use a combination of existing methods for text mining, network analysis, and clustering, and further develop these tools to handle a large dataset of >400 000 publications.

Methods

In our study, we use two different and complementary approaches. A first one detects topics across the collection of publications, counting number of documents, and relations between topics. A second one maps document networks into clusters with an identifiable subject of research. These approaches are described here in brief, with more detail provided in the [Supplementary material online](#).

Data sources

The dataset includes the reference, abstract, address, and citation data for 478 006 cardiovascular publications from 2004 to 2013, including 2014 references to these documents, using an expert informed search strategy and references to core cardiovascular journals, as previously published.⁷ The documents span across >5000 journals, and include cardiovascular publications in leading general journals in medical and life sciences ([Supplementary material online, Table S1](#)). We obtained the data from Clarivate Analytics Web of Science Core Collection (WoS) through a custom data license held by ECOOM, KU Leuven.

Text pre-processing

We took all titles and abstracts of the above publications, and extracted the noun phrases (text fragments of various lengths) using the NLP framework developed at Stanford.¹⁰ [Supplementary material online, Figure S1](#) illustrates the subsequent data flow for the analysis.

Topic modelling

For this approach, we applied latent Dirichlet allocation (LDA)¹¹ to the above-mentioned text fragments from the titles and abstracts of all

publications. This LDA approach groups the text fragments to identify topics and allocates documents to topics. In this approach, a document contributes to several topics. Of note, general terms or terms that are used frequently across the majority of documents are filtered out as part of the methodology, resulting in groups of highly specific text fragments and, consequently, topics, as illustrated in [Supplementary material online, Figure S2](#).

At least three cardiovascular experts (listed in the Acknowledgements section) named each topic based on a set of the top 40 text fragments representing a topic. Further rounds of cross-review validated and consolidated the naming process. A final review of all topics ensured naming consistency across the topics and allowed for additional expert-based classification as clinical, basic, or population research.

We then calculated the number of documents that contributed to a topic, using probability analysis in LDA. Furthermore, we calculated the co-occurrence of topics in the publications, and visualized the outcome of this network analysis using VOSViewer (www.vosviewer.com).¹²

Document clustering

For this second approach, the dataset was reduced to two periods, and we analysed the cardiovascular publications from 2006 to 2008 and those from 2011 to 2013, separately. For each time period, we then calculated the similarities between documents based on the noun phrase text fragments from the titles and abstracts of all publications and based on the references in these publications, using adapted cosine calculations and a hybrid document clustering algorithm, as previously described.¹³ We then applied the Louvain¹⁴ community detection algorithm to identify clusters of similar documents. For this method, each document is only located in one cluster. Subsequently, we applied the DrL/OpenOrd algorithm¹⁵ to map and visualize the documents and clusters. We used R¹⁶ in a high-powered cloud-based parallelized computing environment for all operations.

We identified and described the core documents,¹³ the most common text fragments, as well as, the most highly cited documents and the most productive authors in each cluster, to name the clusters. For each document cluster, we identified the most highly representative topics from the LDA topic model.

Results

Evolution of cardiovascular topics—trends and ‘hot’ topics

We identified 175 topics, listed alphabetically in [Supplementary material online, Table S2](#). This list groups specific topics within areas such as atherosclerosis, coronary artery disease, arrhythmias, heart failure, and their evolution over time.

For a visual and comprehensive overview, we prepared a map of the topics and their interrelation, based on co-occurrence within publications using a network analysis ([Figure 1A](#)). This map identifies different categories of research: population (at the top, blue), clinical (left, green/yellow), and basic research (right, red). Large topics in each category define overarching interests such as *Evidence-guided-treatment* and *Outcomes and prognosis* in clinical research, and *Epidemiology of CVD and risk factors* in population research, topics that have seen large growth in numbers of publications since 2004 ([Figure 1B](#)). *Cell signalling and gene transcription* is a central topic for basic research, with modest growth ([Figure 1B](#)).

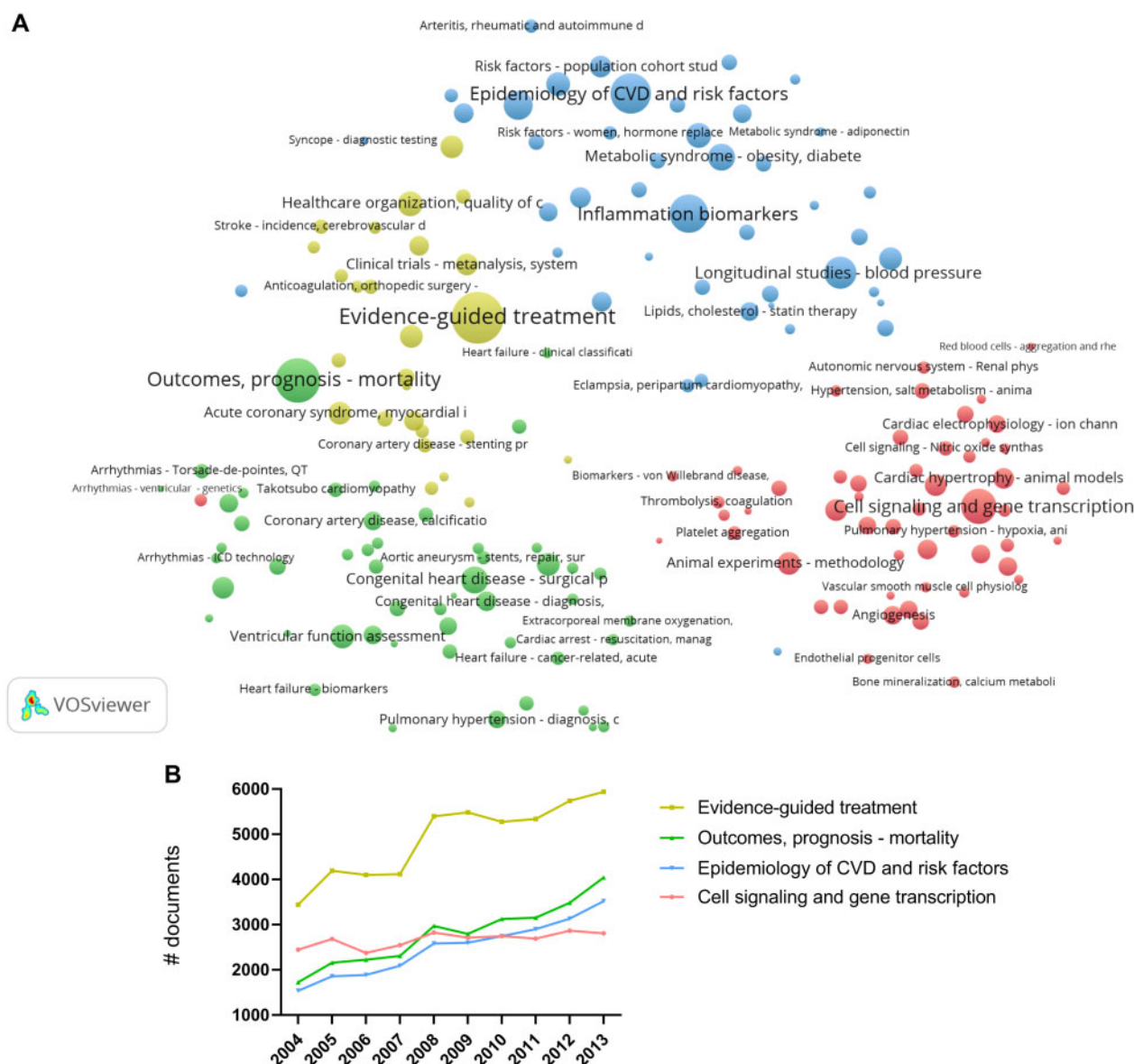


Figure 1 Main areas and organization of research focus. (A) Visual presentation of the topics in 2013 and how they relate to each other, based on how often the topics are included in the same publication. Each circle represents one topic and each group of topics is highlighted in a separate colour; the most similar documents and clusters are located closer to each other based on VOSviewer mapping. (B) Evolution of overarching topics.

More focused ‘hot’ topics that experienced a large growth in number of publications are presented in Figure 2.

In population research, risk factors with research on metabolic syndrome, lipids, diabetes, physical activity, and mental health are prominent. In clinical research, patient management after myocardial infarction (MI) and outside the hospital are leading topics, but the true ‘hot’ topic was aortic valve disease that saw a surge of interest, related to transaortic valve repair, starting 2008. Though still small in numbers, heart failure research and stem cells saw substantial growth. This last clinical topic complements the major hot

topics in basic research, on stem cells and cardiac repair and tissue engineering. In basic research, increasing translational output in metabolic syndrome and diabetes use mostly mouse models. Focused topics are organelle studies on mitochondria and endoplasmic reticulum.

Table 1 complements the fast growing topics of Figure 2 with additional leading 2013 topics. Most of these also have grown since 2004, but two topics, even if large, seem to have lost momentum, i.e. longitudinal studies on blood pressure, and basic research in cardiac electrophysiology.

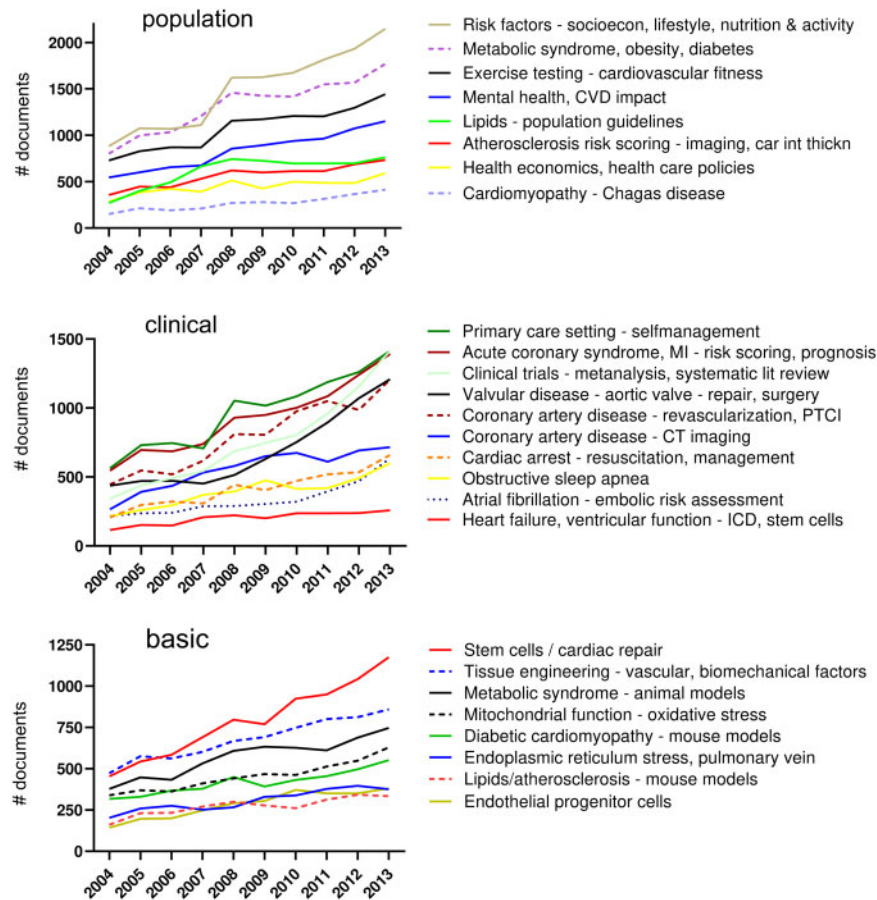


Figure 2 Topics with large growth. For population research, the eight topics that increased more than two-fold in volume are shown; for clinical research, 27 topics increased more than two-fold and 10 of these are presented; for basic research only two topics had more than a two-fold increase, and the top 8 growers are presented. Overarching topics are shown in Figure 1B.

Only four topics in clinical, and none in population research, saw a decrease, whereas seven topics in basic research saw a decline in output (Figure 3A). Across all topics, the growth in publication output, measured as the number of documents in 2013 divided by the number of documents in 2004, was significantly larger in clinical and population research topics than in basic research topics (Figure 3B).

When considering the overall output and growth of publications across the categories of population, clinical and basic research, the data suggest that the share of basic research publications is declining.

Document clusters define large research areas and trends

The size of topics represents the activity within each of these—documents contribute to more than one topic. In a complementary approach, we examined how documents group together based on the similarity of their text and of their references, whereby each document can belong to one cluster only, effectively dividing the total publication output into different areas. The hybrid clustering algorithm was applied to two datasets, i.e. the publications from 2006 to 2008 and 2011 to 2013.

In each period, 10 large clusters emerged, accounting for >90% of all documents.

To identify trends, we compare the two periods and examine the evolution over time (Figure 4). In the graph legends, emerging areas are marked by green triangle, decreasing ones with a red triangle. Risk scoring in the population and related patient management are the leading areas, growing over time (top position). In 2011–13, a large cluster emerges that relates to gene and stem-cell therapy, including research on inducible pluripotent stem cells. Documents within this cluster include research on ischaemic heart disease and arrhythmias. Haemodynamics and biomechanics are another emerging area that includes documents on atherosclerosis and vascular diseases such as aneurysms, but also heart failure and assist devices. Aortic valve disease is a newly defined area in 2011–13. Imaging also becomes very prominent as an area in its own right. Whereas in 2006–08, hypertension was a defined area, this is no longer identifiable in 2011–13.

For the last period, we also examined the structure and interrelation of clusters, using a graphical rendering, giving insight in the size, composition, and presence of subclusters (Figure 5).

Table 1 Large topics in 2013

Topic label	2004 (number of documents)	2013 (number of documents)
Clinical research		
Inflammation biomarkers	1545	2990
Congenital heart disease—surgical procedures	1209	2154
Healthcare organization, quality of care	746	1559
Coronary artery disease, cardiac surgery—peri-operative care	761	1503
Congenital heart disease—diagnosis, surgery, and treatment	739	1496
Ventricular function assessment	824	1464
Basic science		
Inflammation	951	1304
Animal experiments—methodology	1020	1293
Oxidative stress—antioxidants	854	1282
Cardiac hypertrophy—animal models	696	1139
Cardiac electrophysiology—ion channels, calcium homeostasis	1071	1109
Population research		
Longitudinal studies—blood pressure	1929	2196
Cholesterol, PCOS, obesity, and risk	698	1471
Risk factors—diabetes & hypertension	739	1383
Risk factors—population cohort studies	513	1268

PCOS, polycystic ovary syndrome.

In this force-directed DrL graph layout, the documents and clusters are mapped to minimize the distance between the most similar documents and maximize the distance between non-linked documents. This produces a two-dimensional co-ordinate layout where the documents closest to each other share the most similarities since they share common text fragments and references. Conversely, documents and clusters on the edges of the graph have the least similarity to other documents or clusters.

Cluster 2 on gene and stem cells is dense and separate, yet touches and interacts with Cluster 5 [acute coronary syndrome (ACS) and MI]. Cluster 9 on imaging is spread out in subclusters at different locations, including one near Cluster 5 (ACS and MI), and one near Cluster 4 (heart failure). Cluster 8 (arrhythmias) is also split with one part closer to heart failure, another to anticoagulation and atrial fibrillation.

Further naming the subclusters is presently beyond reach, as it would require a lot of expert input and resources. However, linking the clusters and the topics adds granularity to the larger research areas and provides internal methodological validation of the cluster naming.

Table 2 presents the most highly associated topics in the ten largest document clusters in each period. Overall, agreement with the LDA topics is high and provides more detail on the research contained in

the clusters. E.g., the cluster 'Haemodynamics' is now showing different areas of focus, i.e. in congenital disease, aortic, and valvular diseases; the topic 'Arrhythmias' is more populated with device research in the second time period compared to the first.

Discussion

The method for identification of topics in cardiovascular publication output allowed the visualization and evaluation of trends in cardiovascular research. Over a 10-year period significant shifts occur.

Identification of cardiovascular research topics through natural language processing

In cardiovascular research, topics are generally predefined in a taxonomy that can be hierarchical and/or matrix structured. The CardioScape project approach (see Introduction section) was well suited to its purpose of the analysis of 2476 project abstracts in a single time period and using an existing taxonomy has the advantage of recognizable areas of research. The bottom-up approach used here lent itself well to analysis of much larger numbers of documents and generated a topic list that represents the interests from the community during the period under study.

A recent study by the WHO working to identify cardiovascular disease research output from random sets of publications from PubMed required a significant amount of expert-based review of only a small proportion of the published articles.¹⁷ The current approach was more comprehensive in coverage of the field, but despite reliance on advanced automated analysis, experts still had an important role in interpreting and linking concepts to validate the results.

In the current naming of topics and clusters, experts frequently used terms that connect to a classic hierarchical list in the field, including major diseases, and recognizing clinical, population, and basic discovery research. Nevertheless, the approach uncovered specific emerging areas of research such as transcatheter aortic valve implantation (TAVI), topics consistent with broad trends, such as risk stratification and evidence-based guidance, and innovation (gene and stem cell research). Some of these terms would not appear in a classic taxonomy and thus the NLP approach offers novel insights.

The present study was not attempting to classify all research but to capture and identify the most common and evolving topics over time in the cardiovascular field by using a comprehensive set of cardiovascular publications across some 5000 journals.

Emphasis on improving clinical care and risk assessment

The most represented and fast growing topics across the documents are evidence-based guidance for treatment and research on outcomes and prognosis. These result underscore the attention given to guidelines and evidence based medicine (EBM).^{18–23} Part of this research is likely to represent the large number of clinical trials taking place in the cardiovascular field,²⁴ which over time have had a significant effect on the reduction of mortality from CVD due to establishing the effectiveness and safety of a number of drugs and medical interventions in cardiovascular disease.²⁵ The presence of policy

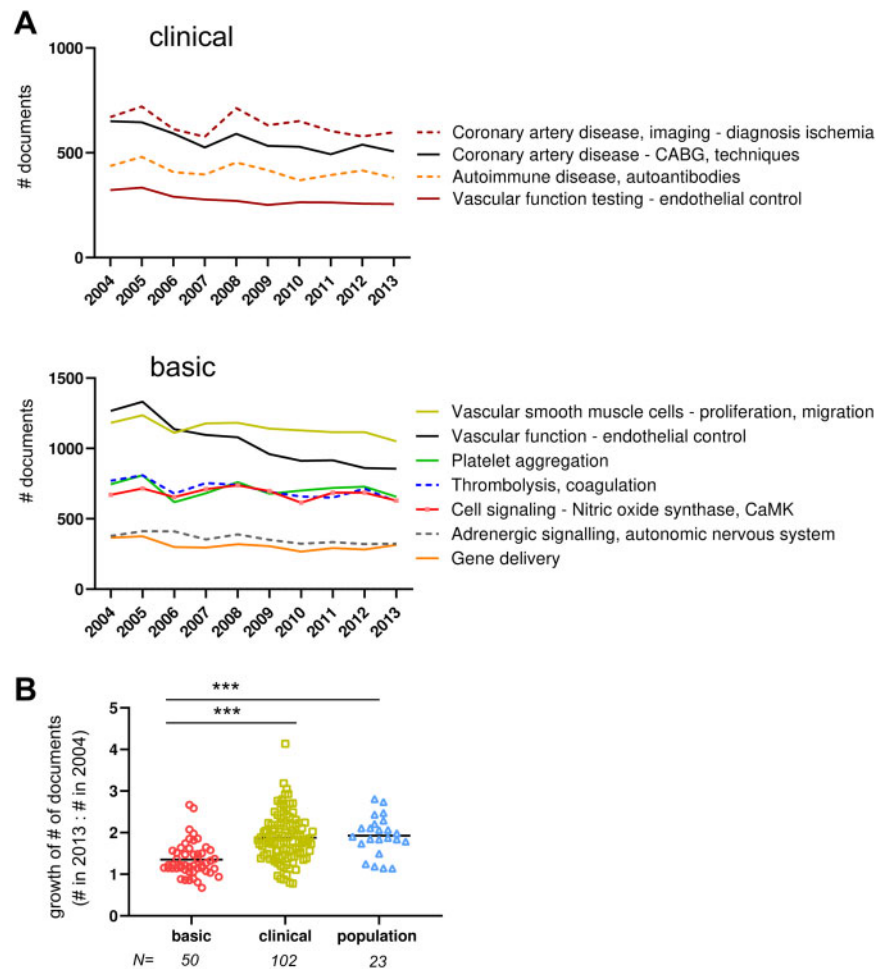


Figure 3 Unequal growth of research output across categories. (A) Topics that saw a decrease of >5%, i.e. 4/102 clinical and 7/50 basic research topics. (B) Average growth in each category. Each dot presents a topic; the values are the fractional growth, i.e. the number of documents in 2013 divided by the number of documents in 2004. Kruskal–Wallis followed by Dunn's test for multiple comparisons; *** $P < 0.0001$ basic vs. clinical and vs. population.

related topics, such as the topics on quality of care and health economics likewise supports the focus on implementation research and a shift of focus from reducing acute mortality to care in chronic disease.

Growth of research on risk factors emphasizes the importance of preventative medicine, evident in both the topics analysis and the document cluster analysis. However, some specific blood pressure studies declined over time, perhaps reflecting the change in focus on the single risk factor of 'blood pressure' to a multivariable spectrum and newly identified risk factors. We have also previously shown that hypertension has moved more closely to clinical cardiovascular research over time.²⁶

Smaller topics illustrate crosstalk with non-cardiovascular diseases, because of shared risk factors or common methods used in research or occurrence of cardiovascular complications. The latter is particularly evident in two topics that focus on cardiovascular complications in pregnancy and in cancer.

Innovation and translation in clinical and basic science

Major diseases such as ischaemic heart disease and arrhythmias, remain present over time but shifts can be seen. There is for example, a larger focus on atrial fibrillation, in particular embolic risk, on novel treatments, such as stem cells in heart failure, and transcatheter aortic valve interventions as a dominant element within the topic of valvular heart disease.¹⁹ Imaging is present in several topics but emerges as a cluster in its own right in the document analysis. Many of these changes are driven by technological innovation and translation.

Basic research as a whole saw its share decline, but with interesting shifts in content. Although the topic analysis and mapping identifies basic research topics as a category, there are complementarities across categories. Stem cell research, tissue engineering, and biomechanical factors saw rapid growth and are also present in clinical

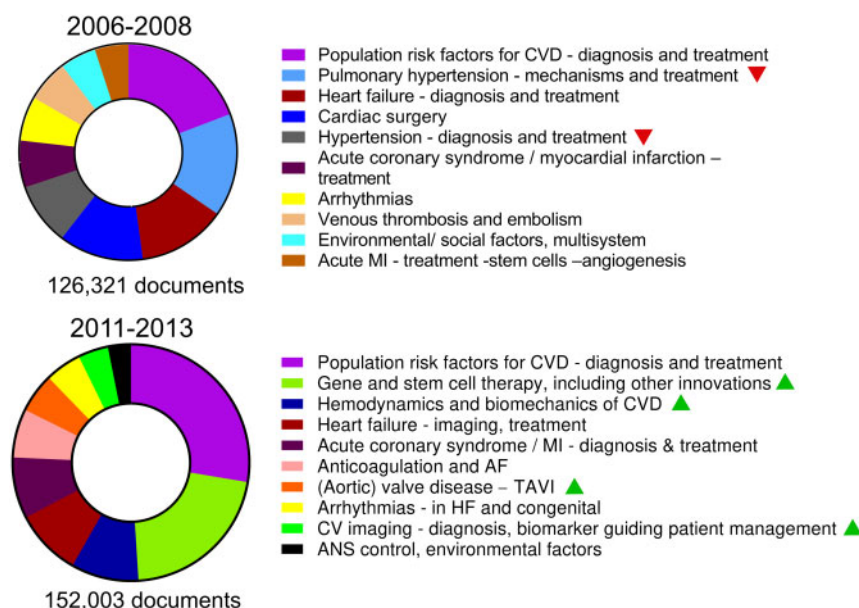


Figure 4 Distribution of document clusters in 2006–08 and in 2011–13. (A) In 2006–08, the 10 largest clusters represent 93% of the total publication output in this period. (B) In 2011–13, the 10 largest clusters represent 92% of the total publication output in this period. The colour codes for similar clusters are maintained across the periods. However, some clusters are present in only one period. The clusters are arranged by size, reading clockwise from the top, and the legends arranged accordingly. Red triangles mark clusters that disappeared and green triangles emerging clusters.

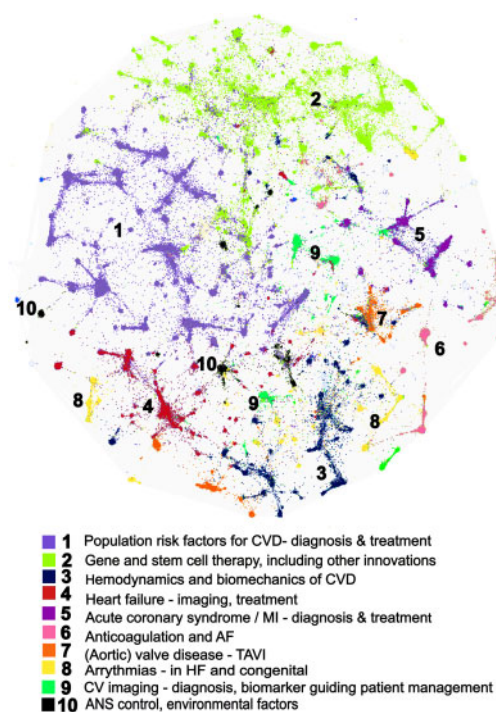


Figure 5 Document clusters' map 2011–13. A visual presentation of documents in clusters and subclusters: the most similar documents and clusters are located closer to each other, based on the DrL two-dimensional mapping layout technique.

topics. This also applies to inflammation and diabetes. Animal models for disease are rapidly growing topics consistent with growth of translational research.

An analysis of the countries of authorship of the publications in the emerging clusters of discovery research shows that the USA leads in the number and share of publications (30%+), followed mostly by Germany, or the UK or Italy. However, for the large document cluster on genes and stem cells in 2011–13, the second most productive country is China, contributing 17.5% of the publications in this cluster (Supplementary material online, [Figure S3](#)).

Interestingly, inflammation, biomarkers, metabolic syndrome, obesity, and lipids are hot topics with growing research output in population, clinical and basic research, indicating integration and crosstalk across the spectrum of cardiovascular research.

Drivers of change

Technology and opportunity-driven scientific interest, but also strategic choices and funding policies are likely to influence trends in research. CardioScope studied public and charity funding in the years 2010–12 and describes major investments in clinical research. Yet the share of publication output globally for clinical research appears to be substantially larger than the share of funding for clinical research reported in CardioScope. This could be explained by clinical research funded by other sources, such as industry or local funding, which are not included in the CardioScope analysis. Also, the present data represent global output. Major research investments in China, and the emphasis on clinical research in the USA, can contribute to some of the global trends.

Table 2 Cluster names and topics present within clusters

Cluster	Cluster name (n = number of documents)	LDA topics (ranked by contribution)
2006–08		
1	Population risk factors for CVD—diagnosis and treatment (n = 24 248)	<ul style="list-style-type: none"> • Epidemiology of CVD and risk factors • Metabolic syndrome—obesity and diabetes • Inflammation biomarkers • Lipids, cholesterol—statin therapy • Risk factors—socioeconomic, lifestyle, nutrition, and activity factors • Evidence-guided treatment
2	Pulmonary hypertension—mechanisms and treatment (n = 19 436)	<ul style="list-style-type: none"> • Cell signalling and gene transcription • Vascular function—endothelial control • Myocardial infarction—ischæmia/reperfusion injury • Vascular smooth muscle cells—proliferation and migration
3	Heart failure—diagnosis and treatment (n = 16 755)	<ul style="list-style-type: none"> • Imaging—echocardiography and cardiac haemodynamics • Ventricular function assessment • Arrhythmias—heart failure and ICD therapy • Evidence-guided treatment • Imaging—magnetic resonance imaging • Cardiac pacemakers—CRT
4	Cardiac surgery (n = 16 033)	<ul style="list-style-type: none"> • Congenital heart disease—diagnosis, surgery, and treatment • Congenital heart disease—surgical procedures • Aortic aneurysm—stents, repair, and surgery • Coronary artery disease, cardiac surgery—peri-operative care • Evidence-guided treatment • Coronary artery disease—CABG technique
5	Hypertension—diagnosis and treatment (n = 11 776)	<ul style="list-style-type: none"> • Hypertension—antihypertensive therapy • Hypertension—BP measurement and diagnosis • Evidence-guided treatment • Vascular disease and arterial stiffness • Epidemiology of CVD and risk factors
6	Acute coronary syndrome/myocardial infarction—treatment (n = 8792)	<ul style="list-style-type: none"> • Coronary artery disease—stents DES—complications • Acute coronary syndromes—antiplatelet treatment • Acute coronary syndrome, myocardial infarction, and STEMI—PTCI and angioplasty • Coronary artery disease—revascularization and PTCI • Evidence-guided treatment • Coronary artery disease—stenting procedures • Acute coronary syndrome and MI—risk scoring and prognosis
7	Arrhythmias (n = 8329)	<ul style="list-style-type: none"> • Atrial fibrillation—ablation treatment • Cardiac electrophysiology—ion channels and calcium homeostasis • Cardiac electrophysiology—action potential and conduction • Arrhythmias—Torsade-de-pointes, QT prolongation, and environmental risk • Arrhythmias—conduction disorders and ECG • Arrhythmias—arrhythmogenic RV cardiomyopathy and long QT—genetics
8	Venous thrombosis and embolism (n = 7883)	<ul style="list-style-type: none"> • Thrombolysis, coagulation • Anticoagulation, orthopedic surgery—NOAC • Cancer—coagulation disorders, LV outflow obstruction • Deep vein thrombosis—complications and treatment • Evidence-guided treatment • Cerebrovascular disease—cognitive dysfunction, risk factors

Continued

Table 2 Continued

Cluster	Cluster name (n = number of documents)	LDA topics (ranked by contribution)
9	Environmental/ social factors, multisystem (n = 6753)	<ul style="list-style-type: none"> • Autonomic nervous system—Heart rate variability • Autonomic nervous system—Renal physiology, blood pressure • Obstructive sleep apnoea • Blood pressure—regulation and autonomic nervous system • Longitudinal studies—blood pressure • Mental health and CVD impact • Low birth weight • Syncope—diagnostic testing
10	Acute MI—treatment—stem cells –angiogenesis (n = 6316)	<ul style="list-style-type: none"> • Angiogenesis • Stem cells/cardiac repair • Cardiovascular development • Tissue engineering—vascular and biomechanical factors • Myocardial infarction—animal studies • Cell signalling and gene transcription
2011–13		
1	Population risk factors for CVD—diagnosis and treatment (n = 42 024)	<ul style="list-style-type: none"> • Epidemiology of CVD and risk factors • Risk factors—socioeconomic, lifestyle, nutrition, and activity factors • Evidence-guided treatment • Inflammation biomarkers • Metabolic syndrome—obesity and diabetes
2	Gene and stem-cell therapy, including other innovations (n = 32 448)	<ul style="list-style-type: none"> • Cell signalling and gene transcription • Stem cells/cardiac repair • Angiogenesis • Vascular smooth muscle cells—proliferation and migration
3	Haemodynamics and biomechanics of CVD (n = 14 033)	<ul style="list-style-type: none"> • Congenital heart disease—diagnosis, surgery, and treatment • Aortic aneurysm—stents, repair, and surgery • Congenital heart disease—surgical procedures • Heart failure—assist devices • Evidence-guided treatment • Vascular function and remodelling—biomechanical factors
4	Heart failure—imaging, treatment (n = 13 876)	<ul style="list-style-type: none"> • Imaging—echocardiography and cardiac haemodynamics • Pulmonary hypertension—diagnosis and causes • Ventricular function assessment • Evidence-guided treatment • Imaging—cardiac echocardiography technique • Exercise testing—cardiovascular fitness
5	Acute coronary syndrome/myocardial infarction—diagnosis and treatment (n = 12 695)	<ul style="list-style-type: none"> • Coronary artery disease—stents DES—complications • Coronary artery disease—revascularization and PTCI • Acute coronary syndrome, MI, and STEMI—PTCI and angioplasty • Acute coronary syndromes—antiplatelet treatment • Coronary artery disease—stenting procedures • Evidence-guided treatment • Acute coronary syndrome and MI—risk scoring and prognosis
6	Anticoagulation and AF (n = 10 140)	<ul style="list-style-type: none"> • Atrial fibrillation—ablation treatment • Anticoagulation, orthopaedic surgery—NOAC • Evidence-guided treatment • Anticoagulation/thrombosis treatment and complications • Atrial fibrillation—embolic risk assessment • Stroke—treatment and outcome

Continued

Table 2 Continued

Cluster	Cluster name (n = number of documents)	LDA topics (ranked by contribution)
7	Aortic valve disease—TAVI (n = 8188)	<ul style="list-style-type: none">• Aortic valve disease—repair and surgery• Coronary artery disease and cardiac surgery—peri-operative care• Valvular heart disease—mitral valve• Coronary artery disease—CABG technique• Congenital heart disease—surgical procedures• Evidence-guided treatment• Cardiopulmonary bypass—blood transfusion• Sepsis and endocarditis
8	Arrhythmias—in HF and ‘congenital’ (n = 7565)	<ul style="list-style-type: none">• Arrhythmias—conduction disorders and ECG• Arrhythmias—ICD technology• Cardiac pacemakers—CRT• Arrhythmias—arrhythmogenic RV cardiomyopathy and long QT—genetics• Arrhythmias—heart failure and ICD therapy• Cardiac electrophysiology—ion channels and calcium homeostasis• Arrhythmias—Torsade-de-pointes, QT prolongation, and environmental risk
9	CV imaging—diagnosis and ‘biomarker’ guiding patient management (n = 6267)	<ul style="list-style-type: none">• Coronary artery disease—CT imaging• Coronary artery disease and calcification—diagnosis and CT angiography• MRI• Coronary artery disease and cardiac imaging—diagnosis myocardial ischaemia
10	ANS control, environmental factors (n = 4767)	<ul style="list-style-type: none">• Autonomic nervous system—Heart rate variability• Obstructive sleep apnoea• Arrhythmias—Torsade-de-pointes, QT prolongation, and environmental risk• Longitudinal studies—blood pressure• Syncope—diagnostic testing• Blood pressure—regulation and autonomic nervous system

AF, atrial fibrillation; ANS, autonomic nervous system; BP, blood pressure; CABG, coronary artery bypass grafting; CRT, cardiac resynchronization therapy; CT, computed tomography; CV, cardiovascular; DES, drug-eluting stent; ECG, electrocardiogram; HF, heart failure; LV, left ventricular; NOAC, new oral anticoagulant; PTCL, percutaneous transluminal coronary intervention; RV, right ventricle; STEMI, ST elevated myocardial infarction.

The slower growth in basic science could reflect a slower growth in investment. This can be absolute or relative towards the increasing costs of advanced research methodology. Another reason could be editorial pressure for more comprehensive papers that may reduce quantity to the benefit of rich content in individual papers.

Finally, growing translational research may blur the boundaries between basic and clinical research and lead to an apparent slower growth in discovery research.

Policy perspectives

Policy development is a forward looking exercise. In health research, medical needs identified by health data and expert opinion, are an important consideration.²⁷ Past research output helps to identify areas that may need more investment. Research funders also use input from society.²⁸ When assessing current priorities in cardiovascular research for the Dutch²⁸ and British²⁹ Heart Foundations we can see that research into heart failure and arrhythmias are common across their top priorities. Focus on healthy lifestyles is a top priority in the Dutch Heart Foundation as well as in the US vision and strategic

agenda.^{4,5} At the macro-level, the data presented here indicate that some of the main issues presented in these research agendas are actively pursued but others less so.

Study limitations

Limitations of studying research topics have been addressed in the bibliometric field.⁸ The reliance of expert input is a limitation and potential source of bias that we tried to minimize by using mixed panels.

The current approach was not sufficiently granular to extract recent emerging topics that contain a limited number of documents. In addition, publication output is somewhat delayed vs. actual research and experts may be aware of ongoing research with still limited output. In this case, the method and dataset can be used to interrogate about specific developments (see [Supplementary material online, Table S3](#) for data on micro-RNA and personalized medicine).

As the data set ends in 2013, very recent developments are not covered. This relates to the methodological complexity. Web of Science data including 2014 references were available mid-2015, the cardiovascular publications dataset was complete in 2016 and

algorithms for analysis including re-iterative expert review required another 18 months. A similar time lag is seen in other studies that rely on data mining and processing.⁹ Congress abstracts could be considered as a source to identify emerging topics but have several limitations. They are of a different nature than papers and the scope of a congress shapes content of selected abstracts. We provide a complementary survey of 3000 abstracts from the 2018 congress of the European Society of Cardiology, illustrating the strong presence of clinical research at this event, within the topics of Clusters 1 and 3–7 of Table 2 (Supplementary material online, Figure S4). Two emerging topics were cardio-oncology and digital health, each representing however <25 abstracts.

In the present analysis, quality and impact of studies in a particular domain were not evaluated, though highly cited papers were part of the cluster identification. In their analysis of poorly cited papers covering 165 000 papers in 1997–2008, Ranasinghe *et al.*³⁰ noted the highest percentage of poorly cited papers in the clinical and population research category. Nevertheless, as they and others³¹ have noted, citations are not the only parameter to assess impact, in particular in clinical medicine.

Conclusions

Identification of leading research topics and trends illustrates the emphasis on improving clinical medicine, and the growing interest in risk stratification and preventive medicine. Translation and innovation re-define cardiovascular research. Linking the present data with the insights of the professional community and of funders and society, may contribute to the building of a future research roadmap.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Acknowledgements

The authors thank to the following experts for their review of the text fragments and input into the names of the topics: Dr Matthew Amoni, Dr Peter Haemers, Prof Sian Harding, Dr Frederik Helsen, Prof Gerd Heusch, Prof Tatiana Kuznetsova, Prof Tobias Op't Hof, Prof Frank Rademakers, Dr Sander Trenson, Dr Bert Vandenberk, and Dr Maarten Vanhaverbeke.

Funding

D.G. had a PhD Fellowship through KU Leuven.

Conflict of interest: K.R.S. is Past Editor-in-Chief of *Cardiovascular Research* (2013–17). W.G. is Editor-in-Chief of *Scientometrics*.

References

- Directorate-General for Research and Innovation. Europe's future: open innovation, open science, open to the world. *Reflections of the Research, Innovation and Science Policy Experts (RISE) High Level Group*. Brussels: European Commission; 2017. p. 228.
- Welcome to DIAMAP. DIAMAP—Roadmap for Diabetes Research in Europe. <http://www.diamap.eu/> (3 December 2013).
- European Respiratory Roadmap—Recommendations for the Future of Respiratory Medicine. <https://www.ersnet.org/publications/european-respiratory-roadmap> (22 January 2018).
- Committee on Global Health and the Future of the United States, Board on Global Health, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. Washington, DC: National Academies Press; 2017. p. 360.
- Fuster V, Frazer J, Snair M, Vedanthan R, Dzau V. The Future Role of the United States in Global Health: emphasis on Cardiovascular Disease. *J Am Coll Cardiol* 2017;**70**:3140–3156.
- ERA-Net on Cardiovascular Diseases—ERA-CVD. <http://www.era-cvd.eu> (22 January 2018).
- Gal D, Glänzel W, Sipido KR. Mapping cross-border collaboration and communication in cardiovascular research from 1992 to 2012. *Eur Heart J* 2017;**38**:1249–1258.
- Velden T, Boyack KW, Gläser J, Koopman R, Scharnhorst A, Wang S. Comparison of topic extraction approaches and their results. *Scientometrics* 2017;**111**:1169–1221.
- Pries AR, Naoum A, Habazettl H, Dunkel M, Preissner R, Coats CJ, Tornada A, Orso F, Van de Werf F, Wood DA, Van de Werf F, Wood DA, O'Kelly S, Craven J, Coats A, Sipido K, De Backer D, Wallentin L, Hasenfuss G, della Sala L, Leggeri I, Wood DA, Van de Werf F, Jaarsma T, Elliott P, Pries AR, Madonna R, Kjeldsen K, Maggioni AP, Franco OH, Hills S, Pugliese F, De Bacquer D. CardioScape mapping the cardiovascular funding landscape in Europe. *Eur Heart J* 2018;**39**:2423–2430.
- Chen D, Manning CD. A fast and accurate dependency parser using neural networks. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: The Association for Computational Linguistics; 2014. pp. 740–750.
- Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. *J Mach Learn Res* 2003;**3**:993–1022.
- van EN, Waltman L. Citation-based clustering of publications using CitNetExplorer and VOSviewer. *Scientometrics* 2017;**111**:1053–1070.
- Glänzel W, Thijs B. Using hybrid methods and 'core documents' for the representation of clusters and topics: the astronomy dataset. *Scientometrics* 2017;**111**:1071–1087.
- Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. *J Stat Mech Theory Exp* 2008;**2008**:P10008.
- Boyack KW, Klavans R. Creation of a highly detailed, dynamic, global model and map of science: creation of a highly detailed, dynamic, global model and map of science. *J Assoc Inf Sci Technol* 2014;**65**:670–685.
- RStudio Team. *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.; 2015.
- Myers L, Mendis S. Cardiovascular disease research output in WHO priority areas between 2002 and 2011. *J Epidemiol Glob Health* 2014;**4**:23–28.
- Neumann F-J, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, Byrne RA, Collet J-P, Falk V, Head SJ, Juni P, Kastrati A, Koller A, Kristensen SD, Niebauer J, Richter DJ, Seferovic PM, Sibbing D, Stefanini GG, Windecker S, Yadav R, Zembala MO; ESC Scientific Document Group. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J* 2019;**40**:87–165.
- Genereux P, Head SJ, Wood DA, Kodali SK, Williams MR, Paradis J-M, Spaziano M, Kappetein AP, Webb JG, Cribier A, Leon MB. Transcatheter aortic valve implantation 10-year anniversary: review of current evidence and clinical implications. *Eur Heart J* 2012;**33**:2388–2398.
- Collins R, Reith C, Emberson J, Armitage J, Baigent C, Blackwell L, Blumenthal R, Danesh J, Smith GD, DeMets D, Evans S, Law M, MacMahon S, Martin S, Neal B, Poulter N, Preiss D, Ridker P, Roberts I, Rodgers A, Sandercock P, Schulz K, Sever P, Simes J, Smeeth L, Wald N, Yusuf S, Peto R. Interpretation of the evidence for the efficacy and safety of statin therapy. *Lancet* 2016;**388**:2532–2561.
- Eerenberg ES, Kamphuisen PW, Sijpkens MK, Meijers JC, Buller HR, Levi M. Reversal of rivaroxaban and dabigatran by prothrombin complex concentrate: a randomized, placebo-controlled, crossover study in healthy subjects. *Circulation* 2011;**124**:1573–1579.
- Razzouk L, Farkouh ME. Imaging outcomes in cardiovascular clinical trials. *Nat Rev Cardiol* 2009;**6**:524–531.
- Brown ML, Gersh BJ, Holmes DR, Bailey KR, Sundt TM III. From randomized trials to registry studies: translating data into clinical information. *Nat Clin Pract Cardiovasc Med* 2008;**5**:613–620.
- Clinical Trials in Cardiology/Vascular Diseases. <https://www.centerwatch.com/clinical-trials/listings/therapeutic-area/1/cardiology-vascular-diseases/> (13 April 2018).
- Solomon SD, Pfeffer MA. The future of clinical trials in cardiovascular medicine. *Circulation* 2016;**133**:2662–2670.
- Gal D, Thijs B, Glänzel W, Sipido KR. A changing landscape in cardiovascular research publication output: bridging the translational gap. *J Am Coll Cardiol* 2018;**71**:1584–1589.

27. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, Howells DW, Ioannidis JPA, Oliver S. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;**383**:156–165.
28. Onderzoeksagenda Hartstichting. <https://www.hartstichting.nl/wetenschappelijk-onderzoek/onderzoeksagenda-hartstichting> (13 April 2018).
29. British Heart Foundation Research Strategy 2015–2020. <http://extras.bhf.org.uk/research-strategy/site/index.html> (13 April 2018).
30. Ranasinghe I, Shojaaee A, Bickdeli B, Gupta A, Chen R, Ross JS, Masoudi F, Spertus JA, Nallamothu BK, Krumholz HM. Poorly cited articles in peer-reviewed cardiovascular journals from 1997–2007: analysis of 5-year citation rates. *Circulation* 2015; **131**:1755–1762.
31. Nallamothu BK, Lüscher TF. Moving from impact to influence: measurement and the changing role of medical journals. *Eur Heart J* 2012;**33**:2892–2896.