CODE-EHR best practice framework for the use of structured electronic healthcare records in clinical research


Big data is central to new developments in global clinical science aiming to improve the lives of patients. Technological advances have led to the routine use of structured electronic healthcare records with the potential to address key gaps in clinical evidence. The covid-19 pandemic has demonstrated the potential of big data and related analytics, but also important pitfalls. Verification, validation, and data privacy, as well as the social mandate to undertake research are key challenges. The European Society of Cardiology and the BigData@Heart consortium have brought together a range of international stakeholders, including patient representatives, clinicians, scientists, regulators, journal editors and industry. We propose the CODE-EHR Minimum Standards Framework as a means to improve the design of studies, enhance transparency and develop a roadmap towards more robust and effective utilisation of healthcare data for research purposes.

In the context of ageing populations and increasing multimorbidity in all disease areas,1,3 large scale, real world data provide an opportunity to better understand the epidemiology of rare and common conditions, and to improve prevention strategies and treatment stratification.4 Tailored management for individual patients has become even more essential to constrain healthcare costs and provide patient centred care that can improve a patient’s quality of life and prognosis. Embedding controlled trials within the real world setting, either within registries or routine clinical practice, is now possible and could provide more generalisable results to the population at large.5

Health data science has undergone rapid development in the past decade, including the common adoption of electronic healthcare record (EHR) systems that condense clinical episodes into a set of coded, structured labels.6 However, concerns over quality, data privacy, transparency, and comparability of these systems have limited the use of the evidence generated with structured healthcare data. These issues have also restricted acceptance by regulators, reimbursement authorities, and guideline task forces. Despite the availability of numerous reporting standards, consensus has not been met on how to realise the Findable, Accessible, Interoperable, and Reusable (FAIR) principles7 in the context of structured healthcare data. Existing reporting checklists ask authors to indicate where in their paper particular

**SUMMARY POINTS**

Research using routinely collected structured healthcare data has the potential for major clinical impact but this requires a clear and transparent approach to describe data sources, linkage protocols, coding definitions, and validation of methods and results

A social license and public mandate are essential components of big data research that can provide societal benefit, addressing the concerns of participants, and ensuring data privacy and integrity

This paper describes the output of international stakeholder meetings for the use of structured healthcare data for research purposes, including patient representatives, clinicians, scientists, regulators, journal editors, and industry representatives

The CODE-EHR checklist provides a minimum standards framework to enhance research design and enable more effective use and dissemination of routine healthcare data for clinical research
design issues have been discussed. For example, STROBE (Strengthening the Reporting of Observational studies in Epidemiology) for observational studies,

RECORD (R-Eporting of studies Conducted using Observational Routinely-collected Data) for routinely collected health data,

and CONSORT-AI (Consolidated Standards of Reporting Trials-Artificial Intelligence) for artificial intelligence interventions. However, these checklists are often lengthy, and adherence does not relate to study quality or even the quality of transparency for that domain. Although checklists can benefit research quality, they are often used for box ticking to facilitate journal publication. In a study of radiology journals, only 15% (120/821) of surveyed authors used the reporting guideline when designing their study. With a proliferation of reporting checklists for every scenario, authors and readers of such reports are increasingly confused about the value of these checklists. As of 14 February 2022, 488 reporting checklists were registered with EQUATOR (Enhancing the QUAlity and Transparency Of health Research) and 111 were in development.

In the case of observational and randomised clinical research using EHRs and other structured data, the source of data, its manipulation, and underpinning governance are of critical importance to extrapolating results. Clarity is needed from a broad stakeholder perspective, providing a quality framework to enhance the design and application of clinical research that increasingly depends on these crucial new sources of data. This article reflects the joint work of a wide range of international stakeholders with a remit to improve the use of structured healthcare data. The programme was coordinated by the European Society of Cardiology, a non-profit organisation of healthcare professionals, and the BigData@Heart Consortium, a public-private partnership funded by the European Union Innovative Medicines Initiative. Our aim was to navigate opportunities and limitations, and to develop a framework for a broad audience of global stakeholders across all disease areas. The CODE-EHR framework seeks to realise the exciting opportunity that digitisation of health data affords to increase efficiency of healthcare systems, and improve the lives and wellbeing of patients.

Stakeholder development of the CODE-EHR framework

A full range of stakeholders participated, including regulators (US Food and Drug Administration, European Medicines Agency), governmental agencies (European Commission, the UK National Institute for Health and Care Excellence, Innovative Medicines Initiative), leading medical journals (The BMJ, European Heart Journal, The Lancet, The Lancet Digital Health), and patient advocacy groups (European Heart Network, ESC Patient Forum), in addition to representatives from the pharmaceutical industry, payers, leading academic institutions, and professional societies (see acknowledgments). Development of the CODE-EHR framework was centred on two stakeholder meetings (7 July 2020 and 26 October 2020), consisting of presentations from key opinion leaders, followed by breakout sessions and plenaries to formulate statements on key topic areas. An iterative process with virtual work was used to achieve consensus positions, with a further meeting on 10 March 2022 to finalise this report.

We aimed to develop pragmatic advice for the use of structured healthcare data within trials and observational studies that is not dependent on particular diseases, and that meets the expectations of stakeholders and the general public. Our objectives were to provide new direction to this increasingly important field in medicine, thereby enhancing the value of routinely collected data to improve future patient wellbeing. Detailed text on the current state-of-the-art for research using healthcare data, in addition to key challenges and limitations, was developed by the stakeholder group, supported by a writing committee. See appendix 1 in which we address the need for common standards to appraise the digital landscape (e.g., coding systems and the vital aspect of linkage), expand on current and future opportunities for the use of structured healthcare data, and show how a social license can lead to co-creation of research with a public health benefit. The key challenges and pathways for improvement are outlined in figure 1, which presents the process of structured healthcare data from initial notation to their potential use to enhance research and subsequently improve clinical practice.

The output of the stakeholder meetings and iterative discussions were condensed into four core central themes: technical process and data stewardship; data security and privacy; publications using structured healthcare data; and addressing the needs of regulators, reimbursement authorities, and clinical practice guidelines. Key statements and advisories from the consensus meetings are summarised in table 1.

Patient and public involvement

The CODE-EHR consensus approach has benefited from patient and public engagement throughout the development process, including representation from the European Society of Cardiology Patient Council and the European Heart Network, an alliance of foundations and associations supporting patients and representing patient interests. We describe a potential method for engagement of the public in future research that can constructively benefit research using big data (fig 2).

CODE-EHR reporting framework

The path from structured healthcare data to clinical research output is complex. To support further development in a transparent way, stakeholder delegates reached consensus of the need for a set of minimum standards that authors could use as a tool to enhance design, reporting, and research output. The CODE-EHR Minimum Standards Framework presented in table 2 allows authors to report on how structured healthcare data were used in their research study
(either in patient identification, disease phenotyping, or outcome derivation). Preferred standards indicate high level attainment of quality and can be used as a tool to improve the future trajectory of research. The checklist was created through an iterative process based on the stakeholder proposals and covers five key areas of enhanced transparency: how and why coding was performed; the process of constructing and linking datasets; clear definitions of both diseases and outcomes; the approach to analysis, including any computational methods; and demonstrating good data governance.

The framework aims to improve the quality of studies using structured healthcare data and to give confidence in their use for clinical decision making. See appendix 2 for a step-by-step approach to completion of the CODE-EHR reporting checklist, with relevant best practice examples. We also present a detailed description of the workflow that led to the checklist in appendix 3. Form versions of the checklist are provided in appendix 4 (word version) and appendix 5 (pdf version).

**Discussion**

Technological progress has led to rapid evolution in health data systems with immediate impact in daily clinical practice. The potential for improving patient care and outcomes are clear, as are the challenges and limitations to achieving this objective. Big data analytics now support large scale (and cost efficient) clinical research, with trials based within registries or the EHR itself now heralding a new era in evidence generation. These processes can be further developed by an accompanying social license and upskilling of knowledge for all stakeholders. Co-creation and shared decision making with patients and the public is an important way to ensure appropriate data stewardship and privacy, leading to clinical impact through robust publications, regulatory decision making, and practice guidelines. In this paper, we have reported on a global multistakeholder process to develop a framework for researchers to use in the design and reporting of studies that include structured or coded healthcare data.

Digital health records are confusing for most researchers, with varying access to a myriad of different coding systems and classifications, and considerable differences across (and within) countries. Linkage of different health sources is often a core component of research based on structured healthcare data, and yet, this aspect is frequently overlooked when reporting such studies. Data privacy and the license for research can be severely compromised if linkage...
Table 1 | Output from the stakeholder consensus meetings

<table>
<thead>
<tr>
<th>Workshop theme</th>
<th>Key consensus statements and advisories</th>
</tr>
</thead>
</table>
| 1. Technical process and data stewardship | - Research using structured healthcare data conducted according to the FAIR data principles of findability, accessibility, interoperability, and reusability. Important considerations are transparency of who performed the coding, the coding system used, and the purpose of coding (reimbursement, diagnosis, etc.).
- Clear and consistent identification and description of the sources of EHR data. Code lists and phenotyping algorithms can be described in detail and published, ideally before a study commences (for example on a coding repository or open-source archive). The minimum data required to meet the definitions will depend on the use case and can be reported to enhance transparency, in addition to the rationale for why certain decisions were made (for example, why one code was chosen over another, or what the effect would be if data collection periods were changed).
- Validation at local, regional and global levels. Evidence demonstrating how algorithms have been externally validated, and also what quality assessment was performed on the research findings, for example on the accuracy, completeness and timeliness of the data. Data quality rules can be used to assess coded data and allow comparisons across institutions and countries.
- Reporting of methods used for data pre-processing and data linkage. This includes the methods used to assess the quality of linkage and the results of any data pre-processing and linkage (with provision of false positive and false negative rates, comparisons of linked and unlinked data, and any sensitivity analyses). A flow diagram showing the processes for cleaning and linking different coding sources and datasets can aid understanding of the study design.
- Reporting of the governance framework underpinning the study from a technical/data stewardship standpoint. This includes a clear purpose for data gathering and the parameters and time limit of consent, clear mechanisms for data processing (“what happens with my data”), and a description of what the data can and cannot be used for (i.e. the mandate given for research).
- Researchers and big data consortia have to be mindful that trustworthiness comes with the duty to act in ethically-responsible ways. This concerns two areas; first is the competence in data handling (meaning that systems are in place to ensure data protection and there is a framework of rules and regulations for data sharing), and second is what motivates the data analysis. Ongoing dialogue can ensure that public values continue to be aligned with the governance structures of health data research projects. Questions arise as to how to measure success at implementing public values into research, and what levels of public support are sufficient to grant a mandate for data usage.
- Complex organisational structures may be less important for this trust than is often asserted. Complicated rules and regulations may do more harm than good in establishing the conditions for public trust in big data health research to flourish, and as a result be counterproductive especially when a social license has not been adequately achieved.
- Embracing values such as transparency, reciprocity, inclusivity and service to the common good. These values can be embedded into the governance framework of big data health research. This calls for constructing a narrative that researchers and research consortia can be held accountable so that patients and the wider public are willing, and consistently willing, to place their trust in health research projects.
- Governance could be aided by developing a framework for accountability. This includes clear distinctions between anonymised, pseudonymised and aggregate data along with plain language explanation to participants and users, and discrimination between primary and secondary use of data sources.

2. Data security and privacy | - Working towards a new, sustainable mandate from the public and patients to use their health data may require moving away from abstract rules and regulations and towards more constructive governance, in which trust is a central concept. The trust of patients and the public in research institutions and in science is pivotal because of the liberties they give to researchers to use their data, which are the product of a social licence based on this trust.
- Gaining this trust would benefit from understanding what society and stakeholders expect from scientists conducting health data research, with engagement of stakeholders from the concept stage. Co-creation of data governance based on inclusion of patient/public communities and dialogue with researchers is crucial for ethical and sustainable governance, and to translate expectations into scientific research and scientific output.

3. Publications using structured healthcare data | - Accountability for the source of data and how the data have been collected (traceability). As with data security, a framework of accountability would enable editorial teams in medical journals to be aware of the technical processes prior to data analysis.
- Sharing of data, codes and algorithms used to analyse datasets. Similar to the requirement for pre-registration of clinical trials and pre-publication of protocols, journals could restrict publication where the coding within a study is not shared.
- Demonstration of data validity and robust analysis. The FDA and EMA already suggest independent checking or accreditation of data sources; this accreditation could be provided to editors to increase their confidence in data quality.
- Balancing the speed of publication against requirements for data validation. Prompt publication (for example of results with immediate public health implication) needs to be balanced against validation of data sources to ensure authenticity.
- Scientific advice committees with experts in big data analytics to aid journal editorial teams. The skill-set required in editors and reviewers for studies using structured healthcare data is not the same as having statistical or clinical trials experience; expertise in EHR data and respective coding systems could add value to the journal review process.
- Widening gap between the knowledge of physicians and the advanced methodologies used in big data papers. Medical/graduate students and practising clinicians, as well as hospital managers and leadership, need training in health data management and analysis. This is important to build a digital workforce with increased capacity and capability to translate publications using new approaches to improve patient care.

4. Addressing the needs of regulators, reimbursement authorities and clinical practice guidelines | - EHR-based trials have the potential to generate reliable and cost-efficient results. Each type of trial and each type of clinical question is considered in an individual context, including under what circumstances a particular type of EHR process could assist in answering questions about a particular intervention, and with what limitations.
- Further research may help explore cases in which EHR studies produce valuable evidence, and when they might be flawed. This will generate confidence in regulators for future EHR studies, and for guideline taskforces to appropriately appraise evidence.
- Quality standards will help to ensure that the information recorded in EHR systems represents real events without bias. This will enable confidence that trials using EHRs can produce reliable results on efficacy and safety, and could include examination of the validity of both data sources and data analyses.
- Source data validation to report on appropriate computational phenotypes. This could be supported by an independent adjudication committee to examine a subset of the EHR and confirm outcome events. The use of AI techniques could facilitate larger validation studies by automated extraction of supporting text from clinical notations. Such validation exercises can be pre-registered, for example in the form of a Study-Within-A-Trial. Another possibility is for researchers to provide consented and anonymised gold standard cases to benchmark against, or for data from devices used to verify codes (such as lead fractures). The value of synthetic datasets for validation, which mimic real data, needs further exploration.
- Mixed model approaches to collect data on particular endpoints. May be valuable for situations where the EHR does not reliably collect relevant data. For example, where patients and/or clinicians are asked for information, or data are collected via wearable devices or telemonitoring. In some cases, parallel monitoring of patients alongside the EHR study may provide additional confidence (for example to identify serious unexpected adverse events). Technological advances in EHR systems will help, such as the ability to retrieve EHR data on a daily basis to support clinical trials.
- Taking advantage of the many real world data initiatives to support new research. Government agencies, regulators, charities and professional bodies have initiated programmes for better use of real-world data that can support further activity and dissemination.
is not secure; hence, our focus is on transparency about how data are coded and linked, and how these approaches are openly discussed and documented. The stakeholder consensus meetings highlighted this area as a key concern for future research, supported by evidence that very few studies have provided sufficient detail to understand the research process. The advent of registry and EHR-based randomised controlled trials reinforces the imperative to see improvements in these areas, and to define new concepts for quality research. With the development of robust analytics supported by machine learning algorithms, similar approaches have already been used to support artificial intelligence in healthcare.

A lack of transparency has a direct impact on the value of research using coded records, with issues arising for medical journals, regulators, clinical guideline writers, and more generally clinicians and the public. Bringing together the full range of these stakeholders, we aimed to take full advantage of recent technical developments to use structured healthcare data for research, to approach limitations directly, and to provide a framework across all medical fields where coded data can be used to improve patient care. A number of other overlapping themes emerged from the discussions, including the generation and retention of public trust and confidence, and the need for coherent plans to deal with data security failures. Forethought about dealing with the harmonisation of data and the requirement for embedded validation methods were highlighted as key factors for future successful research. Similarly, education and communication are crucial for patients, citizens and healthcare professionals to effectively use the results from structured healthcare data studies.

The covid-19 pandemic has illustrated the need for rapid access to routine healthcare data to guide and monitor clinical care, and a clinical trial infrastructure to allow for immediate deployment. The digitalisation of healthcare, in particular the use of EHRs, offered the clinical community a unique opportunity to develop a learning healthcare system that could efficiently address the effects of covid-19. For example, information about the relationship between covid-19 and cardiovascular disease through linked EHR data that has combined primary care data, hospital data, death records, and covid-19 testing in more than 54 million people. However, the pandemic also made clear the obstacles within various systems that restricted the sharing of data in almost real time that could direct care and help design clinical trials. Established governance, security, interoperability (system architecture that spans different EHR systems and healthcare providers), and phenotype definition, among other issues, limited access to routine EHR data especially in the first period of the pandemic.

The CODE-EHR framework is intended to complement available reporting checklists. Although existing checklists are aimed at transparency in the reporting of important methodological components of clinical research, the CODE-EHR framework is designed to ensure that a common set of minimum standards are applied across all research using structured healthcare data. This range includes observational studies and controlled trials, with the preferred standards giving the direction of research design for all future (Continued)
Table 2 | CODE-EHR framework: best practice checklist to report on the use of structured electronic healthcare records in clinical research

<table>
<thead>
<tr>
<th>Date of completion:</th>
<th>Study name:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Objective</td>
</tr>
<tr>
<td>1. Dataset construction and linkage</td>
<td>To provide an understanding of how the structured healthcare data were identified and used.</td>
</tr>
<tr>
<td>2. Data fit for purpose</td>
<td>To ensure transparency with the approach taken, with respect to coding of the structured healthcare data.</td>
</tr>
<tr>
<td>3. Disease and outcome definitions</td>
<td>To fully detail how conditions and outcome events were defined, allowing other researchers to identify errors and repeat the process in other datasets.</td>
</tr>
<tr>
<td>4. Analysis</td>
<td>To fully detail how outcome events were analysed and allow independent assessment of the authenticity of study findings.</td>
</tr>
<tr>
<td>5. Ethics and governance</td>
<td>To provide patients, who might or might not have given consent, and regulatory authorities the ability to interrogate the security and provenance of the data.</td>
</tr>
<tr>
<td>6. Coding manual</td>
<td>DOI of publication or website address: Date published:</td>
</tr>
<tr>
<td>7. Comments</td>
<td></td>
</tr>
<tr>
<td>8. Summary declaration</td>
<td>Choose one from: One or more minimum standards not met OR All minimum standards met</td>
</tr>
</tbody>
</table>

Directions for use:

Research team: To complete the checklist, authors will need to consider these points during the design of the research to ensure that coding protocols and coding manuals are pre-published. Where applicable, it is advisable that all five minimum standards are met for an individual research study, whether observational or a controlled trial. If any component is not applicable to the study, the corresponding author can indicate why this is the case in the comment box. This checklist can accompany the article as a supplementary file on submission to the journal, with the ability for readers to review responses. A comment on the meeting of standards in the text of the method section is suggested, eg. “this study meets all five of the CODE-EHR minimum framework standards for the use of structured healthcare data in clinical research, with two out of five standards meeting preferred criteria (sad reference to this CODE-EHR paper, https://doi.org/10.1136/bmj-2021-069048)”; OR “this study meets four out of five of the CODE-EHR minimum framework standards for the use of structured healthcare data in clinical research, one of the five minimum standards was not met as coding schemes were not specified prior to analysis (sad reference to this CODE-EHR paper, https://doi.org/10.1136/bmj-2021-069048)” Note, easy to complete form versions of this checklist are available in appendix 4 (word version) and appendix 5 (pdf version), and also at https://www.escardio.org/bigdata.

Research appraisers (patients, clinicians, regulators, guideline task forces): Where applicable, it is advisable that all five minimum standards are met for the research study to be considered robust.

EHR studies. Additionally, the framework supports the wider implementation of good quality real world data research based on the FAIR data principles. Researchers are advised to use the checklist during the design phase of their study to ensure that key criteria for successful research and research impact are
already embedded. The process will aid journal editors, regulators, guideline writers, clinicians, and patients to better appreciate the underpinning value, and also the limitations of the study. Dissemination plans for CODE-EHR include discussion with journals to request authors complete the checklist when submitting relevant research, attaining full registration with the EQUATOR network, and outreach via international digital health groups to engage their members and other relevant stakeholder organisations. After publication, the CODE-EHR framework will undergo a two year evaluation, including discussion with researchers using the approach, with a plan for iterative improvements to adapt to this rapidly developing field of medical research.

Conclusion

The CODE-EHR framework was designed by a multistakeholder panel to improve design and reporting of research studies using structured electronic healthcare data. Research using these data sources is a vital component of future healthcare evaluation and delivery and will take an increasingly important role in decisions by regulatory, governmental, and healthcare agencies, as well as clinicians and patients in every medical specialty. The CODE-EHR checklist asks for clarity on reporting and defines a set of minimum and preferred standards on the processes that underpin coding, dataset construction and linkage, disease and outcome definitions, analysis, and research governance. Iterative updates to this framework are expected to enhance research quality and value and to generate new pathways for impact using routinely collected healthcare data.

AUTHOR AFFILIATIONS

1Institute of Cardiovascular Sciences, University of Birmingham, Medical School, Birmingham, UK
2University Hospitals Birmingham NHS Foundation Trust and Health Data Research UK Midlands, Birmingham, UK
3Department of Cardiology, Division of Heart and Lungs, University Medical Centre Utrecht, University of Utrecht, Utrecht, Netherlands
4Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK
5Health Data Research UK and Institute of Health Informatics, University College London, London, UK
6Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany
7Department of Cardiology and Berlin Institute of Health Centre for Regenerative Therapies, German Centre for Cardiovascular Research (DZHK) partner site Berlin; Chanté Universitätmedizin Berlin, Germany
8Department of Cardiology, Oslo University Hospital, Ulleval, Oslo, Norway
9University of Oslo, Institute of Clinical Medicine, Oslo, Norway
10MRC Population Health Research Unit, Nuffield Department of Population Health, Oxford, UK
11Clinical Trial Service Unit and Epidemiological Studies Unit, University of Oxford, Oxford, UK
12University College London Hospitals NHS Trust, London, UK
13European Heart Network, Faculty of Population Health Sciences, University College London, London, UK
14Bayer AB, Stockholm, Sweden
15Division of Cardiovascular Medicine, John Radcliffe Hospital, University of Oxford NIH Oxford Biomedical Research Centre, Oxford, UK
16European Society of Cardiology, Sophia Antipolis, France
17Royal Brompton Hospital, Division of Guy’s St Thomas’ NHS Foundation Trust, London, UK
18School of Cardiovascular Medicine Sciences, King’s College London, London, UK
19Department of Cardiovascular and Pulmonary Sciences, Catholic University of the Sacred Heart, Rome, Italy
20European Heart Journal, Oxford University Press, University of Oxford, Oxford, UK
21Vifor Pharma, Glattbrugg, Switzerland and Ava AG, Zurich, Switzerland
22School of Nursing and Midwifery, Queen’s University Belfast, Northern Ireland
23Late Clinical Development, Cardiovascular, Renal and Metabolism (CVRM), Biopharmaceuticals RD, AstraZeneca, Gothenburg, Sweden
24Leeds Institute of Cardiovascular and Metabolic Medicine and Leeds Institute for Data Analytics, University of Leeds, Leeds, UK
25Leeds Institute for Data Analytics, University of Leeds, Leeds, UK
26Department of Cardiology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
27College of Medical and Dental Sciences, Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, UK
28National Health Care Institute (ZIN), Diemen, Netherlands
29Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, Netherlands
30Department of Clinical Neuroscience, Karolinska Institutet, Solna, Sweden
31Department of Neuroradiology, Karolinska University Hospital Stockholm, Stockholm, Sweden
32Data and Analytics Group, National Institute for Health and Care Excellence, London, UK
33Office of Cardiovascular Devices, US Food and Drug Administration, Silver Spring, MD, USA
34Barts Health NHS Trust and University College London Hospitals NHS Trust
35Centre for Molecular Cardiology, University of Zurich, Zurich, Switzerland
36Research, Education & Development, Royal Brompton and Harefield Hospitals, London, UK
37Faculty of Medicine, Imperial College London, London, UK
38European Society of Cardiology Patient Forum, European Society of Cardiology, Brussels, Belgium
39Central Michigan University College of Medicine, Midland, MI, USA
40Centre for Devices and Radiological Health, US Food and Drug Administration, Silver Spring, MD, USA
41Steinbeisser Project Management, Munich, Germany
42Data Science AI, Biopharmaceuticals RD, AstraZeneca, Gothenburg, Sweden
43Centre for Therapeutic Innovation, Cardiovascular and Metabolic Disease, Institut de Recherches Internationales Servier, Suresnes, France
44Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands
45The Hyve, Utrecht, Netherlands
46Hygeia, Mitera, Hospitals Hellenic Health Group, Athens, Greece
47Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht University, Utrecht, Netherlands
48The Hyve, Utrecht, Netherlands
49Hygeia, Mitera, Hospitals Hellenic Health Group, Athens, Greece
50European Heart Agency, European Society of Cardiology, Brussels, Belgium
51The BMJ, London, UK
52Rudolfstiftung Hospital, Vienna, Austria
53Department of Cardiology, Inselspital, University Hospital Bern, Bern, Switzerland
54Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK
55Department of Epidemiology, University Medical Centre Utrecht, Division Julius Centrum, Utrecht, Netherlands
We are grateful for the work of the CODE-EHR international consensus. Delegates and attendees of the meetings were: Regulatory agencies: European Medicines Agency, Netherlands (Karel Kruiz); US Food and Drug Administration, USA (John Conrato, Robert Kazmerski, Jose Pablo Morales, and Ileana Pfa). Reimbursement authorities: National Health Care Institute, the Netherlands (Wim Goettsch), National Institute for Health and Care Excellence, UK (Adrian Jonas), Dental and Pharmaceutical Benefits Agency (TVE), Sweden (Niklas Hedberg). Medical Journal Editors: European Heart Journal, Switzerland (Filippo Crea and Thomas F Lüscher), The BMJ (UK (Wim Weber and Tiago Villanueva), The Lancet (UK (Stuart Spencer). The Lancet Digital Health, UK (Rupa Sarkar). Industry representatives: AstraZeneca (Martin Fredriksson and Mats Sundgren), Bayer (Andrea Derix, Gunnar Brobert, and Lothar Roessig), Servier (Benot Ty). The Hyde (Kees van Buchove), Vifor Pharma (Maureen Corrin). Funders: EU/EFPIA Innovative Medicines Initiative, Belgium (Colin Carroll), European Commission, Belgium (Ceri Thompson). Patient Advocacy: European Heart Network, Belgium (Birgit Beger, Sageslagt, and Marilena Vran); European Society of Cardiology Patient Forum, France (Paul McGreavy). Clinical Academic and Professional Societies: European Society of Cardiology, France (Barbara Casadei, Stephan Achenbach, and Valentina Tursin); European Heart Agency, Belgium (Panos E Vardas); Oslo University Hospital and University of Oslo, Norway (Dan Atar); University of Oxford, UK (Colin Baigent); University of Leeds, UK (Chris P Gale), Queen’s University Belfast, UK (Donna Fitzsimons); University Hospital Bern outside of this work. Charité-Universitätsmedizin Berlin, Germany (Stefan D Anker), Royal Brompton Hospital and Imperial College London, UK (Martin Cowie); University College London, UK (Imtata Banerjee, Harry Hemingway, RT Tom Lumber, Spies, of this work), University of Utrecht, Netherlands (Folkert W Asselbergs, Rick Groebbe, and Gislaeine Van Theil), University of Birmingham and University Hospitals Birmingham NHS Trust, UK (Dipak Kotecha and George V Gkoutos); University of Cambridge, UK (Angela Wood); Karolinska Institutet and Karolinska University Hospital Stockholm, Sweden (Martin Ingvar). Administrative: BigData@Heart Project Management, Germany (Carl Steinbeissner and Ana Petrova); European Society of Cardiology Project Management, France (Cinzia Cecarelli, Katija Baljevic, and Polyezniy Varami); Medical Writer, UK (Jennifer Taylor).

Contributors: DK and FWA are joint first authors with equal contribution, and act as guarantors of the work. Correspondence to FWA at f.w.asselbergs@umcutrecht.nl. Each subsequent listed author was involved in the CODE-EHR consensus process and wrote or revised sections of text and revised the final manuscript for intellectual content. The corresponding authors attest that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: The BigData@Heart project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 116074. This Joint Undertaking receives support from the European Union’s Horizon 2020 research and innovation programme and European Federation of Pharmaceutical Industries and Associations. The funders had no role in the considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests: All authors have completed the ICMJE uniform disclosure form (www.icmje.org/coi_disclosure.pdf) and declare: DK reports grants from EU/EFPIA Innovative Medicines Initiative (BigData@Heart 116074), during the study; grants from National Institute for Health Research (NIHR CDF-2015-08-074 RATE-AF, NIHR3 20820 DaReThINK, NIHR3 292974 D2T-NeurOvascular), British Heart Foundation (PG/17/157/31087, AN/18/2/34218 and FS/ IDRBT/21/210132), and European Society of Cardiology supported by educational grants from Boehringer Ingelheim, BMS-Pfizer Alliance, Bayer, Daichi Sankyo, Boston Scientific, the NIHR/University of Biomedical Research Centre, and British Heart Foundation/University of Birmingham Accelerator Award (STEER-EF NCT04-3964-18), Amoroxpharm and ICSSC San Raffaele/Menarini (Beta-blockers in Heart Failure Collaborative Group NCT00083244), outside of this work; and advisory board personal fees from Bayer, Amomed, Protherics Medicines Development and Myocardia, outside of this work. FWA reports grants from IMI BigData@Heart, during the study. SDA reports personal fees from Abbott, Abbott, and Abbott Vascular; personal fees from Bayer, Boehringer Ingelheim, and Servier; personal fees from Cardiac Dimensions, Actimed, Astra Zeneca, Amgen, Bioventix, Janssen, Respica, V-Wave, Brahms, Cordis, and Occultech, outside of this work. CB reports grants from Medical Research Council, Boehringer Ingelheim, and NIHR, outside of this work. AB reports grants from Astra Zeneca, outside of this work. BB reports grants from European Commission, during the study.

GB reports grants from European Commission (IMI project support), during the study, and other from Bayer AB, outside of this work. BC reports non-financial support from Roche Diagnostics and rHythm, outside of this work. CC reports grants from European Commission, during the study. MR reports personal fees from AstraZeneca, outside of this work. FC reports personal fees from Amgen, Astra Zeneca, Servier, and BMS; other from GlyCardinal Diagnostics, outside of this work. MC reports personal fees from Vifor Pharma, during the study; and personal fees from Vifor Pharma, outside of this work. AD reports other from Bayer AG, outside of this work. MF reports other from AstraZeneca, and 6 reports personal fees from Amgen, AstraZeneca, Servier, and personal fees from BMS, outside of this work. LD reports grants from World Economic Forum, Swedish Innovation Agency, and European Commission, during the study; and other collaboration with Frisq AB, outside of this work. SL reports grants from European Commission, during the study. TF reports grants from Abbott, Amgen, Novartis, Boehringer Ingelheim, Servier, Vifor, Sanofi, and AstraZeneca, and personal fees from Daichi Sankyo, Pfizer, and Menarini, outside of this work. LR reports other from Bayer AG, during the study and other from Bayer AG, outside this work. CS reports personal fees from Bayer AG, during the study and personal fees from Bayer AG, outside of this work; and advisory board personal fees from Servier, outside of this work. GoV reports grants from IMI BigData@Heart, during the study. PEV reports personal fees from Hygia Hospitals Group, HHG group, European Society of Cardiology, and outside of this work. Other from Novartis, outside of this work. CV is working as an editor at TheBMJ and Acta Médica Portuguesa and is vice president of the European Union of General Practitioners (UEMO). MV reports grants from European Commission, during the study. SW reports grants from Abbott, Amgen, AstraZeneca, BMS, Bayer, Bioventix, Boston Scientific, Cardinal Heart, CardioValve, CSL Behring, Daichi Sankyo, Edwards Lifesciences, Guebert, Infrare, Johnson & Johnson, Medicure, Medtronic, Novartis, Polares, OPPha Suisse, Pfizer, Regeneron, Sanoff-Aventis, Sinomed, Tenumo, and V-Wave, outside of this work, and SW serves as unpaid advisory board member or unpaid member of the steering or executive group of trials funded by Abbott, Abiomed, Amgen, Astra Zeneca, BMS, Boston Scientific, Biotronik, Cardiovalve, Edwards Lifesciences, MedAlliance, Medtronic, Novartis, Polares, Sinomed, V-Wave, and Xeltis, but has not received personal payments by pharmaceutical companies or device manufacturers. SW is also member of the steering or executive committee group of several investigator-initiated trials that received funding by industry without impact on his personal remuneration. SW is an unpaid member of the Pfizer Research Award selection committee in Switzerland and of the Women as One Awards Committee. SW is also member of the Clinical Study Group or the Deutsches Zentrum für Herz-Kreislaufr Forschung, the Women as One Awards Committee. SW is also member of the Clinical Study Group or of the Deutsches Zentrum für Herz-Kreislaufr Forschung, the Women as One Awards Committee. SW is also member of the Clinical Study Group or of the Deutsches Zentrum für Herz-Kreislaufr Forschung, the Women as One Awards Committee.

Provenance and peer review: Not commissioned; externally peer reviewed.

This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: http://creativecommons.org/licenses/by/4.0/.
